

ASTES

Advances in Science, Technology & Engineering Systems Journal

Special Issue

Innovation in Computing, Engineering Science & Technology

2021

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 Innovation in Computing, Engineering Science & Technology (2021) in the *Advances in Science, Technology and Engineering Systems Journal (ASTES Journal)* highlights a pivotal phase in the evolution of modern technological ecosystems. As the boundaries between computing and engineering disciplines continue to dissolve, innovation increasingly emerges from their intersection. This issue brings together a collection of contributions that showcase how advances in computational methods, intelligent systems, and engineering design are collectively shaping the future of science and technology.

A prominent theme throughout this issue is the rapid advancement of intelligent and data-driven technologies. Several papers explore the integration of artificial intelligence, machine learning, and data analytics into engineering applications, demonstrating their transformative impact on system optimization, predictive modeling, and decision-making processes. From smart manufacturing and automation to advanced communication systems and embedded technologies, the contributions reflect a growing reliance on computational intelligence to enhance efficiency, accuracy, and scalability in engineering solutions.

Equally significant is the emphasis on innovation in engineering science, particularly in the development of novel materials, sustainable systems, and high-performance infrastructures. Authors present cutting-edge research on energy-efficient technologies, environmentally conscious design practices, and resilient engineering frameworks that address contemporary global challenges. The convergence of computing capabilities with engineering innovation is evident in areas such as the Internet of Things (IoT), cyber-physical systems, and next-generation networks, where interdisciplinary collaboration plays a crucial role in advancing both theory and application.

The methodological diversity within this special issue further strengthens its contributions. The included studies employ a combination of theoretical modeling, simulation techniques, experimental investigations, and real-world case analyses. This integrative approach ensures that the research not only advances academic understanding but also provides practical insights for industry and policy implementation. Many of the works emphasize scalability and adaptability, offering solutions that can be extended across multiple domains and operational contexts.

Set against the backdrop of a rapidly digitizing world, the 2021 timeframe underscores the urgency of innovation in sustaining economic growth and societal development. The ongoing digital transformation, accelerated by global challenges, has reinforced the importance of robust, flexible, and intelligent systems. The research presented in this issue reflects this momentum, highlighting how emerging technologies are redefining traditional engineering paradigms and enabling new forms of collaboration and productivity.

The editorial team extends its sincere gratitude to the authors for their high-quality contributions and to the reviewers for their rigorous and constructive evaluations. Their collective efforts have ensured the integrity and impact of this special issue, reflecting a shared commitment to advancing knowledge in computing and engineering sciences.

This special issue represents a significant contribution to the discourse on technological innovation, illustrating how the integration of computing and engineering disciplines can drive meaningful progress. By presenting diverse perspectives and forward-looking research, it

provides a valuable platform for exploring the evolving landscape of science and technology and for inspiring future advancements across interconnected fields.

Guest Editor

Prof. Mohamed Abdelaziz Hassan Eleiwa

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Special Issue **December 2020**

CONTENTS

Building Health Monitoring using Visualization of Sound Features Based on Sound Localization
by Mitsuru Kawamoto and Takuji Hamamoto

Exploiting Domain-Aware Aspect Similarity for Multi-Source Cross-Domain Sentiment Classification
by Kwun-Ping Lai, Jackie Chun-Sing Ho and Wai Lam

Light Modulation Enhancement by using an Impedance Matching Scheme for a Subcarrier Multiplexed Light Transmitter
by Seiji Fukushima, Satoshi Yanagihara, Toshio Watanabe and Tsutomu Nagayama

Predicting School Children Academic Performance Using Machine Learning Techniques
by Radwan Qasrawi, Stephanny VicunaPolo, Diala Abu Al-Halawa, Sameh Hallaq and Ziad Abdeen

Lean Six Sigma Implementation in the Food Sector: Nexus between Readiness-Critical Success Factors
by Sarina Abdul Halim-Lim, Nurul Najihah Azalanzazllay, Anjar Priyono, Guven Gurkan Inan and Muhammad Iqbal Hussain

A Reconfigurable Stepped Frequency Continuous Wave Radar Prototype for Smuggling Contrast, Preliminary Assessment
by Massimo Donelli, Giuseppe Espa, Mohammedhusen Manekiya, Giada Marchi and Claudio Pascucci

Value Trace Problems for Code Reading Study in C Programming
by Xiqin Lu, Nobuo Funabiki, Htoo Htoo Sandi Kyaw, Ei Ei Htet, Shune Lae Aung and Nem Khan Dim

SyncBIM: The Decision-Making BIM-Based Cloud Platform with Real-time Facial Recognition and Data Visualization
by Chia-En Yang, Yang-Ting Shen and Shih-Hao Liao

Graph-based Clustering Algorithms – A Review on Novel Approaches
by Mark Hloch, Mario Kubek and Herwig Unger

Service Lifetime Loss Evaluation Method for Transformers Exclusively Serving Solar Power Plants
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

Cyber Incident Handling and the Perceptions of Learners on Cyber Incidents in South African Schools
by Naume Sonhera, Elmarie Kritzinger and Marianne Loock

Assessment of Transformer Cellulose Insulation Life Expectancy Based on Oil Furan Analysis (Case Study: South African Transformers)
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

The Design and Implementation of Intelligent English Learning Chabot based on Transfer Learning Technology
by Nuobei Shi, Qin Zeng and Raymond Shu Tak Lee

Service Robot Management System for Business Improvement and Service Extension
by Hideya Yoshiuchi and Tomohiro Matsuda

Plummeting Makespan by Proficient Workflow Scheduling in Cloud Environment
by Juhi Singh and Shalini Agarwal

Analysis of Grid Events Influenced by Different Levels of Renewable Integration on Extra-large Power Systems
by Christoph Rüeger, Jean Dobrowolski, Petr Korba and Felix Rafael Segundo Sevilla

Antecedents to Learners' Satisfaction with Serious Games: An Investigation Using Partial Least Square
by Ruben Chambilla, Daniel Tomiuk, Cataldo Zuccaro, Michel Plaisent and Prosper Bernard

Remote Patient Monitoring Systems with 5G Networks
by Antonio Casquero Jiménez and Jorge Pérez Martínez

Numeric Simulation of Artificial Antigravity upon General Theory of Relativity
by Yoshio Matsuki and Petro Ivanovich Bidyuk

Advanced Physical Failure Analysis Techniques for Rescuing Damaged Samples with Cracks, Scratches, or Unevenness in Delayering
by Yanlin Pan, Pik Kee Tan, Siong Luong Ting, Chang Qing Chen, Hao Tan, Naiyun Xu, Krishnanunni Menon, Hnin Hnin Win Thoungh Ma and Kyaw Htin

Model Reduction H? Finite Frequency of Takagi-Sugeno Fuzzy Systems
by Rim Mrani Alaoui, Abderrahim El-Amrani and Ismail Boumhidi

Hiragana and Katakana Minutiae based Recognition System
by Piotr Szymkowski, Khalid Saeed and Nobuyuki Nishiuchi

Modelling and Testing Services with Continuous Time SRML
by Ning Yu and Martin Wirsing

Intermediation in Technology Transfer Processes in Agro-Industrial Innovation Systems: State of Art
by Leidy Dayhana Guarin Manrique, Hugo Ernesto Martínez Ardila and Luis Eduardo Becerra Ardila

The Effect of Myocardial Fat's Thickness and Myocardial Impedance on Bipolar Radiofrequency Catheter Ablation Using Computer Simulation
by Yao Sun, Keijiro Nakamura and Xin Zhu

Innovations in Recruitment—Social Media
by Lucie Böhmová and Antonín Pavlíček

Emotion Mining from Speech in Collaborative Learning
by Nasrin Dehbozorgi, Mary Lou Maher and Mohsen Dorodchi

Improving of Heat Spreading in a SiC Propulsion Inverter using Graphene Assembled Films
by Sepideh Amirpour, Torbjörn Thiringer, Yasin Sharifi and Marco Majid Kabiri Samani

A Monthly Rainfall Forecasting from Sea Surface Temperature Spatial Pattern
by Prattana Deeprasertkul and Royol Chitradon

A Summary of Canonical Multivariate Permutation Entropies on Multivariate Fractional Brownian Motion
by Marisa Mohr and Ralf Möller

Associated Risk Factors for the Development of Extensively Resistant Pulmonary Tuberculosis in the First Level of Health Care: From a Public Health Perspective
by Mauricio Mamani, Mario Chauca, Edward Huamani and Richard Gonzales

Machine Learning Algorithms for Real Time Blind Audio Source Separation with Natural Language Detection
by Arwa Alghamdi, Graham Healy and Hoda Abdelhafez

The Gamification Design for Affordances Pedagogy
by Wilawan Inchamnan and Jiraporn Chomsuan

Vibration and Airflow Tactile Perception as Applied to Large Scale Limb Movements for Children
by Hung-Chi Chu, Fang-Lin Chao and Liza Lee

A Novel De-rating Practice for Distributed Photovoltaic Power (DPVP) Generation Transformers
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

Numeric Simulation on the Waves from Artificial Anti-gravity upon General Theory of Relativity
by Yoshio Matsuki and Petro Ivanovich Bidyuk

Hiding Information in DNA Sequence Data using Open Reading Frame Guided Splicing
by Amal Khalifa
Acoustic Scene Classifier Based on Gaussian Mixture Model in the Concept Drift Situation
by Ibnu Daqiqil Id, Masanobu Abe and Sunao Hara

A Hybrid NMF-AttLSTM Method for Short-term Traffic Flow Prediction
by Dejun Chen, Congcong Xiong, Li Guo and Ming Zhong

Development of Miniaturized Monolithic Isolated Gate Driver

by Hsuan-Yu Kuo and Jau-Jr Lin

Use of Unmanned Aerial Vehicles in Aircraft Inspection
by Andrej Novák, Martin Bugaj, Alena Novák Sedláčková, Branislav Kandera, Anna Stelmach and Tomasz Lusiak

iDRP Framework: An Intelligent Malware Exploration Framework for Big Data and Internet of Things (IoT) Ecosystem
by Osaretin Eboya and Julia Binti Juremi

A Survey of Big Data Techniques for Extracting Information from Social Media Data
by Carla Blank, Matthew McBurney, Maria Morgan and Raed Seetan

Mitigation of Nitrous Oxide Emission for Green Growth: An Empirical Approach using ARDL
by Hanan Naser and Fatema Alaali

Boltzmann-Based Distributed Control Method: An Evolutionary Approach using Neighboring Population Constraints
by Gustavo Alonso Chica Pedraza, Eduardo Alirio Mojica Nava and Ernesto Cadena Muñoz

Extraction of Psychological Symptoms and Instantaneous Respiratory Frequency as Indicators of Internet Addiction Using Rule-Based Machine Learning
by Hung-Ming Chi, Liang-Yu Chen and Tzu-Chien Hsiao

Initial Experiments using Game-based Learning Applied in a Classical Knowledge Robotics in In-Person and Distance Learning Classroom
by Márcio Mendonça, Rodrigo Henrique Cunha Palácios, Ivan Rossato Chrun, Diene Eire de Mello, Henrique Cavalieri Agonilha, Elpiniki Papageorgiou and Konstantinos Papageorgiou

A Case Study to Enhance Student Support Initiatives Through Forecasting Student Success in Higher-Education
by Ndiatenda Ndou, Ritesh Ajoodha and Ashwini Jadhav

Leveraging Energy Efficiency Investments: An Innovative Web-based Benchmarking Tool
by Filippos Dimitrios Mexis, Aikaterini Papapostolou, Charikleia Karakosta, Elissaios Sarmas, Diamantis Koutsandreas and Haris Doukas

A Task-based Paradigm for Promoting an Alternative Thinking Style in Teaching Mathematics
by Mikhail Rodionov, Zhanna Dedovets and Irina Akimov

Software Development Lifecycle for Survivable Mobile Telecommunication Systems
by Mykoniati Maria and Lambrinoudakis Costas

Low-Power Primary Cell with Water-Based Electrolyte for Powering of Wireless Sensors
by Dmitry Petrov and Ulrich Hilleringmann

Discover DaVinci: Blockchain, Art and New Ways of Digital Learning
by Marko Suvajdzic and Dragana Stojanovic

An Alternative Approach for Thai Automatic Speech Recognition Based on the CNN-based Keyword Spotting with Real-World Application
by Kanjanapan Sukvichai and Chaitat Utintu

Comparative Analysis and Modern Applications of PoW, PoS, PPoS Blockchain Consensus Mechanisms and New Distributed Ledger Technologies
by Caglar Arslan, Selen Sipahio?lu, Emre ?afak, Mesut Gözütok and Tacettin Köprülü

Performance of Vertical Axis Wind Turbine Type of Slant Straight Blades
by Hashem Abusannuga and Mehmet Özkaymak

Reading Acquisition Software for Portuguese: Preliminary Results
by Ana Sucena, Ana Filipa Silva, Cristina Garrido and Cátia Marques

Evaluation of Information Competencies in the School Setting in Santiago de Chile
by Jorge Joo-Nagata and Fernando Martínez-Abad

Segmentation of Stocks: Dynamic Dimensioning and Space Allocation, using an Algorithm based on Consumption Policy, Case Study
by Anas Laassiri and Abdelfettah Sedqui

Technique to Simulate an Oscillator Ensemble Represented by the Kuramoto Model
by Mark Gourary and Sergey Rusakov

Designs of Frequency Reconfigurable Planar Bow-tie Antenna Integrated with PIN, varactor diodes and Parasitic Elements
by Mabrouki Mariem and Gharsallah Ali

Real Time RSSI Compensation for Precise Distance Calculation using Sensor Fusion for Smart Wearables
by Kumar Rahul Tiwari, Indar Singhal and Alok Mittal

Analyzing the Application of Two Main Microcontrollers in Engineering Education – A Case Study of three IEEE Conferences Focusing on Education
by Arthur James Swart

Recognition of Emotion from Emoticon with Text in Microblog Using LSTM
by Juyana Islam, M. A. H. Akhand, Md. Ahsan Habib, Md Abdus Samad Kamal and Nazmul Siddique

Designing a Model of Consciousness Based on the Findings of Jungian Psychology
by Toshiki Watanabe and Hiroyuki Kameda

Impact Assessment of the Renewable Energy Sources Implementation in Bulgarian Single-Family Houses on the Greenhouse Gas by HOMER Pro Software
by Dilyana Gospodinova and Peter Dineff

Personalized Serious Games for Improving Attention Skills among Palestinian Adolescents
by Malak Amro, Stephanny VicunaPolo, Rashid Jayousi and Radwan Qasrawi

Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques
by Zhiyuan Chen, Howe Seng Goh, Kai Ling Sin, Kelly Lim, Nicole Ka Hei Chung and Xin Yu Liew

A Scheduling Algorithm with RTiK+ for MIL-STD-1553B Based on Windows for Real-Time Operation System
by Jong-Jin Kim, Sang-Gil Lee and Cheol-Hoon Lee

A Survey of FPGA Robotics Applications in the Period 2010 – 2019
by Dimitris Ziouzios, Pavlos Kilintzis, Nikolaos Baras and Minas Dasygenis

Devices and Methods for Microclimate Research in Closed Areas – Underground Mining
by Mila Ilieva-Obretenova

Quantum Secure Lightweight Cryptography with Quantum Permutation Pad
by Randy Kuang, Dafu Lou, Alex He and Alexandre Conlon

Personalized Clinical Treatment Selection Using Genetic Algorithm and Analytic Hierarchy Process
by Olena Nosovets, Vitalii Babenko, Ilya Davydovych, Olena Petrunina, Olga Averianova and Le Dai Zyonh

Factors Affecting the Decision of Selecting Banking to Save Money of Individual Customers – Experimental in Da Nang City
by Le Anh Tuan, Mai Thi Quynh Nhu and Nguyen Le Nhan

Identification of Genetic Variants for Prioritized miRNA-targeted Genes Associated with Complex Traits
by Isabella He, Zhaohui Qin and Yongsheng Bai

A Design of Anthropomorphic Hand based on Human Finger Anatomy
by Zixun He, Yousun Kang and Duk Shin

Application of Piecewise Linear Approximation of the UAV Trajectory for Adaptive Routing in FANET
by Kuzichkin Oleg R., Vasilyev Gleb S., Surzhik Dmitry I. and Kurilov Igor A.

A Recommendation Approach in Social Learning Based on K-Means Clustering
by Sonia Souabi, Asmaâ Retbi, Mohammed Khalidi Idrissi and Samir Bennani

Real-time Target Human Tracking using Camshift and LucasKanade Optical Flow Algorithm
by Van-Truong Nguyen, Anh-Tu Nguyen, Viet-Thang Nguyen, Huy-Anh Bui and Xuan-Thuan Nguyen

Efficient 2D Detection and Positioning of Complex Objects for Robotic Manipulation Using Fully Convolutional Neural Network
by Dominik Štursa, Daniel Honc and Petr Doležel

The Mediating Role of Entrepreneurial Orientation on the Knowledge Creation-Firm Performance Nexus: Evidence from Indonesian IT Companies
by Desman Hidayat, Edi Abdurachman, Elidjen and Yanthi Hutagaol

Numerical Analysis for Feature Extraction and Evaluation of 3D Sickness
by Kohki Nakane, Rentaro Ono and Hiroki Takada

Performance Evaluation of Convolutional Neural Networks (CNNs) And VGG on Real Time Face Recognition System
by Showkat Ahmad Dar and S Palanivel

A New Video Based Emotions Analysis System (VEMOS): An Efficient Solution Compared to iMotions Affectiva Analysis Software
by Nadia Jmour, Slim Masmoudi and Afef Abdelkrim

Enhanced Data Transportation in Remote Locations Using UAV Aided Edge Computing
by Niranjan Ravi and Mohamed El-Sharkawy

Performance Evaluation Reprogrammable Hybrid Fiber-Wireless Router Testbed for Educational Module
by Muhammad Haqeem bin Mohd Nasir, Wan Siti Halimatul Munirah binti Wan Ahmad, Nurul Asyikin binti Mohamed Radzi and Fairuz Abdullah

Factors Impacting Digital Payment Adoption: An Empirical Evidence from Smart City of Dubai
by Anas Najdawi, Zakariya Chabani and Raed Said

Analysis and Evaluation of Competitiveness in Medical Tourism Industry in Taiwan
by Yen-Hung Chen and Tin-Chang Chang

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 01-06 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Building Health Monitoring using Visualization of Sound Features Based on Sound Localization †

Mitsuru Kawamoto*,[1], Takuji Hamamoto[2]

[1]Human Augmentation Research Center, National Institute of Advanced industrial Science and Technology (AIST), Kashiwa, 277-0882, Japan

[2]Tokyo City University, Tokyo, 158-8557, Japan

ARTICLE INFO

Article history:
Received: 21 July, 2021
Accepted: 11 October, 2021
Online: 06 November, 2021

Keywords:
Understanding sound environment
Sound visualization technique
Sound features
Sound map
Monitoring the collapse of buildings

ABSTRACT

This paper describes what can be accomplished by understanding sound environments. Understanding sound environments is achieved by extracting the features of the sound and visualizing the features. The visualization is realized by converting the three features, namely, loudness, continuity, and pitch, into RGB values and expressing the sound with color, where the color is painted in the estimated direction of the sound. The three features can distinguish falling objects within a building and roughly estimate the direction of the generated sounds. The effectiveness of the proposed sound visualization was confirmed using the sounds of cans and stones falling in a building; hence, it is shown that the proposed visualization method will be useful for monitoring the collapse of buildings by sound.

1. Introduction

Techniques for understanding environments are used to obtain sensor-related information in environmental measurements. For example, in terms of monitoring buildings, the analysis of camera and accelerometer sensor data is used to understand the current state of the building and the differences from its past state (e.g., [1]). A camera can capture images of an entire building, and an accelerometer can detect building vibrations. However, the camera only detects events within its field of view and the accelerometer only detects vibrations in its surroundings. Further, these sensors can only be installed in the monitoring area, thereby providing limited coverage, since it is difficult and risky to install sensors under collapse risk. Alternatively, sounds can reach sensors such as microphones for detection. Moreover, if a microphone array is used as the sound sensor, the sound source direction can be estimated. Therefore, sound sensors can allow more flexible monitoring of buildings than cameras and accelerometers.

Considering the benefits of sound sensing, we propose a method for analyzing environmental sounds aiming to evaluate the difference from the past state of buildings, which is called building health monitoring, by extracting their features and estimating the direction of sound sources. As precursory sounds often occur before building collapse, such sounds may be detected and characterized by extracting sound features.

Aiming to perform building health monitoring, a visualization method for sound features based on sound localization to facilitate analysis has been introduced. In most cases, sound monitoring is simply performed by recognizing measured environmental sounds for applications, such as elderly people (e.g., [2], and references therein). On the other hand, the proposed method can provide awareness of changes in buildings by providing a visual representation of environmental sounds. The sound features of loudness, continuity, and pitch are considered. These features are quantified using spectrograms and chromagrams. Then, a visual representation of the sound features along with their estimated source position are obtained. The proposed visualization technique was evaluated using experimental sound signal data generated at a building in Gunkanjima (Hashima), Japan. Our method is a novel paradigm for sound monitoring and a novel contribution to building health monitoring.

*Corresponding Author: Mitsuru Kawamoto, Kashiwa II Campus, University of Tokyo, 6-2-3 Kashiwanoha, Kashiwa, Chiba 277-0882, +81 3 3599 8543 & m.kawamoto@aist.go.jp.

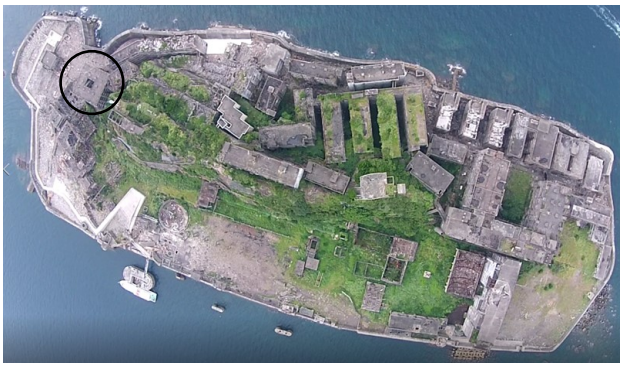
† A preliminary version of this paper was presented at DCOSS2020.

https://dx.doi.org/10.25046/aj060601

2. Materials and Methods

2.1. Data Acquisition

We evaluate the proposed visualization technique which can be used for building health monitoring, using measurements obtained from a sound sensor (microphone array). Figure 1 shows Gunkanjima (Hashima), a World Heritage site in Nagasaki, Japan. The microphone array is installed as shown in Figure 2 on the second floor of building No. 30 (enclosed in black circle of Figure 1(a)) of Gunkanjima. The method is applied to verify the feasibility of determining the building status by analyzing environmental sounds. Building No. 30 is the oldest reinforced concrete building in Japan and is at risk of collapsing. The microphone array is comprised of 16 microphones (each arrow in the righthand side of Figure 2 shows a cluster of four microphones).



(a)



(b)

Figure 1: (a) Bird's eye view and (b) cityscape of Gunkanjima, Japan

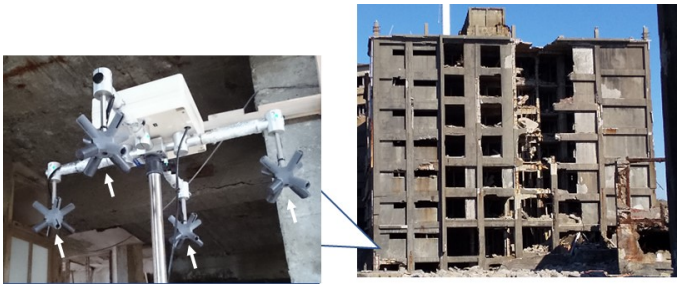


Figure 2: Microphone array installed in Building No. 30 of Gunkanjima (Hashima Island), Japan.

2.2. Sound Model of an Array with M Microphones

The sounds measured with the microphone array (Figure 2) are analyzed in the frequency domain. Then, taking the short-time Fourier transform of each microphone input at time t, the following model is obtained [3]:

$$\mathbf{y}(t,\omega) = \mathbf{A}(t,\omega)\mathbf{s}(t,\omega) + \mathbf{n}(t,\omega), \quad (1)$$

where $\mathbf{y}(t,\omega) = [Y_1(t,\omega), \ldots, Y_M(t,\omega)]^T$ is the input vector of the microphone array, with m-th element $Y_m(t,\omega)$ of the vector, at time t and frequency ω; M (= 16) denotes the number of microphones in the array; and the superscript T denotes the transpose. Additionally, $\mathbf{A}(t,\omega)$ is a matrix of transfer function vectors defined as

$$\mathbf{A}(\omega) = [\mathbf{a}_1(\omega), \ldots, \mathbf{a}_L(\omega)], \quad (2)$$

where $\mathbf{s}(t,\omega) = [S_1(t,\omega), \ldots, S_L(t,\omega)]^T$ is the source spectrum vector of the L sound sources in the measurement environment and $\mathbf{n}(t,\omega) = [N_1(t,\omega), \ldots, N_M(t,\omega)]^T$, a background noise spectrum vector, carries the assumption of following a zero-mean Gaussian distribution. We assume $\mathbf{a}_i(\omega) = [A_{1i}(\omega),\ldots A_{Mi}(\omega)]^T$ to be a transfer function that can be premeasured using time-stretched pulses [4]. We design time-stretched pulses to measure the impulse response, and the energy of the impulse signal is dispersed over time using a filter [4]. This filter is used to advance (or delay) the phase in proportion to the square of the frequency. Therefore, $\mathbf{a}_i(\omega)$ can be obtained by applying the inverse of the filter to the time-stretched pulse response.

2.3. Sound Source Features

In building health monitoring, it is necessary to detect the location, magnitude, and duration of an eventual collapse. Hence, the loudness, continuity, and pitch features are extracted from the sounds observed by the microphone array. To visualize the sound features, a color map is drawn at the estimated sound locations to analyze the types of generated sounds. The proposed method for the building health monitoring displays information in the color map, which we call a sound map.

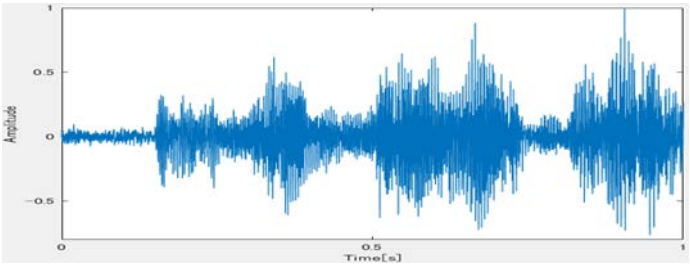


Figure 3: Sample environmental sound.

The observed environmental sounds are divided into 1-second segments, as illustrated in Figure 3. The short-time Fourier transform is applied to each segment to determine the loudness, continuity, and pitch, as detailed below.

1) Loudness Feature

Loudness is important to characterize environmental sounds during building collapse. This information is represented by a sonogram, which we obtain via the MATLAB Music Analysis toolbox [5]. The estimation of the loudness per frequency band is performed using auditory models and the function ma_sone in the MATLAB toolbox, where the specific loudness sensation (in sones) per critical band (in Bark scale) is calculated in six steps. 1) The fast Fourier transform is used to calculate the power spectrum of the audio signal. 2) According to the Bark scale [6], the frequencies are bundled into 20 critical bands. 3) The spectral masking effects are calculated as in [7]. 4) The loudness is

calculated in decibels relative to the threshold of hearing (decibels with respect to sound pressure level—dB-SPL). 5) From the dB-SPL values, equal loudness levels in unit phones are calculated. 6) The loudness is calculated in sones based on [8]. (Regarding detail of the six steps, see [9]). Figure 4 shows the sonogram in the Bark scale of the sound depicted in Figure 3.

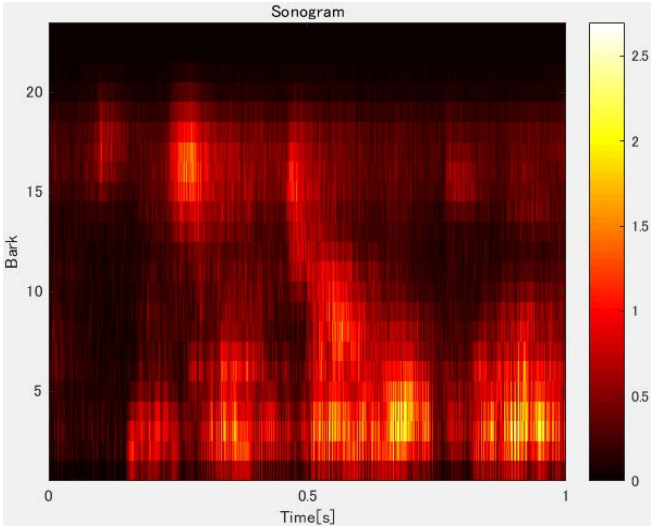


Figure 4: Sonogram in Bark scale of the sound in Figure 3.

The frequency histogram of the sonogram is also computed [5], and the histogram is resampled to 8 bits using the magnitude relation for the median of the histogram. The score expressing the information can then be obtained by converting each 8-bit value into the corresponding decimal value. An environmental sound with several variations in loudness sensation per frequency band is indicated by a higher score.

2) Continuity Feature

We use pitch variations of environmental sounds with respect to time *t* to measure the continuity of sounds. Therefore, continuity is expressed as a score obtained from a feature representing the pitch variation of the environmental sound.

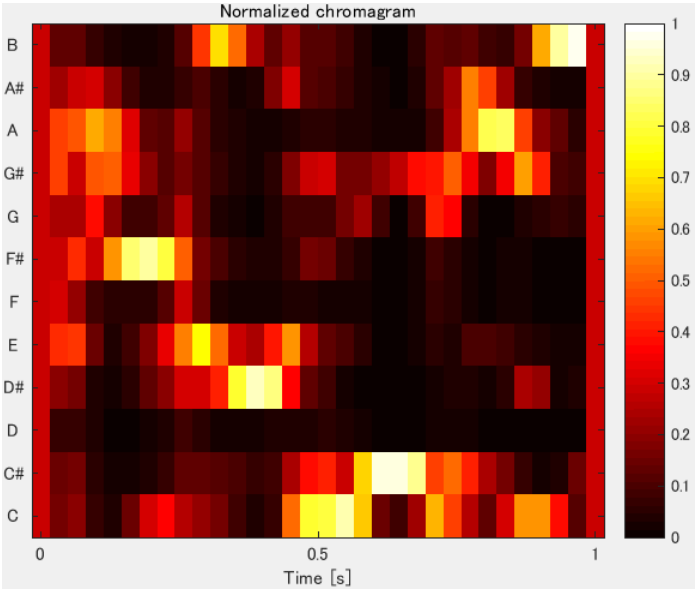


Figure 5: Chromagram of the sound in Figure 3

The corresponding chromagram, like the one shown in Figure 5, can be calculated using pitch features as those illustrated in Figure 3 by applying the method in [10]. Namely, we use the MATLAB Chroma toolbox to calculate the chromagram [11]. Subsequently, a differential chromagram is determined from the initial chromagram (Figure 5) using command diff in MATLAB, and the discrete cosine transform is applied to the differential chromagram (Figure 6). Based on this result, a time-domain histogram is then calculated. Continuity information is thus represented by a score obtained from the time-domain histogram in a method analogous to that used to obtain the loudness.

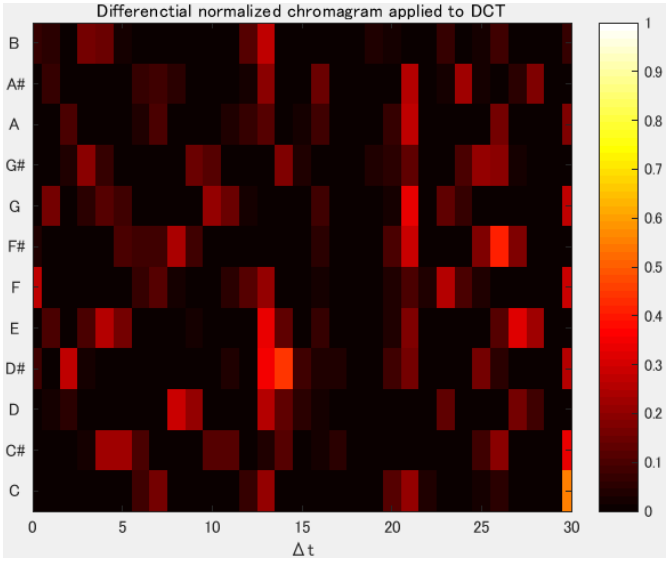


Figure 6: Differential chromagram after applying discrete cosine transform for the sound in Figure 3.

The histogram represents the variation of keys over time *t*. Therefore, a low score suggests that key variations of the environmental sound do not occur frequently, that is, the sound has low continuity. Conversely, a high score indicates frequent key variations of the environmental sound, indicating high continuity. As environmental sounds contain various keys, we use this score to represent continuity.

3) Pitch Feature

Pitch information is represented by the corresponding score, which is used to determine the type of environmental sound.

The spectrogram of the sound depicted in Figure 3 is shown in Figure 7. Edge extraction is applied to the spectrogram allowing the calculation of the number of pixels in its frequency feature areas and centroid frequencies. The improved affinity propagation method [12] is used to categorize the detected frequency characteristic areas of the spectrogram. More details on affinity propagation can be found in [13] and [14].

Each centroid frequency obtained by improved affinity propagation is classified into low-, medium-, or high-frequency groups. Then, a frequency group histogram can be established. The pitch is obtained from the histogram as a score, for which the calculation details can be found in [15]. A low score indicates a low dominant frequency of the environmental sound, whereas a high score indicates the presence of various frequency components.

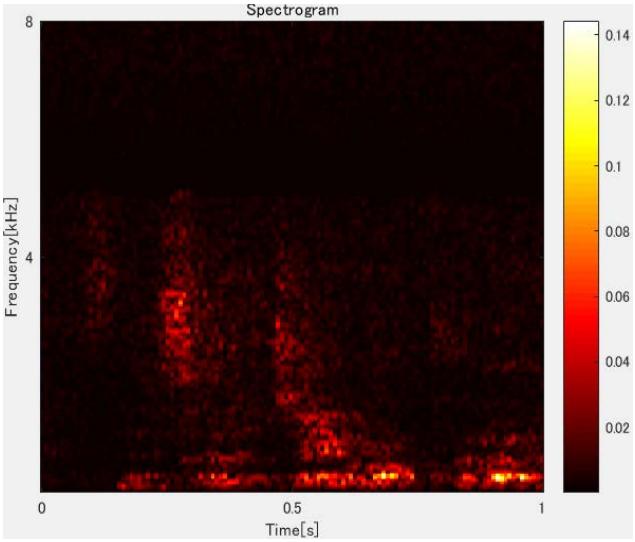


Figure 7: Spectrogram of the sound in Figure 3.

2.4. Sound Localization

Various methods for estimating sound source directions has been proposed (e.g., [16], [17], [18]) . In this paper, using $\mathbf{A}(\omega)$ obtained from (2), the MUSIC [19] is used to estimate the locations of sound sources. Details about the estimation of virtual 3D sound source positions can be found in [10].

2.5. Sound Map from Features

The loudness, continuity, and pitch scores are used to construct a sound map by representing the three scores in a red–green–blue color model. The obtained colors are then overlaid on the estimated sound source positions. Therefore, unlike conventional sound visualization methods such as power spectrum, the proposed sound map reflects not only the pitch but also the loudness and continuity of the sound in a color model, thus establishing a novel representation.

3. Experimental Results

3.1. Experimental Setup

The effectiveness of the proposed method is evaluated by using the sounds of falling cans and stones measured by the microphone array installed as shown in Figure 2.



Figure 8: Experimental scenario using falling cans and stones that hit the second floor of the building as sound sources

The three hitting points for evaluation are marked with an X in Figure 8 and are located on the 2nd floor of the building. The cans are dropped from the 4th and 6th floors of the building to investigate the difference in the sound features caused by the drop height. Figure 8 also shows the placement of the microphone array and the person in charge of dropping the cans and stones.

3.2. Sound Maps

Figure 9 shows the sound maps obtained from the sound sources generated by cans (Figure 9(a)) and stones (Figure 9(b)) hitting on the second floor. The color differences mainly correspond to loudness and continuity, as listed in Table 1, where each feature score is the average across nine trials. There is also a difference in pitch. Hence, the sounds of hitting stones have more frequency components than those of hitting cans. The three extracted features allow to distinguish differences in objects that cause sudden sounds, and Figure 9 also demonstrates the correct localization of the sound sources.



(a)



(b)

Figure 9: Sound maps of sound sources generated by (a) cans and (b) stones hitting the 2nd floor of the building. The red circle indicates the position of the microphone array.

Table 1: The averages of loudness, continuity, and pitch scores for the sounds generated by falling cans and stones hitting on second floor of the building.

	Loudness score	Continuity score	Pitch score
Can	158.2	107.3	19.2
Stone	115.3	179.8.8	32.2



(a)



(b)

Figure 10: Sound maps of the sound sources of dropping cans at the (a) 4th and (b) 6th floors of the building.

Therefore, if there are frequency differences in sounds generated before and after a building collapses, the three features, especially the pitch, can be used to detect them.

Figs. 10 shows the sound maps of the dropping cans from the fourth (Figure 10(a)) and sixth (Figure 10(b)) floors of the building, respectively, averaged across three trials. The X marks in Figure 11 show the corresponding dropping points for the maps in Figure 10. As expected, the colors on the estimated directions of the sound sources for both floors are similar.



(a)



(b)

Figure 11: Dropping points of cans (X marks) from the (a) 4th and (b) 6th floors of the building.

Table 2: Averages of loudness, continuity, and pitch scores for the generated sounds of falling cans at the 4th and 6th floors of the building.

Floor	Loudness score	Continuity score	Pitch score
2	240.0	113.7	25.0
4	240.0	84.3	16.3
6	245.3	28.3	26.0

The colors of the sound maps in Figs. 10(a) and (b) reflect the difference in floors and are mainly related to variations in the continuity score, as listed in Table 2.

The experimental results indicate the feasibility of using the proposed method for visualizing sound features in a building health monitoring system that can roughly determine the locations and types of environmental sounds in buildings.

4. Discussion and Conclusions

Figure 12 shows snapshots of building 30 before (Figure 12(a)) and after (Figure 12(b)) a floor collapse (red area). These snapshots illustrate the difficulty in determining the collapse through the use of images from cameras. Thus, the use of sounds may allow effective monitoring of buildings such as those in Gunkanjima, which is an uninhabited island.



(a)



(b)

Figure 12: Building No. 30 (a) before and (b) after the collapse of the floor (red area).

We propose a method of visualizing sounds which can be used for implementing building health monitoring by calculating the direction and features of environmental sounds. The proposed visualization technique of sound features considering sound localization relies on sound maps that reflect the loudness, continuity, and pitch of multiple sounds.

Experiments considering falling cans and stones were considered to simulate sounds of a collapsing building. The experimental results suggest that, using the proposed method, collapsing floors and walls attributable to the damage and deterioration of buildings can be localized. Moreover, the mechanism of building collapse may be analyzed and clarified by using the proposed method.

The proposed method for building health monitoring using the sound measurement will be used to continuously monitor building No. 30 of Gunkanjima, Japan, and the building monitoring data will be collected. In addition, we will further improve the accuracy of sound localization.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was partly funded by Japan Society for the Promotion of Science (grant number 16H02911 and 17HT0042) and JST CREST under grant JPMJCR18A4.

References

[1] H. Hamamoto, N. Kurata, S. Saruwaari, M. Kawamoto, A. Tomioka, T. Daigo, “Field test of change detection system of building group in preparation for unexpected events in Gunkanjima,” AIJ J. Technol. Des., **24**(57), 553-558, 2018, doi:10.3130/aijt.24.553 (article in Japanese with an abstract in English).

[2] A. Sasou, N. Odontsengel, S. Matsuoka, “An acoustic-based tracking system for monitoring elderly people living alone,” in 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health, **2018-March**, 89-95, 2018, doi:10.5220/0006664800890095.

[3] A Quinlan, M. Kawamoto, Y. Matsusaka, H. Asoh, F. Asano, “Tracking Intermittently Speaking Multiple Speakers Using a Particle Filter,” EURASIP Journal on Audio, Speech, and Music Processing, **2009**(1), 1-11, 2009, doi:10.1155/2009/673202.

[4] Y. Suzuki, F. Asano, H. Kim, T. Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses,” The Journal of the Acoustical Society of America, **97**(2), 1119-1123, 1995, doi: 10.1121/1.412224.

[5] E. Pampalk, Music analysis toolbox [Online]. Available: http://www.pampalk.at/ma/documentation.html.

[6] H. Fastl, E. Zwicker, Psychoacoustics Facts and Models, Springer-Verlag Berlin Heidelberg, 1999.

[7] M. R. Schroeder, D.S. Atal, J.L. Hall, “Optimizing digital speech coders by exploiting masking properties of the human ear,” The Journal of the Acoustical Society of America, **66**(6), 1647-1652, 1979, doi:10.1121/1.383662.

[8] R. A. W. Bladon, B. Lindblom, “Modeling the judgment of vowel quality differences,” The Journal of the Acoustical Society of America, **69**(5), 1414-1422, 1981, doi: 10.1121/1.385824.

[9] E. Pampalk, A. Rauber, D. Dueck, “Content-based organization and visualization of music archives,” in the 10th ACM International Conference on Multimedia 2002, 570-579, 2002, doi:10.1145/641007.641121.

[10] M. Kawamoto, T. Hamamoto, “Building Health Monitoring Using Computational Auditory Scene Analysis,” in 16th Annual International Conference on Distributed Computing in Sensor Systems, DCOSS 2020, 144-146, 2020, doi:10.1109/DCOSS49796.2020.00033.

[11] M. Müller, S. Ewert, “Chroma toolbox: Matlab implementations for extracting variants of chroma-based audio features,” in the 12th International Society for Music Information Retrieval Conference, ISMIR 2011, 215-220, 2011.

[12] M. Kawamoto, “Sound-Environment Monitoring Method Based on Computational Auditory Scene Analysis,” Journal of Signal and Information Processing, **8**(2), 65-77, 2017, doi:10.4236/JSIP.2017.82005.

[13] B. J. Frey, D. Dueck, “Clustering by passing messages between data points,” Science, **315**(5814), 972-976, 2007, doi:10.1126/SCIENCE.1136800.

[14] K. Wang, J. Zhang, D. Li, X. Zhang, T. Guo,” Adaptive Affinity Propagation Clustering,” Acta Automatica Sinica, **33**(12), 1242-1246, 2007.

[15] M. Kawamoto, “Sound-environment monitoring technique based on computational auditory scene analysis,” in 25th European Signal Processing Conference, EUSIPCO 2017, 2017-January, 2516-2524, 2017, doi:10.23919/EUSIPCO.2017.8081664.

[16] T. Nishiura, T. Yamada, S. Nakamura, K. Shikano, “Localization of multiple sound sources based on a CSP analysis with a microphone array,” in 2000 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), **2**, 1053-1056, 2000, doi:10.1109/ICASSP.2000.859144.

[17] H. L. Van Trees, Optimum Array Processing, Wiley Interscience: New York, 2002, doi:10.1002/0471221104.

[18] D. H. Johnson, D. Dudgeon, Array Signal Processing: Concepts and Techniques, New Jersey: Englewood Cliffs: PTR, Prentice Hall, 1993.

[19] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions on Antennas and Propagation, **AP-34**(3), 276–280, 1986, doi:10.1109/TAP.1986.1143830.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 4, 01-12 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Exploiting Domain-Aware Aspect Similarity for Multi-Source Cross-Domain Sentiment Classification

Kwun-Ping Lai*, Jackie Chun-Sing Ho, Wai Lam

Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, 999077, China

ARTICLE INFO

Article history:
Received: 01 May, 2021
Accepted: 15 June, 2021
Online: 10 July, 2021

Keywords:
Domain-aware topic model
Topic-attention network
Adversarial training
Artificial neural networks

ABSTRACT

We propose a novel framework exploiting domain-aware aspect similarity for solving the multi-source cross-domain sentiment classification problem under the constraint of little labeled data. Existing works mainly focus on identifying the common sentiment features from all domains with weighting based on the coarse-grained domain similarity. We argue that it might not provide an accurate similarity measure due to the negative effect of domain-specific aspects. In addition, existing models usually involve training sub-models using a small portion of the labeled data which might not be appropriate under the constraint of little labeled data. To tickle the above limitations, we propose a domain-aware topic model to exploit the fine-grained domain-aware aspect similarity. We utilize the novel domain-aware linear layer to control the exposure of various domains to latent aspect topics. The model discovers latent aspect topics and also captures the proportion of latent aspect topics of the input. Next, we utilize the proposed topic-attention network for training aspect models capturing the transferable sentiment knowledge regarding particular aspect topics. The framework finally makes predictions according to the aspect proportion of the testing data for adjusting the contribution of various aspect models. Experimental results show that our proposed framework achieves the state-of-the-art performance under the constraint of little labeled data. The framework has 71% classification accuracy when there are only 40 labeled data. The performance increases to around 82% with 200 labeled data. This proves the effectiveness of the fine-grained domain-aware aspect similarity measure.

1 Introduction

Online shopping becomes more and more popular during the pandemic. Product reviews serve as an important information source for product sellers to understand customers, and for potential buyers to make decisions. Automatically analyzing product reviews therefore attracts people's attention. Sentiment classification is one of the important tasks. Given sufficient annotation resources, supervised learning method could generate promising result for sentiment classification. However, it would be very expensive or even impractical to obtain sufficient amount of labeled data for unpopular domains. Large pre-trained model, such as the Bidirectional Encoder Representations from Transformers model (BERT) [1], could be an universal way to solve many kinds of problems without exploiting the structure of the problem. In [2], the author apply large pre-trained model to handle this problem task, which has sufficient labeled data only in source domain but has no labeled data in target domain, with fine tuning on source domain and predicting on target domain. In [3], the author train the large pre-trained model using various sentiment related tasks and show that the model could directly apply to the target domain even without the fine-tuning stage. However, these large pre-trained models do not consider the structure of the problem and they have certain hardware requirement that might not be suitable in some situations. We focus on smaller models, which have a few layers, in this work in order to handle the constraint of little labeled data[1]. Besides using the gigantic pre-trained model, domain adaptation (or cross-domain) [4, 5] attempts to solve this problem by utilizing the knowledge from the source domain(s) with abundant annotation resources and transfers the knowledge to the target domain. This requires the model to learn transferable sentiment knowledge by eliminating the domain discrepancy problem. Domain adversarial training [6, 7] is an ef-

*Corresponding Author: Kwun-Ping Lai, Email: kplai@se.cuhk.edu.hk

[1]To give a brief comparison of our proposed framework and the large pre-trained model, we present the performance of the standard BERT-Large model in the experiment section. We ignore other variants of the large pre-trained models as they are not the major focus of this work.

fective method to capture common sentiment features which are useful in the target domain. Various works using domain adversarial training [8]–[11] achieve good performance for single-source cross-domain sentiment classification. It could be also applied to the large pre-trained model to further boost the performance [12]. Moreover, it is quite typical that multiple source domains are available, the model might be exposed to a wider variety of sentiment information and the amount of annotation requirement for every single domain would be smaller. A simple approach is to combine the data from multiple sources and form a new combined source domain. Existing models tackling single-source cross-domain sentiment classification mentioned above could be directly applied to this new problem setting after merging all source domains. However, the method of combining multiple sources does not guarantee a better performance than using only the best individual source domain [13, 14]. Recent works measure the global domain similarity [15]–[17], i.e. domain similarity between the whole source and target domain, or instance-based domain similarity [18]–[21], i.e. domain similarity between the whole source domain and every single test data point. We observe that these approaches are coarse-grained and ignore the fine-grained aspect relationship buried in every single domain. Domain-specific aspects from the source domain might have negative effect in measuring the similarity between the source domain and the target domain, or the single data point. For instance, we would like to predict the sentiment polarity of some reviews from the Kitchen domain and we have available data from the Book, and the DVD domain. Intuitively, the global domain similarity might not have much difference as both of them are not similar to the target. However, reviews related to the cookbook aspect from the Book domain, or reviews talking about cookery show from the DVD domain might contribute more to the prediction of Kitchen domain. Discovering domain-aware latent aspects and measuring the aspect similarity could be a possible way to address the problem. Based on this idea, we introduce the domain-aware aspect similarity measure based on various discovered domain-shared latent aspect topics using the proposed domain-aware topic model. The negative effect of domain-specific aspects could be reduced.

Existing models measuring domain similarity have another drawback. They usually train a set of expert models with each using a single source domain paired with the target domain. Then, the domain similarity is measured to decide the weighting of each expert model. Another way is to select a subset of data from all source domains which are similar to the target data. We argue that these approaches are not suitable under the constraint of little labeled data as each single sub-model is trained using a small portion of the limited labeled data which might obtain a heavily biased observation. The performance under limited amount of labeled data is underexplored for most of existing methods as they require considerable amount of labeled data for training. In [22], the author study the problem setting applying the constraint. However, they assume equal contribution for every source domain. We study the situation under the constraint of little labeled data and at the same time handling the contribution of source domains using fine-grained domain-aware aspect similarity.

To address the negative effect of domain-specific aspects during the domain similarity measure, and also the limitation of the constraint of little labeled data, we propose a novel framework exploiting domain-aware aspect similarity for measuring the contribution of each aspect model representing the captured knowledge of particular aspects. It is capable of working under the constraint of little labeled data. Specifically, the framework consists of the domain-aware topic model for discovering latent aspect topics and inferring the aspect proportion utilizing a novel aspect topic control mechanism, and the topic-attention network for training multiple aspect models capturing the transferable sentiment knowledge regarding particular aspects. The framework makes predictions using the measured aspect proportion of the testing data, which is a more fine-grained measure than the domain similarity, to decide the contribution of various aspect models. Experimental results show that the proposed domain-aware aspect similarity measure leads to a better performance.

1.1 Contributions

The contributions of this work are as follows:

- We propose a novel framework exploiting the domain-aware aspect similarity to measure the contribution of various aspect models for predicting the sentiment polarity. The proposed domain-aware aspect similarity is a fine-grained measure which is designed to address the negative effect of domain-specific aspects existing in the coarse-grained domain similarity measure.
- We present a novel domain-aware topic model which is capable of discovering domain-specific and domain-shared aspect topics, together with the aspect distribution of the data in an unsupervised way. It is achieved by utilizing the proposed domain-aware linear layer controlling the exposure of different domains to latent aspect topics.
- Experimental results show that our proposed framework achieves the state-of-the-art performance for the multi-source cross-domain sentiment classification under the constraint of little labeled data.

1.2 Organization

The rest of this paper is organized as follows. We present related works regarding cross-domain sentiment classification in Section 2. We describe the problem setting and our proposed framework in Section 3. We conduct extensive experiments and present results in Section 4. Finally, we talk about limitations and furture works in Section 5, and summarize our work in Section 6.

2 Related Works

Sentiment analysis [23]–[25] is the computational study of people's opinions, sentiments, emotions, appraisals, and attitudes towards entities [26]. In this work, we focus on textual sentiment data which is based on review of products, and the classification of the sentiment polarity of reviews. We first present the related works of single-source cross-domain sentiment classification. Next, we further extend to multiple-source case.

2.1 Single-Source Cross-Domain Sentiment Classification

Early works involve the manual selection of pivots based on predefined measures, such as frequency [27], mutual information [5, 28] and pointwise mutual information [29], which might have limited accuracy.

Recently, the rapid development of deep learning provides an alternative for solving the problem. Domain adversarial training is a promising technique for handling the domain adaptation. In [8], the author make use of memory networks to identify and visualize pivots. Besides pivots, [9] also consider non-pivot features by using the NP-Net network. In [10], the author combine external aspect information for predicting the sentiment.

Large pre-trained models attract people's attention since the BERT model [1] obtains the state-of-the-art performance across various machine learning tasks. Researchers also apply it on the sentiment classification task. Transformer-based models [2, 12, 3] utilize the amazing learning capability of the deep transformer structure to learn a better representation for text data during the pre-training stage and adapt themselves to downstream tasks (sentiment classification in our case) using fine tuning. However, we argue that the deep transformer structure has been encoded with semantic or syntactic knowledge during the pre-training process which makes the direct comparison against shallow models unfair. It also has certain hardware requirement which hinders its application in some situations.

Methods mentioned above focus on individual source only and they do not exploit the structure among domains. Although we can still directly apply these models to solve the problem by either training multiple sub-models and averaging predictions, or merging all source domains into a single domain, having a performance better than using only the single best source is not guaranteed. Therefore, exploring the structure or relationship among various domains is essential.

2.2 Multi-Source Cross-Domain Sentiment Classification

Early works assuming equal contribution for every source domain [30]–[32] could be a possible approach to handle the relationship between source domains and the target. Other solutions try to align features from various domains globally [33]–[22]. However, the source domain with higher degree of similarity to the target domain contributing more during the prediction process is a reasonable intuition. These methods fail to capture the domain relation. Recent works try to measure domain contribution in order to further improve the performance.

Researchers propose methods to measure the global domain similarity [15]–[17], i.e. the domain similarity between the whole source and target domain, or the instance-based domain similarity [18]–[21], i.e. the domain similarity between the whole source domain and every single test data point. In [15], the author measure the domain similarity using the proposed sentiment graph. In [17], the author employ a multi-armed bandit controller to handle the dynamic domain selection. In [18], the author compute the attention weight to decide the contribution of various already trained expert models. [20] also utilize the attention mechanism to assign importance weights. They incorporate a Granger-causal objective in their mixture of experts training. The total loss measuring distances of attention weights from desired attributions based on how much the inclusion of each expert reduces the prediction error. Maximum Cluster Difference is used in [19] as the metric to decide how much confidence to put in each source expert for a given example. In [21], the author utilize the output from the domain classifier to determine the weighting of a domain-specific extractor.

These methods measure the coarse-grained domain relation and ignore the fine-grained aspect relationship buried in every single domain. In addition, these methods do not consider the constraint of limited labeled data, which is the main focus of this work.

3 Model Descriptions

3.1 Problem Setting

The problem setting consists of the source domain group D_s and the target domain D_t. The source domain group has m domains $\{D_{s_k}\}|_{k=1}^{m}$ while there is only one target domain. For each source domain, we have two sets of data: i) the labeled data $L = \{x_i^l, y_i^l\}|_{i=1}^{n_L}$ and ii) the unlabeled data $U = \{x_j^u, d_j\}|_{j=1}^{n_U}$ where n_L and n_U are the number of data of labeled and unlabeled data respectively, and d_j is the augmented domain membership indicator. Note that y_i is the sentiment label for the whole review x_i and we do not have any fine-grained aspect-level information. The kth source domain can be written as $D_{s_k} = \left\{L_{s_k} = \{x_i^{l,s_k}, y_i^{s_k}\}|_{i=1}^{n_{L_{s_k}}}, U_s = \{x_j^{u,s_k}, d_j^{s_k}\}|_{j=1}^{n_{U_{s_k}}}\right\}$. The data of the target domain has similar structure except that we do not have the sentiment label, i.e. $D_t = \left\{\{x_i^t\}|_{i=1}^{n_t}, U_t = \{x_j^{u,t}, d_j^t\}|_{j=1}^{n_{U_t}}\right\}$ respectively. $n_{L_{s_k}}$ is the number of labeled data and they are the same for all k. We set all $d_*^{s_k}$ to k and all d_*^t to $m+1$. The objective of the multi-source cross-domain sentiment classification is to find out a best mapping function f so that given the training data $T = \{D_{s_1}, D_{s_2}, ..., D_{s_m}, D_t\}$, the aim is to predict the label of the target domain labeled data $\overline{\mathbf{y}}^{\mathbf{t}} = f(\mathbf{x}^{\mathbf{t}})$.

3.2 Overview of Our Framework

We describe our proposed framework exploiting domain-aware aspect similarity. Specifically, there are two components: i) the domain-aware topic model discovering domain-aware latent aspect topics, ii) the topic-attention network identifying sentiment topic capturing the transferable aspect-based sentiment knowledge. The first component captures both domain-specific and domain-shared latent aspect topics, and infers the aspect distribution of each review. It is an unsupervised model that utilizes only the unlabeled data. It is analogous to the standard topic model which discovers latent topics as well as topic distributions. However, the standard topic model is not capable of controlling discovered latent topics. Our proposed domain-aware topic model is capable of separating discovered latent topics into two groups: we name them as domain-specific aspect topics and domain-shared aspect topics. The topic control is achieved by using the domain-aware linear layer described in the latter subsection. Specifically, the model discover $n_{\texttt{spec}}$ domain-specific aspect topics for every domain, and $n_{\texttt{share}}$ domain-shared

aspect topics which are shared among all domains. Each review has a $n_{\text{spec}} + n_{\text{share}}$ dimensional aspect distribution with the first n_{spec} dimension corresponding to domain-specific aspect topics and the last n_{share} dimension corresponding to domain-share aspect topics. Discovered aspect topics and inferred aspect distributions have three important functions:

- By considering only domain-shared aspect topics, the negative effect of domain-specific aspect topics could be minimized for measuring the contribution during the inference process.
- The overall aspect distribution of the testing data reveals the importance of each discovered aspect topic following the assumption that the topic appearing more frequent is more important for the target domain.
- The aspect distribution of the unlabeled data could be used for picking reviews with a high coverage of a particular set of aspect topics.

Based on the domain-shared aspect distribution of the target domain, we divide discovered domain-shared aspect topics into groups with each group having unlabeled reviews from all domains with high aspect proportion forming the training dataset for the second component. Specifically, we divide domain-shared aspect topics into groups based on the overall aspect distribution of the target domain. We aim at separating aspect topics and train an expert model for each group of aspects. Each aspect model focuses on a particular set of aspects so as to boost the learning capability of that set of fine-grained aspect topics. Therefore, we need to construct the dataset carrying the information related to selected aspect topics. We select the unlabeled data from all domains with high aspect proportion of a particular set of aspect topics to form the aspect-based training dataset.

Each of the aspect-based training dataset guides the next component to focus on the corresponding aspect group and identify the related transferable sentiment knowledge. The obtained training dataset is jointly trained with the limited labeled data using the topic-attention network to generate an aspect model for each aspect-based training dataset. The topic-attention network is a compact model which is designed to work effectively under limited training data. The topic-attention network captures two topics simultaneously: i) the sentiment topic and ii) the domain topic. The sentiment topic captures the transferable sentiment knowledge which could be applied to the target domain. The domain topic serves as an auxiliary training task for constructing a strong domain classifier which helps the sentiment topic to identify domain-independent features by using domain adversarial training. These two topics are captured by the corresponding topical query built in the topic-attention layer. These topical queries are learnt automatically during the training process. The limited labeled data works with the sentiment classifier to control the knowledge discovery related to sentiment (sentiment topic captures sentiment knowledge while domain topic does not), while the unlabeled data works with the domain classifier to control the knowledge discovery related to domain. Finally, the framework makes predictions using various aspect models with contribution defined by the aspect distribution of the testing data. For example, if the testing data has a higher coverage regarding aspect group 1, then naturally the prediction made by the aspect model of group 1 should contribute more to the finally prediction as intuitively that aspect model would have more related sentiment knowledge to make judgement. We believe this fine-grained latent aspect similarity would provide a more accurate sentiment prediction than the traditional coarse-grained domain similarity due to the fact that we eliminate the negative effect of domain-specific aspects when measuring the similarity between the testing data and the expert models.

We first describe the architecture of the two components. Then, we describe the procedure of inferring the sentiment polarity of reviews of the target domain.

3.3 Domain-Aware Topic Model

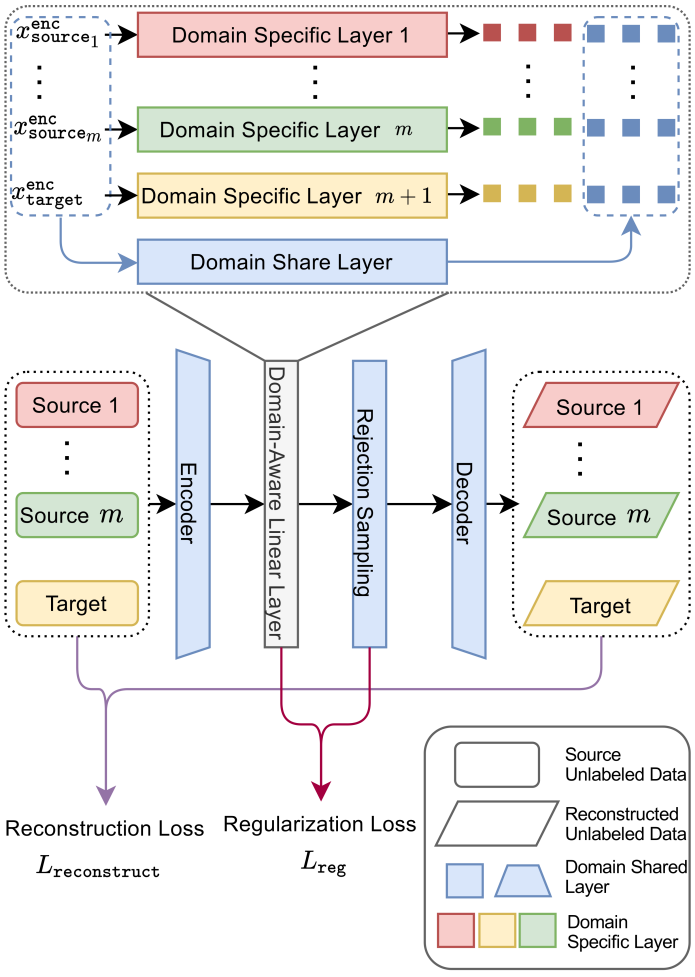


Figure 1: Diagram depicting the proposed domain-aware topic model. The middle part provides a high-level overview of the proposed domain-aware topic model. The model aims at inferring the dense representation of the unlabeled data from all domains in terms of aspect topic proportion. The model discovers domain-specific aspect topics and domain-shared aspect topics utilizing the domain-aware linear layer which is illustrated in the upper part of the figure. The model is trained by minimizing the reconstruction loss calculated by using the input data and the reconstructed data, and the regularization loss based on the inferred α and the predefined Dirichlet prior.

The domain-aware topic model follows the mechanism of the variational autoencoder framework (VAE) [35] which utilizes the encoder for inferring the latent variable (the Dirichlet prior α in our case representing the expected aspect distribution) and the de-

coder for reconstructing the input. Researchers try to apply the VAE network for achieving functionalities of standard topic model in a neural network way, such as inferring the topic proportion of the input and the word distribution of each topic. This provides some advantages such as reducing the difficulty of designing the inference process, leveraging the scalability of neural network, and the easiness of integrating with other neural networks [36]. However, the standard VAE using Gaussian distribution to model the latent variable might not be suitable for text data due to the sparseness of the text data. The Dirichlet distribution used in the topic model [37] has a problem of breaking the back-propagation. Calculating the gradient for the sampling process from the Dirichlet distribution is difficult. Researchers propose approximation methods [38, 39, 40, 41] in order to apply Dirichlet distribution to the neural topic model. We follow the rejection sampling method [42] in this work. Although discovered topics might carry extra information which might be helpful for identifying the hidden structure of the text data, it is not intuitive for applying this information to help the sentiment classification task. We introduce the domain-aware linear layer for controlling the formation of domain-specific and domain-shared aspect topics. To the best of our knowledge, we do not find any similar aspect topic control layer applied for multiple-source cross-domain sentiment classification in related works. The domain-aware linear layer identifies both domain-specific aspect topics and domain-shared aspect topics. We utilize domain-shared aspect topics only which could provide a more accurate measure for calculating the similarity. In addition, the inferred aspect topic proportion is used for constructing the aspect-based training dataset, and determining the level of contribution of each aspect model. Details of the architecture of the model are described below.

3.3.1 Encoder

The input of the encoder is the bag of words of the review. Specifically, we count the occurrence of each vocabulary in each review and we use a vector of dimension V to store the value. This serve as the input representing the review. The encoder is used to infer the Dirichlet prior of the aspect distribution of the input. The bag-of-words input is first transformed using a fully connected layer with RELU activation followed by a dropout layer.

$$\texttt{Layer}^{\texttt{enc}}(x) = \texttt{Dropout}\Big(\texttt{RELU}(W^{\texttt{enc}}x + b^{\texttt{enc}})\Big) \tag{1}$$

3.3.2 Domain-Aware Linear Layer

Next, the output is fed into the domain-aware linear layer for obtaining domain-specific and domain-shared features. The domain-aware linear layer has $m + 1$ sub-layers including m domain-specific sub-layers handling the feature extraction of the corresponding domain and 1 domain-shared sub-layer handling all domains as follows:

$$\texttt{Layer}^{\texttt{DL}}_{d_x}(x) = [W^{\texttt{DL}}_{d_x}x + b^{\texttt{DL}}_{d_x}; W^{\texttt{DL}}_{\texttt{shared}}x + b^{\texttt{DL}}_{\texttt{shared}}] \tag{2}$$

where d_x is the domain ID of the input x, and [;] represents the operation of vector concatenation. The output $x^{\texttt{DL}}$ is batch normalized and passed to the `SoftPlus` function to infer the Dirichlet prior α of the aspect distribution. To make sure each value in α is greater than zero, we set all values smaller than $\alpha_{\texttt{min}}$ to $\alpha_{\texttt{min}}$.

$$\alpha = \texttt{max}\Big(\texttt{SoftPlus}\Big(\texttt{BatchNorm}(x^{\texttt{DL}})\Big), \alpha_{\texttt{min}}\Big) \tag{3}$$

We use the rejection sampling method proposed in [42] to sample the aspect distribution z and at the same time it allows the gradient to back-propagate to α.

3.3.3 Decoder

The decoder layer is used for reconstructing the bag-of-word input. The sampled aspect distribution z is transformed by the domain-aware linear layer as follows:

$$\texttt{Layer}^{\texttt{dec}}(x) = [W^{\texttt{dec}}_{d_x}x; W^{\texttt{dec}}_{\texttt{shared}}x] \tag{4}$$

The output $x^{\texttt{dec}}$ is batch normalized and passed to the log-softmax function representing the log probability of generating the word.

$$y = \texttt{ln}\Big(\texttt{Softmax}\Big(\texttt{BatchNorm}(x^{\texttt{dec}})\Big)\Big) \tag{5}$$

3.3.4 Loss Function

The loss function includes the regularization loss and the reconstruction loss. The regularization loss measures the difference of the log probability of generating the aspect distribution z between two prior, α and $\overline{\alpha}$ as follows:

$$L_{\texttt{reg}} = lnP(z|\alpha) - lnP(z|\overline{\alpha}), \quad P(y|x) \sim Dir(x) \tag{6}$$

where α is inferred by the model and $\overline{\alpha}$ is the predefined Dirichlet prior. The reconstruction loss is the log probability of generating the bag-of-word input calculated as follows:

$$L_{\texttt{reconstruct}} = -\sum_{i=1}^{V} y_i x_i \tag{7}$$

where V is the vocabulary size, y_i is the log probability of the ith word generated by the model, and x_i is the count of the ith word in the input.

3.4 Topic-Attention Network

The topic-attention network aims at capturing the transferable sentiment knowledge from the limited labeled data of various source domains. To achieve this goal, the network is designed to capture two topics simultaneously: i) the sentiment topic, and ii) the domain topic. The sentiment topic identifies the transferable sentiment knowledge from the input data while the domain topic helps to train a strong domain classifier. We use the technique of domain adversarial training [6, 7, 43] to maintain the domain independence of the sentiment topic. However, instead of using the standard gradient reversal layer, we use the adversarial loss function [22] to achieve the same purpose with a more stable gradient and a faster convergence. The model has two training tasks: i) the sentiment task for identifying the sentiment knowledge, and ii) the auxiliary domain task for training a strong domain classifier. The adversarial loss function is applied to the domain classifier output of the sentiment topic and the sentiment classifier output of the domain topic to hold the indistinguishability property of these two topics. Details of the architecture of the model is described below.

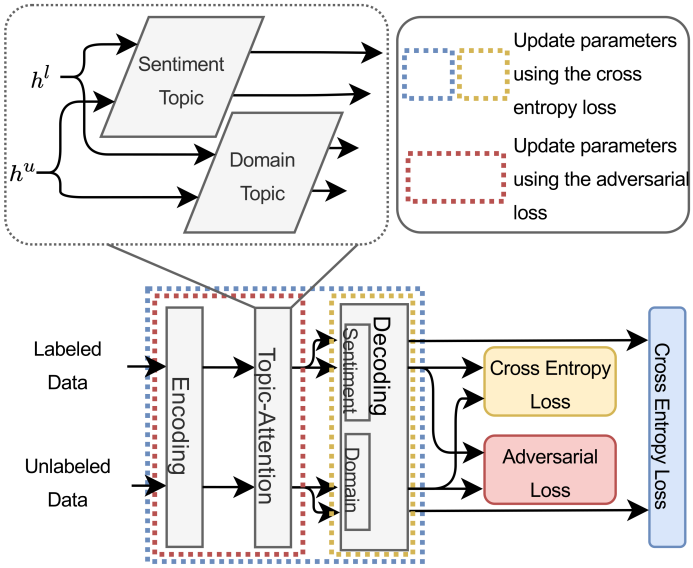


Figure 2: Diagram depicting the proposed topic-attention network. The bottom part provides a high-level overview of the proposed topic-attention network. The network captures two topics, i.e. the sentiment topic and the domain topic, from the review data and classifies the sentiment polarity and the domain membership. The adversarial loss maintains the indistinguishability of topics (domain indistinguishability of the sentiment topic and sentiment indistinguishability for the domain topic). Therefore, the sentiment knowledge captured by the sentiment topic could be transferred to the target domain. The colored dashed boxes show the scope of updating parameters for the corresponding loss.

3.4.1 Encoding Layer

Each word is mapped to the corresponding embedding vector and then transformed by a feed-forward layer with tanh activation for obtaining the feature vector h.

$$h = \texttt{tanh}\left(W^{\text{enc}}\,\texttt{Embedding}(x) + b^{\text{enc}}\right) \tag{8}$$

3.4.2 Topic-Attention Layer

The feature vector h_i of the ith word is re-weighted by the topical attention weight β_i^k calculated as follows:

$$\beta_i^k = \frac{m_i e^{q_k^\top h_i}}{\sum_{i'=1}^{k_w} m_{i'} e^{q_k^\top h_{i'}}} \times n_m \tag{9}$$

where k indicates the topic (either sentiment or domain topic), m_i is the word-level indicator indicating whether the ith position is a word or a padding, n_m is the number of non-padding words, and q_k is the topical query vector for topic k learnt by the model. Note that we have two topical query vectors representing two topics. The topical feature vector t_i^k of the topic k and the review i is obtained by summing feature vectors weighted by the corresponding topical attention weight β_*^k as follows:

$$t_i^k = \sum_{j=1}^{W_i} \beta_j^k h_j \tag{10}$$

where W_i is the number of words in review i. t_i^k represents extracted features of the review by topic k.

3.4.3 Decoding Layer

This layer consists of two decoders with each handling one training task, namely the sentiment decoder and the domain decoder for classifying the sentiment polarity and the domain membership respectively. Note that the review feature vector of labeled data is passed to the sentiment decoder while the unlabeled data of the aspect groups is passed to the domain decoder. Although we use the same t^k to represent the input feature vector in the following two equations, they are actually representing the review features captured from the labeled data, and unlabeled data respectively. Specifically, the review feature vector is linearly transformed and passed to the `Softmax` function for obtaining a valid probability distribution.

$$p^{\text{sen},k} = \texttt{Softmax}(W^{\text{sen}} t^k + b^{\text{sen}}) \tag{11}$$

$$p^{\text{dom},k} = \texttt{Softmax}(W^{\text{dom}} t^k + b^{\text{dom}}) \tag{12}$$

Note that there are four outputs generated by the decoding layer, including two outputs generated by the captured features of two topics passing to the sentiment decoder, and similarly the remaining two generated by the domain decoder. The two topics are sentiment and domain topic, i.e. $k = \{\texttt{sen}, \texttt{dom}\}$. Therefore, the four outputs are: $p^{\text{sen,sen}}$ and $p^{\text{sen,dom}}$ coming from the labeled data passing to the sentiment decoder (the first superscript) having specific features captured by the sentiment and domain topic (the second superscript) respectively, and $p^{\text{dom,sen}}$ and $p^{\text{dom,dom}}$ coming from the unlabeled data passing to the domain decoder having specific features captured by the corresponding topic.

3.4.4 Loss Function

We use the standard cross entropy loss to measure the classification performance:

$$L^{\text{sen},k} = -\frac{1}{n_L} \sum_{i=1}^{n_L} \ln p_{i,s_i}^{\text{sen},k} \tag{13}$$

$$L^{\text{dom},k} = -\frac{1}{n_U} \sum_{i=1}^{n_U} \ln p_{i,d_i}^{\text{dom},k} \tag{14}$$

where s_i and d_i are the class indicator specifying the sentiment polarity or the domain membership of the ith training data, and $p_{i,c}^*$ is the predicted probability regarding the cth class. Therefore, we have four cross entropy losses. The loss generated by the sentiment decoder from the sentiment topic and the loss generated by the domain decoder from the domain topic are used to update all parameters of the model using back-propagation. The remaining two are used to update the parameters of the decoding layer only. We introduce the adversarial loss function for doing adversarial training for both tasks as follows:

$$f_{adv}(p) = \sum_{i=1}^{c} (p_i - \frac{1}{c})^2 \tag{15}$$

where c is the number of classes and p_i is the predicted probability for the class i. Note that c for sentiment task is 2 while it is $m + 1$ for the domain task. We use the probability distributions generated by the sentiment decoder from the domain topic $p^{\texttt{sen,dom}}$, and by the domain decoder from the sentiment topic $p^{\texttt{dom,sen}}$, to calculate the adversarial losses, which are used to update the parameters of the encoding layer and the topic-attention layer.

3.5 Training Strategy

We first train the domain-aware topic model using the unlabeled data X^u from all domains. The model is then used for predicting the aspect proportion of the unlabeled data X^u and testing data X^t to obtain α^u and α^t. Note that the domain-aware topic model is an unsupervised model that does not utilize any labeled data from source domains nor target domain. The aspect score θ^t of the target domain is calculated using the mean value of the domain-shared aspect part of α^t over all testing data:

$$\theta^t = \frac{1}{n_t} \sum_{i=1}^{n_t} \alpha_i^t[-n_{\texttt{share}} :] \tag{16}$$

where $\alpha_i^t[-n_{\texttt{share}} :]$ represents the last $n_{\texttt{share}}$ dimensions of the vector α_i^t. Therefore, θ^t is a $n_{\texttt{share}}$ dimensional vector with each value representing the importance score of the corresponding aspect topic for the target domain. We divide the domain-shared aspect topics into k groups based on their importance score using θ_t in descending order. The set $\texttt{topic}_{g_{k'}}$ contains the topic indices of the k'th aspect group. For each group $g_{k'}$, we select top n unlabeled data from all domains based on the aspect topic score of the k'th group $\omega_{k'}^u$, which is the sum of the corresponding domain-shared aspect proportion of the k'th group for the uth review using its discovered aspect proportion:

$$\omega_{k'}^u = \sum_{i \in \texttt{topic}_{g_{k'}}} \alpha^u[i] \tag{17}$$

where $\alpha^u[i]$ represents the value in the ith dimension of α^u.

Next, we train k aspect models using the topic-attention network. For each aspect model, the limited labeled data X^l, Y^l is used for training the sentiment task while the group of selected unlabeled data $g_{k'}$ is used for training the auxiliary domain task. The last step is to utilize the obtained models for predicting the sentiment polarity of all testing data x^t. Let $AM_{k'}$ be the aspect model trained by using the dataset $\{X^l, Y^l, g_{k'}\}$, we denote the sentiment prediction of the sentiment topic generated by the model as $p_{k'}^{t'}$ for the target review $x^{t'}$ as follows:

$$p_{k'}^{t'} = AM_{k'}(x^{t'}) \tag{18}$$

Finally, we combine the sentiment predictions of the sentiment topic generated by all aspect models having each contributes according to the aspect proportion of the testing data to obtain the final prediction:

$$p^{t'} = \sum_{i=1}^{k} \omega_i^{t'} p_i^{t'} \tag{19}$$

where $\omega_i^{t'}$ is the contribution of the ith aspect model to the final prediction.

4 Experiment

4.1 Experiment Settings

We use the Amazon review dataset [5] for the evaluation of our proposed framework. The Amazon review dataset is a common benchmark for sentiment classification. We use 5 most common domains, namely Book, DVD, Electronics, Kitchen and Video. For each experiment cross, we reserve one domain as the target domain and use others as source domains. There are 5 combinations in total and we conduct experiments on these 5 crosses. For each domain, we follow the dataset setting in [9] collecting 6000 labeled data, with half positive and half negative polarity. We do further sampling to select a subset of the labeled data to fulfill the constraint of little labeled data. We first construct two lists with each having 3000 elements representing the index of the labeled data of positive and negative class respectively. We randomly shuffle the lists and pick first n indices. Next, we select the labeled data based on these indices. In order to have a comparable result for different size of labeled data, we fix the seed number of the random function so that the runs with different size of labeled data would obtain a same shuffle result. Therefore, the run with 20 labeled data contains the 10 labeled data from the run with 10 labeled data, and also another 10 new labeled data. Similarly, the run with 30 labeled data contains the 20 labeled data from the run with 20 labeled data. With this setting, we can directly estimate the effect of adding additional labeled data and compare the performance directly. We continue the process for other source domains. Finally, we construct 5 datasets having 10 to 50 labeled data for each target domain (there are 40 to 200 labeled data in total as there are 4 source domains). The unlabeled dataset includes all unlabeled data from all domains (including the target domain). All labeled data from the target domain is served as the testing data. We run every single run for 10 times and present the average accuracy with standard deviation in order to obtain a reliable result for model comparison.

4.2 Implementation Details

4.2.1 Domain-Aware Topic Model

The Dirichlet prior is set to 0.01. The minimum of inferred prior is set to 0.00001. We set the number of domain-specific and domain-shared topics to 20 and 40 respectively. We divide the domain-shared aspect topics into 5 groups. The domain-aware topic model is trained for 100 warm-up epochs, and stopped after 10 epochs of no improvement.

4.2.2 Topic-Attention Network

We use word2vec[2] embedding [44] to represent each word. We do not further train them to prevent overfitting. The batch size is

[2]It is a distributed representations of words in vector space. It helps various natural language processing task by putting similar words in a closer location.

[3]It is an optimization algorithm with adaptive learning rate. It considers the momentum of the gradient by using the moving average of the gradient. It also uses the moving average of the squared gradient to scale the learning rate of each individual parameter.

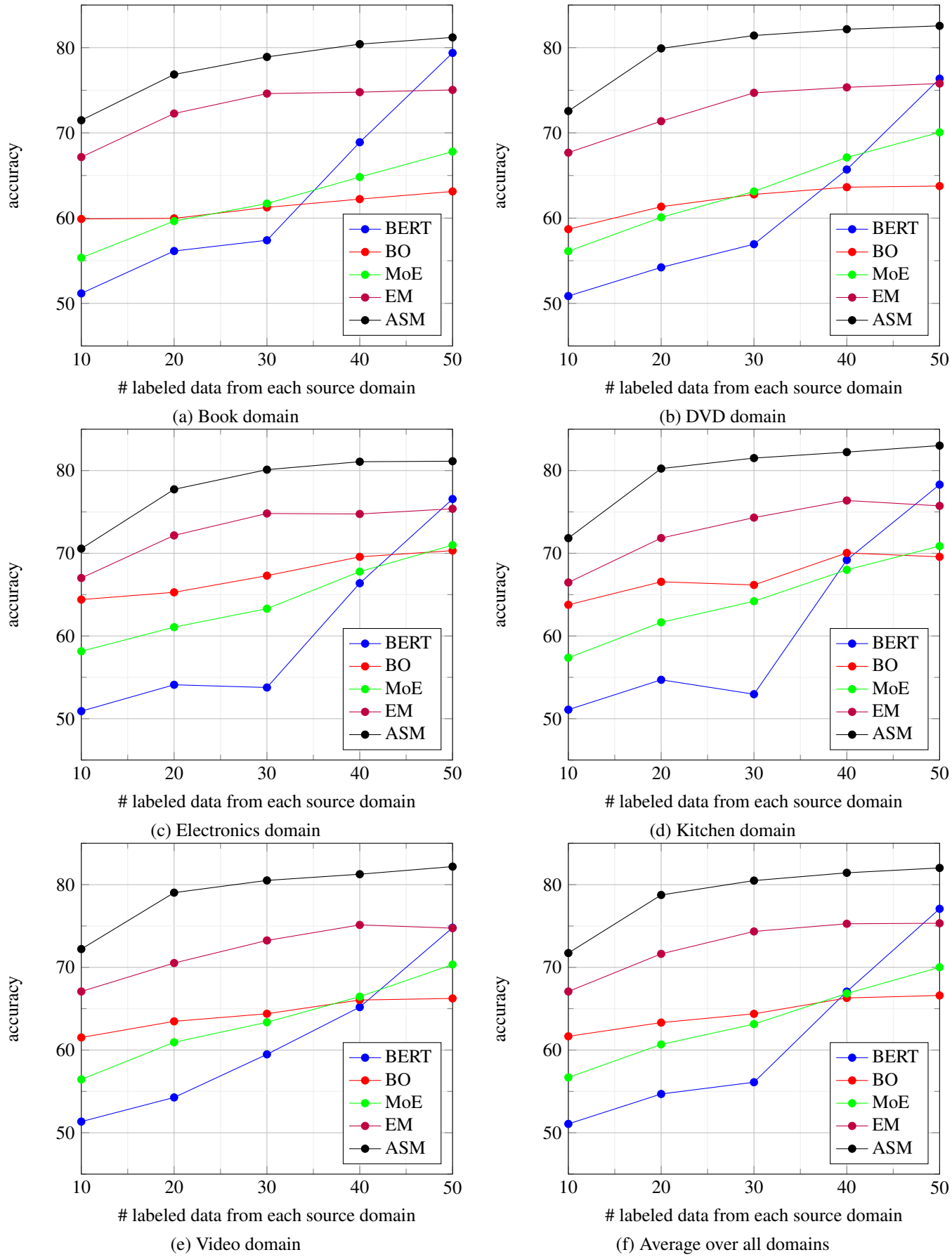


Figure 3: Figure showing the experimental results. Graph (a) to (e) shows the performance of Book, DVD, Electronics, Kitchen and Video domain as target domain respectively. Graph (f) shows the average accuracy of all domains.

set to the number of available labeled data. The topic-attention network is trained for 20 epochs. We use Adam[3] optimizer [45] for back-propagation for both models.

4.3 Evaluation Metric

We use accuracy to measure the evaluate the performance of various models. The target is a binary class. Therefore, correct cases involve the true positive (TP) and true negative (TN). Incorrect cases involve the false positive (FP) and the false negative (FN). The accuracy is calculated as follows:

$$\texttt{accuracy} = \frac{\texttt{TP} + \texttt{TN}}{\texttt{TP} + \texttt{TN} + \texttt{FP} + \texttt{FN}} \tag{20}$$

The average accuracy is calculated by taking the average of accuracy scores of multiple runs.

4.4 Main Results

Models used for performance comparison are as follows:

- BERT [1]: This is the **B**idirectional **E**ncoder **R**epresentations from **T**ransformers model, which is the popular pre-trained model designed to handle various text mining tasks. We use the BERT-Large model with fine tuning using the labeled data to obtain the prediction.
- BO [46]: This model employs **B**ayesian **o**ptimization for selecting data from source domains and transfer the learnt knowledge to conduct prediction on the target domain.
- MoE [19]: This is the **m**ixture **o**f **e**xpert model. It measures the similarity between single test data to every source domain for deciding the contribution of the expert models.
- EM [22]: This is the **e**nsemble **m**odel. It uses various base learners with different focuses on the training data to capture a diverse sentiment knowledge.
- ASM: This is the proposed framework exploiting the domain-aware **a**spect **s**imilarity **m**easure for obtaining a more accurate measure to adjust the contribution of various aspect models focusing on different aspect sentiment knowledge.

Results are presented in Figure 3 and Table 1. We use the classification accuracy as the metric to measure the performance. The proposed framework achieves the best average accuracy among all crosses. Its average performance is 71.73%, 78.75%, 80.49%, 81.43%, and 82.02% for 10, 20, 30, 40 and 50 labeled data cases respectively, or 40, 80, 120, 160 and 200 labeled data cases in total respectively.

4.5 Discussions

Our proposed framework performs substantially better than the comparison models. The proposed framework has an average of 4%, 7%, 6%, 6% and 6% absolute improvement over the second best result for 10, 20, 30, 40 and 50 labeled data cases respectively, or 40, 80, 120, 160 and 200 labeled data cases in total respectively. The variance of the proposed model is comparable to or better than the second best models. The result proves that our proposed framework is very effective for conducting multi-source cross-domain sentiment classification under the constraint of little labeled data. The model can capture transferable sentiment knowledge for predicting the sentiment polarity of the target reviews.

We also do comparative analysis to test the effectiveness of the proposed fine-grained domain-aware aspect similarity measure. It is based on the discovered aspect topics and also the aspect topic proportion for adjusting the contribution of various aspect models. We try to remove these two components to test the performance of the variants. The results are presented in Table 2. The first variant is *rand. select data + avg. pred.*, which means using the unlabeled data selected in a random way instead of using the aspect-based training dataset constructed by the domain-aware topic model, and combining the predictions of various aspect models by averaging them. In other words, the first variant removes both components. The second variant is *avg. pred.*. It keeps the first component (train the aspect models using the aspect-based training dataset) and only removes the second component. Therefore, it assumes equal contribution from various aspect models, just like the first variant. The last one is the proposed framework equipped with both components. Results show that the proposed fine-grained domain-aware aspect similarity measure improves the performance in general except the case having very few labeled data. We think the reason is that the aspect model could not locate the correct aspect sentiment knowledge from the limited data. Thus, the simply averaging the prediction of these biased aspect models would be better than relying on some models. Although the second variant (*avg. pred.*) has a better performance than the full framework in 10 labeled data case, the difference is very small (around 0.18%). Therefore, this comparative analysis could show that the proposed fine-grained domain-aware aspect similarity measure is effective for adjusting the contribution from different discovered aspects.

When comparing with the EM model [22] with similar network architecture but having an equal contribution for the source domains, the result shows that varying the contribution based on the domain-aware aspect similarity leads to a better performance.

We observe that our proposed framework has a small performance gain when giving more labeled training data, besides the case from 10 to 20. The EM model also has similar problem as mentioned in [22]. However, the BERT model [1] has an opposite behavior, which has a steady performance gain. We believe that the reason is due to the compact architecture of the topic-attention network which prevents overfitting the limited labeled data in order to have a better domain adaptation. Increasing the learning capability of the model and at the same time handling domain adaptation could be a future research direction.

5 Limitations and Future Works

The proposed framework involves two separate models handling their own jobs. These models do not share any learning parameters. Many works report that the single model handling various tasks would have a better generalization and thus leads to a better performance. One possible future work might consider integrating both

Table 1: Sentiment classification accuracy of different models

# Labeled Data	Model	Book	DVD	Electronics	Kitchen	Video	Average
10 (40)	BERT	51.17 ± 1.25	50.86 ± 1.06	50.91 ± 1.52	51.09 ± 2.54	51.35 ± 2.21	51.07
	BO	59.90 ± 1.88	58.70 ± 3.66	64.40 ± 2.30	63.77 ± 2.14	61.52 ± 3.28	61.66
	MoE	55.35 ± 3.65	56.12 ± 3.94	58.15 ± 4.87	57.37 ± 4.32	56.45 ± 3.82	56.69
	EM	67.16 ± 5.03	67.68 ± 4.55	67.01 ± 5.29	66.47 ± 5.40	67.08 ± 3.67	67.08
	ASM	**71.48 ± 4.70**	**72.56 ± 5.97**	**70.56 ± 4.44**	**71.83 ± 3.97**	**72.21 ± 4.27**	**71.73**
20 (80)	BERT	56.14 ± 6.14	54.22 ± 5.73	54.10 ± 4.70	54.70 ± 5.27	54.27 ± 5.60	54.69
	BO	59.97 ± 1.85	61.34 ± 2.65	65.28 ± 3.40	66.55 ± 2.54	63.47 ± 3.03	63.32
	MoE	59.65 ± 4.99	60.09 ± 5.66	61.07 ± 5.38	61.65 ± 5.09	60.94 ± 5.24	60.68
	EM	72.27 ± 2.67	71.37 ± 3.97	72.16 ± 2.74	71.84 ± 2.60	70.52 ± 1.38	71.63
	ASM	**76.85 ± 2.41**	**79.91 ± 1.85**	**77.73 ± 3.51**	**80.24 ± 1.58**	**79.03 ± 1.85**	**78.75**
30 (120)	BERT	57.40 ± 5.87	56.94 ± 6.88	53.77 ± 5.40	52.96 ± 2.05	59.48 ± 7.99	56.11
	BO	61.26 ± 2.03	62.78 ± 2.32	67.29 ± 2.89	66.17 ± 3.11	64.39 ± 2.51	64.38
	MoE	61.71 ± 5.47	63.13 ± 5.68	63.30 ± 6.43	64.20 ± 6.06	63.36 ± 5.92	63.14
	EM	74.61 ± 2.54	74.70 ± 1.35	74.81 ± 2.03	74.31 ± 1.10	73.25 ± 1.97	74.34
	ASM	**78.91 ± 2.13**	**81.42 ± 1.07**	**80.11 ± 1.58**	**81.51 ± 1.06**	**80.51 ± 1.75**	**80.49**
40 (160)	BERT	68.90 ± 8.55	65.70 ± 9.25	66.38 ± 8.45	69.19 ± 8.69	65.19 ± 9.77	67.07
	BO	62.23 ± 1.25	63.63 ± 2.45	69.57 ± 2.07	70.04 ± 2.22	66.05 ± 1.75	66.30
	MoE	64.82 ± 4.47	67.12 ± 5.22	67.78 ± 6.05	68.00 ± 5.89	66.46 ± 5.36	66.84
	EM	74.78 ± 1.06	75.34 ± 2.24	74.75 ± 1.89	76.38 ± 1.16	75.13 ± 1.77	75.27
	ASM	**80.41 ± 1.13**	**82.16 ± 0.86**	**81.07 ± 1.38**	**82.23 ± 1.02**	**81.26 ± 1.71**	**81.43**
50 (200)	BERT	79.38 ± 4.79	76.36 ± 8.25	76.56 ± 7.01	78.30 ± 8.08	74.79 ± 9.68	77.08
	BO	63.13 ± 2.63	63.76 ± 2.14	70.32 ± 1.93	69.57 ± 2.56	66.24 ± 1.67	66.60
	MoE	67.80 ± 2.24	70.06 ± 2.59	70.99 ± 2.59	70.87 ± 2.68	70.34 ± 2.76	70.01
	EM	75.04 ± 1.85	75.79 ± 1.96	75.38 ± 2.34	75.73 ± 2.27	74.74 ± 1.42	75.33
	ASM	**81.20 ± 0.86**	**82.56 ± 0.88**	**81.13 ± 1.30**	**83.02 ± 0.87**	**82.18 ± 1.01**	**82.02**

Table 2: Comparative analysis of the proposed framework.

# Labeled Data	Model	Avg. Accuracy
10 (40)	rand. select data + avg. pred.	70.98
	avg.	**71.91**
	ASM	71.73
20 (80)	rand. select data + avg. pred.	76.47
	avg.	78.18
	ASM	**78.75**
30 (120)	rand. select data + avg. pred.	78.03
	avg.	79.75
	ASM	**80.49**
40 (160)	rand. select data + avg. pred.	79.25
	avg.	80.72
	ASM	**81.43**
50 (200)	rand. select data + avg. pred.	79.63
	avg.	81.08
	ASM	**82.02**

models together forming an unified model to take the advantage of multi-task learning. This might further improve the performance for the sentiment classification task.

6 Conclusion

We study the task of multi-source cross-domain sentiment classification under the constraint of little labeled data. We propose a novel framework exploiting domain-aware aspect similarity to identify the contribution of discovered fine-grained aspect topics. This fine-grained similarity measure aims at addressing the negative effect of domain-specific aspects appearing in the existing coarse-grained domain similarity measure, and also the limitation caused by the constraint of little labeled data. Aspect topics are extracted by the proposed domain-aware topic model in an unsupervised way. The topic-attention network then learns the transferable sentiment knowledge based on the selected data related to discovered aspects. The framework finally makes predictions according to the aspect proportion of the testing data for adjusting the contribution of various aspect models. Extensive experiments show that our proposed framework achieves the state-of-the-art performance. The framework achieves a good performance, i.e. around 71%, even though there are only 40 labeled data. The performance reaches around 82% when there are 200 labeled data. This shows that our proposed fine-grained domain-aware aspect similarity measure is very effective under the constraint of little labeled data.

References

[1] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv preprint arXiv:1810.04805, 2018, doi:10.18653/v1/N19-1423.

[2] B. Myagmar, J. Li, S. Kimura, "Cross-Domain Sentiment Classification With Bidirectional Contextualized Transformer Language Models," IEEE Access, 163219–163230, 2019, doi:10.1109/ACCESS.2019.2952360.

[3] J. Zhou, J. Tian, R. Wang, Y. Wu, W. Xiao, L. He, "SentiX: A Sentiment-Aware Pre-Trained Model for Cross-Domain Sentiment Analysis," in Proceedings of the 28th International Conference on Computational Linguistics, 568–579, 2020, doi:10.18653/V1/2020.COLING-MAIN.49.

[4] S. Ben-David, J. Blitzer, K. Crammer, F. Pereira, "Analysis of Representa-tions for Domain Adaptation," in Advances in Neural Information Processing Systems, **19**, 2007, doi:10.7551/mitpress/7503.003.0022.

[5] J. Blitzer, M. Dredze, F. Pereira, "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in Proceedings of the 45th annual meeting of the association of computational linguistics, 440–447, 2007.

[6] H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, "Domain-adversarial neural networks," NIPS 2014 Workshop on Transfer and Multi-task learning: Theory Meets Practice, 2014, doi:10.1007/978-3-319-58347-1 10.

[7] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, "Domain-adversarial training of neural networks," The Journal of Machine Learning Research, **17**(1), 2096–2030, 2016, doi:10.1007/978-3-319-58347-1 10.

[8] Z. Li, Y. Zhang, Y. Wei, Y. Wu, Q. Yang, "End-to-End Adversarial Memory Network for Cross-domain Sentiment Classification," in Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2237–2243, 2017, doi:10.24963/IJCAI.2017/311.

[9] Z. Li, Y. Wei, Y. Zhang, Q. Yang, "Hierarchical attention transfer network for cross-domain sentiment classification," in Thirty-Second AAAI Conference on Artificial Intelligence, 5852–5859, 2018, doi:10.1609/AAAI.V33I01.33015773.

[10] K. Zhang, H. Zhang, Q. Liu, H. Zhao, H. Zhu, E. Chen, "Interactive Attention Transfer Network for Cross-Domain Sentiment Classification," in Proceedings of the AAAI Conference on Artificial Intelligence, 5773–5780, 2019, doi:10.1609/aaai.v33i01.33015773.

[11] Q. Xue, W. Zhang, H. Zha, "Improving domain-adapted sentiment classification by deep adversarial mutual learning," in Proceedings of the AAAI Conference on Artificial Intelligence, 9362–9369, 2020, doi:10.1609/AAAI.V34I05.6477.

[12] C. Du, H. Sun, J. Wang, Q. Qi, J. Liao, "Adversarial and Domain-Aware BERT for Cross-Domain Sentiment Analysis," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 4019–4028, 2020, doi:10.18653/v1/2020.acl-main.370.

[13] S. Zhao, B. Li, C. Reed, P. Xu, K. Keutzer, "Multi-source Domain Adaptation in the Deep Learning Era: A Systematic Survey," arXiv preprint arXiv:2002.12169, 2020.

[14] S. Zhao, Y. Xiao, J. Guo, X. Yue, J. Yang, R. Krishna, P. Xu, K. Keutzer, "Curriculum CycleGAN for Textual Sentiment Domain Adaptation with Multiple Sources," in The Web Conference (WWW), 2021, doi:10.1145/3442381.3449981.

[15] F. Wu, Y. Huang, "Sentiment Domain Adaptation with Multiple Sources," in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (1: Long Papers), 301–310, 2016, doi:10.18653/v1/P16-1029.

[16] M. Yu, X. Guo, J. Yi, S. Chang, S. Potdar, Y. Cheng, G. Tesauro, H. Wang, B. Zhou, "Diverse Few-Shot Text Classification with Multiple Metrics," in Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, **1** (Long Papers), 1206–1215, 2018, doi:10.18653/v1/N18-1109.

[17] H. Guo, R. Pasunuru, M. Bansal, "Multi-Source Domain Adaptation for Text Classification via DistanceNet-Bandits," in The Thirty-Fourth AAAI Conference on Artificial Intelligence, 7830–7838, 2020, doi:10.1609/AAAI.V34I05.6288.

[18] Y.-B. Kim, K. Stratos, D. Kim, "Domain Attention with an Ensemble of Experts," in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (1: Long Papers), 643–653, 2017, doi:10.18653/v1/P17-1060.

[19] J. Guo, D. Shah, R. Barzilay, "Multi-Source Domain Adaptation with Mixture of Experts," in Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 4694–4703, 2018, doi:10.18653/v1/D18-1498.

[20] M. Yang, Y. Shen, X. Chen, C. Li, "Multi-Source Domain Adaptation for Sentiment Classification with Granger Causal Inference," in Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 1913–1916, 2020, doi:10.1145/3397271.3401314.

[21] Y. Dai, J. Liu, X. Ren, Z. Xu, "Adversarial Training Based Multi-Source Unsupervised Domain Adaptation for Sentiment Analysis," in The Thirty-Fourth AAAI Conference on Artificial Intelligence, 7618–7625, 2020, doi:10.1609/AAAI.V34I05.6262.

[22] K. Lai, J. C. Ho, W. Lam, "Ensemble Model for Multi-Source Cross-Domain Sentiment Classification with Little Labeled Data," in 2020 IEEE/WIC/ACM International Conference on Web Intelligence (WI), IEEE, 2020, doi:10.1109/WIIAT50758.2020.00038.

[23] K. Shaukat, T. M. Alam, M. Ahmed, S. Luo, I. A. Hameed, M. S. Iqbal, J. Li, M. A. Iqbal, "A Model to Enhance Governance Issues through Opinion Extraction," in 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 0511–0516, 2020, doi:10.1109/IEMCON51383.2020.9284876.

[24] H. Gupta, S. Pande, A. Khamparia, V. Bhagat, N. Karale, et al., "Twitter Sentiment Analysis Using Deep Learning," in IOP Conference Series: Materials Science and Engineering, 012114, 2021, doi:10.1088/1757-899X/1022/1/012114.

[25] A. Badgaiyya, P. Shankarpale, R. Wankhade, U. Shetye, K. Gholap, S. Pande, "An Application of Sentiment Analysis Based on Hybrid Database of Movie Ratings," International Research Journal of Engineering and Technology (IRJET), **8**, 655–665, 2021.

[26] L. Zhang, S. Wang, B. Liu, "Deep learning for sentiment analysis: A survey," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, **8**(4), e1253, doi:10.1002/widm.1253.

[27] J. Blitzer, R. McDonald, F. Pereira, "Domain adaptation with structural correspondence learning," in Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, 120–128, 2006, doi:10.3115/1610075.1610094.

[28] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, Z. Chen, "Cross-domain sentiment classification via spectral feature alignment," in Proceedings of the 19th International Conference on World Wide Web, 751–760, 2010, doi:10.1145/1772690.1772767.

[29] D. Bollegala, T. Mu, J. Y. Goulermas, "Cross-domain sentiment classification using sentiment sensitive embeddings," IEEE Transactions on Knowledge and Data Engineering, **28**(2), 398–410, 2015, doi:10.1109/TKDE.2015.2475761.

[30] K. Crammer, M. Kearns, J. Wortman, "Learning from Multiple Sources," Journal of Machine Learning Research, **9**(57), 1757–1774, 2008.

[31] S. Li, C. Zong, "Multi-domain Sentiment Classification," in Proceedings of ACL-08: HLT, Short Papers, 257–260, 2008.

[32] P. Luo, F. Zhuang, H. Xiong, Y. Xiong, Q. He, "Transfer learning from multiple source domains via consensus regularization," in Proceedings of the 17th ACM conference on Information and knowledge management, 103–112, 2008, doi:10.1145/1458082.1458099.

[33] H. Zhao, S. Zhang, G. Wu, J. M. Moura, J. P. Costeira, G. J. Gordon, "Adversarial multiple source domain adaptation," in Advances in Neural Information Processing Systems, 8568–8579, 2018.

[34] X. Chen, C. Cardie, "Multinomial Adversarial Networks for Multi-Domain Text Classification," in Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, **1** (Long Papers), 1226–1240, 2018, doi:10.18653/v1/N18-1111.

[35] D. P. Kingma, M. Welling, "Auto-Encoding Variational Bayes," in 2nd International Conference on Learning Representations, ICLR, 2014.

[36] H. Zhao, D. Phung, V. Huynh, Y. Jin, L. Du, W. Buntine, "Topic Modelling Meets Deep Neural Networks: A Survey," arXiv preprint arXiv:2103.00498, 2021.

[37] D. M. Blei, A. Y. Ng, M. I. Jordan, "Latent Dirichlet Allocation," Journal of Machine Learning Research, **3**, 993–1022, 2003, doi:10.1016/B978-0-12-411519-4.00006-9.

[38] M. Figurnov, S. Mohamed, A. Mnih, "Implicit Reparameterization Gradients," in Advances in Neural Information Processing Systems, 439–450, Curran Associates, Inc., 2018.

[39] W. Joo, W. Lee, S. Park, I.-C. Moon, "Dirichlet Variational Autoencoder," CoRR, **abs/1901.02739**, 2019, doi:10.1016/j.patcog.2020.107514.

[40] H. Zhang, B. Chen, D. Guo, M. Zhou, "WHAI: Weibull Hybrid Autoencoding Inference for Deep Topic Modeling," in International Conference on Learning Representations, 2018.

[41] C. Naesseth, F. Ruiz, S. Linderman, D. Blei, "Reparameterization Gradients through Acceptance-Rejection Sampling Algorithms," in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 489–498, 2017.

[42] S. Burkhardt, S. Kramer, "Decoupling Sparsity and Smoothness in the Dirichlet Variational Autoencoder Topic Model," Journal of Machine Learning Research, **20**(131), 1–27, 2019.

[43] Y. Ganin, V. Lempitsky, "Unsupervised Domain Adaptation by Backpropagation," in Proceedings of the 32nd International Conference on International Conference on Machine Learning, 1180–1189, 2015.

[44] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, "Distributed representations of words and phrases and their compositionality," in Advances in Neural Information Processing Systems, 3111–3119, 2013.

[45] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[46] S. Ruder, B. Plank, "Learning to select data for transfer learning with Bayesian Optimization," in Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 372–382, 2017, doi:10.18653/v1/D17-1038.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 07-11 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Light Modulation Enhancement by using an Impedance Matching Scheme for a Subcarrier Multiplexed Light Transmitter

Seiji Fukushima*, Satoshi Yanagihara, Toshio Watanabe, Tsutomu Nagayama

Graduate School of Science and Engineering, Kagoshima University, Korimoto 1, Kagoshima-shi, 890-0065, Japan

ARTICLE INFO

Article history:
Received: 30 July, 2021
Accepted: 15 October, 2021
Online: 06 November, 2021

Keywords:
Radio-on-Fiber
Laser Diode
Subcarrier Multiplexing
Impedance Matching

ABSTRACT

We propose and analyze new impedance matching schemes to enhance applied voltage to an optical modulator and light modulation for a subcarrier multiplexed light transmitter or a radio-on-fiber transmitter that carries radio-frequency signal through an optical fiber. Our proposal includes two methods using a quarter-wavelength impedance transformer and a tapered microstrip line for impedance matching between a driver circuit and an electro-absorption modulator. Simulation results for both two schemes show that large enhancement is observed for 20 GHz and lower frequency and that some resonant boosts are observed for higher frequency as well. Discussions are described to design a circuit with improved performances. It is shown that our schemes can drive the electro-absorption modulator at a voltage higher than 1 V that is sufficient to drive the modulator.

1. Introduction

The use of optical fiber network has become more popular to send both broadband digital data [1] and modulated radio frequency (RF) signal [2]. The latter is referred to as radio-on-fiber (RoF) technology, where the signal is transmitted as subcarrier-multiplexed (SCM) laser light. The digital network requires optical devices such as a laser diode (LD), a light modulator, and a photodiode to have broadband response. In contrast, the RoF network does not require such broadband response but requires only a few spot-frequency or segmented-frequency responses. Such RoF systems are employed in a network between a data center and an access point for cellular services, where radio carrier frequency becomes higher, especially at the fifth generation cellular services using a Ka-band [3]. In addition to these services, some satellites employ a Ku-band so that the use of optical fiber link is practical between an antenna and a control site in an earth station. For example, some small cube satellites, developed by universities and consortiums, sent the Ku-band signals to earth stations in its downlink system [4]. A RoF link was employed in the earth station to reduce RF signal loss in conventional coaxial cables [5-7]. In addition to the low link loss, the optical fiber will benefit us in many aspects such as weight and cost.

To achieve a higher frequency operation and to avoid chromatic dispersion in an optical fiber, an electro-absorption modulator integrated laser diode or an externally modulated laser (EML) would be one of the strong candidates to develop a SCM transmitter. The EML is monolithic integration of an electro-absorption modulator (EAM) and a distributed feedback laser diode (DFB-LD), as shown in Figure 1. A current source is connected to the DFB-LD to obtain continuous laser light and a digital or a RF source to the EAM to obtain modulated light. The EML emits SCM light when the RF signal is applied to the EAM. Both the optical devices can concentrate on either high-speed modulation or single-mode emission of laser light [8, 9]. Most EMLs are designed for 1.5-μm wavelength networks and are widely used in commercial optical networks due to their performances, compactness, and reliability. In addition to the commercial optical networks, they are applied to the shorter-length, high-throughput Ethernet services [10-12], which makes an optical transmitter smaller in size and compatible with optical transmitter standards. Challenges have been done to develop a 1.3-μm wavelength device [11].

A link length limit due to chromatic dispersion has been analyzed for the RoF link in the satellite earth station that employs the EML [13], which is different from the digital network systems. Effort has been made to achieve a more-functional device such as radio frequency conversion in the SCM light [14].

One of the interesting approaches to enhance the response of EMLs is the use of via connection in three-dimensional substrate to mount the EML [12]; however, this technique needs the state-of-the-art manufacturing process. On the other hand, this paper focuses on a passive transmission line between a RF source and the EAM, as shown in Figure 2. The EAM is a reversely biased

*Corresponding Author: Seiji Fukushima, fukushima@eee.kagoshima-u.ac.jp

https://dx.doi.org/10.25046/aj060602

diode with an external resistive load of 50 Ω in general. The conventional 50-Ω matching may be a bottle neck for driving the device. Our idea to overcome this issue is a simple matching scheme such as a quarter-wavelength impedance transformer (QIT) and a tapered microstrip line (TML) that can match the driver and the modulator with different impedance [15-17]. Hence, we can increase the external load resistance to increase the applied voltage to the modulator and its modulation index.

Our proposed scheme is described in Section 2 and some results in Section 3. We discuss characteristics and features of the QIT and TML in Section 4. We will conclude this paper in Section 5.

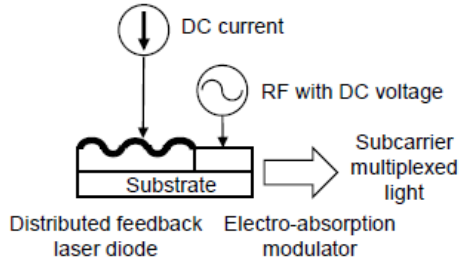


Figure 1: Cross Section of an Electro-Absorption Modulator Integrated Laser Diode (EML) and Power Sources

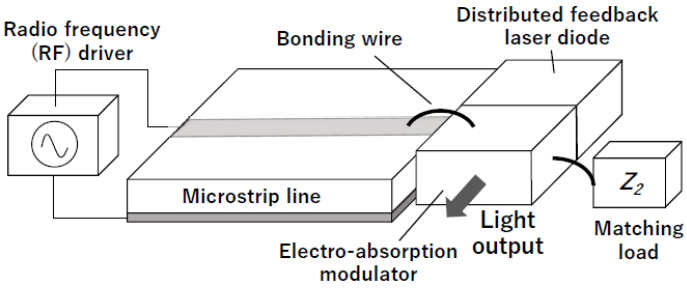


Figure 2: Schematic Model Including a RF Source, a Transmission Line, and an EML

2. Proposed Structure and Circuit Model

A circuit around the EML is shown in Figure 2. A RF signal is transferred from a RF driver to the EAM in the EML through a microstrip line. Impedance of the driver, the EML, and the microstrip line is 50 Ω in general. We should notice the reversely biased EML has very high impedance but it can be impedance-matched by using an external parallel resistance of approximately 50 Ω. We could choose favorite impedance if the entire circuit can be matched by other methods. We propose the use of QIT and TML that are shown in Figure 3 (a) and (b), respectively. The QIT must follow equation below

$$Z_0^2 = Z_1 Z_2 \quad (1)$$

where Z_0 is microstrip line impedance, and Z_1 and Z_2 impedance of the RF driver and the parallel load to the EAM, respectively, and the microstrip line length must be a quarter of a wavelength (λ/4) or multiplied length by odd numbers (3λ/4, 5λ/4…). If the TML is long enough and if the impedance at the both end matches with adjacent circuits, no electric reflection occurs at the boundaries.

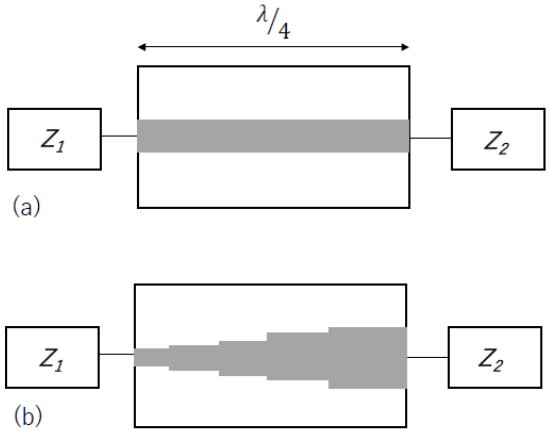


Figure 3: Structure of the Transmitter with (a) a Quarter-Wavelength Impedance Transformer (QIT) and (b) a Tapered Microstrip Line (TML).

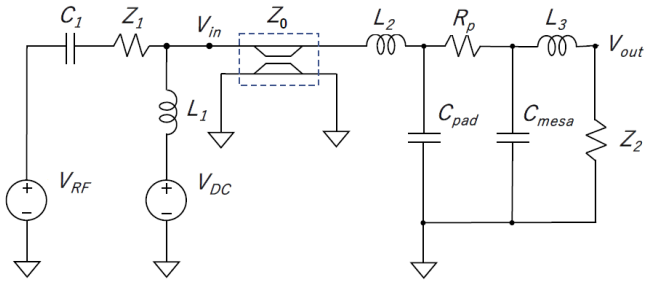


Figure 4: Equivalent Circuit Consisting of a RF Source, a Transmission Line, and an EAM.

Figure 4 shows the entire equivalent circuit of the driver, the RF transmission line and, the EAM. A symbol, V_{RF}, on the left hand with its load impedance Z_1 is the RF driver and the next the reverse voltage source, where the both sources are connected to a RF line Z_0 through a bias-tee of C_1 and L_1. The RF line Z_0 is placed on the center to connect the RF source and the EAM. All the electric signals, dc and ac, pass through Z_0 to the EAM section consisting of two inductors, one resistor, and two capacitors; L_2 and L_3: inductances of bonding wires to the microstrip and to the external load, respectively, R_p: a resistance of electrode, C_{mesa}: a capacitance of the modulator diode, and C_{pad}: a capacitance of an electrode pad. An external matching resistive load Z_2 is connected to the EAM in parallel. In our simulation, we used the values as follows: V_{RF}=1 V, R_p =5 Ω, L_1=79.4 μH, L_2=0.11 nH, L_3=0.04 nH, C_1=51 nF, C_{pad}=0.029 pF, C_{mesa}=0.45 pF. These values have been obtained from measurements or theoretical calculations. Other parameters will be described when they are needed. We assumed the central operation frequency to be 28 GHz, one of the fifth generation cellular service bands. We analyzed the proposed circuits by using a circuit simulator LTspice. A tapered microstrip line circuit model was not available so that cascaded, stepped microstrip line (Tline) models, shown, in Figure 3(b), were employed instead. The number of steps was five. As a result, we were able to use the LTspice circuit models including Tline for all the components. Any commercially available substates were not assumed except for MEGTRON7 to have general discussions in this paper.

If we assume that the DFB-LD in the EML is operated with an automatic-power-control current supply, the DFB-LD emits

continuous single-mode laser beam. The EAM modulates the laser beam into SCM laser light that carries information of RF signal.

3. Simulation Results

3.1. Results for Quarter-wavelength Impedance Transformer

Assuming a dielectric constant of the substrate as 3.4 and the output impedance of EML, Z_2, as 50 (gray), 70 (blue), or 100 Ω (orange), we calculated terminal voltage spectra for circuits with QIT, as shown in Figure 5. Here, the impedance of the RF driver, Z_1, can be calculated as 50, 35.7, or 25 Ω by using Eq. (1). The terminal voltage is defined by the voltage across the EAM. The voltage at the lower frequency or at the left edge in Figure 5 is higher for the higher impedance and the enhancement ratio is nearly proportional to Z_2. This tendency can be found at most of the frequency ranges between 1 and 20 GHz. The enhancement is obtained simply due to higher voltage division in Z_2. The voltage curves decrease as the frequency becomes higher in a case, Z_2 = 50 Ω, which is the same with the conventional structure, while some peaks are found for Z_2 of 70 and 100 Ω. The peaks are obtained at 6.1, 30, 60, and 92 GHz for Z_2 of 100 Ω. In addition to the simple boost of the voltage in the lower frequency range, we obtained response peaking in the higher frequency range, which is due to resonance by some inductive and capacitive elements.

We obtained powerful voltage boost by using the QIT microstrip line between the driver and the EML with the load impedance higher than 50 Ω, as expected. The resonant peaks enable the applied voltage boost at three or four frequencies.

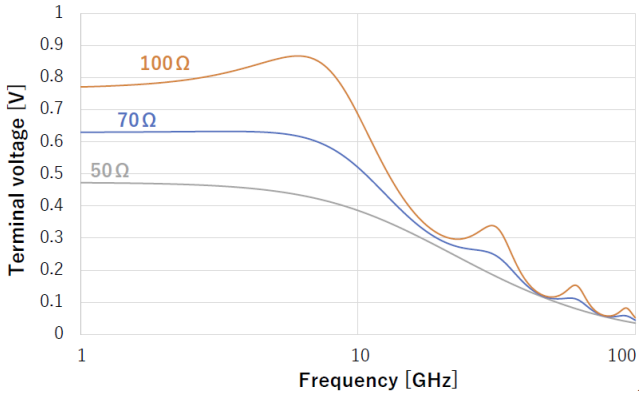


Figure 5: Spectra of the Terminal Voltage across the EAM when the QIT is Employed.

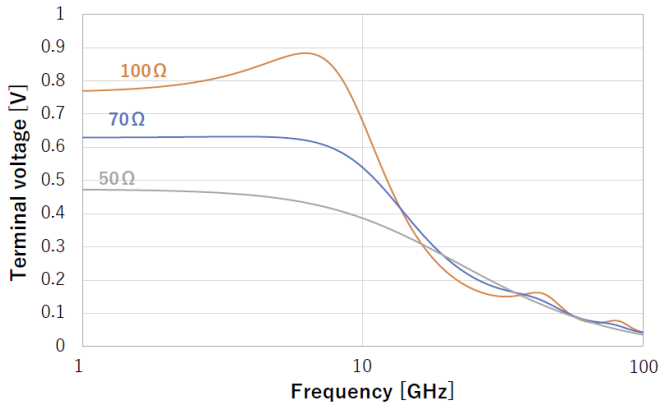


Figure 6: Spectra of the Terminal Voltage across the EAM when the TML is Employed.

3.2. Results for Tapered Microstrip Line

Figure 6 shows the calculated terminal voltage spectra for the circuit including the TML. The results are the same with the results of QIT circuit qualitatively. In the lower frequency range, the terminal voltage is amplified with magnification of Z_2/Z_1 approximately, which means that we can obtain higher voltage from the higher impedance of Z_2. In the higher frequency range, fewer number of resonant peaks are found and their amplitudes are smaller even if they exist. The peak at 6.4 GHz is observed as very similar to the results of QIT while two lower peaks are observed at 41.7 and 79.8 GHz for Z_2 of 100 Ω.

When we have to determine which circuit we should adopt, we have to note that both will function at the same performance for frequencies lower than 20 GHz but that the details differ between the QIT and the TML. The QIT exhibits a higher voltage or a high light modulation index when we manufacture all the devices and a circuit, as specified and as designed. If we are not able to manufacture the devices as specified nor as designed, the TML would be a good choice since it allows some margin in the resonant frequencies to provide lower and broader peaks.

For clearer comparison, we show terminal voltage dependence on the load impedance Z_2 for a frequency at 1 GHz (blue), the first resonant frequency near 6 GHz (orange), and the second resonant frequency near 30 GHz (gray) in Figure 7. The plots and lines are overlapped for the two models of QIT and TML at 1 GHz and the first resonant frequency while they differ at the second resonant frequency.

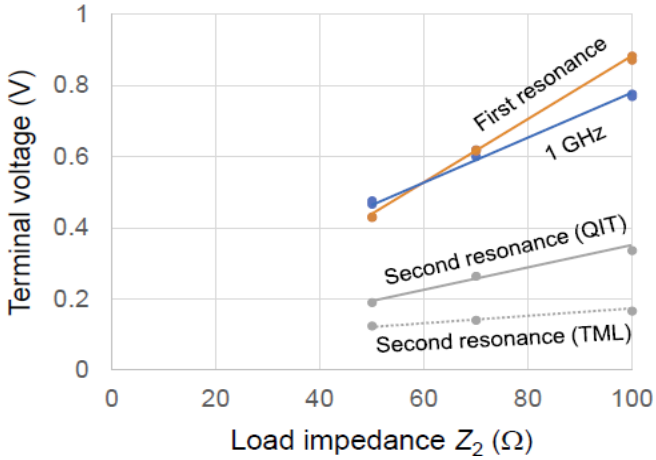


Figure 7: Terminal Voltage Dependence on the load impedance Z_2.

The terminal voltages always increase as the load impedance becomes higher. All the lines are linear in a Z_2 range between 50 and 100 Ω; however, the number of calculated points of the load impedance is limited to only three. Thus, we can double the voltage at 1GHz and the first resonant frequency when we employ a load impedance of 100 Ω other than the original 50 Ω. When we pay our attention to the second resonance, we should choose the QIT model (solid line) since it provides much higher voltages than the TML model (broken line). In both cases, the terminal voltages are even higher with the 100-Ω load than the voltages with the original 50-Ω load. Our ideas are worth applying to the optical transmitter especially when we consider that the second resonant

frequency, 30 GHz, is within the fifth-generation cellular service band.

4. Discussion

Difference of the improvement between the two proposed structures, QIT and TML, is small at a 20-GHz frequency and lower. However, we should analyze the results in the previous section for better performance at higher frequencies. Resonant frequency dependences on the order of the resonance are shown in Figures 8 and 9 for the QIT and TML, respectively. We plotted these figures, assuming dielectric constants to be 3.4, 10, and 15, where some parameters of MEGTRON7 were adopted for the constant of 3.4. We took account of dielectric losses due to the substrates. The load impedance to the EML, Z_2, was assumed to be 100 Ω. The number of the orders is larger within the plotted frequency range as the dielectric constant becomes larger. The number of plotted points are three, four, and, five for dielectric constants of 3.4, 10, and, 15, respectively, as shown in Figure 8. The relationship between the resonant frequency and the order is close to a linear function. The number of the points for the QIT is larger than that for the TML. We obtained the dependency on the dielectric constant, the order of resonant peak, and the type of the matching structure, which provides us more flexibility in designing and manufacturing the matching circuits.

The results described above focus only on the resonant frequency; however, the terminal voltage at a resonant frequency is of importance as well. Figures 10 and 11 show a voltage across the EAM versus a resonant frequency for the QIT and TML, respectively. The employed parameters are the same with those for Figures 8 and 9. The figures include results for the conventional 50-Ω-matching circuit. The curves are overlapped for all the three dielectric constants and the voltages are inversely proportional to the logarithmic plot of the frequencies, as shown in Figure 10. At a 30-GHz frequency near the fifth generation cellular band, the terminal voltage of 0.34 V is achieved for all the dielectric constants. We obtained only 0.2 V from the conventional structure in contrast. Similar results are obtained for the TML, as shown in Figure 11. Here, we have to note the terminal voltages are dependent on a dielectric constant at the TML model. The lower dielectric constant substrate provides us higher voltage.

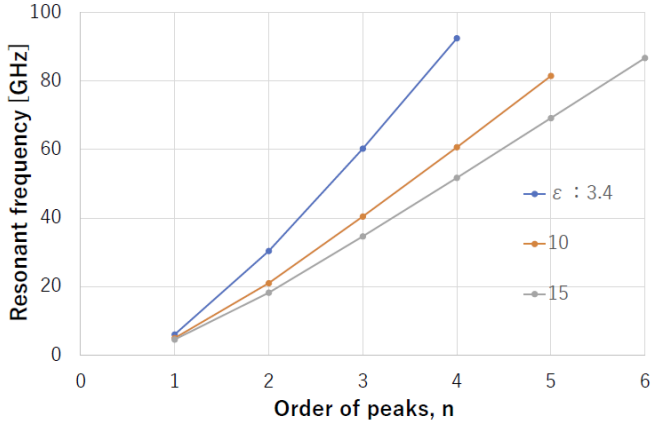


Figure 8: Resonant Frequency Dependence on the Order of Peak for the QIT Model.

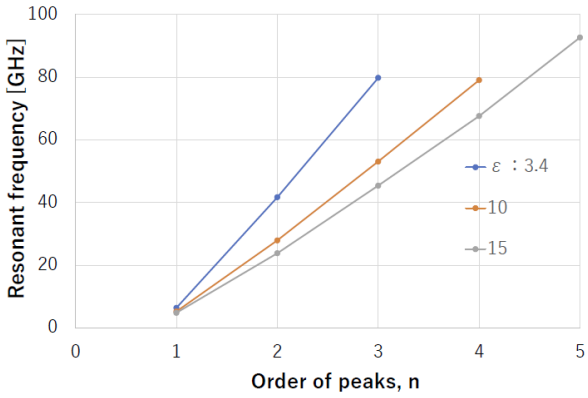


Figure 9: Resonant Frequency Dependence on the Order of Peak for the TML Model.

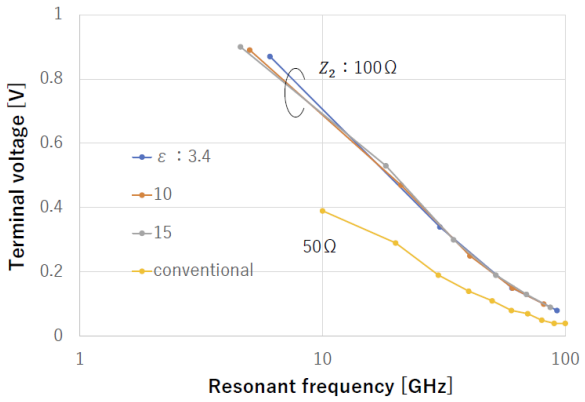


Figure 10: Terminal Voltage across the EAM versus Resonant Frequency for the QIT Model.

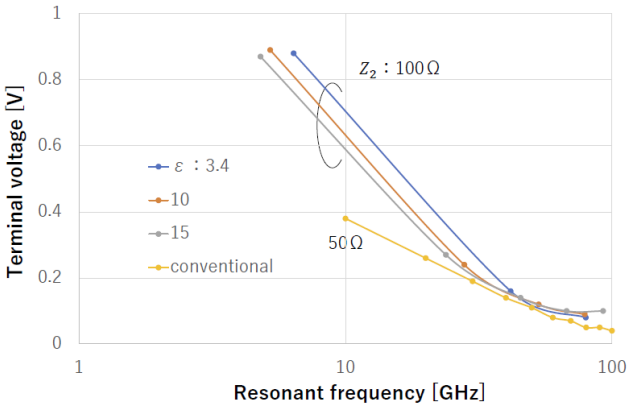


Figure 11: Terminal Voltage across the EAM versus Resonant Frequency for the TML Model.

Numeric analysis has been successfully done for the proposed structure that employs a higher impedance load in parallel to the EAM. In addition, we analyzed the behavior at higher impedance, Z_2, for the QIT. Figure 12 shows a terminal voltage dependence on a resonant frequency for the QIT circuit with the impedance Z_2 of 500 Ω. The tendency is similar qualitatively with the previous results. It should be noted that the terminal voltage becomes approximately 5.5 times those obtained from the conventional circuit. The voltages at resonant frequencies 30 and 60 GHz are 1.05 and 0.52 V, respectively, which are sufficient to drive the EML and to transmit the SCM laser light. This trial may not be applied to the TML model unfortunately since such a high-

impedance microstrip line cannot be fabricated on a typical substrate.

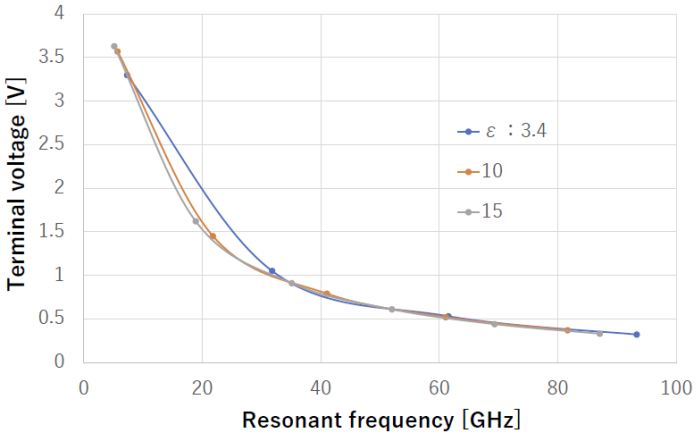


Figure 12: Terminal Voltage across the EAM versus Resonant Frequency for the QIT Model, Assuming the Load Impedance, Z_2, is 500 Ω.

5. Conclusions

We have proposed new structures to boost applied voltage to the EAM and its modulation index in a SCM or RoF transmitter that carries radio-frequency signal through an optical fiber. The two proposed models employ a QIT and a TML, both are capable of impedance matching for two adjacent devices with different impedance. Our analysis has been done numerically by using an electronic circuit simulator. Comparing with the conventional circuit, the voltage applied to the modulator increases approximately at the ratio of the load impedance to the transmission line impedance at 20 GHz and lower frequency in both the structures. Due to the capacitive and inductive elements in a circuit, some resonant peaks were found at higher frequency in both the models. The applied voltage to the EAM is enhanced at the resonant frequencies.

To further improve performances, we have analyzed characteristics of our models with parameters of dielectric constant of substrates and load impedance. The results are similar between the two models. The impedance of the QIT microstrip line is constant and a standard one of 50 Ω. On the other hand, the impedance must be varied in a wide range from Z_1 to Z_2 on a TML substrate, which might be difficult to fabricate in some cases. We should note the matter described above in designing a circuit.

It is finally concluded our proposal does not require any active device but that it can enhance the voltage across the EAM and modulation index of the SCM light. This technique would be easily applied to most of the light transmitters such as a direct modulation laser and a Mach-Zender light modulator. In addition, this kind of RoF link would be applied to higher-frequency applications [18].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 17K06432.

References

[1] E. B. Basch, R. Egorov, S. Gringeri, S. Elby, "Architectural tradeoffs for reconfigurable dense wavelength-division multiplexing systems," IEEE Journa; of Selected Topics in Quantum Electronics, **12**(4), 615-626, 2006. DOI: 10.1109/JSTQE.2006.876167

[2] S. Torihata, H. Tsuchiya, K. Haeiwa, T. Yamashita, S. Tanizawa, Y. Ozaki, "The optical remote antenna for the digital TV broadcasting relay networking," International Topical Meeting on Microwave Photonics, Awaji, Japan, 375-378, 2002.

[3] M. Sung, J. Kim, E.-S. Kim, S.-H. Cho, Y.-J. Won, B.-C. Lim, S.-Y. Pyun, H. Lee, J. K. Lee, J. H. Lee, "RoF-based radio access network for 5G mobile communication systems in 28 GHz millimeter-wave," Journal of Lightwave Technology., **38**(2), 409-420, 2020/doi 10.1109/JLT.2019.2942636.

[4] Y. Sakamoto, N. Nishio, "Determination using radio interefreometer of small-diameter antennas for LEO satellites," IEEE Transaction on Aerospace and Electronic Systems, **47**(3), 2111-2118, 2011.

[5] S. Fukushima, N. Miura, T. Shimaki, T. Funasako, K. Yamashita, T. Hachino, Y, Igarashi, "Electro-absorption modulator integrated laser application to a cube satellite earth station," 2013 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR 2013), paper TuPO-11, 2013..

[6] S. Fukushima, T. Shimaki, T. Funasako, N. Miura, K. Yamashita, T. Hachino, Y. Igarashi, "Cube satellite earth station based on bidirectional WDM radio-on-fiber link," 2013 Asia-Pacific Radio Science Conference, paper DCa-6, 2013.

[7] S. Fukushima, T. Shimaki, K, Yamashita, T. Funasako, T. Hachino, "Frequency division multiplexed radio-on-fiber link employing an electro-absorption modulator integrated laser diode for a cube sattelite erath station," IEICE Transaction on Electronics, **E99-C**(2), 212-218, 2016.

[8] H. Kamioka, M. Fukuda, T. Tohmori, Y. Itaya, "Realiability of an electro-absorption modulator integrated with a distributed feedcak mlaser," Pacific Rim. Conference on Lasers and Electro-Optics (CLEO-PR '99), **4**, 1202-1203, 1999.

[9] K. Wakita, K. Sato, K. Kotaka, M. Yamamoto, T. Ktaoka, "20-Gbit/s, 1.55-mm strained-InGaAsP MQW modulator integrated DFB laser module," Electronics Letters, **30**(4), 302-303, 1993.

[10] H. Takahashi, T. Shimamura, T. Sugiyama, M. Kubota, K. Nakamura, "High-power 25-Gb/s electroabsorption modulator integrated with a laser diode," IEEE Photonic Technology Letters, **21**(10), 633-635, 2009.

[11] T. Fujisawa, K.Takahata, W. Kobayashi, T. tadokoro, N. Fujisawa, S. Kanazawa, F. Kano, "1.3-mm, 50-Gbit/s EADFB lasers for 400GbE," 2011 Optical Fiber Communication Conference and Exposition, paper OWD4, 2011.

[12] X. Zhang, X. Liu, A. Rainier van Dommele, M. K. Matters-Kammerer, "Dual-channel 56 Gb/s PAM-4 electro-absorption modulator driver for 3D wafer scale packaging," 2018 Asia-Pacific Microwave Conference (APMC 2018), paper WE4-K-03, 2018.

[13] T. Uezono, S. Fukushima, T. Watanabe, T. Nagayama, "Link length analysis of a radio-on-fiber system for a cube satellite earth station," Annual Conference on Engineering and Applied Science (ACEAT), 18-23, 2017.

[14] S. Fukushima, Y. Doi, Y. Matsuoka, H. Takeuchi, "EA-modulator/LD integrated device application to mm-wave optical link," IEICE Scoiety Conference, paper B-5-219, 1998.

[15] D. M. Pozer, Microwave Engineering, 4th Ed., Wiley, 2011.

[16] A. Ueda, T. Noguchi, H. Iwashita, Y. Sekimoto, M. Ishiguro, S. Takano, T. Nagatsuma, T. Ito, A. Hirata, T. Ishibashi, "W-band waveguide photomixer using a uni-trabeling-carrier photodiode with 2-mW output," IEEE Transaction on Microwave Theory and Techniques, **51**(5), 1455-1459, 2003. DOI: 10.1109/TMTT.2003.810154

[17] S. Yanagihra, S. Fukushima, T. Watanabe, T. Nagayama, "Impedance-matched feedline fro a smiconductor laser emitting subcarrier multiplexed light," 25th Optoelectronics and Communications Conference, paper T4-4.5, 2020.

[18] M. Nishio, Q. Liu, T. Miyazaki, M. Hirata, Y. Kuroki, M. Kusuhata, N. Iwashita, C. Minamitake, S. Yasuda, N. Iino, T. Omodaka, O. Kameya, N. Kawano, T. Suzuyama, Y. Shibuya, N. Kurihara, "Observation site atmospheric phase fluctuations observed by three-element VLBI, IEEE Transaction on Antennas and Propagation, **55**(7), 2056-2063, 2007.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 08-15 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Predicting School Children Academic Performance Using Machine Learning Techniques

Radwan Qasrawi[1,*], Stephanny VicunaPolo[2] , Diala Abu Al-Halawa[3], Sameh Hallaq[4], Ziad Abdeen[3]

[1]*Department of Computer Science, Al Quds University, Jerusalem, 9103401, Palestine*

[2]*Center of Innovation Technology, Al Quds University, Jerusalem, 9103401, Palestine*

[3]*Faculty of Medicine, Al Quds University, Jerusalem, 9103401, Palestine*

[4]*Department of Economics, Al Quds University, Jerusalem, 9103401, Palestine*

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 21 August, 2021
Online: 10 September, 2021

Keywords:
Data mining
Academic Performance
Machine Learning
Educational Data Mining
Decision Tree
Random Forest
Neural Networks
Logistic Regression
Support Vector Machine

ABSTRACT

The study aims to assess the machine learning techniques in predicting students' associated factors that affect their academic performance. The study sample consisted of 5084 middle and high school students between the ages of 10 and 17, attending public and UNRWA schools in the West Bank. The 'Health Behaviors School Children' questionnaire for the 2013-2014 academic year was used for data collection, and was then analyzed through machine learning techniques in order to evaluate their relationship with student academic outcomes. Six machine learning techniques (Random Forest, Neural Network, Support Vector Machine, Decision Tree, Naïve Bayes, and Logistic Regression) were used for prediction. The results indicated that the logistic regression and Naïve Bayes models had the highest accuracy levels (94.3%, 94%) respectively, followed by a decision tree, Neural Network, Random Forest, and Support Vector Machine (93.3%,91.9%,91.7%, and 80.2%) respectively. Thus, the Logistic Regression and Naïve Bayes had the best performance in classifying and predicting student academic performance with the associated factors. Furthermore, Decision Tree, Random Forest, and Neural Network had better predictive performance than Support Vector Machine. The results indicated that perception, Smoking, Depression, PTSD, Healthy Food Consumption, Age, gender, Grade Level, and Family income are the most important and significant factors that influence student academic performance. Overall, machine learning techniques prove efficient tools for identifying and predicting the features that influence student academic performance. The deployment of machine learning techniques within schools' information systems will facilitate the development of health prevention and intervention programs that will enhance students' academic performance.

1. Introduction

The prediction of students' academic performance is of great concern to educational institutions, governments, parents, and other stakeholders as it would enhance the development of curriculum, teaching methodologies, and planned interventions [1]. Several studies have evidenced that given the multi-causal nature of the concept of academic performance, a wide range of variables, from socio-economic to nutritional, health, and environmental, are directly correlated with it at all levels of education [2]. Good health, nutrition, and physical activity, for instance, are proven to enhance students' cognitive functions and therefore achieve better academic outcomes[3] [4].

Although the correlation between lifestyle habits and academic performance has been established [5], few studies have investigated the uses of Data Mining, namely Machine Learning Techniques among educational environments [6]-[10].

DM is used in various fields of research as it analyzes and extracts useful information and hidden patterns from large data

*Corresponding Author: Radwan Qasrawi. Email: radwan@staff.alquds.edu

https://dx.doi.org/10.25046/aj060502

sets. The implications of DM, particularly ML techniques, in education are to provide new knowledge to help educational practitioners make predictions, informed decisions, as well as boost the shift of educational institutions from a traditional learning model towards automation[10]. Furthermore, ML techniques such as Clustering, Neural Networks (NNs), Decision Trees (DT), Random Forest (RTs), Support Vector Machine (SVMs), Logistic Regression (LRs), Naïve Bayes among others, have proven effective in predicting multicausal concepts such as academic performance [11]-[13].

2. Literature Review

2.1. Academic Performance-Associated Risk Factors

Abundant literature exists regarding Machine Learning and educational outcomes, although studies have found that certain variables do indeed significantly impact academic performance, there is no consistent agreement on a specific set of predictors or risk factors [14]. Nonetheless, some of the common variables yielding correlations include gender, parents occupation, family income, parents education, ethnicity, family size, mental health, post-traumatic stress disorder (PTSD), lifestyle habits, physical activity, and nutrition [3], [15].

In [16], the author demonstrates three key variable categories for improved academic performance: 1) Health behaviors, 2) Nutrition and Cognition, and 3) Obesity. The authors describe health behaviors as being influenced by risk associated factors such as physical activity, aerobic fitness, body mass, and energy consumption [13], [14]. Furthermore, relevant nutrition-associated risk factors, such as protein-energy consumption, breakfast consumption [19], short-term fasting, and sugar intake are among the key variables substantially influencing the development of brain functions.

In [10], the study provides a framework for researchers to use DM and ML models for the prediction of student academic performance. The guidelines consist of six stages: 1) data collection, 2) data initial preparation, 3) statistical analysis, 4) data preprocessing, 5) data mining implementation, and 6) result evaluation. The study has also found that early student performance prediction is not a widely studied area, yet it possesses the potential in aiding universities provide timely actions to improve students' academic outcomes. The study concluded that the most essential factors in predicting academic success are "prior-academic achievement, student demographics, e-learning activity, and psychosocial attributes" [10].

In [20], the authors aimed to predict middle and high-school students' academic achievement with a focus in Science and Arts classes by making use of the Chi-squared Automatic Interaction Detection (CHAID) model. The results yielded that students' performance in arts and science is strongly correlated to psychosocial features such as PTSD, self-reported health perception, age, and instances of depression [21]. However, variables such as good nutrition [15], family support, and gender weighed more on students' science performance. The study found that data mining, namely machine learning, possesses great potential for discovering hidden information that would otherwise be difficult by using conventional analysis [20].

2.2. Machine Learning Techniques

Distinct studies have approached the analysis and prediction of academic achievement through a variety of methods. In [1], the authors review the ML techniques applied to the prediction of students' academic achievement through a study of published research, concluding the following:

- Most literature (around 70% of articles in their review) seeks to study student performance at the university level.
- The most widely used technique for predicting students' behavior is supervised learning.
- Support Vector Machine, DT, Naïves Bayes, and Random Forests are well-studied algorithmic proposals generating positive results.
- Neural Networks as less used, yet predict students' performance more accurately.
- Unsupervised learning has low accuracy in predicting students' behavior based on studies reviewed.

In [12], the authors seek to predict higher education student performance using the Random Forest Classification algorithm on a study of the UCI student performance dataset consisting of 8000 data. The study concludes that the RF algorithm predicts results that are 15% more accurate than the SVM algorithm.

The study in [22] analyzed the methodologies for the prediction of academic performance and the main variables impacting student outcomes. The study found that most experimental datasets focused on undergraduate Science, Technology, Engineering, and Mathematics (STEM) courses from a single educational entity. The majority of datasets were relatively small with a sample size of fewer than 1000 students. Furthermore, the prediction models were commonly developed as stand-alone modules instead of an ensemble technique despite it having better accuracy rates. Most ML techniques used are supervised through regression, neural networks, and tree-based models. Finally, the study concludes that the most effective predictive models are the Hybrid Random Forest, Feedforward 3-L Neural Network, and Naïve Bayes [22].

Seeking to introduce a prediction model for academic performance, the authors in [13] used supervised machine learning techniques like SVM and Logistic Regression. The results show that the Support Vector Machine algorithm yields more accurate results than logistic regression.

In [14], the authors employed a multi-dimensional methodological approach to predict first-year student academic performance at Babcock University with a sample size of 7,500 data. The study used classification algorithms such as Random Forest, Random Tree, REPtree, J48, OneR, Decision stump, JRip, ZeroR, Decision Table, PART, and Artificial Neural Networks. The study developed a Recommender System (RS) to enhance academic performance prediction. The results found that Random Forest, Reptree, J48, JRip, PART, Decision Table, and Artificial Neural Networks outperformed other classifiers with an accuracy rate of 96.78% (the lowest). Nonetheless, Random Tree, with an accuracy rate of 99.908%, proved ideal for this study's objectives [14].

In [9], the authors propose a prediction model for student performance based on a Decision Tree algorithm and through the application of an ensemble method on a UCI dataset for students in the Alentejo region of Portugal. The study used three supervised algorithms, J48, NN, and MLP. The results show that J48, with an accuracy rate of 96.78%, yields the most accurate results.

Thus, several studies have investigated academic performance from distinct perspectives, including teaching and educational factors, physical health, and educational environment [4], [5], [16]. The literature concludes that early prediction of risk factors affecting academic performance is essential in aiding struggling students in self-regulating the factors that can impact their cognitive development [3], [16], [19]. Nonetheless, in this study, we seek to assess machine learning models to predict students' academic status based on three levels of health factors (physical, mental, and social).

3. Research Methodology

In this research, we have used machine learning techniques for the classification and prediction of the influence of health factors on students' academic performance. Five machine learning models were utilized for data analysis. The methodology follows the guidelines outlined in [10] through data collection, data initial preparation, data analysis, data preprocessing, machine learning implementation of the models studied, and finally results evaluation through a confusion matrix and other visualization tools.

3.1. Data Collection

The data used for this research was collected from the National Health and School Behavior study conducted by Al-Quds University and the Ministry of Education (MoE) in the year 2013-2014. The data contains student's health, health behavior, and academic performance indicators. The survey was administered via the 'self-reporting' method of data collection while supervised by appointed school social workers. The research considers features that relate to the social, physical, and mental health of Palestinian children in school grades 5-9 (10-17 years). A representative sample of 5084 students was used in this study. The data includes associated risk factors variables related to mental health, physical health, social health, nutrition, and academic performance. The associated risk factors variables are listed in Table 1 below.

Table 1: Machine Learning Models' variables

Variable Name	Description	Value
Gender	Gender	Boys, Girls
Age	Age	Age (11-16 Years)
FAS	Economic Status	Low, Medium, High
ST	School Type	Public or Refugee
LP	Living Place	Urban or non-Urban)
FatherEdu	Father Education	≤ Secondary or > Secondary
MotherEdu	Mother Education	≤ Secondary or > Secondary
Healthy	Healthy Food Consumption	Low or High
Unhealthy	Unhealthy Food Consumption	Low or High
BMI	Body Mass Index	Normal, Overweight, or Obese
Anemia	Anemia hemoglobin blood test	Non-Anemic (+12g/dl), Anemic (<12 g/dl)
Smoking	Tobacco Risk	Yes, No
PA	Physical Activity	Low Active or Active
LA	Leisure Time Activity	Low Active or Active
FS	Family Support	Low, Medium, High
PS	Peer Support	Low, Medium, High
SS	School Support	Low, Medium, High
PTSD	Posttraumatic stress symptoms	Low, Moderate, Severe
Depression	Depression Symptoms	Low, Moderate, Severe
Psychosomatic_Symp	Psychosomatic symptoms	Low, Moderate, Severe
Positive_Health	Positive health perceptions	Low, Medium, High
Life_Satisfaction	Life Satisfaction	Low, Medium, High
Average_Score	Student Total Average Grades	Excellent/Very Good, Good, Weak/Fail

Socio-demographic variables: Include variables such as age, living place, gender, family income, parent's education, and school type.

Physical activity: Physical activity was measured in a two-scale system, the activity scale and the leisure scale. The activity scale was composed of three question items:

- In the last week, how many days were you physically active for more than 60 minutes
- Number of hours playing sports outside school
- Number of hours exercising per week

The leisure time scale similarly comprised three question items:

- Number of hours watching TV
- Number of hours playing video games
- Number of hours using the internet

All six items considered only weekdays (5 days a week). Respondents' answers were analyzed in quartiles for both scales. Physically active students were identified as those in the upper quartiles, and likewise, low and non-active students were identified as those in the upper quartile of the leisure scale.

Nutrition: Nutrition data were collected though and in-person self-reported 24h-food recall of one-day food intake, as well as through a food frequency questionnaire (FFQ). The food items were scaled by classifying them into 8 food categories (vegetables, fruits, dairy products, sweets, soft drinks, sweet beverages, and energy drinks) [23]. Response choices ranged from (1) never, (2) 1-2 times a week, (3) 3-4 times a week, to (4) 5-7 times a week (almost daily).

Furthermore, respondents were grouped into two classes based on consumption habits, healthy and unhealthy consumers. Respondents were classified into the *healthy* group if they formed part of the upper 2 quartiles and indicated they didn't consume unhealthy foods. Respondents were identified as *unhealthy* if they had indicated that they did not consume any of the healthy food groups, or if they were in the upper 2 quartiles on the unhealthy foods consumption frequency.

Anemia: Anemia was measured through the collection of blood samples for the assessment of hemoglobin levels. An instantly calibrated finger-prick was used for the collection of blood samples. Respondents with anemia were identified as those with a hemoglobin levels below 12 g dl^{-1} for children younger

than 14 years old, and those with levels below 13 g dl^{-1} for respondents over 14 years of age.

Social Support: Social support was measured through three variables: 1) family support, 2) school support, and 3) peer support. Each variables asked to what extent respondents felt they were able to talk to their family, teachers, and peers.

Positive health perception: Self-reported health perception was measured through six question items:

- I like the way things are going for me
- my life is going well
- I have a good life
- I feel good about what is happening to me
- I would like to change many things in my life
- I wish I had a different kind of life—often or almost always.

Life satisfaction. The life satisfaction variable was measured through a 1-item scale based on the evaluative aspects of subjective well-being gathered from psychological research. The scale ranges from 0 to 10, where 0 represents the worst possible life, and 10 the best possible life satisfaction.

Posttraumatic stress disorder symptoms (PTSD)— PTSD symptoms were measured through 20- question items where respondents had to select their degree of association with the givens statement on a scale from 0 (not at all) to 4 (very much).

Academic performance: Respondents grades were obtained from prior school records, the average grades obtained in six courses (Arabic language, English language, Religion, Social studies, Science, and Mathematics) were collected and included in the total average estimation. Academic performance was considered by categorizing the average total score into Excellent/Very Good, Good, Weak/Fail.

3.2. Data Preparation

The raw data variables were selected, and cleaned the data prior to analysis. This step was taken to ensure the elimination of data redundancy, removed missing values, and derived new variables when applicable. The initial dataset consisted of 5084 student records. The key features in the dataset are balanced, whereas the gender variable is unbalanced.

3.3. Data Pre-Processing

The data has passed through a preprocessing phase before being ran in the machine learning algorithms. The data preprocessing includes techniques for organizing the data for machine learning models, such as data transformation, coding the categorical data, and feature selection, among others. The majority of the studied variables are composed of multiple questions, and thus need to be merged into a single variable with a preserved distance between the values. Preprocessing computes the variables from the scale of variables and converts them into meaningful numerical values. Common distance values were used depending on variable category levels: The three levels set as variable categories are 1) Low =0, 2) Moderate =1, and 3) High =2, and the two levels are set as 1) No =0, and 2) yes=1. The machine learning models are then used to analyze the students' academic achievement, as well as to predict the associated risk factors. The cross-validation technique that repetitively splits data into training and testing sets was used to overcome the problem of over fitting.

3.4. Statistical Analysis and Machine Learning Implementation

Six ML predictive models were built and compared to each other depending on their prediction accuracy from the given dataset. The ML algorithms used are described in the table below (Table 2).

Table 2: Description of Machine Learning Techniques

ML algorithm	Description
Artificial Neural Network (ANN or NN)	The Neural Networks algorithm is a computational model built up of many small aggregators that are clustered into groups, or layers that feed-forward from the input data to an output prediction [24]. Thus, the NN algorithm maps the predictor variables to an outcome variable through intermediate hidden layers.
Random Forest (RF)	Random Forest is an ensemble model that comprises decision trees, and then outputs the plurality of votes from the trees [25]. Each tree is exposed to a data subset and independently evaluates the features available to arrive in a conclusion [24]. RFs must be trained and are sensitive to the training data.
Support Vector Machine (SVM)	The SVM classifies linear and non-linear data; the model starts with a small number of the most similar cases across classes "support vectors", and the remaining cases are then classified based on their fall in the side of the support vectors [26].
Logistic Regression (LR)	Logistic regression is a statistical technique for classifying records based on values of input fields. LR is similar to linear regression, but it helps to find the probability that a certain item belongs in a specific class, thus the outcome lies between 0 and 1 [26].
Decision Tree (DT)	Decision trees are structures with multiple levels and a hierarchy of observations about an object in order to predict its class [24]. The tree "leaves" represent the classes the object could belong to and its "branches" reference the features that could lead the object to a certain class. DTs are used to predict and classify certain objects [27].
Naïve Bayes Classification	NB is analyzes training sets across the set of features and finds the probabilities of features' ability to predict the target variable. Naïve Bayes gives the probability of an event based on prior knowledge conditions without assuming that features are related even though they could be interdependent [26].

The machine learning algorithms were applied to predict the students' academic performance. The Orange data mining software [28] was used in testing and validating the machine learning models.

3.5. Results evaluation

The study found that Logistic Regression and Naïve Bayes had the highest accuracy rates (94.3% and 94% respectively), and are the most suitable machine-learning algorithms to predict student academic performance. The other algorithms similarly reported high accuracy rates and performed very well in predicting the associated factors related to academic performance. The results have been evaluated through a confusion matrix describing the performance of all machine learning models used (found under section 4.1- Table 4), as well as graphical representations of the data, and importance ranking of variables.

4. Results

A descriptive analysis was performed for describing the data distribution and for understanding the data set before the evaluation of machine learning techniques. The data set is composed of 5084 students with a mean age of 13±1.5, ranging from 10-17 years old. Among these students, approximately 32.8%

are boys and 67.2% are girls. Data in figure 1 shows the academic performance levels distributed by grade levels. The study found that the overall distribution of students' academic performance levels is almost similar in all grades. Results show that 46.7% of students obtained a low-average score, 20.6% obtained a medium-average score, and 32.7% obtained a high-average score.

Results in figure 2 show the average academic performance score by grade and gender. Results indicated that most of the students reported a good average score (70-75) in all grades for both girls and boys. We observed no significant difference found between the average performance score and grade or gender.

Figure 3 demonstrates the comparison between machine learning algorithms' accuracy rates for the algorithms used in predicting students' academic performance. Besides LR and NB models, which had the highest accuracy rates, the other ML algorithms had high-performance accuracy while SVM reported the lowest performance accuracy 80.2%. Logistic regression has 94.3% correct instances and 5.7% incorrect instances. For the further representation of prediction accuracy, the instances classification accuracy of the six models is shown in Table 2, while results in Table 4 show the algorithms' performance features through confusion matrices description.

Table 5 describes the details of the algorithms' performance. The classification results show that the Logistic regression and Naïve Bayes models are the best algorithms for predicting students' academic achievement. The logistic regression precision and area under curve (AUC) results (99.3% and 94.3% respectively) reflected in Table 4 indicate that most of the tested machine learning algorithms perform effectively in predicting positive values. Furthermore, the support vector machine algorithm reported the lowest acceptable performance, yet its significance is still valid for predicting students' academic performance.

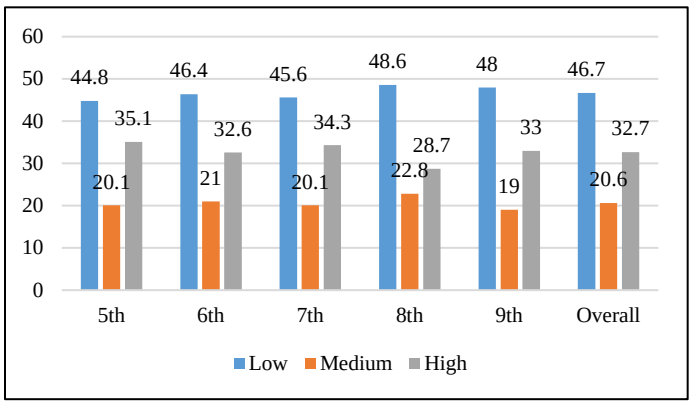


Figure 1: Students' academic performance by grade

The logistic regression ROC curve for the three academic performance classes is presented in figures 4.5, and 6. Three numerical categories for respondents' grades are included: 1) low, 2) medium, and 3) high. The ROC curve resides in the upper-left corner, thus the LR algorithm has a better prediction of positive value with an AUC of 99%. Furthermore, the Mathew Correlation Coefficient has been calculated for the 6 machine learning algorithms as shown in Table 3. Results indicated that the different classifiers had constant and high-performance levels. Most of the models reported MCC values closed to one in the [-1,+1] interval, which means that the machine learning models performed very well in predicting the students' academic performance.

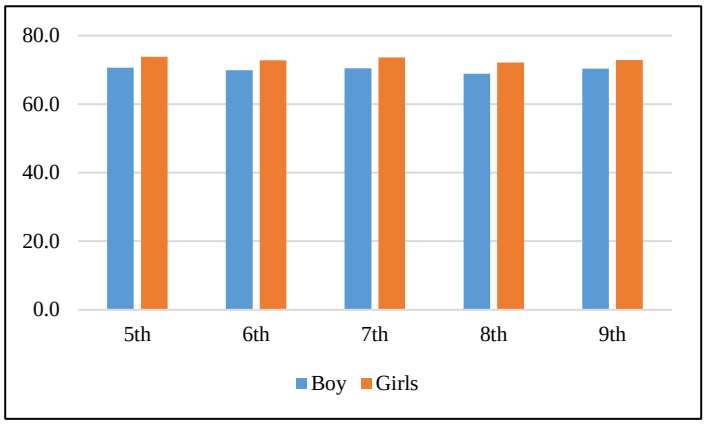


Figure 2: The students' average academic performance by grade and gender

Table 3: The estimated Mathew correlation Coefficient for the 6 ML models

Machine Learning Model	Mathew Correlation Coefficient
Logistic Regression	0.93
Naïve Bayes	0.927
Random Forest	0.92
Decision Tree	0.92
Neural Network	0.915
SVM	0.78

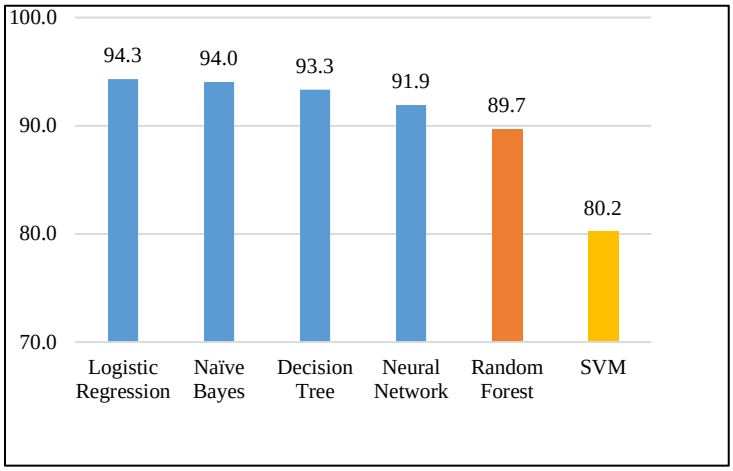


Figure 3: Comparison of Prediction Accuracy Percentage among the 6 ML Models

The machine learning models features importance ranking results are shown in Table 6. The ranking analysis includes Information Gain, Gain ratio, Gini and Chi-Square algorithms. Results in Table 6 indicated that positive health perception, Smoking, Depression, PTSD, Healthy Food Consumption, Age, gender, Grade Level, and Family income are the most important features for predicting students' academic performance.

Table 3: The confusion matrix that describes the performance of Machine Learning models.

Random Forest Predicted Value				
Actual Value	**L**	**M**	**H**	∑
L	94.2	5.5	0.5	**2523**
M	5.3	86.5	7.3	**1257**
H	0.5	8.0	92.3	**1304**
∑	**2604**	**1185**	**1295**	**5084**
Actual Value	**Support Vector Machine Predicted Value**			∑
L	88.3	21.5	0.0	**2523**
M	11.5	60.4	16.8	**1257**
H	0.2	18.1	83.2	**1304**
∑	**2555**	**1237**	**1292**	**5084**
Actual Value	**Decision Tree Predicted Value**			∑
L	94.3	4.7	0.0	**2523**
M	5.6	90.3	6.1	**1257**
H	0.0	5.0	93.9	**1304**
∑	**2618**	**1139**	**1327**	**5084**
Actual Value	**Logistic Regression Predicted Value**			∑
L	96.5	6.4	0.0	**2523**
M	3.5	88.8	4.6	**1257**
H	0.0	4.8	95.4	**1304**
∑	**2532**	**1249**	**1303**	**5084**
Actual Value	**Neural Network Predicted Value**			∑
L	95.5	8.1	0.0	**2523**
M	4.5	83.6	7.1	**1257**
H	0.0	8.3	92.9	**1304**
∑	**2535**	**1259**	**1290**	**5084**
	Naïve Bayes Predicted Value			
L	95.8	6.1	0.0	**2523**
M	4.2	87.9	3.4	**1257**
H	0.1	6.0	96.6	**1304**
∑	**2554**		**1260** **1270**	**5084**

Table 4: Performance of algorithmic models

Model	**AUC**	**CA**	**F1**	**Precision**	**Recall**
Logistic Regression	0.993	0.943	0.943	0.943	0.943
Naïve Bayes	0.975	0.940	0.940	0.941	0.940
Decision Tree	0.939	0.933	0.932	0.932	0.933
Neural Network	0.986	0.919	0.919	0.919	0.919
Random Forest	0.970	0.897	0.895	0.895	0.897
SVM	0.938	0.802	0.802	0.801	0.802

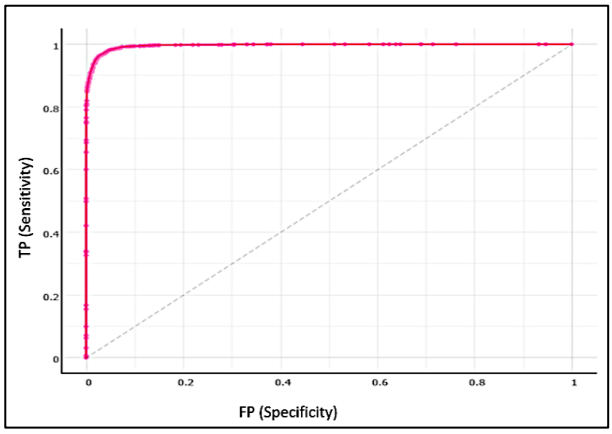


Figure 4: Logistic Regression ROC curve for low academic

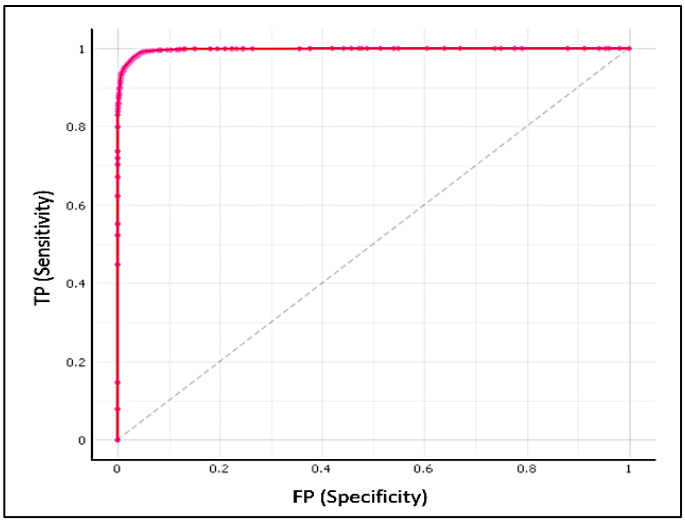


Figure 5: Logistic Regression ROC curve for moderate academic performance

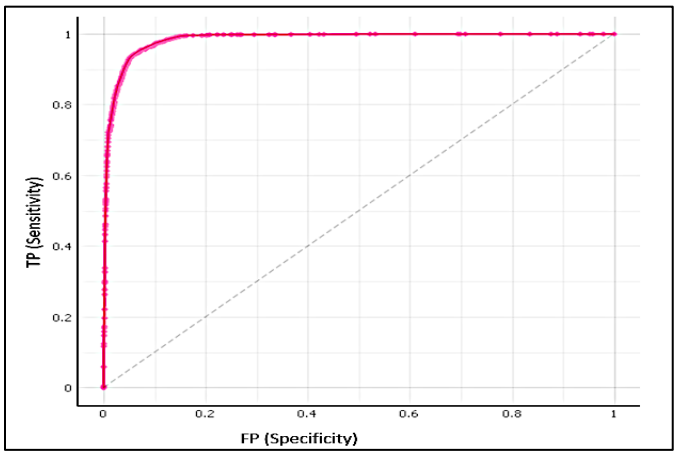


Figure 6. Logistic regression ROC curve for high academic performance

5. Discussion

The study aimed to predict the association between health factors (mental, physical, and social) with student academic performance. With this in mind, the results have identified that healthy students indeed perform better at school.

Twenty-two relevant features were utilized after performing feature selection using machine learning algorithms. These features were employed as input variables, while the average academic score was considered as the target variable. Among the tested models, logistic regression showed the best performance results with sensitivity =94.3%, specificity=96.3%, and AUC=99.3%. Out of the 22 features selected, 11 features are related to academic performance including physical, mental, and social health indicators.

The data has shown that positive health-associated factors are significantly interrelated with students' physical growth, cognitive development, and academic performance [7]. Furthermore, it has been found that unhealthy factors, such as obesity, depression, anxiety, and PTSD, have a direct negative impact on student's mental health and cognitive development,

these findings were found to be consistent with other related studies [29], [30].

Table 5: Features' Importance Ranking

	Features Ranking Algorithms				
Features	**Info. gain**	**Gain ratio**	**Gini**	**χ^2**	**Relief**
PTSD	0.040	0.025	0.020	43.8	0.006
Smoking	0.025	0.038	0.012	138.4	0.026
positive health perception	0.024	0.042	0.012	139.1	0.009
Depression	0.016	0.016	0.008	49.7	0.034
Healthy Food Consumption	0.010	0.010	0.005	30.6	0.029
Family Support	0.008	0.005	0.004	5.0	0.057
Age	0.006	0.003	0.003	27.9	0.011
Gender	0.006	0.006	0.003	12.9	0.010
Grade Level	0.004	0.002	0.002	16.8	0.013
Psychosomatic Symptom	0.004	0.002	0.002	5.7	0.015
Family Income	0.003	0.002	0.001	10.7	0.065
Mother Education	0.002	0.002	0.001	5.7	0.007
School Support	0.002	0.001	0.001	8.3	0.030
Residency	0.002	0.001	0.001	7.4	0.039
Unhealthy Consumption	0.001	0.001	0.001	2.7	0.012
Physical Activity	0.001	0.001	0.001	1.0	0.007
Anemia	0.001	0.001	0.000	5.5	0.003
Life Satisfaction	0.001	0.001	0.000	5.1	0.021
Peer Support	0.001	0.000	0.000	0.8	0.014
BMI	0.000	0.001	0.000	2.5	0.014
Leisure Time Activity	0.000	0.000	0.000	0.6	0.040
father Education	0.000	0.000	0.000	0.5	0.013

The prediction results indicated that the negative features had more effect on students with low academic performance. Students with low family income, smoking, anemia, depression, and unhealthy food consumption reported a significant association with low academic performance. Parents' education, life satisfaction, positive health perception, and physical inactivity had less significant effects on low academic performance than other features.

In the present paper, we evidence that several machine learning algorithms were able to predict academic performance associated with risk factors. All tested algorithms were successfully managed to predict the target variable. Nonetheless, we have found that the Logistic Regression and Naïve Bayes algorithms are the most effective in predicting academic performance when health (mental and physical) and social risk factors are considered. These findings strengthen previous findings that discuss the validity of data mining and ML techniques in predicting students' academic performance. Although the common ML model with highest accuracy in several literature is Naïve Bayes, we have found that when accounting for only key subject areas and excluding commonly high-performance subjects (such as Arts and Physical Education), Logistic Regression is able to provide similar accuracy levels.

6. Conclusion

The study assessed the accuracy and performance of machine learning models in predicting the associated health factors on students' academic performance. Six ML models were used for this study (Random Forest, Logistic Regression, Neural Network, Support Vector Machine, Naïve Bayes, and Decision Tree). Although all models yielded valid and positive results, the logistic regression model proved to provide the best accuracy rate and performance, closely followed by the Naïve Bayes model. The study's findings demonstrate that other ML algorithms, such as Random Forest, Decision Tree, and Neural Networks are also able to satisfactorily predict students' academic performance, while support vector machine yielded the results with the least accuracy.

Furthermore, the obtained results show that the indicators that most affect academic performance include smoking, depression, PTSD, family support, anemia, leisure time, unhealthy food consumption, place of residence, and grade level. On the other hand, gender, physical activity, friends' support, school support, life satisfaction, mother's education, and positive health do not significantly influence students' academic performance.

Therefore, this research concludes that ML algorithms, particularly (yet not exclusively) Logistic Regression and Naïve Bayes, could be integrated into schools' information systems for the automatic prediction of students' academic performance based on key variables. In this manner, students, families, and school staff, and administration will be able to tackle issues that might affect students' performance through the use of the obtained prediction results. Likewise, through the use of accurate machine learning techniques, such as Random Forest, public health professionals, healthcare providers, and decision-makers will be able to predict rising issues, and implement relevant intervention programs to enhance students; health and education

7. Recommendations

Based on the findings presented in this paper, future research will benefit from expanding the study by adding additional associated factors, including cognitive development skills, in-school student behavior, social activities, and digital media activities. Further variables to consider that would enhance the aim of the research are factors related to the school environment and teaching methods, in particular those that are student-centered, motivational, and personalized as we believe they might positively influence academic performance.

Moreover, in a future study, we seek to deploy the above-mentioned machine learning algorithms by integrating them within schools' information management systems in Palestine to test the effectiveness of ML algorithms in a practical school environment. The results could aid researchers globally in developing frameworks for ML-led, automated school performance indicators.

Abbreviations

Data Mining (DM), Machine Learning (ML), Random Forest (RF), Artificial Neural Networks (ANN or NN), Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression (LR), Naives Bayes (NB), Post-Traumatic Stress Disorder, (PTSD), Chi-squared Automatic Interaction Detection (CHAID), Science, Technology, Engineering and Mathematics (STEM), United Nations Relief and Works Agency (UNRWA), Food Frequency Questionnaire (FFQ), Area Under Curve (AUC), Classification Accuracy (CA), Features (F), Intelligent Recommender System (IRS), Ministry of Education (MoE).

Conflict of Interest

The authors declare no conflict of interest

Acknowledgments

The authors wish to thank the Palestinian Ministry of Education, as well as Al Quds University for allowing us to carry out the study, as well as the children involved and their parents for their cooperation and patience. We further wish to acknowledge all field workers whose hard work made this research possible. The authors further wish to thank the ministry of the education team, mainly Mr. Mohammad Remawi and Ms. Suzan Tutah for their support in data collection. The Sources of Support is German Research Foundation (DFG) under grant number JU 2769/2.

References

[1] J.L. Rastrollo-Guerrero, J.A. Gómez-Pulido, A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," Applied Sciences (Switzerland), **10**(3), 2020, doi:10.3390/app10031042.

[2] M.M. Dubuc, M. Aubertin-Leheudre, A.D. Karelis, "Relationship between interference control and working memory with academic performance in high school students: The Adolescent Student Academic Performance longitudinal study (ASAP)," Journal of Adolescence, **80**, 204–213, 2020, doi:10.1016/j.adolescence.2020.03.001.

[3] R. Qasrawi, "Links between nutrition, life style habits and academic acheivment in Palestinian schoolchildren: A cross-sectional study," Al-Quds Journal for Academic Research, **01**(1), 90–102, 2021, doi:10.47874/2021p6.

[4] E.L. Faught, J.P. Ekwaru, D. Gleddie, K.E. Storey, M. Asbridge, P.J. Veugelers, "The combined impact of diet, physical activity, sleep and screen time on academic achievement: A prospective study of elementary school students in Nova Scotia, Canada," International Journal of Behavioral Nutrition and Physical Activity, **14**(1), 1–13, 2017, doi:10.1186/s12966-017-0476-0.

[5] B. Asmare, M. Taddele, S. Berihun, F. Wagnew, "Nutritional status and correlation with academic performance among primary school children, northwest Ethiopia," BMC Research Notes, **11**(1), 1–6, 2018, doi:10.1186/s13104-018-3909-1.

[6] R.R. Halde, A. Deshpande, A. Mahajan, "Psychology assisted prediction of academic performance using machine learning," in 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 431–435, 2016, doi:10.1109/RTEICT.2016.7807857.

[7] H.F. Golino, C.M.A. Gomes, D. Andrade, "Predicting Academic Achievement of High-School Students Using Machine Learning," Psychology, **05**(18), 2046–2057, 2014, doi:10.4236/psych.2014.518207.

[8] V.L. Uskov, J.P. Bakken, A. Byerly, A. Shah, "Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education," in 2019 IEEE Global Engineering Education Conference (EDUCON), 1370–1376, 2019, doi:10.1109/EDUCON.2019.8725237.

[9] M. Imran, S. Latif, D. Mehmood, M.S. Shah, "Student academic performance prediction using supervised learning techniques," International Journal of Emerging Technologies in Learning, **14**(14), 92–104, 2019, doi:10.3991/ijet.v14i14.10310.

[10] E. Alyahyan, D. Düştegör, "Predicting academic success in higher education: literature review and best practices," International Journal of Educational Technology in Higher Education, **17**(1), 2020, doi:10.1186/s41239-020-0177-7.

[11] A. Blanch, A. Aluja, "A regression tree of the aptitudes, personality, and academic performance relationship," Personality and Individual Differences, **54**(6), 703–708, 2013, doi:10.1016/j.paid.2012.11.032.

[12] A. Makwe, A.S. Rathore, An Empirical Study of Neural Network Hyperparameters, 2021, doi:10.1007/978-981-15-5788-0_36.

[13] D.I.F.S. Engr. Sana Bhutto, Dr. Qasim Ali Arain, Maleeha Anwar, "Through Supervised Machine Learning," Predicting Students' Academic Performance Through Supervised Machine Learning, 2020.

[14] M. Goga, S. Kuyoro, N. Goga, "A Recommender for Improving the Student Academic Performance," Procedia - Social and Behavioral Sciences, **180**(May), 1481–1488, 2015, doi:10.1016/j.sbspro.2015.02.296.

[15] T. Burrows, S. Goldman, K. Pursey, R. Lim, "Is there an association between dietary intake and academic achievement: a systematic review," Journal of Human Nutrition and Dietetics, **30**(2), 117–140, 2017, doi:10.1111/jhn.12407.

[16] T.M. Burkhalter, C.H. Hillman, "A narrative review of physical activity, nutrition, and obesity to cognition and scholastic performance across the human lifespan," Advances in Nutrition, **2**(2), 201–206, 2011, doi:10.3945/an.111.000331.

[17] I.K. Desai, A. V. Kurpad, V.R. Chomitz, T. Thomas, "Aerobic fitness, micronutrient status, and academic achievement in Indian school-aged children," PLoS ONE, **10**(3), 1–13, 2015, doi:10.1371/journal.pone.0122487.

[18] F.M. Asigbee, S.D. Whitney, C.E. Peterson, "The Link Between Nutrition and Physical," Journal of School Health, **88**(6), 407–415, 2018.

[19] K. Adolphus, C.L. Lawton, L. Dye, "The effects of breakfast on behaviour and academic performance in children and adolescents," Frontiers in Human Neuroscience, **7**(JUL), 1–28, 2013, doi:10.3389/fnhum.2013.00425.

[20] R. Qasrawi, Z. Abdeen, H. Taweel, M.A. Younis, D.A. Al-Halawa, "Data mining techniques in identifying factors associated with schoolchildren science and arts academic achievement," in Proceedings - 2020 International Conference on Promising Electronic Technologies, ICPET 2020, 2020, doi:10.1109/ICPET51420.2020.00023.

[21] P.A. Banerjee, "A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools," Cogent Education, **3**(1), 1–17, 2016, doi:10.1080/2331186X.2016.1178441.

[22] A. Namoun, A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," Applied Sciences (Switzerland), **11**(1), 1–28, 2021, doi:10.3390/app11010237.

[23] G.C. Frank, T.A. Nicklas, L.S. Webber, C. Major, J.F. Miller, G.S. Berenson, "A food frequency questionnaire for adolescents: defining eating patterns.," Journal of the American Dietetic Association, **92**(3), 313–318, 1992.

[24] H. Phwkrg, R.I. Frppxqlfdwlrq, W.S. Olyhv, "9 – 11 April , 2019 - American University in Dubai , Dubai , UAE 2019 IEEE Global Engineering Education Conference (EDUCON) Page 1370 9 – 11 April , 2019 - American University in Dubai , Dubai , UAE 2019 IEEE Global Engineering Education Conference (ED," 2019 IEEE Global Engineering Education Conference (EDUCON), 1370–1376, 2019.

[25] Y.L. Pavlov, "Random forests," Random Forests, 1–122, 2019, doi:10.1201/9780429469275-8.

[26] S. Uddin, A. Khan, M.E. Hossain, M.A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," BMC Medical Informatics and Decision Making, **19**(1), 1–16, 2019, doi:10.1186/s12911-019-1004-8.

[27] P. Geurts, A. Irrthum, L. Wehenkel, "Supervised learning with decision tree-based methods in computational and systems biology," Molecular BioSystems, **5**(12), 1593–1605, 2009, doi:10.1039/b907946g.

[28] J. Demšar, T. Curk, A. Erjavec, Č. Gorup, T. Hočevar, M. Milutinovič, M. Možina, M. Polajnar, M. Toplak, A. Starič, "Orange: data mining toolbox in Python," The Journal of Machine Learning Research, **14**(1), 2349–2353, 2013.

[29] C.A.C. Montanez, P. Fergus, A. Hussain, D. Al-Jumeily, B. Abdulaimma, J. Hind, N. Radi, "Machine learning approaches for the prediction of obesity using publicly available genetic profiles," Proceedings of the International Joint Conference on Neural Networks, **2017-May**, 2743–2750, 2017, doi:10.1109/IJCNN.2017.7966194.

[30] S. Roy, A. Garg, "Predicting academic performance of student using classification techniques," 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics, UPCON 2017, **2018-Janua**, 568–572, 2017, doi:10.1109/UPCON.2017.8251112.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 12-21 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Lean Six Sigma Implementation in the Food Sector: Nexus between Readiness-Critical Success Factors

Sarina Abdul Halim-Lim*,[1], Nurul Najihah Azalanzazllay[1], Anjar Priyono [2], Guven Gurkan Inan[3], Muhammad Iqbal Hussain[4]

[1]*Department of Food Technology, Faculty of Food Science and Technology, Universiti Putra Malaysia, 43400, Malaysia*

[2]*Department of Management, Universitas Islam Indonesia, Yogyakarta, 55000, Indonesia*

[3]*Industrial Engineering Department Samsun University, Samsun, 55020, Turkey*

[4]*Faculty of Mechanical Engineering Technology, Universiti Malaysia, Perlis, 02400, Malaysia*

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 24 October, 2021
Online: 09 November, 2021

Keywords:
Lean Six Sigma
LSS Readiness
Critical Success Factors
Change Management
Food sector

ABSTRACT

Lean Six Sigma (LSS) is a renowned approach for boosting operational excellence and competitive advantage through integrated core objectives of value creation and variation reduction. Despite its proven benefits in many leading companies, LSS implementation in the food sector is still behind compared with other sectors. LSS implementation is costly, and most businesses have failed due to a lack of preparation and an unsupportive organizational culture. Therefore, there is a need to identify LSS readiness factors that suit the food sector to minimize the risk of implementation failure in the industry. The current study concentrates on the LSS pre-implementation phase to determine the competency criteria to adopt LSS customized for the food business. This study will explore the LSS readiness criteria during the pre-implementation stage and critical success factors (CSFs) during the implementation stage in the food sector through Lewin's Change Theory. Twelve food sector employees who were associated with quality management activities were interviewed using a semi-structured approach. The interview was recorded, transcribed and the transcription was analyzed using content analysis. The results showed six readiness themes in the food manufacturing sector with twenty-nine LSS readiness attributes, while seventeen factors out of thirty-one CSFs for the LSS at the implementation stage. The identified readiness factors are management commitment and leadership (ten attributes), organizational culture (nine attributes), employee involvement (six attributes), process management (four attributes), project management (four attributes) and external factors (three attributes). Through Pareto analysis, the most prioritized CSFs are from top management and leadership and employee involvement themes, with the training program being identified as the most important LSS CSFs (85%). This study will serve as a foundation for a benchmarking tool for managers to improve the effectiveness of an LSS implementation in the food sector.

1. Introduction

A strategic business plan is required in the food sector to minimize the process errors that would impact the product quality. This study which, is an extension of work initially presented in the 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), addresses Lean Six Sigma (LSS) application as the approach. Recently, there was an increase in quality and lower-cost products demand. The food business experienced challenges due to high customer expectations, strict food standards, and a competitive market [1]. Furthermore, the food businesses have been combating waste that can occur at any point in the supply chain [2]. But waste elimination requires a structured way of thinking, where new manufacturing initiatives should be considered a valuable innovation for the food businesses. Since 2000, the model for integrating Lean and Six Sigma has been sought by the organization globally to improve process performance, maximize profit and minimize operating costs [3]. Lean has a role to reduce non-value-added activities, which is also considered waste in the

*Corresponding Author: Sarina Abdul Halim-Lim, sarinalim@upm.edu.my

https://dx.doi.org/10.25046/aj060603

customer's eyes. Whilst Six Sigma is viewed as a problem-solving methodology focusing on making a process effective by reducing process variations [4]. The hybrid of LSS depicted the mixture of best values of these two effective approaches, enabling both variations and waste reduction across the process [2, 3]. In [4], the author confirmed the prominent advantage of LSS projects in this sector mostly is cost saving. This is what drives and motivates quality executives in the food sector to continue their CI efforts, however, there is another motivation such as the need to abide by the laws and regulations. By reducing overfilling confectionary cream that can be expensive [2], LSS enabled food companies not to distribute underweight food products which can cause them to be penalized. LSS activities also encourage circular economy activities by decreasing waste of overfilling cream [5].

Studies have highlighted that the food business has been slower to adopt the LSS mindset and has remained passive compared to other industries [4] and [6]. The reason is that the cost of products is relatively low compared to the cost structure of the business [7], prioritization on the strict regulations that characterize the food sector [8], the nature of the food sector being conservative towards change.

The perceived management difficulties that the businesses in this sector encounters, for instance resistance to change, may be mitigated through LSS preparedness evaluation outlined by organizational change theory [9]. Resistance to the new effort such as LSS, occurred at various levels within an organization. It considered as a significant issue that could result in the failure of LSS implementation [1], [10] and [11]. Thus, the importance of conducting an assessment on the state of LSS readiness prior the implementation need to be highlighted, because preparedness is one of the LSS implementation success factor by [11]. It also will facilitate the effort to integrate LSS with management innovation such as digital technologies [12]. The pre-implementation stage receives insufficient attention compared to the LSS implementation phase elements, such as crucial success factors (CSFs) and barriers.

Thus, this paper aims to investigate the factors that contribute to LSS preparedness customized to the food context. Variables contributing to the preparedness of LSS and their impact on food sector readiness for LSS implementation must be determined to achieve this research aim.

2. Literature Review

2.1. Lewin's change theory

Readiness study was pioneered by [13] to technological adoption, in which he established that the term is semantically contrary to the resistance to change. [14] associated readiness for a practice that promotes organizational transformation by minimizing inhibitors of success or uncovering the important points to improve the success of the change. Based on Lewin's theory of change model, it is appropriate to explore the factors contributing to unfreezing and freezing phases to drive any kind of change [15]. Unfreezing refers to accepting the need to make the change and preparing the people ready for the change [15]. By adopting the theory, this study intends to understand two phases of the theory which are the unfreezing, followed by the change phase. in the context of LSS as depicted in the research framework in Figure 1. Several studies have identified the importance of readiness outside of the implementation process and encouraged the longevity of the LSS system [10]. Due to the large venture capital and continuous commitments involved with the adoption of LSS, the company should tactically plan to avoid misspending on CI activities [6] and [16]. Furthermore, the understudied LSS issue especially reluctance towards change, high turnover and complacency of current process performance, are highly involved in the two phases of Lewin's change theory.

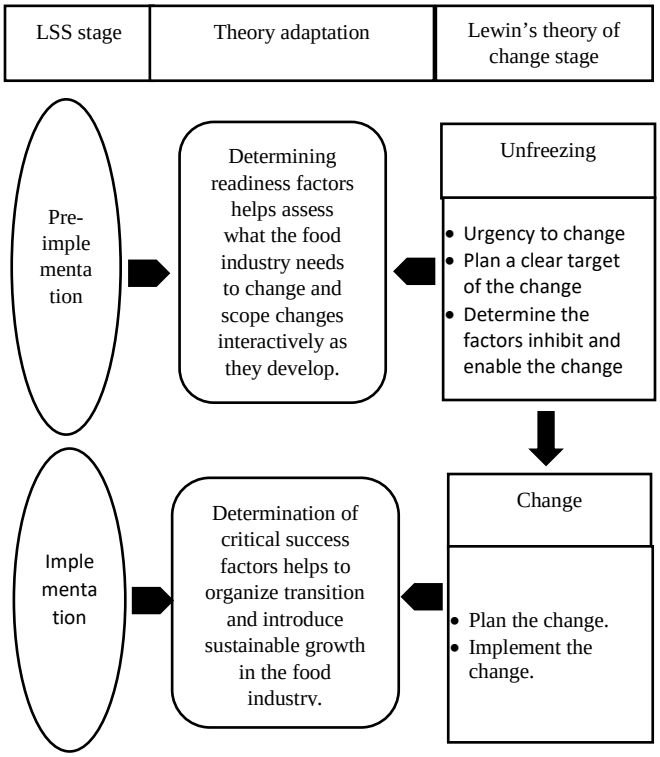


Figure 1: Research framework

3. Methodology

Based on the current body of knowledge, there is no research addressing the preparedness of LSS in the context of the food business. This study is primarily an inductive qualitative investigation based on semi-structured interviews comprising practitioners from twelve different food firms familiar with LSS. The inductive approach used in this study appreciates the insights and views of the interviewees as data [17]. This study is filling the gap of knowledge on LSS preparedness in the food business, necessitating a qualitative semi-structured interview rather than a quantitative technique commonly employed in readiness studies in other industries [3]. The research process was guided by the iterative qualitative data analysis model by [18], which comprises data collection, data reduction, data coding and analysis and conclusion, contribution, limitation and future research agenda.

3.1. Sampling

This study applies purposive sampling to choose a suitable sample using the criterion technique. Respondents involving food practitioners that are familiar with LSS and involved in CI projects [19] as depicted in Table 1. The size of a food business can have a causal relationship to LSS implementation effectiveness [6] and [20]. Hence, one of this study's sampling criteria is the experts working in a big company [6]. The interview guideline was created to not only to minimize bias but direct the interview sessions. Then, a pilot interview was carried with three respondents, which was shown by [21]. The interview is inductive in nature, maintaining an adaptive approach to give the chance for probing the questions based on the answers given.

Table 1: Details of the experts

Respondents	Current industry	Position	Experience in the food sector (year)	Role in LSS
E1	Cooking oil	Operational Excellence Manager	9	LSS Black Belt
E2	Dairy	Production Manager	14	SS Green Belt Lean Black Belt
E3	Dairy	Process Engineer	7	LSS Green Belt
E4	Fats	Assistant Manager Quality	9	LSS Green Belt
E5	Beverages	Production Engineer	7	LSS Green Belt
E6	Butter	Continuous Improvement Manager	1	LSS Black Belt
E7	Dairy	Operation Manager	15	LSS Black Belt
E8	Dairy	Operation Manager	15	LSS Black Belts
E9	Beverages	Quality Assurance Executive	2	Certified Quality Engineer
E10	Confectionary	Value Stream Manager	13	LSS Green Belt
E11	Oils	Supply Chain Manager	9	LSS Green Belt
E12	Meat and poultry	Quality Assurance Manager	10 years	Six Sigma Green Belt

3.2. Data collection

There are two steps of data collection involved in this study. The first step is three pilot interview sessions were conducted as outlined by [21] in the literature, followed by the twelve semi-structured interviews using revised interview protocol. The interviews lasted ranges 60 to 180 minutes and were conducted face-to-face in the English language. Interviews were performed using a checklist, with notes and commentaries inserted throughout the session. With the participants' permission, the researcher recorded the interview sessions and comprehensive notes taken during the sessions. In grounded theory, the sample size cannot be established in advance because it is dependent on the emerging theoretical categories [22]. The theory becomes evident when gathering new data no longer sparks new theoretical insights of the theoretical categories [8]. The twelve interviews with practitioners resulted in data saturation due to data overlap and no new insight materials being offered after the ninth interview, exhibiting the aspects of data saturation [21] and [22]. Despite reaching data saturation with the tenth respondent, the researcher conducted two additional interview sessions to confirm that no new information emerged, as shown in Figure 2.

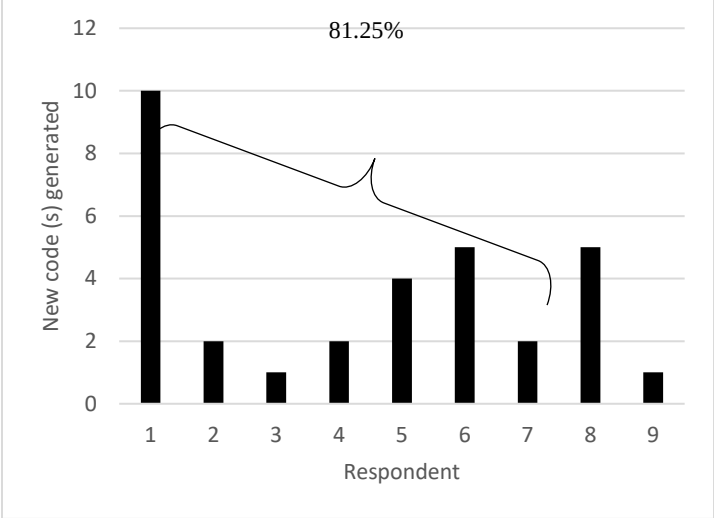


Figure 2: New LSS readiness codes generated at each interview

3.3. Data coding and analysis

Shortly after the interview, the interview sessions were familiarized and verbatim transcribed before the process of data reduction. The data was analyzed following the abductive approach, which combined interpretative content analysis and deductive themes identification [18]. The data was transcribed using software QSR Nvivo 12, and a field note was used to complement the transcription. The transcripts of the interviews were familiarized during data reduction, and the coding procedure was carried out to identify appropriate data pertinent that appropriate with the topic, as described by [18].

There are three fundamental coding modes in grounded theory: open, axial, and selective [22]. In open coding, the transcribed interviews are accurately conceptualized into appropriate readiness themes. The researchers were able to infer new meanings from the data collected using inductive coding, while the research objectives guided the emergent coding. During the axial coding phase, the content from the open coding themes was then assessed to develop linkages between the themes. The themes that resulted from the axial coding describe the emergent pattern in the LSS implementation and organizational readiness to change theory. Selective coding was conducted as the final step, where the purpose is to identify the core themes of the LSS readiness framework.

Comprehensive data analysis and explanation were performed to verify and clarify previously developed logical findings. The four components of [23] were applied to this investigation to guarantee the data presented was reliable. The causal map depicting the causal relation of readiness factors is developed based on inductive data analysis [24], [25] and [26]. The map depicts connected links by arrows, that showed the reviewers' inferred causality [25]. Descriptive analysis through Pareto analysis was conducted to identify the most important LSS CSFs.

4. Results

The determination of the key factors of the LSS pre-implementation stage will be conducted with consideration at both the 'unfreezing' and 'freezing' stages outlined by several readiness studies [27]. The findings revealed CSFs that could help coordinate LSS implementation in the context of the food sector during transitional periods [4].

4.1. Readiness factors in the food sector

Through the qualitative evaluation of employees in the food sector that has been practicing LSS, this study gathered a lot of data about the influence of LSS readiness criteria. When 29 readiness factors were identified through the interview process, the results of the interviews were described in terms of the highest number of mentioned readiness factors by those who took part in interviews.

The result shows that management support has topped all the identified LSS readiness factors. The respondents stated leadership and management are both have a critical role in LSS deployment. Top management is highly suggested to communicate their vision on LSS adoption and set realistic targets for the LSS program in the company. In the food sector, effective communication by the top management is one of the critical ingredients in encouraging the employees' involvement in the LSS related activities under their strict quality assurance programme.

The result shows that, although the regulations and industry standards outline the most quality programmes in the food sector, the top management has a critical role in developing the urgency towards the LSS adoption at the organizational level. According to respondent E7, senior management should keep a consistent cadence when encouraging the industry to adopt LSS. Top management support can be viewed by their overall knowledge on the "why" and "how" LSS implementation can contribute to their operational performance and company's bottom line and subsequently oversee towards business excellence.

The second critical readiness factor is financial support allocated for the program's implementation and its continuation in the company. Financial investment is the most important type of support from upper management, as it has a cascading effect on other readiness factors. For example, as reflected in Figure 3.0, each dimension may correlate with each other, and these findings may suggest there is a correlation between the financial resources, training, reward and recognition, and investment in essential resources (trainers, machines, and software) under top management support theme. The respondents highly suggested the top management should be eager to spend in LSS activities including appreciation programs, soft and hard technologies, training and contracting advisors to facilitate LSS implementation.

Two emerging readiness factors in the context of the food manufacturing industry. Respondents E3, E5, E7, E8, and E9 pointed out the appointing a change agent to lead the LSS readiness phase is particularly critical in the food business. The purpose is to ensure the LSS initiation is guided and have a smooth transition between the pre-implementation and implementation stages. The assignation of the LSS champion as the change agent was implemented at the implementation phase reflected the initiation of the LSS journey of the business [28].

Another newly identified readiness factor in the study is top-level management awareness of the importance of LSS implementation. Top management should understand the demand for LSS implementation in their business and fully comprehend the reasons for embarking on the LSS. The top management was well-informed, primarily about the importance of a compulsory food safety programme in the food businesses depending on the type of food commodity. However, a brief LSS awareness session for the top management is suggested in the food sector to obtain the buy-in, commitment, and engagement in the LSS programme.

According to the vast majority of respondents, the ability to identify and maintain skillful human resources in LSS is essential for the successful implementation of LSS initiatives. Companies that initiate LSS with poor understanding and skills are deemed to fail to sustain the programme. The need for external experts is critical in educating theoretically and practically the team members because the training programme in class would be insufficient given the LSS project duration of six months, which would necessitate project consultation with the coach. According to E1and E2, it is common for the food companies to have the same staff becoming the internal expert for both, food safety and continuous improvement activities.

According to E3 and E8, the food industry has been characterized as a complex industry, that could be a factor in the effectiveness of LSS. The industry's complexity included industrial norms and standards, raw material types, procedures involved, and critical-to-quality performance monitoring. As respondents advocate updating machinery and software to examine the capability of measuring variations and performances of critical-to-quality characteristics, common concerns in production (e.g. long waiting duration when devices break down) must be considered.

Respondent E2 advocated that the LSS should be implemented within a realistic, feasible, and achievable time frame, with awards and recognition included in the continuous improvement programme. As a result, respondents recommended that relevant guidelines be developed as a handbook to direct the organization's implementation of the LSS.

The involvement of qualified personnel in the quality management project is a significant indicator of preparedness at the LSS level. Identifying and training the right applicant would result in the addition of extrinsic value to the organization. However, several respondents emphasized that senior management has an ultimate say to choose the change agent/champion and let the decision of the project team's appropriate members to the champion. Typically, team members appointment happened at the implementation stage of the LSS. However, the respondents suggested that the existence of nurturing the top talent in the company should already have existed.

"It is critical to choose the suitable employees for the LSS belt training to see the return of investment of the training" (E6).

In the LSS implementation, having top talent is believed to have a better result, attracting more talent where it motivates the other employees to get involved [28].

Respondents E6, E7, E9, and E11 asserted that increasing LSS awareness would improve employees' willingness to adapt in favour of LSS implementation. It is agreed by the respondents that the employees will embrace better on the program when their understanding of the principal and advantage of LSS towards the business and themselves. LSS should be introduced as early as in tertiary education, where E5 stated that LSS can be embedded under related subjects in food bachelor degree program such as Operation Management, Quality Management and Supply Chain Management.

Additionally, several respondents indicated that specific tactics should focus on different projects. Implementing LSS should not be viewed as the answer for all production process issues. The importance of selecting the correct process and project contribute to the subsequent LSS adoption direction in the company. According to [3], project prioritization should be done in considering the best products in the business, contribution of the project to bottom-line saving and profit, contribution of the project to growth and impact of the project on the sustainability enhancement.

E9 warned that lack of LSS awareness would risk the success of the LSS adoption. In the food sector, interest in LSS can be nurtured through the other quality-related activities in the businesses, for instance, quality assurance and food safety activities. Another significant issue identified is the process for forming teams for CI activities, as they are considered as the main personnel accountable for quality improvement projects. Typically, food companies commonly will have a food safety team that adhered to both food safety and quality improvement issues.

Process management's fundamental notion is to understand quality criteria from the eyes of the customer. The guidelines and suggestions from LSS projects should be aligned with customer demands and the characteristics of each commodity in the food sector. E6 indicated that government assistance, such as establishing LSS awareness campaigns for the business, can be crucial in its readiness. The government can assist by engaging industry players through experience-sharing workshops and learning visits to LSS projects that have been successful as benchmarks and case studies.

4.2. LSS CSFs in the food sector

The findings indicated that 31 identified factor nodes contributed to the success of LSS implementation. The number of the codes was analyzed using Pareto analysis to determine their frequency and criticality under 80/20 rules. In [1], the author suggested that Pareto analysis helps classify "vital few" success factors from "trivial many," which focus on the factors to accentuate the desired goals.

Figure 3.0 shows that 17 CSFs were identified as the prioritized factors for the success of LSS implementation through 80/20 rules. The food sector context identified training programmes as the most critical factors, followed by top management support, financial resources, communication structure employee involvement and as the top five main success factors for LSS implementation. Some of the CSFs are correlated and linked from the "unfreezing" stage to the "freezing stage" of LSS implementation. All of the LSS readiness factors in Figure 3.0, such as external factors (EF), employees involvement (EI), organizational culture (OC), management support and leadership (MSL) and project management (PM) are the common factor identified at the readiness and implementation phase of LSS.

Based on the prioritized CSFs, a causal mapping in Figure 4 developed based on the interview sessions showed that the potential of some of the factors have an impact on the emerging of other factors as the critical factor for the LSS program to be successful.

The respondents mentioned the training program has the most causal impact on successful LSS adoption, especially in this industry. Three distinct factors arose from the training: coaching and mentoring, LSS awareness and understanding level of LSS.

"Our LSS training programmes will include coaching and mentoring session as our company has hired in-house trainers to train the Green Belt/Black Belt programme for our employees."(E1)

"LSS basic training is crucial to create awareness among employees and their critical roles in LSS implementation." (E5)

"I would say continuous training for the employees is critical to increasing their understanding of the function and benefits of each LSS tool." (E12)

One of the most critical factors following the training programme was the support and commitment of senior management on LSS projects. Respondent E12 mentioned that the requirement of LSS project approval highly depends on the top management vision and buy-in of the project.

"If there is new management, there must be massive support from the top to ensure they acknowledged and understand the benefits of current initiatives. When management is willing to invest in LSS projects, I would say they are supportive." (E7)

"I think that management support is influenced by their awareness on the advantages of LSS projects." (E10)

Another important factor to consider is the availability of financial resources and the budget allocation communicated by the top management to implement LSS initiatives. Another emerging factor that is relatable to the allocation of funding is the rewards and recognition. Respondents stated that their company allocates a budget for rewarding the employees and investment in cost-saving projects, which serve as a driving factor in making it a successful project. Another respondent responded on this issue:

"Having a strong financial capability to invest in systems, training, the hiring of trainers, and software is critical for any organization, in my opinion." (E2)

"LSS projects necessitate investment on the part of the management, particularly in terms of staff training and programme maintenance." (E6)

Another major component that contributed to the success of the LSS initiative in the food business was the participation of employees. In response to respondent R1, he stated that his company engages their employees by empowering them through hands-on training to provide the best exposure to them in the workplace. E7 stated that employee involvement led to awareness and an increased understanding of the LSS concept and benefits among those who worked for the company. Respondent E11 revealed that his company allows shop-floor employees to participate in decision-making, motivating and building trust among employees and management.

Medium of communication was identified as one of the most important CSFs that can be implemented to achieve a good result in the food business for successful LSS implementation. An effective communication program is essential, especially in convincing top management. He stated that top management roles in communicating the LSS strategic thinking to all departments are critical. Besides, it was to ensure the awareness of employees on LSS initiatives is under the same vision of the management. Other respondents stated that:

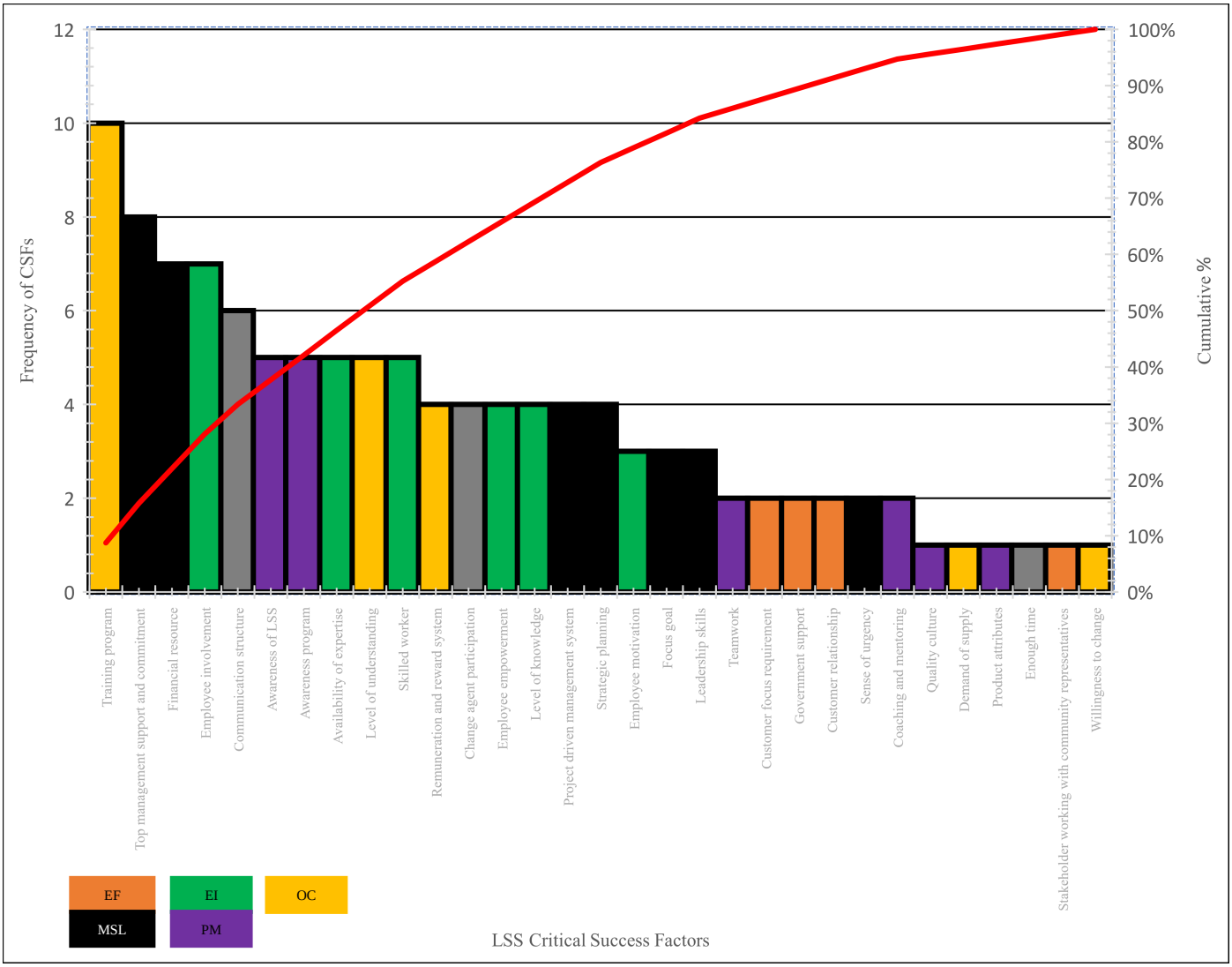


Figure 3: Pareto chart of CSFs of LSS implementation in the food sector

"The acquisition is having a simple briefing to the shop floor to deliver the standard procedure of the LSS projects and ensuring that they are aware of how these projects will impact them is critical" (E8).

The Pareto chart in Figure 3 also revealed that the LSS enablers for the food sector mainly comprised of soft factors. The change process has an impact on any form of factors that could lead to the LSS implementation success in the food sector, including the "support from the government" and customer focus business. Some of the quotes being mentioned by the respondents:

"The government should provide help and support by rewarding those who complete successful LSS projects in a competition or by providing financial assistance to those who attend LSS training. Sources of coaches from other successful industries should be able to be an eye-opener" (E10)

"The LSS concept is driven by the ability to accrue the trust of the customer through value creation and variation reduction. The project should be tailored to the specific customer to ensure that it meets all of the customer requirements." (E1, E5).

5. Discussions

5.1. Readiness factors and their causal relationship

It was claimed that there is a necessity to explore critical and readiness factors for each sector because the nature and operations of various businesses may differ in terms of their factor, applicability, and value to the business [29]. The food sector's distinguishing traits, including the obligation to comply with strict food legislation rules, necessary cleaning of machinery and production sites, and the rapid changeover of various products, can impede the success and readiness of adopting LSS [2] and [30].

Venturing into LSS implementation necessitates high budget and continuous efforts, mandating strategic planning focused on the essential criteria during the pre-implementation stages to avoid overspending or underspending on CI operations [6, 11].

The findings revealed that management support and leadership, organizational culture readiness, employee involvement, process management, project management, and external interactions all impacted the preparedness of the LSS program. Figure 5 depicts the identified preparedness criteria to demonstrate the relationship between them.

5.1.1. Management support and leadership (MSL)

Management support and leadership was identified as the most critical factor at the pre-implementation phase. If any attempts at continuous improvement are to be made in a long time and with unwavering funding, this indicates a need for management support [11]. The organization's top management is capable of incorporating LSS as a business strategy for improving performance and business excellence. Implementing LSS demands significant capital expenditures and high commitments, demanding strategic planning to avoid cutting CI operations budget [6] and [16]. Management support implies that top management is taking action and developing plans to facilitate and steer the business's implementation of the LSS [1].

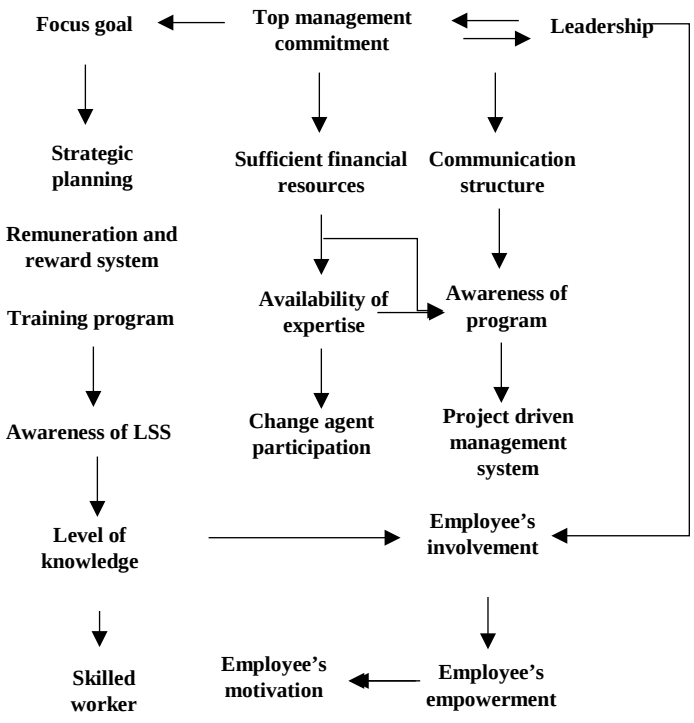


Figure 4: Causal diagram of the prioritized LSS CSFs in the food sector

5.1.2. *Organizational culture readiness*

Following the interview, it was determined that in this industry, an organization's preparedness to adopt LSS relies on its organizational culture's willingness to adapt. According to a previous study, quality-driven culture should be promoted among employees in the organization. The foundation of success can be achieved with the corporate culture change [31] and [32] and [33]. Regular training enables the company to produce talented personnel with sufficient tools and techniques to be used by staff [32].

5.1.3. *Employee involvement*

Employee engagement is referring to the level of understanding of their roles within the company indicated the readiness of LSS activities. Employees who have adequate skills to apply the tools of LSS and an awareness of the principle of LSS are thought to be less reluctant to change. Operating as a team has been demonstrated to decrease the unwillingness towards the change, which validates the view articulated in the literature that management may appoint members and educate the team to tackle issues in process improvement [34]. By assigning employees to the problem-solving activity, the organization can help them feel more confident in their abilities. This would foster a sense of trust between top management and the employees.

5.1.4. *Project management*

The result shows that project management is a prominent LSS readiness factor as LSS is a project-based program. When administering the LSS project, prioritization should be made to deliver the most financial benefits to the firm [27], [32] and [35]. LSS project management entails determining the companywide level of understanding about LSS, which can inform the management on the company preparedness for LSS implementation [32]. Participation in the LSS awareness activities will increase the familiarity of the employees and minimize the unwillingness towards the change in the program's direction.

5.1.5. *Process management*

Considering the nature of the business, food businesses should consider integrating the LSS with customer relation activities. This will enable the voice of consumers will always be considered in the process and critical to quality attributes. The integrated components should be made as to the business's major priority for the future. Consumer requirements must be considered for the LSS projects to consider and increase the "consumer voice" while taking into account the particular features of the food sector. Gaining operational excellence is stated to be critical in the survival of a business [3]. Efficiency can be improved by identifying root causes, developing improvement solutions, and preventing them from occurring again [27], [34], [35], [36] and 37]. Aside from that, when a reliable system for measurement is accessible in the process, organizations are better prepared for LSS.

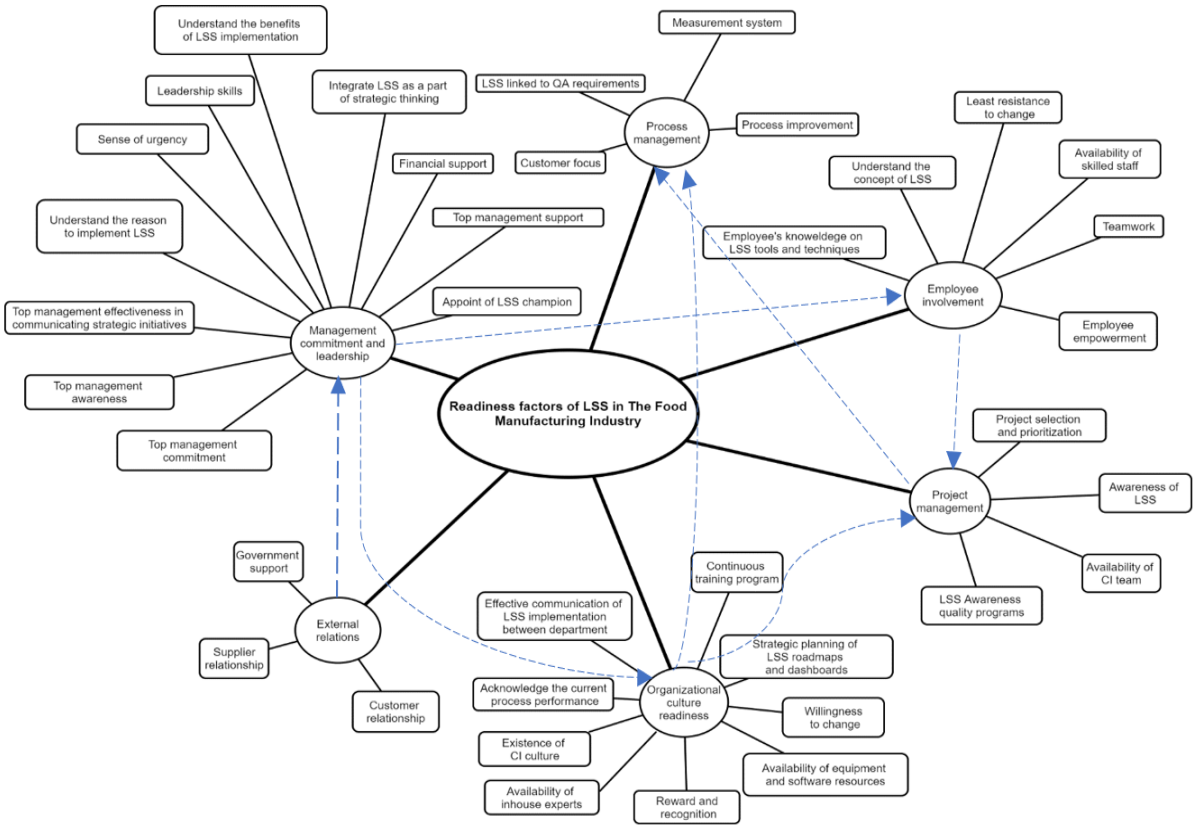


Figure 5: The conceptual LSS readiness factors in the food sector and its constructs.

5.1.6. External factors

Most of the respondents discussed the importance of government support in initiating LSS in the food businesses. In the previous studies on LSS readiness, this factor is not commonly found. However, this component was identified as affecting the leanest implementation in the red meat business. The authorities play the policymaker's role and are responsible for strategizing effective adoption strategies [37]. However, cooperation between the external factors is defined as their events with their suppliers and customers [38]. Upon examination of the factors, it was discovered that management support has become the most vital determinant of LSS readiness. The management support entails the highest number of tasks in ensuring the preparedness of a food business to uptake LSS.

6. Conclusions

The implementation of a successful LSS can result in significant financial advantages for a company. However, most companies fail to implement LSS because they are pressured into doing so as part of a fad of business excellence approaches without taking the time to assess whether the organization is prepared to execute it successfully. The key research objective is discovering relevant factors that need to be assessed before investing in the LSS adoption customized for the food business, as well as the CSFs that must be considered throughout this process, as it is the linking phase following the readiness phase as according to Lewin's change theory.

According to the findings of this study, the emerging LSS readiness factors determined are management support and leadership, the readiness of organizational culture, employee involvement, process and project management, and finally external factors. The results show there is a proposition that the discovered factors for LSS implementation in this sector are causally related. The findings show 17 CSFs, where some of these factors are common factors under the readiness stage. In the food business, the most prevalent CSFs of the LSS are training, top management commitment and leadership, financial capacity, employee engagement, and communication.

There are two ways to look at the significance of this study. First, the determined LSS readiness factors provided a simple and easy approach as a decision-making tool to assist food sector practitioners in understanding their current practices, evaluating if their organizational environment supports LSS and finally deciding whether to invest in LSS Implementation. The managers also will be informed on the factors that are important in the pre-implementation stage. Since some of the readiness themes were found will be an enabling factor in the implementation phase, the managers will be able to plan the implementation strategically. Whilst, for the food business that has implemented LSS, the result of this study opens the window to identify factors that require more attention to reach the targeted level of LSS readiness. Second, this paper extends to the limited body of knowledge about LSS implementation in the food sector and organizational readiness to change theory. Additionally, the findings of this study can be evaluated by comparing to those of similar studies conducted in various industries or countries, determining the impact of LSS readiness in different countries. Additionally, the findings of this study can be evaluated by comparing to the other similar studies conducted in various industries or countries, evaluating the impact of LSS readiness in different countries.

There were several limitations concerning the generalizability of this work, which provide the opportunity for future research. This study produced results that are described by the features of the interpretive type of study. While analyzing the data sets, it was found that hypothesis development must be built on a solid foundation, allowing for an empirical study to be conducted. Because the study used a small and focused sample, it was not possible to determine whether the findings had any external validity. Future research should consider that the result from grounded theory can be developed to hypothesis to be tested through quantitative study to establish a generalized LSS readiness factor with a larger sample size for the food sector. Future studies can also focus on the size of the company in the food production business such as big enterprise, medium or small. The identified factors can be used as the readiness measurement instrument in other industrial settings. Finally, the result of this study can be used as a vital theory for the construction of an LSS readiness diagnostic tool for the food sector.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

This work is supported by the Universiti Putra Malaysia under research grant Putra Young Initiative (IPM/2017/9552400) and Malaysia Ministry of Education (MOE) for providing research grant under Fundamental Research Grant Scheme (FRGS/1/2018/TK03/UPM/02/7).

References

[1] S.A.H. Lim, J. Antony, "Statistical process control readiness in the food sector: Development of a self- assessment tool," Trends in Food Science and Technology, **58**, 133-139, 2016, doi:10.1016/j.tifs.2016.10.025.

[2] M. Dora, X. Gellynck. "Lean six sigma implementations in a food processing SME: A case study," Quality and Reliability Engineering International, **31**, 1151–1159, 2015, doi:10.1002/qre.1852

[3] V. R. Sreedharan, R. Raju, M. V. Sunder, J. Antony, "Assessment of lean six sigma readiness (LESIRE) for manufacturing industries using fuzzy logic," International Journal of Quality and Reliability Management, **36**(2), 137–161, 2019, doi:1108/IJQRM-09-2017-0181.

[4] L. B. M. Costa, M. G. Filho, L. D. Fredendall, F. J. Gómez Paredes, "Lean, six sigma and lean six sigma in the food sector: A systematic literature review," Trends in Food Science and Technology, **82**, 122– 133, 2018, doi:10.1016/j. tifs.2018.10.002.

[5] L. Marrucci, M. Marchi, T. Daddi, "Improving the carbon footprint of food and packaging waste management in a supermarket of the Italian retail sector," Waste Management, **105**, 594-603, 2020, doi:10.1016/j.wasman.2020.03.002.

[6] L.B.M. Costa, M. G. Filho, L. D. Fredendall , G.M.D Ganga, "The effect of lean six sigma practices on food sector performance: Implications of the Sector's experience and typical characteristics," Food Control, **112**, 2020, doi:10.1016/j.foodcont.2020.107110

[7] J. A. Garza-reyes, I. E. Betsis, V. Kumar, M.A.R. Al-Shboul, "Lean readiness-the case of European pharmaceutical manufacturing industry,"

International Journal of Productivity and Performance Management, **67**(1), 20-44, 2018, doi: 10.1108/IJPpm-04-2016-0083

[8] B.V. Chowdary, D. George, "Improvement of manufacturing operations at a pharmaceutical company: a lean manufacturing approach," Journal of Manufacturing Technology Management, 23 (1), 56-75, 2012, doi: 10.1108/17410381211196285.

[9] A. A. Armenakis, S. G. Harris, K. Mossholder, "Creating readiness for organizational change," Human Relations, **46**, 681–703, 1993,doi:10.1177/001872679304600601.

[10] Z. Radnor, 2011. "Implementing lean in health care: making the link between the approach, readiness and sustainability," International Journal of Industrial Engineering and Management, **2**(1), 1-12, 2011, doi: 10.1016/j.socscimed.2011.01.011.

[11] J. Antony, "Readiness factors for the lean six sigma journey in the higher education sector," International Journal of Productivity and Performance Management, **63**(2), 257–264, 2014, doi:10.1108/IJPPM-04-2013-0077

[12] C.Andrea, and M. Kumar. "Lean Six Sigma and Industry 4.0 integration for Operational Excellence: evidence from Italian manufacturing companies." Production Planning & Control, **32**(13), 1084-1101, 2020, doi:10.1080/09537287.2020.1784485

[13] E.H. Jacobson, "The effect of changing industrial methods and automation on personnel," Paper presented at the Symposium on Preventive and Social Psychology, Washington, DC, 1957.

[14] S. Al-Balushi, P.J. Sohal, P.J. Singh, A. Al Hajri, Y.M. Al Farsi,R, Al Abri, "Readiness factors for lean implementation in healthcare settings – a literature review," Journal of Health Organization and Management, **28**(2), 135-153, 2014, doi: 10.1108/JHOM-04-2013-0083.

[15] B.J. Weiner, "A theory of organizational readiness to change," Implementation Science, **4**(1), 67, 2009, doi: \10.1186/1748-5908-4-67.

[16] E. Drohomeretski, S. E. G. da Costa, E. P. de Lima, P. A. Garbuio, "Lean, six sigma and lean six sigma: An analysis based on operations strategy," International Journal of Production Research, **52**(3), 804–824, 2014, doi:10.1080/00207543.2013.842015.

[17] C. Robson, K. McCartan, Real world research (4th ed.), John Wiley & Sons, 2006.

[18] M. B. Miles, A. M. Huberman, Qualitative data analysis: An expanded sourcebook., Sage Publications, 1994

[19] M. Q. Patton, Qualitative research and evaluation methods, Thousand Oaks, Sage, 2002.

[20] MG Filho, GMD. Ganga, A. Gunasekaran, "Lean manufacturing in Brazilian small and medium enterprises: Implementation and effect on performance," International Journal of Production Research, **7543**, 1–23, 2016, doi: 10.1080/00207543.2016.1201606.

[21] M. Dikko, "Establishing construct validity and reliability: Pilot testing of a qualitative interview for research in takaful (islamic insurance)," The Qualitative Report, **21**, (3), 521- 528, 2016, doi: 10.46743/2160-3715/2016.2243.

[22] R. Bogdan, S.K. Biklen, Qualitative research for education: An introduction to theories and methods, Pearson A & B, 2007.

[23] E. Guba, "Criteria for assessing the trustworthiness of naturalistic inquiries," Educational Technology Research and Development, **29**(2), 75-91, 1981, doi:.

[24] M.A. Abernethy, M. Horne, A.M. Lillis, M.A. Mallina, F.H. Selto, "A multi-method approach to building causal performance maps from expert knowledge," Management Accounting Research, **16**(2), 135-155, 2005, doi: 10.1016/j.mar.2005.03.003.

[25] I. Pyrko, V. Dörfler, "Using causal mapping in the analysis of semi-structured interviews," In Academy of Management Proceedings , 2018(1), 2018, doi: 10.5465/ambpp.2018.14348abstract.

[26] A. J. Scavarda, T. Bouzdine-Chameeva, S. M. Goldstein, J.M. Hays, A. V. Hill, "A methodology for constructing collective causal maps," Decision Sciences, **37**(2), 263-283, 2006, doi: 10.1111/j.1540-5915.2006.00124.x.

[27] K. Jayaraman, T. Leam Kee, K. Lin Soh, "The perceptions and perspectives of lean six sigma (LSS) practitioners," The TQM Journal, **24**(5), 433–446, 2012, doi: 10.1108/17542731211261584.

[28] V. F., Vallejo, J. Antony, Douglas, J. A., Alexander, P., & Sony, M. (2020). Development of a roadmap for Lean Six Sigma implementation and sustainability in a Scottish packing company. The TQM Journal.

[29] J. Rockart, "Chief executives define their own data needs," Harvard Business Review, **57**(2), 1979.

[30] A. Panwar, B.P. Nepal, R. Jain, A. P. S. Rathore, "On the adoption of lean manufacturing principles in process industries," Production Planning & Control, **26**(7), 564–587, 2015, doi: 10.1180/09537287.2014.936532.

[31] A. Gurumurthy, P. Mazumdar, S. Muthusubramanian, "Graph theoretic approach for analyzing the readiness of an organization for adapting lean thinking: a case study," International Journal of Organizational Analysis, **21**(3), 96-427, 2013, doi:10.1108/ IJOA-04-2013-0652.

[32] Y. Lagrosen, R. Chebl, M. Rios Tuesta, "Organizational learning and six sigma deployment readiness evaluation: A case study," International Journal of Lean Six Sigma, **2**(1), 23–40, 2011, doi: 10.1108/20401461111119431.

[33] SC. Cross, "Lean cuisine." Industrial Engineer, **41**(1), 46-47, 2009

[34] P.B. Keliji, B.S.D Abadi, M. Abedini, "Investigating readiness in the Iranian steel industry through six sigma combined with fuzzy delphi and fuzzy DANP," Decision Science Letters, **7**, 465–480, 2018, doi: 10.5267/j.dsl.2018.1.001.

[35] M. N. Mishra, "Identify critical success factors to implement integrated green and lean six sigma," International Journal of Lean Six Sigma, 2018, doi:10.1108/IJLSS-07-2017-0076.

[36] T. Y. Lee, W. K. Wong, K. W. Yeung, "Developing a readiness self-assessment model (RSM) for six sigma for China enterprises," International Journal of Quality and Reliability Management, . **28**(2),169–194, 2011, doi: 10.1108/02656711111101746.

[37] A. Cox, D. Chicksand, M. Palmer, "Stairways to heaven or treadmills to oblivion?: Creating sustainable strategies in red meat supply chains," British Food Journal, **109**, 2007, doi:10.1108/00070700710780689

[38] C. A. Moya, D. Galvez, L. Muller, M. Camargo, "A new framework to support lean six sigma deployment in SMEs," International Journal of Lean Six Sigma, **10**(1), 58–80, 2019, doi:10.1108/IJLSS-01-2018-0001.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 4, 13-20 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

A Reconfigurable Stepped Frequency Continuous Wave Radar Prototype for Smuggling Contrast, Preliminary Assessment

Massimo Donelli[*,1,2], Giuseppe Espa[2,3], Mohammedhusen Manekiya[1], Giada Marchi[1] , Claudio Pascucci[4]

[1]*Department of Information Engineering and Computer Science, University of Trento, Trento, 38123, Italy*

[2]*Center for Security and Crime Sciences (CSSC), University of Trento and Verona, Trento, 38120, Italy*

[3]*Department of Economy and Management, University of Trento, Trento, 38120, Italy*

[4]*Guardia di Finanza, Department, Comando Regionale Trentino Alto Adige, Trento, 38120, Italy*

ARTICLE INFO

Article history:
Received: 27 April, 2021
Accepted: 21 June, 2021
Online: 10 July, 2021

Keywords:
Stepped frequency continuous wave radar (SFCW)
Electromagnetic propagation
Microwave systems

ABSTRACT

A reconfigurable Stepped Frequency Continuous Wave (SFCW) radar prototype for supporting the italian financial police to contrast smuggling, is proposed in this work. In particular, the proposed radar can provide information related to the container contents and the presence of false bottoms speeding up the control operations at borders and ports. Moreover, it is able to reveal the presence of people hidden behind reinforced concrete hiding places. Radar resolution is improved by using suitable post-processing method such as the MUltiple Signal Classification (MUSIC) algorithm. Numerical as well as experimental results obtained considering realistic operative scenarios demonstrated the potentialities and capabilities of this system as an effective tool for smuggling contrast. The preliminary experimental results have been obtained using a compact radar prototype equipped with high gain and directivity antennas to cover all the different frequency bands.

1 Introduction

In recent years, globalisation has led to a sharp increase in the smuggling of goods and illegal materials such as drugs, tobacco, and weapons. To contrast smuggling, Guardia di Finanza (GdF), the Italian financial police, increases the controls at customs and ports. However, checking the contents of thousands containers or vehicles requires a lot of time, human resources, and money. New technologies and tools can effectively to support the GdF work and be effective to contrast smuggling. In such a scenario, radar technologies and, particularly, ground penetration (GPR) and through the wall radars (TTR) could represent effective tools for GdF. A scan-A (1D) analysis of a container can provide accurate information concerning the contents, presence of false bottoms or hidden humans without the need to empty the container or vehicle, saving time, money and resources. In order to be useful for these operations, it is necessary to reach a high resolution of about tens centimetres and an unambiguous range of less than ten meters. These performances can be easily obtained with an SFCW radar [1]. Due to their complexity and the high cost of microwave components, SFCW radars had less diffusion for through the wall and ground penetration radar applications than pulsed radars [2]. However, in the last decade, microwave components strongly improved their performances and reduced their dimensions and cost, making the realization of compact SFCW radars more feasible. Recently SWCW radars have been successfully adopted for GPR applications [3, 4], landmine detection [5], surveillance [6], through the wall inspection [7], the monitoring of the structural integrity of engineering structures [8]–[10] or for biomedical applications such as the monitoring of human or animal life signals [11]–[13]. The last application can be very useful for police operation since it can be used to detect fugitives hidden inside improvised hiding places such as bunkers and wall false bottoms or human traffic. SFCW radar capabilities to characterize the electric properties of materials as in [10] can be very useful to identify hidden goods, such as tobacco, drugs, and hidden weapons. To improve the radar performances, such as resolution, detection capabilities, clean the clutter and reduce the background noise, suitable post-processing algorithms can be applied [14]–[18]. This paper is an extension of work originally presented in 2nd Global Power, Energy and Communication Conference (GPECOM) [19] where only preliminary numerical results have been reported. In this work, a low-cost SFCW radar prototype as support, the GdF work

[*]Corresponding Author: Massimo Donelli, Email: massimo.donelli@unitn.it

is developed, fabricated, numerically and experimentally assessed in realistic scenarios. The device is based on the ability to operate at different frequency bands enclosed in a range 100MHz-12GHz.. Based on a Direct Digital Synthesis (DDS) generator, the proposed system is easily reconfigurable in terms of frequency bands, number of steps, and spatial resolution. The device can be programmed as ground, through the wall and life detection radar, and it is suitable for smuggling contrast and rescue operations. The prototype is based on the system reported in [19], and it is capable of taking a single scan-A deep profile measurements using 600 frequency steps in less than 2 seconds with a minimum spatial resolution of 0.005 m, simulating a pulsed radar or as continuous-wave radar (CW) to work like a Doppler radar. The device is too slow for moving targets, but it is perfect for GPR and TTW applications where the targets are fixed. The obtained numerical and experimental preliminary results, related to realistic scenarios and directly provided by the GdF, demonstrated the capabilities and potentialities of the proposed system.

2 Mathematical Formulation

This section reports the mathematical formulation of a stepped frequency continuous wave radar and the formulation of post-processing MUSIC algorithm. A typical SFCW radar schema is shown in Fig. 1. It consists of a sweep generator directly connected with a wideband directive antenna. The receiving section can receive the backscattered electromagnetic wave and aims to provide the in-phase and quadrature (I/Q) signals. The control system directly process the frequency domain signal into a synthesized time-domain signal. In particular, the received baseband I/Q signals are combined into a complex number as follows:

$$\Gamma_m = I_m + jQ_m = A_i e^{j\phi_m} \tag{1}$$

where $\phi_m = -2\pi f_m \tau$, whit f_m is the m-th frequency step and τ is the propagation delay, $I_m = A_i cos(\phi_m)$, and $Q_m = A_m sin(\phi_m)$.

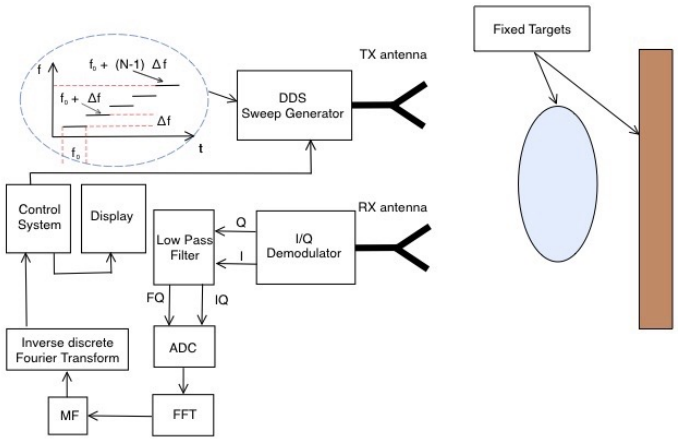


Figure 1: Schema of the SFCW radar.

The sweep generator generates M different sinusoidal tones. The receiver is aimed at collecting all the backscattered signals from the targets . In particular, the I_m/Q_m signals are organized into a complex vector $\hat{\Gamma} = [\Gamma_0, \ldots, \Gamma_{M-1}]$. The inverse discrete Fourier transformer reported in the following equation is then applied to the $\hat{\Gamma}$ vector in order to obtain the conversion from frequency to time domain and better reveal the presence of echoes produced by targets.

$$T_m = \frac{1}{M}\sum_{m=0}^{M-1}\Gamma_m = \frac{1}{M}\sum_{m=0}^{M-1} A_m e^{j\left(\frac{2\pi m_i}{M} - 2\pi\Delta f \frac{2R_t}{v}\right)} \tag{2}$$

SFCW radar strength is its ability to set the resolution properly and unambiguously adapt to the different operative scenarios. The resolution range is provided by the following equation:

$$\Delta R = \frac{v}{2M\Delta f} \tag{3}$$

where v is the electromagnetic wave propagation velocity in the medium. As it can be noticed from (3) the resolution range only depends by the frequency steps number M and the Δf. The following relation can estimate the unambiguous range:

$$R_{max} = \left(\frac{M}{2} - 1\right)\Delta R \tag{4}$$

similarly to equation (3), the unambiguous range can be changed by acting on the step number M and on the resolution range. In most of practical scenarios a A-scan (1D) range profile is enough and it can be obtained by processing the received signal for a given fixed position of the RX and TX antennas. However, it is worth noticing that for all the applications that require a B-scan (2D) representation, such as GPR applications, it is possible repeating the procedure for A-scan (1D) by using different antenna positions.

2.1 The MUSIC post processing algorithm

To properly resolve close targets, suitable post-processing algorithms must be applied to the SFCW radar signals. These signal are very weak and enveloped by the background or by other noise sources. Super-resolution algorithms such as the well-known MUSIC algorithm has been recently successfully used for different practical applications [20]–[24] and the MUSIC algorithm can be very useful to extract the echoes of targets from the background noise. In particular, the I/Q signals collected by the receiver can be expressed by the following vectorial relation:

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ . \\ . \\ . \\ s_n \end{bmatrix} = \begin{bmatrix} P & e^{-j2\pi f_{-1}\gamma_2} & . & . & . & e^{-j2\pi f_{N-1}\gamma_n} \\ e^{j2\pi f_1\gamma_1} & P & . & . & . & e^{-j2\pi f_{N-2}\gamma_n} \\ e^{j2\pi f_2\gamma_1} & e^{j2\pi f_1\gamma_2} & . & . & . & e^{-j2\pi f_{N-3}\gamma_n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ e^{j2\pi f_{N-1}\gamma_1} & e^{j2\pi f_{N-2}\gamma_2} & . & . & . & P \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ . \\ . \\ y_n \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ . \\ . \\ . \\ n_n \end{bmatrix} \tag{5}$$

that in compact form is $[\boldsymbol{s}] = [\boldsymbol{A}][\boldsymbol{y}] + [\boldsymbol{n}]$, where $[\boldsymbol{n}]$ is the noise vector, $[\boldsymbol{y}]$ is the reflection coefficient vector whose elements are related to the f_N frequency steps, $[\boldsymbol{A}]$ is the $n \times n$ delay matrix. The $n \times n$ signal covariance matrix is given by:

$$[\boldsymbol{R}] = \boldsymbol{S}\boldsymbol{S}^* = ([\boldsymbol{A}][\boldsymbol{y}] + [\boldsymbol{n}])([\boldsymbol{A}][\boldsymbol{y}] + [\boldsymbol{n}])^* \tag{6}$$

where $*$ represent the complex conjugate transpose. The signal covariance matrix $[\boldsymbol{R}]$ can be expressed as:

$$[\boldsymbol{R}] = [\boldsymbol{R}_s] + [\boldsymbol{R}_n] \tag{7}$$

where $[\boldsymbol{R}_n] = \sigma^2 [\boldsymbol{I}]$ is the noise autocorrelation matrix, σ is the noise variance, $[\boldsymbol{I}]$ is the identity matrix and $[\boldsymbol{R}_s]$ is the signal covariance matrix, which can be written as $[\boldsymbol{R}_s] = |A_n|^2 [\boldsymbol{e}_i] \left[\boldsymbol{e}_i^H\right] = P [\boldsymbol{e}_i] \left[\boldsymbol{e}_i^H\right]$, where $|A_n|^2 = P$ is the power of complex exponentials, $[\boldsymbol{e}_i] = \left[1, e^{j\omega_1\gamma_1}, e^{j\omega_1\gamma_2}, ..., e^{j\omega_1\gamma_n}\right]$, with $\omega = 2\pi f$. Since $[\boldsymbol{R}_s]$ is a hermitian matrix then the remaining eigenvectors υ_n will be orthogonal to $[\boldsymbol{e}_i]$ obtaining $\left[\boldsymbol{e}_i^H\right] \upsilon_n = 0; \quad n = 1, 2, ..., N$. The peak position in the time-domain can be obtained by searching the maximum value of the following function.

$$Fmusic(\gamma) = \frac{\boldsymbol{a}(\gamma)^* \boldsymbol{a}(\gamma)}{\boldsymbol{a}(\gamma)^* \boldsymbol{N}_n \boldsymbol{N}_n^* \boldsymbol{a}(\gamma)} \tag{8}$$

where $\boldsymbol{a}(\gamma)$ is the mode vector, obtained from the columns of matrix $[A]$ and $\boldsymbol{N}_n^*$ are the noise eigenvectors.

3 Prototype description

The SFCW radar prototype consists of a programmable digital signal generator (DDS), namely the TG124A (Signal Hound company), with a frequency range from 100KHz up to 12.5GHz and a power range from $-12dB$ up to $-3dB$. The receiver is a digital spectrum analyzer, the SA124B (Signal Hound company), with the following characteristics: frequency range 100KHz, 12.5 GHz, minimum detectable power $-150dBm$. The DDS generator is synchronized with the receiver by means of a coaxial cable that reports the reference local oscillator signal mandatory to obtain the baseband I/Q signals. The DDS generator is directly connected with a transmitting directive radiator, namely a log-periodic wideband antenna for the frequency band comprised between 500.0 MHz-6.0GHz and a high gain horn antenna for the frequency range 8.0GHz-13GHz, the antenna must be manually changed for different operations. The receiver can receive the backscattered electromagnetic wave and provide the I/Q signal that the control system can directly process. Moreover, the SA124B receiver can be easily programmed to sample the I/Q signals and transmit them to an elaboration unit (a high-performance laptop) with a high-speed USB connection. The radar prototype's RF sections are connected with three shielded semi-rigid coaxial cables equipped with sub-miniature type A (SMA) connectors. The SFCW radar prototype photo is shown in Fig. 2. The radar prototype can operate both in bistatic as well as monostatic configuration, as reported in Figs. 1 and 2 respectively. When the radar operates in monostatic configuration, a circulator is connected between the generator, receiver and antenna. The monostatic configuration is considered to reduce the weight of the prototype when operator is involved in through the wall operations (TTW). To reduce the perturbations on the electromagnetic field produsded by the mechanical support, the device was assembled on Teflon dielectric support and placed on a mechanical pedestal to steer the antenna properly. To properly set the radar functionalities, a graphical user interface (GUI) in Matlab language, has been developed. With the GUI, it is possible to customize the radar characteristics to fit with different operative scenarios. In particular, Fig. 3 reports a GUI screen snapshot, representing the tool for through the wall operations. On the left side panel, the radar signals are presented both in time domain and with their spectrogram.

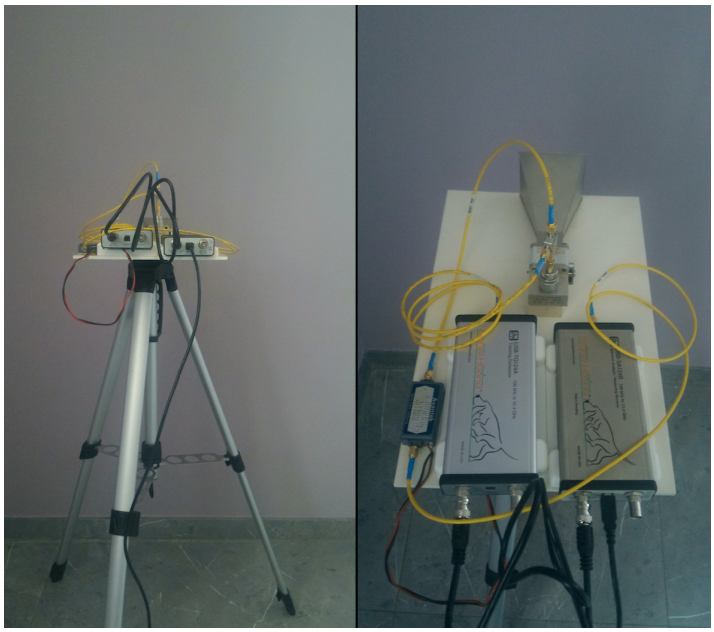


Figure 2: Photo of the assembled SFCW radar prototype.

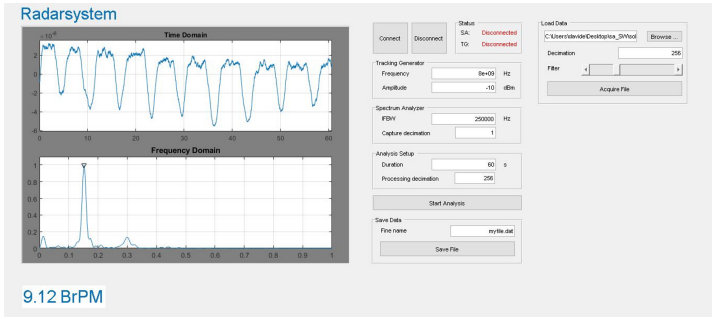


Figure 3: The SFCW radar graphical user interface (GUI).

4 Numerical and experimental assessment

This section is aimed at assessing the capabilities of the proposed SFCW radar. In particular, in sub-section 3.1 are reported the simulated results presented in [19] while sub-section 3.2 reports a preliminary experimental assessment aimed at detecting the human presence behind a wall. This experiment has been carried out to demonstrate the capability of this system to detect fugitives.

4.1 Numerical assessment

In this sub-section, the proposed SFCW radar has been assessed in realistic operative scenarios suggested by the GdF. In particular, the contents of different containers equipped with the false bottom are checked with a fast scan-A (1D) analysis. The radar system, the antenna and related scenarios are accurately simulated with a

customized numerical time-domain electromagnetic engine (FDTD). All the simulations have been performed with quad-cores personal computer, 16 GByte RAM. A gaussian white noise has been added to the original data to simulate better a realistic scenario. The signal to noise ratio of this experiment was $SNR = 5dB$. Different SNR ratio has been tested in a controlled environment, a noise above 5 dB produced very low effects on the simulations demonstrating the robustness of the proposed system. With reference to Figs. 4, 6, and 10, the blue colour represents the container's metallic walls, grey colour represents air-filled, brown colour relates to a homogeneous content of woods or granite blocks. The small orange lines in Fig. 6 represents empty space between wood blocks. The light pink colour in Fig. 10 represents a homogeneous block of tobacco.

4.1.1 *Empty container with false bottom at different distance*

Smugglers usually modify containers stored in ports by inserting false bottoms of different material inside them, the goal is to hide goods and illegal material. In the first experiment, an empty standard container of length 6.058 m, width 2.438 m, and height 2.591 m is considered. The container doors are opened, and the radar pointed precisely in the middle of the aperture at position $x_r = 1.219m$, $y = 1.2945m$.

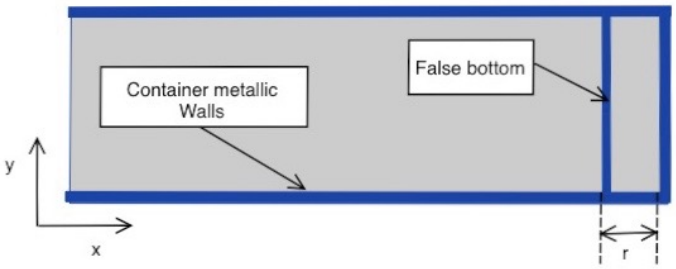


Figure 4: Empty container equipped with a metallic false bottom placed at different distances, A 0.353 m, B 0.453 m, C 0.554 m, D 0.653 m, and E 0.754 m.

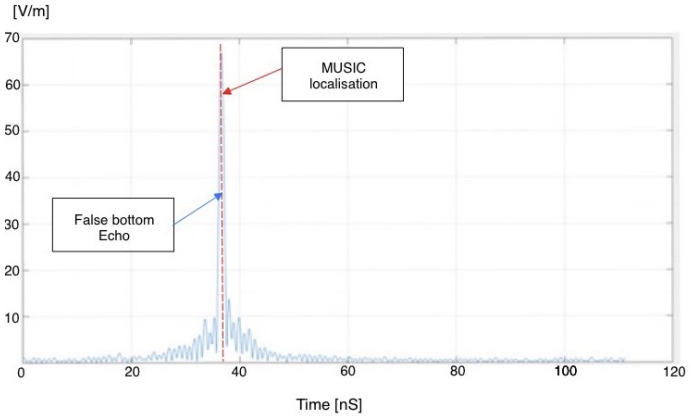


Figure 5: Empty container equipped with a metallic false bottom placed at different r=0.554 m, test case C. Time domain retrieved signal, blue line, MUSIC peak detection dotted red line.

Figure 10 shows the container section equipped with false metallic bottom placed at different distances. The radar parameters are the following: $f_{min} = 1.0$ GHz, $f_{max} = 2.071$ GHz, $M = 120$, $\Delta f = 90MHz$, bandwidth $B = 1.071$GHz. With these parameters an unambiguous range $R_{max} = 16.8m$ and a resolution of $\Delta R = 0.14m$ can be reached. To obtain the position of the false bottom, it is enough a single scan A radar trace. The synthetic data were obtained with an FDTD electromagnetic simulator, namely the gprMax, a well-known gnu software for ground penetration radar (GPR) simulations. To consider the defocusing effects, due to the reflections on container metallic walls, a set of simulations have been performed, considering an empty container and different frequency bands. The goal is to evaluate the effects of reflections and reduce the defocusing effect which can afflict the detection of the peak. Different false bottom positions were considered as reported in Tab. 1 and the synthetic data obtained with gprMax, for each frequency step, are combined and post-processed with the same procedure reported in [25] to obtain the response of a stepped frequency continuous wave radar. Tab. 2 reports the positions of false bottom obtained with a simple peak detection of the inverse discrete Fourier transformer in the time domain. For the sake of comparison, the false bottom position were also estimated with the MUSIC algorithm. Tab. 2 also reports the position errors obtained with the simple peak detection using the MUSIC algorithm. As it can be noticed from the data of Tab. 2 the localization error is very small. The MUSIC algorithm simplifies the data reading and interpretation, keeping the errors below five millimetres. An example of post-processed I/Q data converted into a scan-A time-domain diagram is reported in Fig. 5. As it can be noticed from the data reported in Fig. 5 the peak detected with the standard post processing procedure and with the music algorithm overlaps since both methods localize the peak with a high degree of accuracy. The next experiment considers an empty container with a non-metallic false bottom. This experiment's scope is to assess the capabilities of radar to detect low reflective materials such as wood and plastic wall. Three different commonly used materials, (dry wood, plexiglas, and pressed board wood, whose dielectric characteristics are summarized in Tab. 3) are considered for the wall false bottom realization. The geometry of the considered scenario is shown in Fig. 6. The scan-a diagram's simulation results obtained positioning the electromagnetic source in the middle of container aperture are reported in Figs. 7,8 and 9 respectively. The echo of the container end wall was filtered to visualize the false bottom's echo better. This filtering procedure is quite simple since the real dimensions of a standard container are known. For all considered scenarios, the false bottom made of dielectric material was successfully identified and localized with a high degree of accuracy. In particular, the wood dry false bottom (Fig. 7) is localized with a very low error $\gamma = 2.7mm$. The localization of plexiglas false bottom (Fig. 8) presents a slightly high error $\gamma = 6.8mm$ with respect to the wood dry false bottom wall. The last scenario (Fig. 9) provide the worst results: the localization error $\gamma = 16.1mm$, even though is quite high with respect to the two previous scenarios; still satisfactory.

Table 1: False bottom positions.

	A	B	C	D	E
r [m]	0.353	0.453	0.554	0.653	0.754

Table 2: Retrieved false bottom positions with peak detection and MUSIC algorithm.

Scenario	Peak Pos	MUSIC Pos.	Err. Peak	Err. MUSIC
A	5.700 m	5.701m	5.0 mm	3.5 mm
B	5.600 m	5.602 m	4.9 mm	3.3 mm
C	5.499 m	5.500 m	3.9 mm	3.4 mm
D	5.399 m	5.402 m	4.4 mm	3.3 mm
E	5.302 m	5.301 m	3.3 mm	3.0 mm

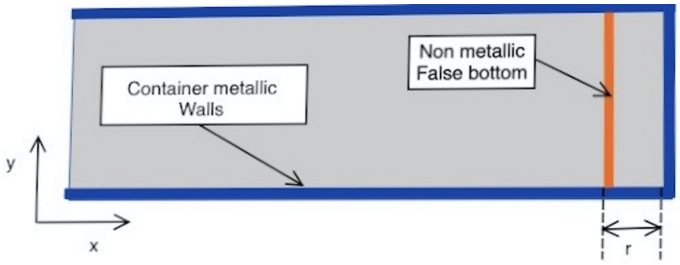


Figure 6: Empty container equipped with a non metallic false bottom placed at distance $r = 5.40m$.

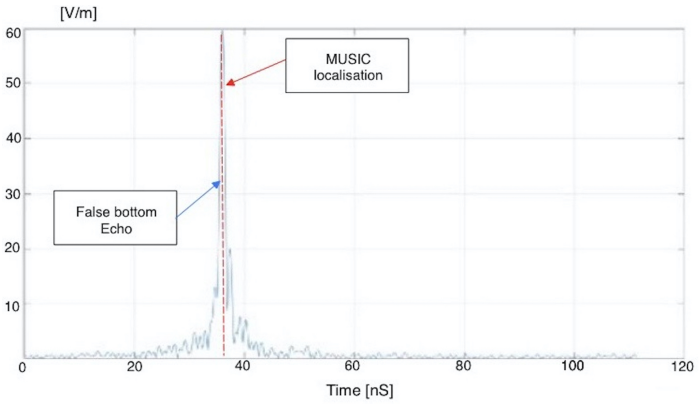


Figure 7: Empty container with a wood dry false bottom placed at distance $r = 5.40m$. Retrieved echo radar signal and MUSIC peak estimation.

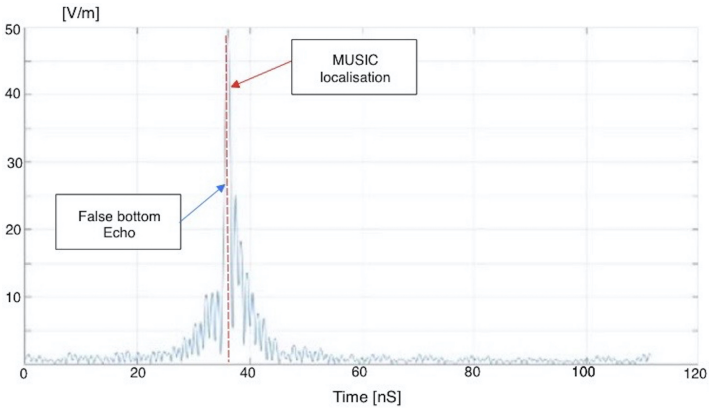


Figure 8: Empty container with a plexiglas false bottom placed at distance $r = 5.40m$.. Retrieved echo radar signal and MUSIC peak estimation.

Table 3: Dielectric characteristics of wall materials used for the false bottom.

Scenario	Material	ε_r
A	Wood, dry	4.5
B	Plexiglas	3.5
C	Wood, pressed board	2.0

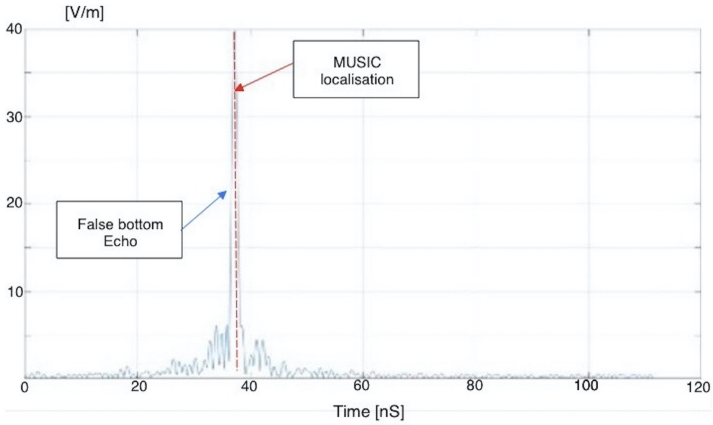


Figure 9: Empty container with a wood pressed board false bottom, placed at distance $r = 5.40m$.. Retrieved echo radar signal and MUSIC peak estimation.

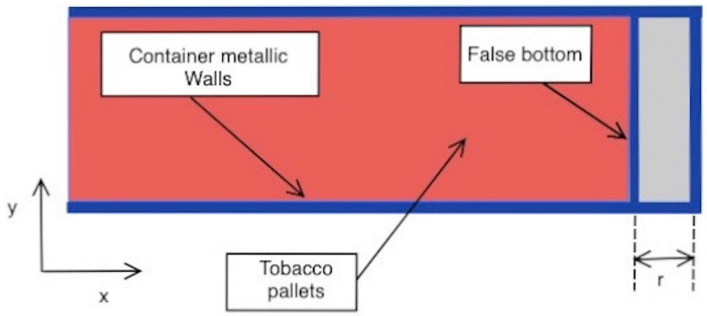


Figure 10: Container filled with tobacco pallets and a metallic false bottom placed at a distances $r = 5.30m$.

In the following experiments, different goods fill the container. The objective is to assess the radar capability to successfully detect the presence of false bottoms or other anomalies in goods without the necessity of unloading the container. This aspect is essential for police operations because it permits to check off more containers in less time. The shipper's shipping manifest permits estimating an electromagnetic wave's velocity and finding anomalies in the load. The first experiment considers a container filled with tobacco pallets and a metallic false bottom placed at a distance $r = 5.3m$ from the container door. The goal is to assess the capability to detect the false bottom despite the load presence. Scan-a result of the above scenario is reported in Fig. 11, as it can be noticed the echo is quite noisy due to the load presence. However, thanks to the MUSIC localization algorithm, the false bottom is successfully localized with a high level of accuracy with an error $\gamma = 12.5mm$. In the last experiment, the container is filled with blocks of granite. The dielectric permittivity of granite is $\varepsilon_r = 7.0$ quite high. This experiment is of particular interest because unload a container filled with granite

blocks requires much time and specific mechanical facilities. The scan-a results obtained from the simulation performed with gprMax are reported in Fig. 13. The echo is very well defined and presents a sharp peak easily detected by the MUSIC algorithm. The metallic false bottom is successfully detected with an error $\gamma = 2.3mm$. It is worth noticing that the results reported in Fig. 13are less noisy with respect to the data related to 11 a little bit noisier.

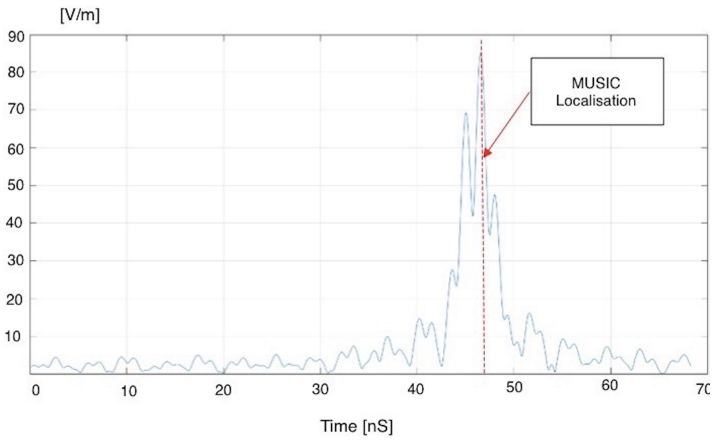


Figure 11: Container filled with tobacco pallets and a metallic false bottom placed at a distances $r = 5.30m$. Retrieved echo radar signal.

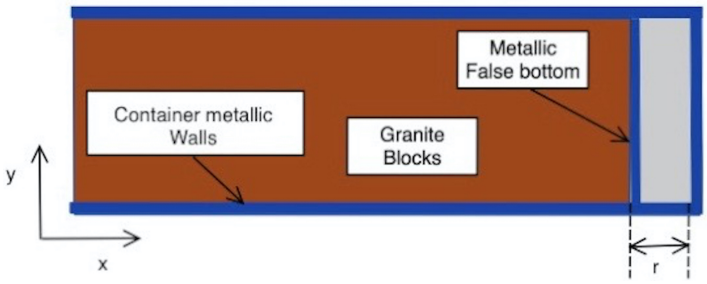


Figure 12: Container filled with granite blocks and a metallic false bottom placed at a distances $r = 5.30m$.

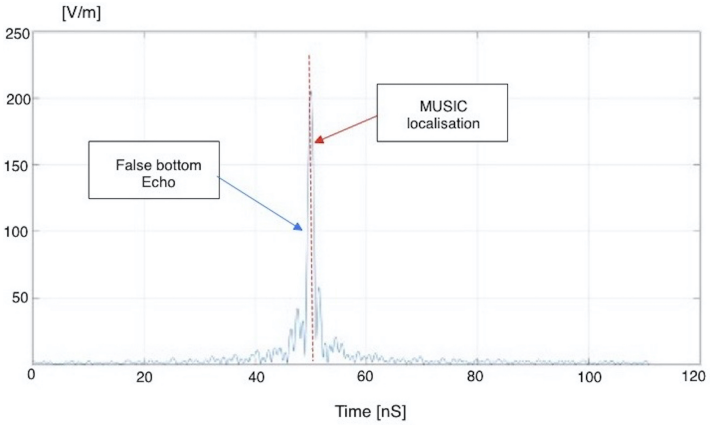


Figure 13: Container filled with granite blocks and a metallic false bottom placed at a distances $r = 5.30m$. Retrieved echo radar signal.

4.2 Experimental assessment

In this section, the SFCW radar prototype has been assessed in an operative scenario. The experimental setup refers to a typical through the wall application where a fugitive hides in a hiding place created behind a concrete wall. To detect humans hide behind walls, the radar prototype operates in continuous wave (CW) mode. The small chest and heart movements are detected by considering the Doppler effect carried on by the scattered electromagnetic wave. The wall thickness is $25cm$, the distance of the fugitive's chest from the wall is $d = 1.5m$, and he is sitting on a chair. The radar prototype is manually steered to the other side of the wall at a distance of $d_r = 1.0m$ with the goal of detecting the fugitive by revealing his signals life such as the breathing rate and the heartbeat. A photo of the considered through the experimental wall setup is reported in Fig.14.

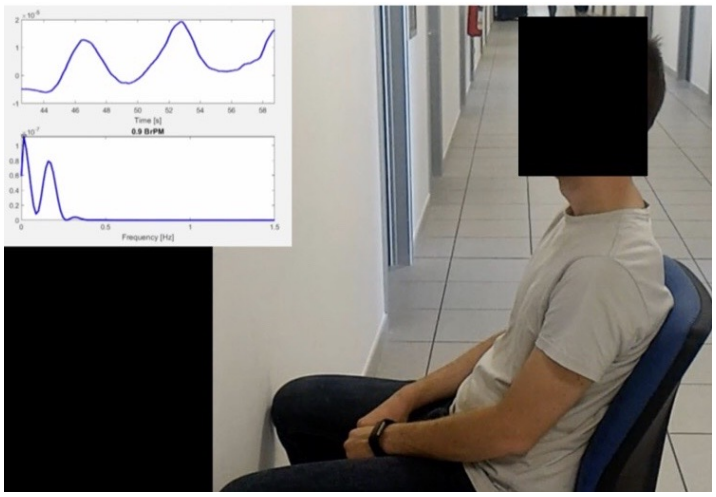


Figure 14: Experimental set-up. Radar in through the wall modality. Life signal detection, breathe rate and heartbeat, of a subject placed behind a 25 cm thickness wall.

The tool for the detection of human life signals has been selected in the radar GUI. The mechanical pedestal has been set at hight $h = 1.5m$ and then horizontally moved along the wall with a step size of about $s = 0.50m$, the total length of the wall was $2.0m$. The DDS generator's power has been set to the minimum value of $P_{tx} = -13dBm$. The electromagnetic wave power that reaches the fugitive chest is very low, and it cannot create health or interference problems with other devices. The following Fig. 15 reports the fugitive breathing rate detected at the second scan $1.0m$ far away from the sidewall. As it can be noticed from the data reported in Fig. 15, the fugitive shows normal breathing conditions, and it is quite relaxed. When the radar is moved on the first, third or fourth positions at $0.5m$, $1,5m$ and $2.0m$ respectively, far away from the sidewall, the system reports no human activities. In the next experiments, the fugitive simulated an accelerated breathing rate increasing from eight, with reference to Fig. 15, up to twenty breath cycles for minutes reported in Fig. 16. The radar prototype's resolution capabilities are quite accurate, and it can detect the small movements of the fugitive's heart. Fig. 17 (red line) reports the heart movement's radar signal. For the sake of comparison, the data reported in Fig. 17 has been compared with the data obtained

with a commercial optical heartbeat detector (blue line), as it can be noticed from Fig. 17 the agreement is quite accurate, and it demonstrate the capabilities of this prototype not only for police operations but also for the rescue operation.

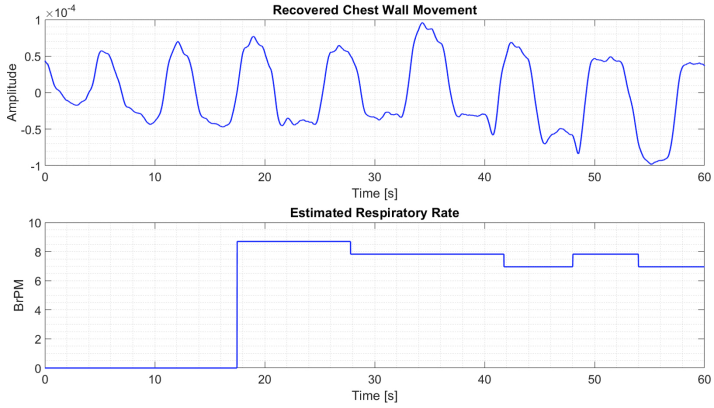


Figure 15: Experimental set-up. Life signals, breathing rate, of the fugitive detected behind the wall. Normal breathing conditions.

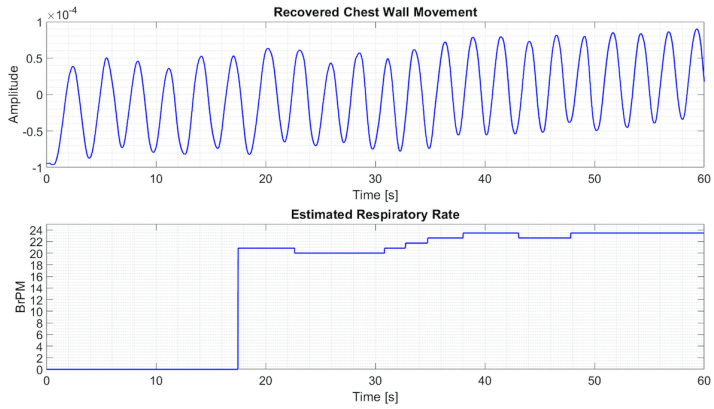


Figure 16: Experimental set-up. Life signals, breathing rate, of the fugitive detected behind the wall. Simulated accelerated breathing conditions.

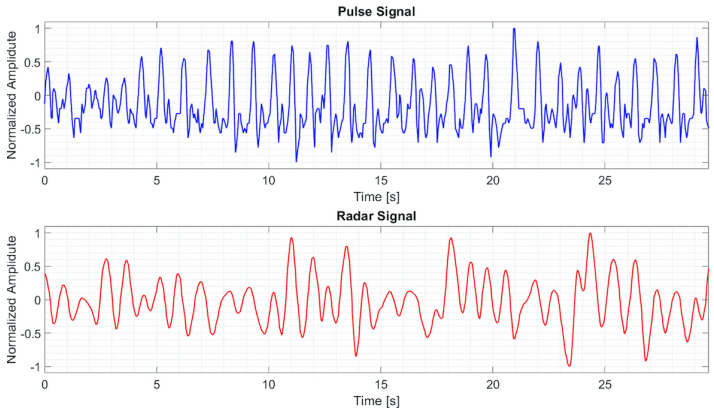


Figure 17: Experimental set-up. Life signals, heartbeat rate of the fugitive detected behind the wall with a more focused beam.

5 Conclusion

In this work, a Stepped Frequency Continuous Wave (SFCW) radar operating in the L band, suitable for fast inspection of foods and goods stored in metallic containers, has been proposed and numerically assessed in realistic scenarios. In particular, the proposed SFCW radar can estimate the dielectric characteristics of materials contained in a container and easily detect metallic or dielectric false bottoms commonly used for smuggling. A numerical assessment campaign, carried on considering different materials and simulating realistic scenarios, demonstrated the effectiveness and potentialities of the proposed system as a helpful tool for the contrast of smuggling.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors would like to thank Dr. E. Bortolotti for management the project and for the revision of manuscript, and the Autostrade Brennero S.P.A. A22 for funding the radar prototype.

References

[1] C. Nguyen, P. Joongsuk, Stepped-frequency radar sensors, Springer, 2016.

[2] G. Tronca, I. Tsalicoalou, S. Lehner, G. Catanzariti, "Comparison of pulsed and stepped frequency continuous wave (SFCW) GPR systems," in 17th International Conference on Ground Penetrating Radar (GPR), 1–4, 2018, doi: 10.1109/ICGPR.2018.8441654.

[3] A. Langman, M.-R. Inggs, "Pulse versus stepped frequency continuous wave modulation for ground penetrating radar," in Proceeding of International Geoscience and Remote Sensing Symposium, IGARSS 2001, 1533–1535, 2001, doi:10.1109/IGARSS.2001.976902.

[4] G. Gok, Y. K. Alp, O. Arikan, "A New Method for Specific Emitter Identification With Results on Real Radar Measurements," IEEE Trans. Inform. Forensic Secur., **15**, 3335–3346, 2020, doi:10.1109/TIFS.2020.2988558.

[5] L. Nicolaescu, P. V. Genderen, K. Palmer, "Calibration procedures of a stepped frequency continuous wave radar for landmines detection," in Proceedings of the International Conference on Radar (IEEE Cat. No.03EX695), 412–417, 2003, doi:10.1109/RADAR.2003.1278777.

[6] A. S. Turk, A. Kizilay, M. Orhan, A. Caliskan, "High resolution signal processing techniques for millimeter wave short range surveillance radar," in 17th International Radar Symposium (IRS), 1–4, 2016, doi:10.1109/IRS.2016.7884697.

[7] A. Kusmadi, A. Munir, "Simulation design of compact stepped-frequency continuous-wave through-wall radar," in International Conference on Electrical Engineering and Informatics (ICEEI), 332–335, 2015, doi:10.1109/ICEEI.2015.7352521.

[8] M. Pieraccini, M. Fratini, F. Parrini, G. Macaluso, C. Atzeni, "High-speed CW step-frequency coherent radar for dynamic monitoring of civil engineering structures," Electron. Letters, **40**, 907–908, 2004, doi:10.1049/el:20040549.

[9] M. Pieraccini, M. Fratini, F. Parrini, C. Atzeni, "Dynamic Monitoring of Bridges Using a High-Speed Coherent Radar," IEEE Trans. Geosci. Remote Sensing, **44**, 3284–3288, 2006, doi:10.1109/TGRS.2006.879112.

[10] S. S. Lambot, E. C. Slob, I. Van-den Bosch, B. Stockbroeckx, M. Vanclooster, "Modeling of ground-penetrating Radar for accurate characterization of subsurface electric properties," IEEE Trans. Geosci. Remote Sensing,, **42**, 2555–2568, 2004, doi:10.1109/TGRS.2004.834800.

[11] W.-C. Su, M.-C. Tang, R. E. Arif and, T.-S. Horng, F.-K. Wang, "Stepped-Frequency Continuous-Wave Radar With Self-Injection-Locking Technology for Monitoring Multiple Human Vital Signs," IEEE Trans. Microwave Theory Techn., **67**, 2555–2568, 2019, doi:10.1109/TMTT.2019.2933199.

[12] M. Caruso, M. Bassi, A. Bevilacqua, A. Neviani, "A 2–16 GHz 65 nm CMOS Stepped-Frequency Radar Transmitter With Harmonic Rejection for High-Resolution Medical Imaging Applications," IEEE Trans. Circuits Syst., **62**, 413–422, 2015, doi:10.1109/TCSI.2014.2362332.

[13] M. Donelli, F. Viani, "Life signals detection system based on a continuous-wave X-band radar," Electronics Letters, **52**, 1903–1904, 2016, doi:10.1049/EL.2016.2902.

[14] I. Nicolaescu, "Improvement of Stepped-Frequency Continuous Wave Ground-Penetrating Radar Cross-Range Resolution," IEEE Trans. Geosci. Remote Sensing, **51**, 85–92, 2013, doi:10.1109/TGRS.2012.2198069.

[15] M. Scherhaufl, F. Hammer, M. Pichler-Scheder, C. Kastl, A. Stelzer, "Radar Distance Measurement With Viterbi Algorithm to Resolve Phase Ambiguity," IEEE Trans. Microwave Theory Techn., **68**, 3784–3793, 2020, doi:10.1109/TMTT.2020.2985357.

[16] G. Cui, L. Kong, X. Yang, "Reconstruction Filter Design for Stepped-Frequency Continuous Wave," IEEE Trans. Signal Process., **60**, 4421–4426, 2012, doi:10.1109/TSP.2012.2197206.

[17] S. R. J. Axelsson, "Analysis of Random Step Frequency Radar and Comparison With Experiments," IEEE Trans. Geosci. Remote Sensing, **45**, 804–904, 2007, doi:10.1109/TGRS.2006.888865.

[18] A. N. Gaikwad, D. Singh, M. J. Nigam, "Application of clutter reduction techniques for detection of metallic and low dielectric target behind the brick wall by stepped frequency continuous wave radar in ultra-wideband range," IET Radar Sonar Navig., **5**, 416–425, 2011, doi:10.1049/iet-rsn.2010.0059.

[19] M. Donelli, M. Manekiya, G. Marchi, I. Maccani, C. Pascucci, "High Resolution L-band Stepped Frequency Continous Wave Radar for Smuggling Contrast," in 2020 2nd Global Power, Energy and Communication Conference (GPECOM), 327–332, 2020, doi:10.1109/GPECOM49333.2020.9247932.

[20] T. Yamakura, H. Yamada, Y. Yamaguchi, "Resolution improvement of the MUSIC algorithm utilizing two differently polarized antennas," in IEICE Transaction on Communications, 1827–1832, 1996, doi:10.1155/2019/8907685.

[21] S.-M. Shrestha, I. Arai, T. Miwa, Y. Tomizawa, "Signal processing of ground penetrating radar using super resolution technique," in IEEE Radar Conference, 300–305, 2001, doi:10.1109/NRC.2001.922995.

[22] K.-T. Kim, D.-K. Seo, H.-T. Kim, "Efficient radar target recognition using the MUSIC algorithm and inveriant features," IEEE Trans. Antennas and Propagation, **50**, 325–337, 2002, doi:10.1109/8.999623.

[23] S.-M. Shrestha, I. Arai, T. Miwa, "Signal processing of ground penetrating radar combining FFT and MUSIC algorithm for high resolution," in Tech. Rep. IEICE SANE2000 SAT2000-130, 31–43, 2003, doi:10.1155/S1110865703307036.

[24] S.-M. Shrestha, I. Arai, "Signal processing of ground penetrating radar using spectral estimation techniques to estimate the position of buried targets," EURASIP Journal on Applied Signal Processing, **12**, 1198–1209, 2003, doi:10.1155/S1110865703307036.

[25] V. Kafedziski, S. Pecov, D. Tanevski, "Target detection in SFCW ground penetrating radar with C3 algorithm and Hough transform based on gprMax simulation and experimental data," in 2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP), 2018, doi:10.1109/IWSSIP.2018.8439227.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 1, 14-26 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Value Trace Problems for Code Reading Study in C Programming

Xiqin Lu[1], Nobuo Funabiki[*,1], Htoo Htoo Sandi Kyaw[1], Ei Ei Htet[1], Shune Lae Aung[2], Nem Khan Dim[2]

[1]*Department of Electrical and Engineering, Okayama University, Okayama, 700-8530, Japan*

[2]*Department of Computer Studies, University of Yangon, Yangon, 11041, Myanmar*

ARTICLE INFO

Article history:
Received: 05 July, 2021
Accepted: 02 January, 2022
Online: 13 January, 2022

Keywords:
C programming
Value trace problem
Code reading
Self-study
Grammar concept
Algorithm
Pointer

ABSTRACT

C programming is taught in a lot of universities across the world as the first computer programming language. Then, for novice students, it is important to read many simple C source codes and understand their behaviors to be familiar to the programming paradigm. Unfortunately, effective tools to support independent code reading study at home have not been well designed. Heretofore, we have proposed the value trace problem (VTP) for Java programming. A VTP instance consists of one source code, several questions, and the correct answers to them. Each question asks the value of a critical variable or output message in the source code. The correctness of any student answer is checked instantly by string matching at the answer interface for self-study. In this paper, we present the value trace problem (VTP) for code reading self-study of C programming. 42 VTP instances are generated using simple C source codes on basic grammar concepts and fundamental data structures & algorithms in textbooks and websites. In addition, for hard instances on pointer and algorithms, the devices of hints, multiple choice questions, and references are provided to improve their solution performances. For evaluations, we requested 49 *undergraduate students in Japan, China, and Myanmar to independently solve them at home. Their average correct answer rate reached* 94.29%*, where our devices for hard instances improved it by* 33.26%*. Thus, the effectiveness of our proposal is confirmed in motivating self-study of C programming to novice students.*

1 Introduction

Presently, *C programming* is widely taught in many universities across the world as the first computer programming language. In addition to *information technology (IT)* departments, many departments including science, agriculture and mechanical/electrical engineering are teaching *C programming* courses. Actually, the study of *C programming* can offer basic knowledge to study more advanced and practical languages such as *Java*, *JavaScript*, and *Python*. Besides, students in IT departments should study *C programming* in parallel with computer architecture, because they can learn accesses to memories or registers through it. Furthermore, *C* is still the third most popular programming language, despite the long time passed from the appearance. Then, for such novice students, it is important to read and understand a lot of simple source codes to be familiar to the *C programming* structure. Unfortunately, effective tools to support independent *code reading self-study* at home have not been well designed within our knowledge.

Heretofore, we have proposed and implemented the *Java programming learning assistant system (JPLAS)* for *Java programming* study [1]. As the objected-oriented programming language, *Java* is now widely used in IT societies. For assisting *Java programming* studies at different learning levels, JPLAS provides several types of *exercise problems* with different difficulties. For any problem type, the student answer to a question is marked automatically at the system to support programming self-studies.

Among them, the *value trace problem (VTP)* is presented for novice students to study basic concepts of programming grammar and skills through *code reading study* [2]. A VTP instance consists of one source code, several questions, and the correct answer strings to them. Each question asks the value of a critical variable or an important output in the given code. The correctness of any student answer is marked through *string matching* with the correct answer string at the answer interface using the web browser instantly.

The manual collection of C source codes and the manual selection of the variables or output messages from each source code for questions can be the limitation of *VTP*. When a teacher uses *VTP* in the programming course, he/she needs to carefully collect the

[*]Corresponding Author: Nobuo Funabiki, 3-1-1 Tsushimanaka, Okayama University, funabiki@okayama-u.ac.jp

source codes and select the variables or messages with their timing whose actual values or contents should be traced in the questions, so that they can be synchronized with the course progress. One way to generate VTP instances by a teacher efficiently is the use of sample source codes that will be provided with the textbook in the course. Usually, the teacher can download them from the website. Then, he/she can select the variables or output messages that are closely related to the teaching topics in each class of the course. The impossibility of practicing source code writing by a student will be another limitation of *VTP*. It should be offered by other problem types such as the *code writing problem* [3].

In this paper, we present the *value trace problem (VTP)* for *code reading self-study* of *C programming* by extending our previous works for Java programming. To assist the study of novice students, 42 simple C source codes on *basic grammar concepts* and *fundamental data structures & algorithms* are collected from textbooks and websites. Here, *fundamental data structures & algorithms* are often taught using *C* programs in IT departments. Then, by analyzing critical variables/messages and adding necessary standard output statements in them, the corresponding VTP instances are generated manually.

The memory management using *pointer* in *C programming* can be a hard topic for novice students to study, although it is the remarkable feature in implementing fast and efficient programs. Besides, some fundamental algorithms can be difficult, because novice students are often not familiar to them. Therefore, we additionally provide the devices of hints, multiple choice questions, and references to improve the solving performances by the students for the hard VTP instances on *pointer* and *algorithms*.

For evaluations of the proposal, we asked 49 students in Japan, China, and Myanmar to independently solve the generated 42 VTP instances at home. Then, the results showed that the average correct answer rate among them reached 94.29% that is sufficiently high, where the devices for hard instances improved it by 33.26%. Thus, the effectiveness of our proposal is confirmed in motivating self-study of *C programming* by novice students [4]–[10].

VTP can be applied for studying other programming languages such as *JavaScript* and *Python* in addition to *Java* in a straightforward way. Besides, *VTP* can be used for studying a natural language such as English or Japanese as the potential application. For natural language study, the instance may give a sentence or a paragraph and ask the appearing keywords in the questions. Here, the variable for programming study becomes the key concept in the sentence or paragraph for natural language study, and the value becomes the keyword.

The remaining part of this paper is organized as follows: Section 2 discusses related works in literature. Section 3 presents the *value trace problem (VTP)* for *C programming*. Section 4 presents devices to improve the solving performances for hard instances on *pointer* and *algorithms*. Section 5 shows their evaluation results. Finally, Section 6 concludes this paper with future works.

2 Related Works in Literature

In this section, we introduce related works in literature. Some papers presented programming study tools, and some discussed the importance of code reading in programming study.

2.1 Programming Study Tools

In [11], the author reviewed some tools to support teaching and learning of programming, and categorized them into four groups: 1) tools that include simple reduced development environments, 2) example-based environments, 3) tools that are based on visualizations and animations, and 4) simulation environments. Our proposal can be categorized to 2).

In [12], the author presented a tool called *QuizPACK*, which is similar to the *value trace problem (VTP)*, and showed that this tool significantly improved the knowledge of semantics. A question in *QuizPACK* asks the value of a particular expression (usually a variable) in a fragment of a program. A teacher needs to prepare a set of problems using source code fragments and expressions to be questioned manually. If compared with *VTP* in this paper, the advantage of *QuizPACK* is that the constants related to the questioned variables in the code and their correct answers can be generated dynamically so that different students can solve the same question with the different correct answers. On the other hand, *VTP* fixes the constants and the correct answers. The disadvantage of *QuizPACK* is that it can only ask the ending values of the variables in the code fragment. It cannot ask the values at the different timing of the program. Thus, *QuizPACK* is not suitable for studying *algorithms*, where different algorithms can give the same variable values.

In [13], the author presented a problem-solving environment named *LECGO (Learning Environment for programming using C using Geometrical Objects)* for learning *C programming* by beginners. *LECGO* emphasized: (a) multiple external representations in student learning, (b) motivation through performing problem-solving activities from the familiar and meaningful context, (c) the active participation of students by using hands-on experience, (d) appropriate feedback to aid self-corrections, and (e) holistic, activity-based, multi-media, multi-representational and multi-layered content for learning basic concepts.

In [14], the author presented *Gidget*, which is a game such that the eponymous robot protagonist is cast as a fallible character that blames itself for not being able to correctly write code to complete its missions. Players of *Gidget* can learn programming by debugging the problematic code.

In [15], the author presented a game-based learning support environment for novice students to learn programming. This environment exploits game composition tasks to make the elementary programming more intuitive to be learnt, and comprises concept visualization techniques, to allow the students to learn key concepts in programming through game object manipulations.

In [16], the author proposed the *web-based programming assisted system for cooperation (WPASC)* for facilitating cooperative programming learning, and investigated cooperative programming learning behaviors of students and its relationships with learning performances.

In [17], the author presented systematic literature review results on assessment tools for programming assignments, to help instructors make their selections in programming courses. They identified three specialties in assessment tools, namely, *contests*, *quizzes*, and *software testing*. In *contests*, a tool compiles the source code of a

student, runs it with a set of test cases for the input, and informs whether the code is accepted or not. In *quizzes*, it gives a set of questions, where for each question, a student enters a fragment of a code in the interface of the tool. In *software testing*, it checks the correctness of the source code of a student by comparing its output with the output of the model code, or by running the test code against the source code. VTP in this paper is different from any of them.

In [18], the author presented a system that combines the dual objectives of automated grading and program repairing for introductory programming courses called *GradeIT*. It grades the submitted source code with the number of passed test cases, the inverse of time spent to solve, and the fraction of successful compilations. For the last one, the compilation score is defined by 1 – (number of compilation errors in the code) / (maximum number of compilation errors among students). For repairing, it uses simple re-writing rules to fix simple but frequent compile time errors.

In [19], the author proposed a plug-in system to *Moodle* called *LAPLE (Learning Analytics in Programming Language Education)* to provide a learning dashboard to capture the behaviors of the students in the classroom and identify the difficulties faced by different students looking at different knowledge. *LAPLE* asks the students to write full source codes, and collects and analyzes the compiling logs of the students every five minutes to yield real-time visualizations for feedbacks during the class. The authors analyzed the distribution of the classified 36 error types, and encouraged the students to pay attentions to the frequent error types. *LAPLE* is designed for use at programming classes, while *VTP* is for use at homes for self-studies.

2.2 Code Reading in Programming Study

In [20], the author concluded that through literature reviews and interviews, *code reading* is connected to comprehending programs and algorithms, or algorithmic ideas, as well as details, and is needed in many aspects of learning programming, but at the same time, there is not much knowledge about the reading and comprehension process of learners. They claimed that a possible means to foster programming learning is to teach code reading directly, including reading strategies. The value trace problem in this paper can be one way to achieve it for *C programming* study, which supports the novelty of this paper.

In [21], the author concentrated on finding out whether reading programs before writing programs is more efficient for students to learn programming. He asked software developers and found that computer science graduates should understand existing codes, since most developers need to maintain and extend existing systems. He concluded that if students get through the syntax of programming more quickly by conducting code reading exercises, class time could be spent on issues such as the efficiency and readability of software. Code reading problems are important for beginners to learn a new programming language.

In [22], the author concluded that programming learning should focus on algorithmic thinking, rather than programming language features and skills. Programming learning could help improve problem solving skills through algorithmic thinking. The generated VTP instances in this paper cover important algorithmic topics such as stack, queue, and sorting algorithms, which will be helpful for developing algorithmic thinking.

In [23], the author indicated that source code reading behaviors are important in teaching programming skills as well as designing IDEs and programming languages. A source code is a different type of text than a natural-language text. It is highly formal and structured, and has a very limited vocabulary of keywords, operators, and separators, yet, tremendous combination possibilities for literals and identifiers. They recorded programmers' eye movements when reading short source codes. The results show that most attention is oriented towards understanding of identifiers, operators, keywords, and literals, relatively little reading time is spent on separators.

3 Value Trace Problem for C programming

In this section, we present the *value trace problem (VTP)* for *C programming*.

3.1 Definition of Value Trace Problem

The mission of *VTP* is to let students trace the value of a critical variable or an important output in the given source code by carefully reading and understanding its behaviors. To generate a VTP instance, one *source code*, a set of *questions* with the *answer forms*, and the *correct answer strings* to them need to be prepared. A question requests the current value of the important variable or message in the source code that is specified by the teacher. Therefore, the source code should contain the *standard output* statements that correspond to them, so that the students can easily find which values are asked in the questions.

3.2 Design Goals of Value Trace Problem

VTP for *C programming* have the following design goals:

1) A variety of C source codes are given to beginners for *code reading study*.

2) A student can solve the questions in a VTP instance by reading the given source code and understanding its behavior carefully.

3) Any answer from a student will be checked through *string matching* at the interface automatically, and the result will be returned to the student instantly.

4) A teacher can choose proper source codes in terms of contents and difficulty levels for the VTP instance that will be assigned at each class in the programming course.

3.3 VTP Instance Generation Procedure

A new VTP instance can be generated by the following procedure:

1. A proper *source code* for studying a *basic grammar concept* or a *fundamental data structure/algorithm* is selected.

2. The important *variables* or *output messages* to be traced in the source code are selected.

3. The corresponding *standard output statements* are added in the code.

4. The questions of asking the values/messages at the standard outputs and their *correct answers* are prepared.

5. The source code, the corresponding questions, and the correct answers are put together into one *input text file*.

6. The program in [4] is executed with the *input text file* to generate the *CSS/HTML/JavaScript files* for the *answer interface* on the web browser.

7. The generated VTP instance is registered as the assignment to students.

3.4 Answer Interface

The *answer interface* using the web browser in [4] is adopted for the students to solve the generated VTP instances in this paper. The necessary functions such as *string matching* for answer marking and the answer data storage for submissions are implemented in *JavaScript* so that this interface can be used without the Internet connection.

To use this answer interface, the correct answers to the questions need to be distributed to the students, so that their answers can be marked instantly at offline. Then, to prevent the students from looking at the correct answers before solving the questions, the hash values of them are taken using *SHA256* before the distributions. Actually, each correct answer is concatenated with the assignment and problem IDs before hashed, so that the same correct answers for the different questions are converted to the different hash values. It can avoid knowing the correct answer from the same hash value for a different question. At the marking, any answer from a student is concatenated and hashed using *SHA256* for *string matching* with the hashed correct answer.

3.5 Example VTP Instances

Two VTP instances generated in this paper are illustrated here as the typical examples for the *variable trace study* and the *pointer trace study* in *C programming*.

3.5.1 VTP Instance for Variable Trace Study

Figure 1 shows the *input text file* for the *if in for-loop* instance at *ID=12* in Table 1. This file contains the source code, the questions with four answer forms, and their correct answers, which must be done manually. Each correct answer is separated by ",". This instance asks to trace the value of *num* and the logic implemented by two *if*. A student needs to read and understand the source code, and then fill in the forms with correct answers.

```
int main(void){
   int max = 9;
   int n;
   for (n = 2; n <= max; n++){
      if (n == max){
         printf("%d is prime number \n", max);
      }
      else if(max % n == 0){
         printf("%d is not prime \n", max);
         break;
      }
   }
   return 0;
}

What is the value of max? _1_
What is the output? _2_ is _3_ _4_

9,9,not,prime
```

Figure 1: Input text file for VTP instance at *ID=12*.

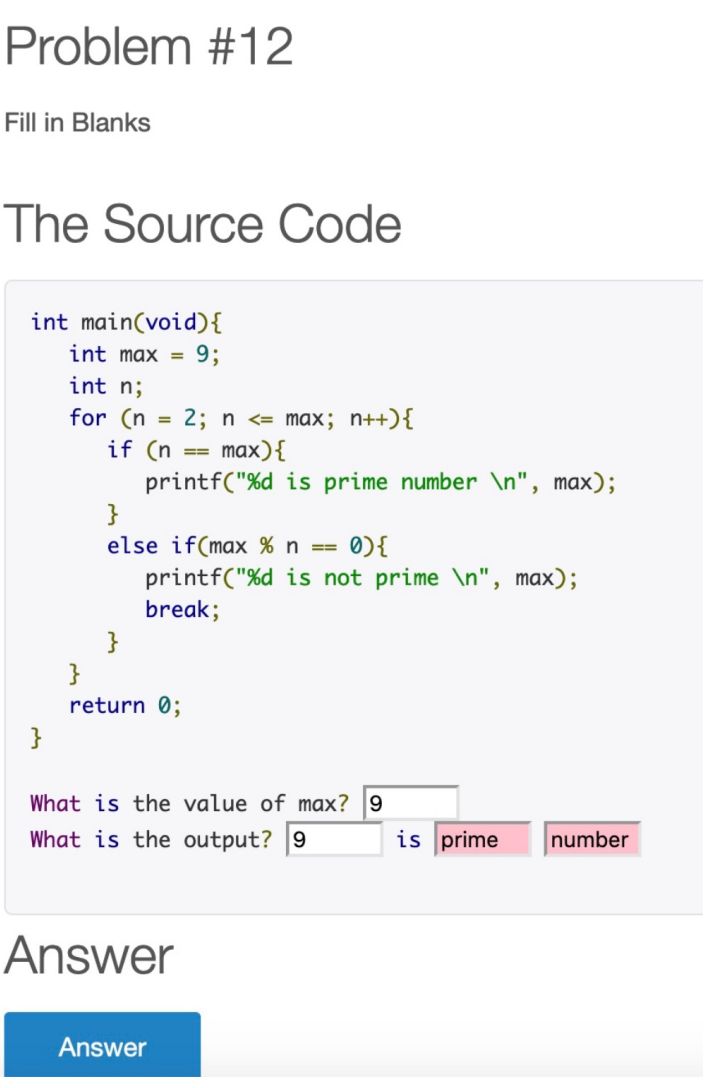


Figure 2: Answer interface for VTP instance at *ID=12*.

```
int main(void)
{
   int test[5] = {80,60,55,22,75};

   printf("test[0] is %d \n", test[0]);
   printf("test[1] is %d \n", test[1]);
   printf("The address of test[0] is %p \n", &
       test[0]);
   printf("test is %p \n", test);
   printf("The address of test[1] is %p \n", &
       test[1]);
   printf("That is *test is %d \n", *test);
   return 0;
}

hint:The address of test[0] is 0028FF20
     1 word 4 bit in this system

What is the output of this program?
test[0] is  _1_
test[1] is  _2_
The address of test[0] is  _3_
test is  _4_
The address of test[1] is  _5_
That is *test  _6_

80,60,0028FF20,0028FF20,0028FF24,80
```

Figure 3: Input text file for VTP instance at *ID=9*.

Figure 2 illustrates the answer interface on a web browser for this VTP instance. After a student fills in the answer forms, he/she should click the blue button "Answer". If the student answer is incorrect, the form color will be *red*. For the correct one, it becomes *white*. Using this interface, a student can fix the mistakes and submit the answers instantly, until all the answers become correct.

3.5.2 *VTP Instance for Pointer Trace Study*

Figure 3 shows the *input text file* for the *memory address with pointer for integer array* instance at *ID=9* in Table 1. This instance asks to trace the memory address or the value given by *pointer* including the one for the array *test*. Figure 4 depicts the answer interface for this sample VTP instance.

3.6 *Generations of Value Trace Problem Instances*

Now, we present the generated 42 *value trace problem (VTP)* instances with 586 answer forms for studying *basic grammar concepts* and *fundamental data structures & algorithms* in *C programming*.

Table 1 shows the topic, the number of lines (#of lines) in the source code, the number of questions, and the number of answer forms in each VTP instance. Here, to encourage novice students solving all of them without giving up halfway, we took the following stances at their generations:

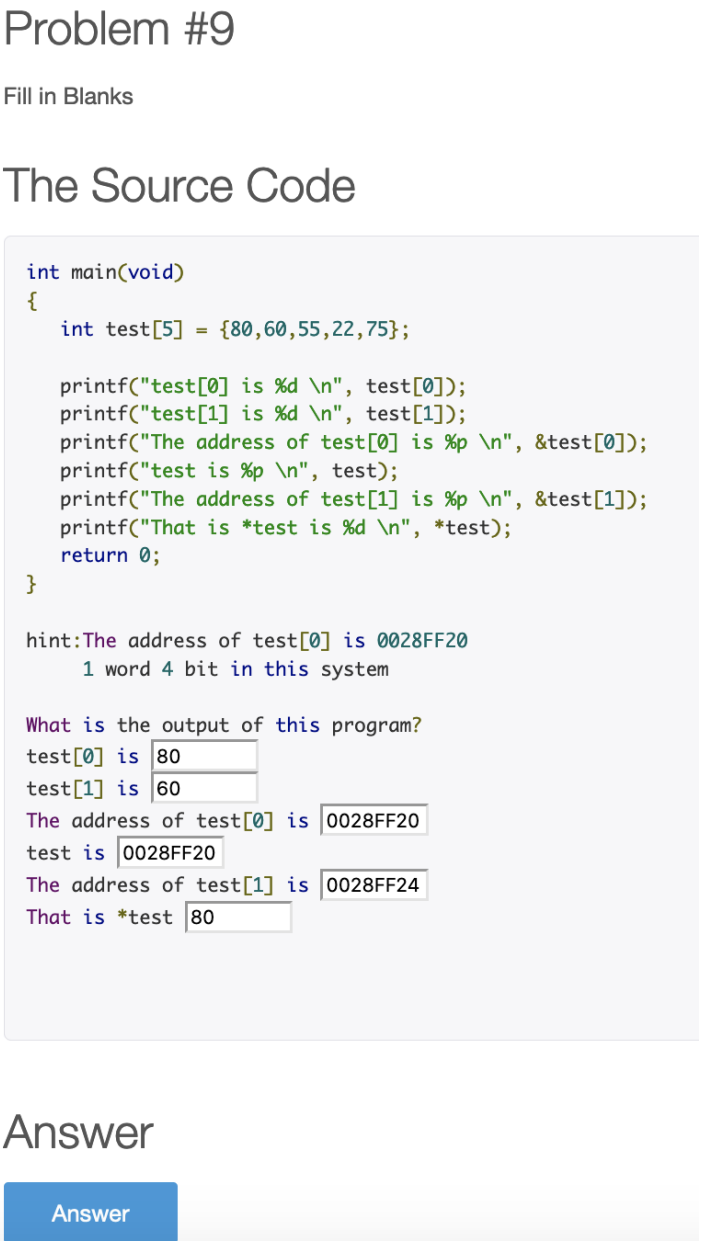


Figure 4: Answer interface for VTP instance at *ID=9*.

1. A short source code with less than 30 lines is basically selected for each instance. Actually, the average number of lines among the source codes becomes 24.5. However, source codes for *fundamental data structures & algorithms* are longer due to their natures.

2. The number of questions is limited to at most four for each instance. Actually, the average number of questions becomes 1.26.

3. The difficulty level of each instance is gradually increased from the first to the last roughly, by following *C programming* textbooks.

3.6.1 VTP Instances for Basic Grammar Concepts

For studying *basic grammar concepts*, 26 VTP instances were generated using the C source codes in the websites [5]-[7]. The selected grammar concepts are essential in fundamental *C programming*.

3.6.2 VTP Instances for Fundamental Data Structures & Algorithms

To study *fundamental data structures & algorithms* of C programming, 16 VTP instances were generated using the source codes in the textbook [8]. They can be hard subjects for novice students in IT departments. To let them read the implemented algorithm procedure in each source code step-by-step and understand the behaviors, the hallway actual values of the important variables in the code are asked in the questions. For example, the actual values of the array variables to store the sorted data and the variable for *pivot* at each iteration are asked in the VTP instance for *quick sort*. These questions can also avoid answering them without correctly understanding the algorithm implementation in the source code.

4 Devices for Hard VTP Instances

In this section, we present our devices to improve the solution performances for hard VTP instances on *pointer* and *algorithms*.

4.1 Hard VTP Instances

In our preliminary application of our generated VTP instances to limited students, the instances at *ID=22*, *ID=28*, *ID=31*, *ID=38*, and *ID=41* in Table 1 showed low correct answer rates compared with the others. These results suggest that not a few students do not understand the topics of *memory address*, *linked list*, *reverse polish notation*, *quick sort*, and *data size of pointer* well. They are related to memory managements by *pointer* or *algorithms*.

To improve the comprehensions on these topics and the solution performances of students, the following devices are presented in this paper:

1) We add *tips* in questions on pointers.
2) We give *hints* in the answer interface on memory addresses.
3) We prepare *multiple choice questions* on linked lists.
4) We add *links to reference websites* in the answer interface on algorithms.

4.2 Tips in Questions

On *tips* for pointers, Figures 5 and 6 show the source code and the questions for the VTP instance at *ID=41*. In the questions, the size of one word in this system is described by *(64-bit size system)*, and the targeting data in each question is described by *Data size of normal variables*, *Data size of pointer*, and *Data size of pointer reference*.

Table 1: Generated 42 VTP instances.

ID	topic	#of lines	#of questions	#of forms
1	two-dimensional array	18	1	12
2	max function	7	1	5
3	arithmetic	6	1	6
4	data type	12	1	16
5	for-loop	10	1	1
6	for-loop with character	11	1	9
7	function call	12	1	10
8	function for swap	21	1	4
9	memory address with pointer for integer array	11	1	56
10	car structure	12	1	2
11	output in for-loop	9	1	8
12	if in for-loop	15	2	4
13	book structure	14	1	3
14	pointer for one-D array	8	1	2
15	if in if with integer	24	2	2
16	arithmetic with if	11	2	2
17	if in if with character	27	2	4
18	while loop	19	1	8
19	structure data setting	18	1	7
20	output format specifier	11	1	4
21	memory address with pointer for character array	10	4	11
22	memory address with pointer for various array	24	1	13
23	four arithmetic operations	11	1	4
24	greatest common divisor (GCD)	21	1	4
25	inner product operation	16	1	1
26	linear search	23	1	5
27	binary search	29	1	17
28	linked list	67	1	7
29	tree traversal algorithms	81	3	63
30	stack data structure	47	1	8
31	reverse polish notation	67	2	12
32	queue data structure	55	1	4
33	insertion sort algorithm	25	1	30
34	selection sort algorithm	28	1	18
35	bubble sort algorithm	27	1	37
36	shell sort algorithm	24	1	23
37	merge sort algorithm	34	1	16
38	quick sort algorithm	38	1	56
39	heap sort algorithm	44	1	42
40	pointer	12	1	6
41	data size of pointer	20	1	9
42	list structure	47	1	12
maximum		81	4	63
minimum		6	1	1
average		24.50	1.26	13.95
total		1029	53	586

```
#include <stdio.h>
main()
{
    char vch, *pch;
    int vin, *pin;
    double vdo, *pdo;
    printf("sizeof(vch) = %d\n", sizeof(vch));
    printf("sizeof(vin) = %d\n", sizeof(vin));
    printf("sizeof(vdo) = %d\n", sizeof(vdo));

    printf("sizeof(pch) = %d\n", sizeof(pch));
    printf("sizeof(pin) = %d\n", sizeof(pin));
    printf("sizeof(pdo) = %d\n", sizeof(pdo));

    printf("sizeof(*pch) = %d\n", sizeof(*pch));
    printf("sizeof(*pin) = %d\n", sizeof(*pin));
    printf("sizeof(*pdo) = %d\n", sizeof(*pdo));
}
```

Figure 5: Source code for VTP instance at *ID=41*.

```
(64-bit size system)
What is the output of this program?
Data size of normal variables
sizeof(vch) = _1_
sizeof(vin) = _2_
sizeof(vdo) = _3_

Data size of pointer
sizeof(pch) = _4_
sizeof(pin) = _5_
sizeof(pdo) = _6_

Data size of pointer reference
sizeof(*pch) = _7_
sizeof(*pin) = _8_
sizeof(*pdo) = _9_
```

Figure 6: Questions for VTP instance at *ID=41*.

4.3 Hints in Answer Interface

On *hints* for memory addresses, Figures 7-9 show the source code, the hints, and the questions for the VTP instance at *ID=22*. The hints describe the byte size of each data type with the word size that is necessary to calculate the increase of the memory address, and the memory addresses of the first two elements in the array variables, so that students can answer the question on them.

```
#include <stdio.h>
int main(void){
    int vch, vsh, vin, vfl, vdo;
    vch = sizeof(char);
    vsh = sizeof(short);
    vin = sizeof(int);
    vfl = sizeof(float);
    vdo = sizeof(double);

    printf("%d %d %d %d %d \n", vch, vsh, vin,
        vfl, vdo);//hintA

    char arr1[] = {'o','k','a','y','a','m','a'};
    int arr2[5] = {1,2,3,4,5};
    float arr3[5] = {1.1,2.2,3.3,4.4,5.5};
    double arr4[5] = {1.0,2.2,3.5,22.4,10.0};

    printf("arr1[4] is %c \n", arr1[4]);
    printf("arr2[3] is %d \n", arr2[3]);
    printf("arr3[2] is %f \n", arr3[2]);
    printf("arr4[1] is %f \n", arr4[1]);

    printf("The address of arr1[2] is %p \n", &
        arr1[2]);//hintB
    printf("The address of arr2[2] is %p \n", &
        arr2[2]);//hintC
    printf("The address of arr3[3] is %p \n", &
        arr3[3]);//hintD
    printf("The address of arr4[2] is %p \n", &
        arr4[2]);//hintE
    return 0;
}
```

Figure 7: Source code for VTP instance at *ID=22*.

```
hintA: Data types and sizes(32-bit system)
char: 1 byte
short: 2 bytes
int 4: bytes
float: 4 bytes
double: 8 bytes

hintB: The address of arr1[0] is 0028FE30
       The address of arr1[1] is 0028FE31

hintC: The address of arr2[0] is 0028FE20
       The address of arr2[1] is 0028FE24

hintD: The address of arr3[1] is 0028FEF4
       The address of arr3[2] is 0028FEF8
hexadecimal: the symbols "0"~"9" to represent
    values zero to nine, and "A"~"F" to represent
     values ten to fifteen.

hintE: The address of arr4[0] is 0028FE10
       The address of arr4[1] is 0028FE18
```

Figure 8: Hints for VTP instance at *ID=22*.

```
What is the output of this program?
_1_ _2_ _3_ _4_ _5_
arr1[4] is _6_
arr2[3] is _7_
arr3[2] is _8_
arr4[1] is _9_
The address of arr1[2] is _10_
The address of arr2[2] is _11_
The address of arr3[3] is _12_
The address of arr4[2] is _13_
```

Figure 9: Questions for VTP instance at *ID=22*.

4.4 Multiple Choice Questions

On *multiple choice questions* for linked lists, Figures 10 and 11 show the source code and the questions for the VTP instance at *ID=28*. The multiple choice questions of the correct answers are used at Q1 and Q3 as the hints to understand the pointer and structure in the linked list.

```
#include <stdio.h>
#include <errno.h>
#include <string.h>
#include <time.h>
#include <stdlib.h>
struct cell{
    struct cell *next;
    int data;
};
typedef struct cell cell_t;
cell_t *list_alloc(int data){
    cell_t *new = NULL;
    new = (cell_t *)malloc(sizeof(cell_t));
    if(new == NULL){
        fprintf(stderr, "ERROR: list_alloc(): %s\
           n", strerror(errno));
        return(NULL);
    }
    new->next = NULL;
    new->data = data;
    return(new);
}
int list_add(cell_t *header, int data){
    cell_t *next = NULL;
    cell_t *prev = header;
    next = list_alloc(data);
    if(next == NULL) return(-1);
    while(prev->next != NULL){
        prev = prev->next;
    }
    prev->next = next;
    return 0;
}
void list_free(cell_t *header){
    cell_t *temp = header;
    cell_t *swap = NULL;
    while(temp != NULL){
        swap = temp->next;
        free(temp);
        temp = swap;
    }
}
static void list_print(cell_t *header){
    cell_t *p = header;
    printf("list{ ");
    while(p != NULL){
        printf("%d ", p->data);
        p = p->next;
    }
    printf("}\n");
}
int main(void){
    int cnt = 0;
    cell_t *header = list_alloc(0);
    if(header == NULL) return -1;
    for(cnt = 1; cnt < 5; cnt++){
        list_add(header, cnt*7%10);
    }
    list_print(header);
    list_free(header);
    return 0;
}
```

Figure 10: Source code for VTP instance at *ID=28*.

```
Q1: Select one word for each blank from a list of
     choices.
What is data type of *header? _1_
What is data type of header? _2_
(A) pointer
(B) data
(C) cell
(D) cell_t

Q2: What is "data" of the cell pointed by "header
    " in the code?  _3_
What is "next" of the cell pointed by "header" in
     the code?  _4_

Q3: Select one word for each blank from the list
    of choices given below from list_free method.
What is the name of cell that will be freed this
    time? _5_
What is the name of cell that will be freed next
    time? _6_
(A) swap
(B) temp

Q4: In list_add method, what is the name of new
    cell? _7_

Q5: What is the output of this programming?
list { _8_ _9_ _10_ _11_ _12_ }
```

Figure 11: Questions for VTP instance at *ID=28*.

4.5 Links to Websites in Answer Interface

On *links to reference websites* for algorithms, the links to the websites that explain *reverse polish notation* in [9] and *quick sort algorithm* in [10] are included in the answer interface. By clicking these links, the students can easily access to the necessary information to solve them.

5 Evaluations

In this section, we evaluate the generated 42 VTP instances with 586 answer forms by applying them to 49 undergraduate students in Myanmar (21 students), Japan (18 students), and China (10 students) who are currently studying or have studied *C programming* and *fundamental data structures & algorithms*.

5.1 Solution Results for Individual VTP Instances

First, the solution results for the 42 individual VTP instances by the 49 students are analyzed. Figure 12 shows the total number of answer submission times and the average correct answer rate (%) for each VTP instance submitted by the students. As the summary, Table 2 shows the maximum, minimum, average, and standard deviation (SD) of the total number of submissions and the average correct answer rate among the 42 VTP instances. The average correct rate 94.29% and the average number of submissions 139.17 by all the students (2.84 by one student) for each instance suggest that the generated VTP instances are of moderate difficulty for novices to study *C programming*.

Table 2: Summary of solution results for individual VTP instances.

\	total# of submissions	average correct rate
maximum	628	100 %
minimum	55	82.32 %
average	139.17	94.29%
SD	106.14	4.84%

5.1.1 VTP Instances with Low Results

The VTP instances for *pointer* or *fundamental data structures & algorithms* are generally difficult for the students. For the eight instances at *ID=21*, *24*, *28*, *29*, *30*, *32*, *36*, and *37*, which are on *pointer*, *stack*, *linked list*, and *sorting*, the average correct answer rate is under 90%.

For the VTP instance *ID=41 linked list*, the correct answer rate 82.32% is lowest and the number of answer submissions 628 is highest among the 42 instances. This topic is still hard for the students, regardless of the presented devices in Section 4. However, it is observed that the students tried to solve this instance by referencing to the devices.

For the VTP instance for *merge sort* at *ID=37*, both the correct rate 85.87% and the number of submissions 101 are relatively low. Many students gave up solving it correctly at early time, which suggests their insufficiency in understanding this algorithm. Like for *quick sort*, the device of adding the link to the reference website in the answer interface will be necessary for this instance.

5.1.2 Distribution of Correct Answer Rates

Figure 13 shows the distribution of the correct answer rates among the 42 VTP instances. 34 instances achieved more than the 90% correct rate and seven did 100%, which indicates that these instances are moderate for novice students. However, the rates of eight instances were less than 90%, which need to be improved.

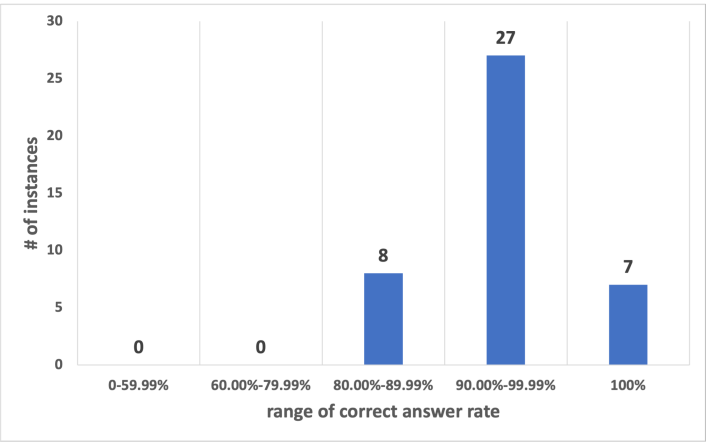


Figure 13: Distribution of correct answer rates among instances.

5.1.3 Distribution of Submission Times

Figure 14 shows the distribution of the number of answer submission times among the VTP instances. 21 instances were solved within 100 submissions. 30 instances were within 150 submissions by all the students or 3.57 submissions by one student. However, for the three instances at *ID=22*, *28*, and *29*, the answers were submitted more than 300 times, which will need to be improved.

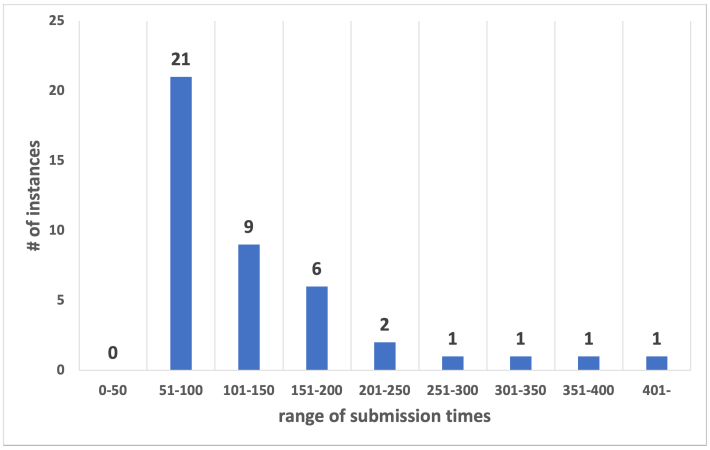


Figure 14: Distribution of submission times among instances.

5.2 Solution Results for Individual Students

Next, the solution results for the individual students to the 42 VTP instances are analyzed. Figure 15 shows the total number of answer

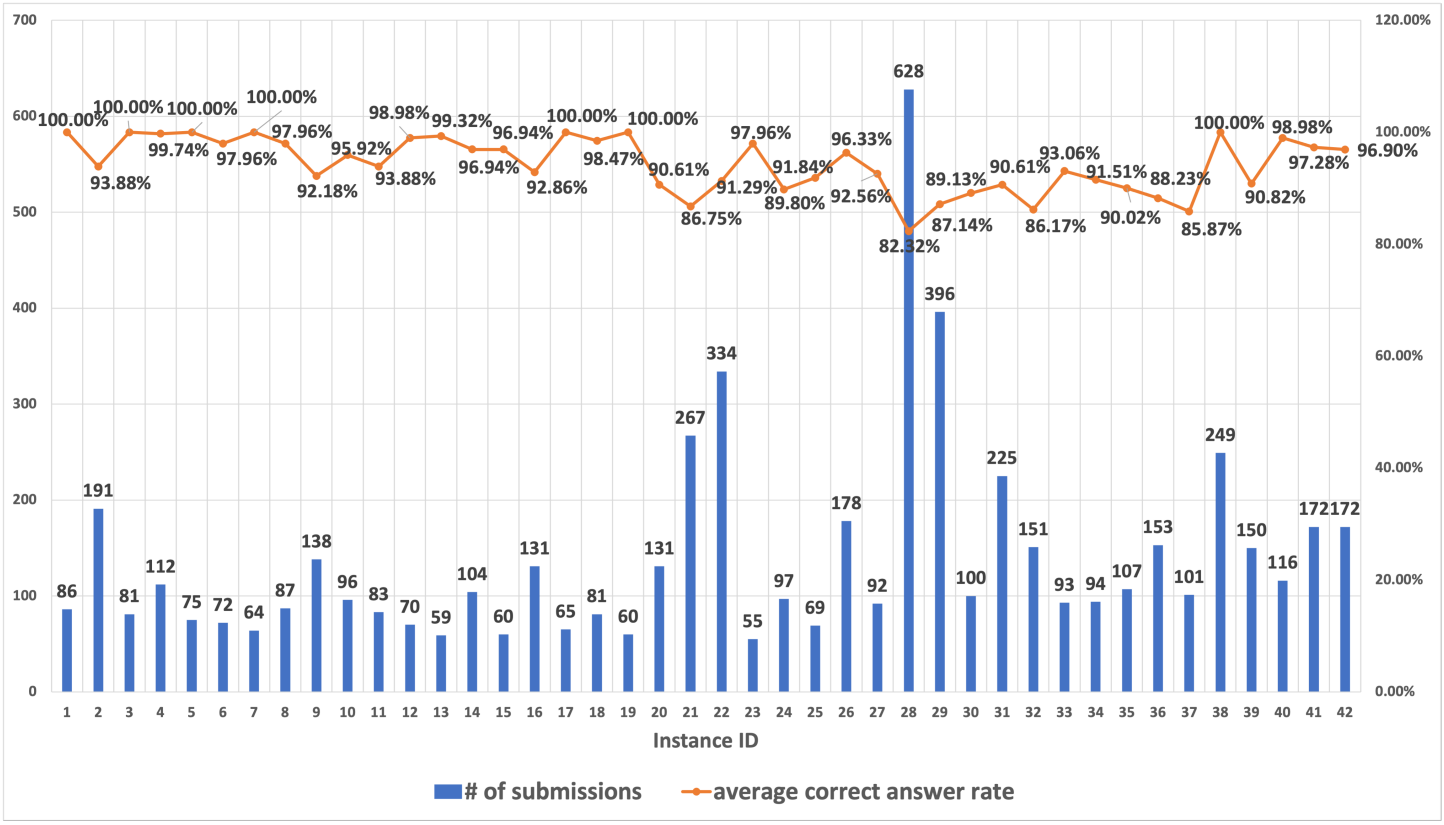


Figure 12: Solution results for individual VTP instances.

submissions and the average of the correct answer rates among all the instances for each student. Table 3 shows the summary results for each student. The results suggest that most of the students well solved the VTP instances.

Table 3: Summary of solution results for individual students.

	total # of submissions	correct answer rate
maximum	394	100%
minimum	38	58.75%
average	119.29	94.29%
SD	89.61	8.25%

5.2.1 *Students with Low Results*

For 11 students at *ID=8*, *10*, *11*, *12*, *13*, *15*, *18*, *19*, *33*, *46*, and *49*, the average correct answer rate among the 42 VTP instances is smaller than 90%.

For the student at *ID=18*, both the correct answer rate 58.75% and the number of answer submissions 38 are lowest among the students. Even, this student did not try to solve some instances. The teacher should find the reason and help this student studying *C programming*. The five students at *ID=10*, *12*, *33*, *46*, and *49* also have the similar proclivities.

For the two students at *ID=8* and *15*, the number of submissions is relatively high. Thus, they made sufficient efforts to solve the VTP instances. It can be expected that the proper guidance by the teacher can quickly improve their performances.

5.2.2 *Distribution of Correct Answer Rates*

Figure 16 shows the distribution of the correct answer rates among the 49 students. 38 students achieved over 90% correct answer rates. Eight students at *ID=22*, *24*, *26*, *27*, *28*, *32*, *39*, and *44* achieved 100%.

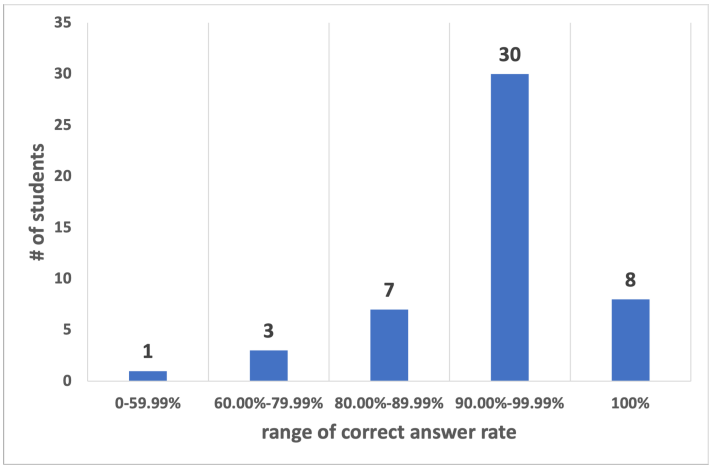


Figure 16: Correct answer rate distribution of students.

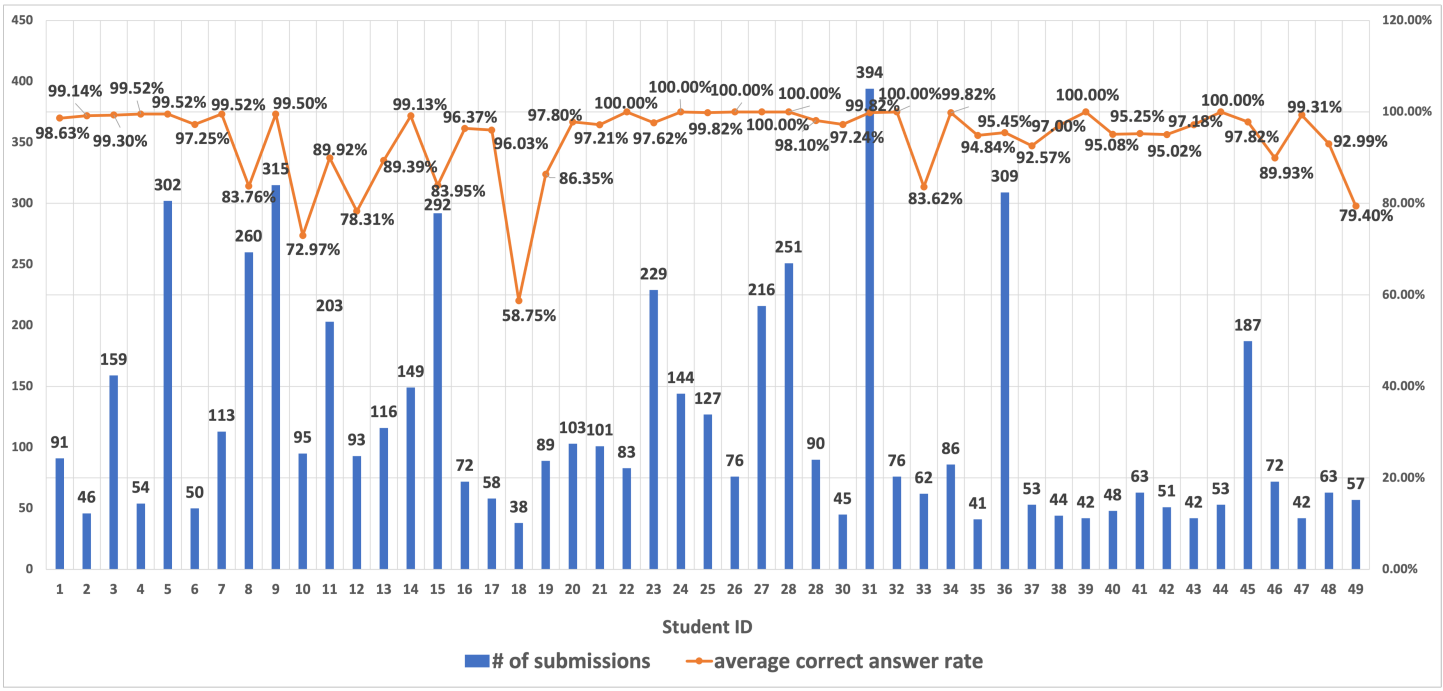


Figure 15: Results for each student.

5.2.3 *Distribution of Submission Times*

Figure 17 shows the distribution of the number of answer submission times among the students. For 30 students, the total number of submissions was smaller than 100, where the excellent student at *ID=39* correctly solved any instance except one instance by submitting the answers only one time. For the four students at *ID=5*, *9*, *31*, and *36*, the total number of submissions exceeded 300, and the answer rate is larger than 90%. It was found that these students did not give up solving them.

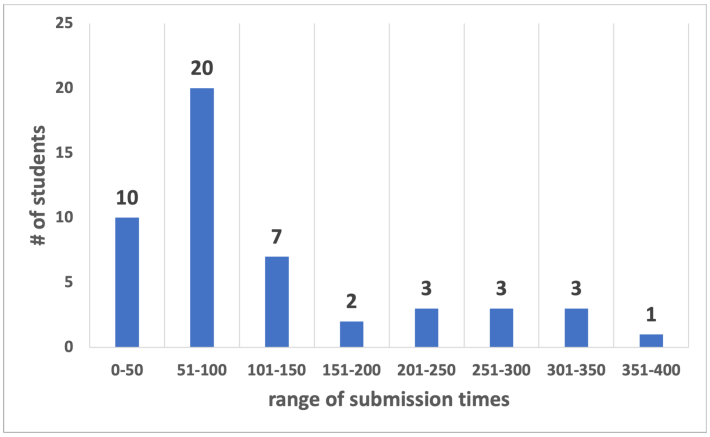


Figure 17: Submission times distribution among students.

5.3 *Evaluations of Devices for Hard Instances*

To evaluate the effectiveness of the presented devices in Section 4, we compare the solution results *with* and *without* using them for the four hard VTP instances at *ID=22*, *28*, *31*, and *38*.

5.3.1 *Comparison Results*

In our preliminary application, we applied the four instances to 21 students among 49 *without* using the proposed devices. Then, to avoid the solution improvements by solving the same instances twice by the same students, the results of the remaining 28 students are used as the results *with* them.

Table 4 compares the average number of answer submission times and the average correct answer rate for the four instances by the students between the *with* group and the *without* group. In the *without* group, the average correct answer rate for any instance was smaller than 70%, which suggests that many students suffered from solving them. Then, in the *with* group, it was improved to higher than 80%, which can be considered as the minimum rate that every student should exceed. The average number of submissions was increased except for *ID=31*, because the students were motivated to continue solving them due to the devices.

Table 4: Comparisons of solution results for hard instances.

ID	topic	average # of submissions		average correct answer rate	
		with	*without*	*with*	*without*
22	memory address with pointer for various array	7.95	7.17	91.29%	54.40%
28	linked list	14.95	1.92	82.32%	56.00%
31	reverse polish notation	5.36	6.88	90.61%	52.00%
38	quick sort algorithm	5.93	1.71	100.00%	68.79%
average		8.55	4.42	91.06%	57.80%

5.3.2 t-test *Analysis*

t-test is applied to the solution results to check the significant difference between the average correct answer rates of the two groups. The *p-value* is $1.179E-16$ and is smaller than 0.05, which means that there is the significant difference between them. Thus, it is confirmed that the proposed devices for the hard VTP instances are effective in improving the solution performances of novice students.

5.4 *Discussions*

At solving the VTP instances, some students may reach the correct answers without understanding them well. They may randomly submit possible answers or copy the answers of other students. Fortunately, the answer interface can record any submitted answer of a student. By analyzing the records, it may be possible to find the undesirable behaviors of students, which will be in our next study.

6 Conclusion

This paper presented the *value trace problem (VTP)* for independent *code reading study* of *C programming*. 42 VTP instances were generated using simple C source codes for *basic grammar concepts* and *fundamental data structures & algorithms* in textbooks or websites. Besides, the devices of tips, hints, multiple choice questions, and links to reference websites were presented for hard instances on *pointer* and *algorithms* to improve the solution performances.

For evaluations, the generated 42 VTP instances were assigned to 49 undergraduate students in Japan, China, and Myanmar to be solved at home using the answer interface that can run offline. The average correct answer rate among all the VTP instances reached 94.29%, where by our proposed devices, the average rate for the four hard instances was improved by 33.26% from 57.80% before improvements. Thus, the effectiveness of our proposal was confirmed in motivating *self-study* of *C programming* to novice students at home.

A limitation of *VTP* is the manual selection of variables or output messages in the source code for questions, in addition to the manual collection of C source codes. An algorithm should be investigated to automatically generate a new VTP instance from a given source code. Another limitation is the impossibility of practicing source code writing by students.

As future works, we will study the automatic VTP generation algorithm, make new VTP instances using C source codes for other grammar concepts and algorithms, such as file I/O, error handing, recursions, and graph algorithms, and apply them to students in *C programming* courses. Then, more comprehensively and formally, we will study the methodologies of organizing and conducting pedagogical experiments to verify the effectiveness of the proposal in increasing motivations of novice students to study *C programming*, and the methods of processing the results of pedagogical experiments.

References

[1] Sio-Iong Ao, IAENG Transactions on Engineering Sciences - Special Issue for the International Association of Engineers Conferences 2016 Volume II, 2018.

[2] N. Funabiki, K.K. Zaw, W.-C. Kao, "A proposal of value trace problem for algorithm code reading in Java programming learning assistant system," Information Engineering Express, 1(3), 2015, doi:10.52731/iee.v1.i3.39.

[3] K.K. Zaw, W. Zaw, N. Funabiki, W.C. Kao, "An informative test code approach in code writing problem for three object-oriented programming concepts in java programming learning assistant system," IAENG International Journal of Computer Science, **46**(3), 2019.

[4] N. Funabiki, H. Masaoka, N. Ishihara, I.W. Lai, W.C. Kao, "Offline answering function for fill-in-blank problems in Java Programming Learning Assistant System," in 2016 IEEE International Conference on Consumer Electronics-Taiwan, ICCE-TW 2016, 2016, doi:10.1109/ICCE-TW.2016.7521045.

[5] Pointer and gcc, https://ubuntuforums.org/archive/index.php/t-858030.html.

[6] FCP 2 If else.ppt, https://dokumen.tips/documents/fcp2ifelseppt.html.

[7] Loop and break, https://www.loopandbreak.com.

[8] T. Hikita, Algorithms by C, Science Pub., 1995.

[9] Reverse polish notation, http://www.cs.man.ac.uk/ pjj/cs212/fix.html.

[10] Quick Sort, https://www.programiz.com/dsa/quick-sort.

[11] M. G`ómez-Albarr`án, The teaching and learning of programming: A survey of supporting software tools, Computer Journal, **48**(2), 2005, doi:10.1093/comjnl/bxh080.

[12] P. Brusilovsky, S. Sosnovsky, "Individualized Exercises for Self-Assessment of Programming Knowledge: An Evaluation of QuizPACK," ACM Journal on Educational Resources in Computing, **5**(3), 6, 2005, doi:10.1145/1163405.1163411.

[13] M. Kordaki, "A drawing and multi-representational computer environment for beginners' learning of programming using C: Design and pilot formative evaluation," Computers and Education, **54**(1), 69–87, 2010, doi:10.1016/j.compedu.2009.07.012.

[14] M.J. Lee, A.J. Ko, "Personifying programming tool feedback improves novice programmers' learning," in ICER'11 - Proceedings of the ACM SIGCSE 2011 International Computing Education Research Workshop, 2011, doi:10.1145/2016911.2016934.

[15] F.W.B. Li, C. Watson, "Game-based concept visualization for learning programming," in MM'11 - Proceedings of the 2011 ACM Multimedia Conference and Co-Located Workshops - ACM International Workshop on Multimedia Technologies for Distance Learning, MTDL'11, 2011, doi:10.1145/2072598.2072607.

[16] W.Y. Hwang, R. Shadiev, C.Y. Wang, Z.H. Huang, "A pilot study of cooperative programming learning behavior and its relationship with students' learning performance," Computers and Education, **58**(4), 2012, doi:10.1016/j.compedu.2011.12.009.

[17] D.M. Souza, K.R. Felizardo, E.F. Barbosa, "A systematic literature review of assessment tools for programming assignments," in Proceedings - 2016 IEEE 29th Conference on Software Engineering Education and Training, CSEEandT 2016, Institute of Electrical and Electronics Engineers Inc.: 147–156, 2016, doi:10.1109/CSEET.2016.48.

[18] S. Parihar, R. Das, Z. Dadachanji, A. Karkare, P.K. Singh, A. Bhattacharya, "Automatic grading and feedback using program repair for introductory programming courses," in Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, Association for Computing Machinery: 92–97, 2017, doi:10.1145/3059009.3059026.

[19] X. Fu, A. Shimada, H. Ogata, Y. Taniguchi, D. Suehiro, "Real-time learning analytics for C programming language courses," in ACM International Conference Proceeding Series, Association for Computing Machinery: 280–288, 2017, doi:10.1145/3027385.3027407.

[20] T. Busjahn, C. Schulte, "The use of code reading in teaching programming," in ACM International Conference Proceeding Series, 3–11, 2013, doi:10.1145/2526968.2526969.

[21] T. Vandegrift, "Reading before writing: can students read and understand code and documentation?," SIGCSE '05 Cojoined Meeting, 2005.

[22] D. Kwon, I. Yoon, and W. Lee, "D.Y. Kwon, I.K. Yoon, W.G. Lee, "Design of programming learning process using hybrid programming environment for computing education," KSII Transactions on Internet and Information Systems, **5**(10), 1799–1813, 2011, doi:10.3837/tiis.2011.10.007.

[23] T. Busjahn, R. Bednarik, C. Schulte, "What influences dwell time during source code reading? Analysis of element type and frequency as factors," in Eye Tracking Research and Applications Symposium (ETRA), 2014, doi:10.1145/2578153.2578211.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 16-22 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

SyncBIM: The Decision-Making BIM-Based Cloud Platform with Real-time Facial Recognition and Data Visualization

Chia-En Yang*, Yang-Ting Shen, Shih-Hao Liao

National Cheng Kung University, architecture, Tainan, 701401, Taiwan

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 24 August, 2021
Online: 10 September, 2021

Keywords:
BIM
Data visualization
Face recognition
Facility management
Maintenance engineering
Real-time systems

ABSTRACT

In this research we developed an BIM-based system to monitor and visualize the real-time building users information. Concentrating on building in-use stages, advantages in tracking facial recognition should be revealed through the availability of real-time information. In this way could explore the possibility of how BIM and IoT could improve data-oriented facility management. The integration system called SyncBIM. The five system elements of the construction platform are also proposed to allow more efficient management and data transmission of the building O&M system. This study integrates the three pieces of technology respectively known as BIM platform, Internet of Things, and computer vision to explore the architecture and technology required by the building O&M system, and establishes a "Decision-Making BIM-Based Cloud Platform" to allow the opportunities of information integration and collaborative operations for the application of BIM and the introduction of computer vision technology.

1. Introduction

The relevant definitions and scopes of Building Information Modelling (BIM) are constantly evolving over time. The American Institute of Architects (AIA) defines BIM as a multi-dimensional model presentation technology that adopts the database approach to integrate information of construction projects. In recent years, as the "Digital Twin" has been repeatedly regarded as one of the top ten global technology trends of the year by the information technology research organization [1], it could be understood that such trend is highly connected with BIM as it integrates multiple proposed applications in the technological and related fields.

In addition to implementing strategies in compliance with government policies, the BIM TASK GROUP, established under the promotion of the UK government, has its major goal of promoting the various duties of the building production and construction supply chain. However, after the government agencies having formulated various collaborative work specifications, starting from April 2016, it could be foreseen that building information models with highly collaborative functions would even become one of the basic elements of future architectures [2].

BIM plays an important role in all stages of the building life cycle, making it infrastructure of the future as one of the trends. Because BIM not only serves as a beneficial tool or database in the construction phase, but also provides various flexibility and integration functions that can offer various requirements for building O&M, and exerts powerful application integration capabilities(see Figure 1).

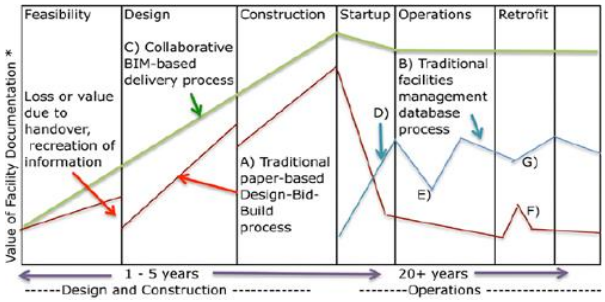


Figure 1: The value of documentation in the facility lifecycle (Source: Eastman, 2008)

For the current application of building O&M, the importance of facility management (FM) and the improvement of building efficiency and functionality in facility management have gradually become prominent. Nowadays, increasingly complex building functions are making the building space and equipment more diversified and sophisticated. Building management is no longer simply checking whether it is operating normally, but finding out its unique usability based on the huge volume of building usage

*Corresponding Author: Chia-En Yang. Email: gaeunt@gmail.com

https://dx.doi.org/10.25046/aj060503

information generated, real-time spatial analysis, and retrieval. The changeable[3] complexity and real-time nature of modern computer technology can assist decision-making and provide optimized applied artificial intelligence (AI). For a building space involving a wide range of levels, the integration of software and hardware and the deployment of various systems, such as environmental data, the number of users, and the analysis of crowds, etc., would have a certain degree of crucial impact on building O&M. In response to such impact, this study believes that the usage status of these buildings must be actively displayed through machine learning and the application of AI technology, and such relationship of changes requires to be reflected in the O&M process. In order to solve the problem of BIM in the O&M of building facilities, this study proposes the "BIM management platform for decision-making", which is hereby referred to as "SyncBIM"[4].

2. Related Work

In this Section, we present two popular aspects of adopting BIM in the building life cycle, which are Digital Twin and AI applications. The first lays the foundation for the cloud platform architecture, and diversified AI applications drive BIM to become the "sense" of architectures.

2.1. Application of Digital Twin to BIM

AEC includes buildings&architectures, engineering, and construction industries. At present, BIM technology has reached certain achievements under the promotion of government agencies of various countries. Meanwhile, with the evolution of information and communication technology, more application technologies such as computer vision, AI, and various sensors and environmental sensing devices such as LiDAR could be able to complement the possibility of collecting real-time and on-site data in a real environment reality, which would allow the BIM model to predict and plan the allocation of resources and transportation, and supervise the safety of the site environment. The introduction of Digital Twin at an earlier stage can further enhance the application of BIM advantages of the whole building life cycle.

An indicative example of a Digital Twin application is the "Virtual Singapore" developed by the French software company Dassault Systèmes for five years. On this virtual three-dimensional platform, one can freely browse all corners of Singapore City, regardless of the data of buildings, roads, public infrastructure, or the natural environment[5]. More importantly, it provides four main functions, namely "Virtual Experimentation", "Virtual Test-Bedding", "Urban Planning and Decision-Making", and "Research and Development" that allow decision-makers of urban affairs to predict the impact of different decisions and policy directions on the city. Therefore, there would be wider application in the fields of urban disaster prevention such as logistics and environmental conservation, etc.

Figure 2 shows the UI of Virtual Singapore. With real-time information and the large volume of data provided by sectors from public departments and communities, Virtual Singapore is able to solve the problem of livability in areas with limited resources and unlimited potential. For example, after importing annual data of insolation areas, temperatures, and humidity, the system would be able to provide space decision-makers with the information for making a more comfortable and humanized environment.



Figure 2: Interface of Virtual Singapore (Source: nrf.gov.sg)

From the perspective of the "smart city", infrastructure is the element that composes a city. The system developed by Dassault Systèmes allows people to see the digitization of urban infrastructure, which is constructed in the digital domain of computers and adds value to or optimizes the decision-making for the city to shape the vision of smart city life of the future[6]. The BIM model plays the role of an information carrier, and after importing the information, it serves as a bridge between the real and virtual worlds. In the next stage, the health and safety issues at the construction site, which has the widest impact and draws the most attention during the construction phase, will be further reviewed and discussed in the scope and application focused on by this study, and the development of the research proposed by this study will be applied to the connection of the construction site to make trigger more innovative and progressive linkage[7].

2.2. Application of Artificial Intelligence in the Building Life Cycle

The current major technology of artificial intelligence (AI) aims at analyzing and realizing professional human techniques of various fields through extensive computer science training [8]. With its quality management and supervision functions, AI could replace a large amount of work previously requiring human judgment, which would achieve accurate and effective predictive capability [9]. For example, for a construction site facing the pressure from both construction schedule and cost while requiring the high integration of the construction specifications and interface of various fields, the introduction of AI and BIM technology for the integration of various information would provide a certain degree of efficiency improvement and added value[10].

According to the research conducted by the Science & Technology Policy Research and Information Center of the National Applied Research Laboratories, The assistance provided by AI to the construction industry can be roughly divided into four categories as shown in Table 1, which are respectively known as (1) Planning and Design, (2) Safety, (3) Automation Equipment, and (4) Monitoring and Maintenance.

The following four AI application areas proposed by the Science & Technology Policy Research and Information Center of the National Applied Research Laboratories are all able to be highly connected with BIM:

- Planning and Design

Table 1: Application of AI in the Construction Industry (Source: Science & Technology Policy Research and Information Center of the National Applied Research Laboratories, 2018)

Application of AI in the Construction Industry	
Categories	Contents
Planning and Design	The BIM model uses AI to automatically complete pipeline configuration, analysis, and machine learning to manage construction risks
Safety	AI with sound and image recognition system to automatically mark the construction site
Automation Equipment	Smart building management to make users feel comfortable
Monitoring and Maintenance	AI with robots and drones to improve construction efficiency

The construction site requires to consider structural, electrical and mechanical, and pipeline engineering issues, etc., which involve different contractors using different appliances and interfaces that may often cause delays and repetitive work in the final integration stage. Therefore, Building System Planning developed GenMEP by using BIM modeling software Autodesk Revit. By adopting Generative Design, the configuration of the electrical system in the building model can be automatically built using the GenMEP plug-in while also taking into account the design specifications and the spatial complexity of different buildings (see Figure 3).

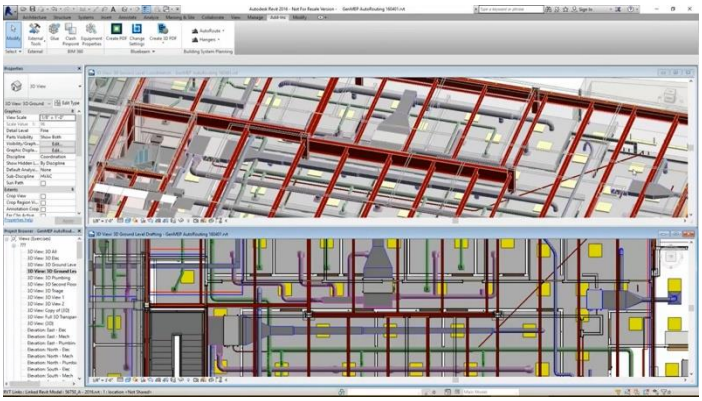


Figure 3: Interface of GenMEP (Source: BuildingSP, Inc.)

- Safety

AI technology such as data analysis and machine learning would be adopted to manage construction risks. BIM 360 Project IQ launched by Autodesk is an example in which the BIM 360 Field is shaped by various images, audio, and other construction management information collected on site, which is analyzed to provide suggestions and reduce risks for relevant QA staff on site (see Figure 4).

- Automation Equipment

Jetson AI platform constructed by Nvidia and the Japanese company Komatsu is designed to perform 3D visualization and analysis of real-time interactive modes such as tracking the on-site staff, machinery, and objects, which could allow AI to take over tasks of on-site monitoring and management to provide round-the-clock protection for site safety, as well as improving the quality of construction management.

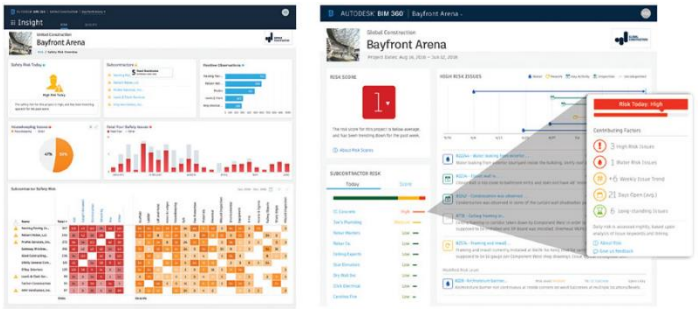


Figure 4: Interface of BIM 360 Project IQ (Source: BIM360)

- Monitoring and Maintenance

Suffolk Construction based in Boston adopts a virtual model constructed by VR to monitor the progress of on-site construction projects to cope with the rising construction costs and on-site safety management, as well as improving construction efficiency. In addition, Figure 5 shows that Doxel AI from Silicon Valley combines exploration robots and drones with photography equipment and LiDAR to scan and construct the real-time situation of the site, and adopts visualization data and deep learning algorithm processing to allow users and managers to automatically eliminate various errors and monitoring risks at the construction site by using the on-site construction progress analysis adopting AI technology.

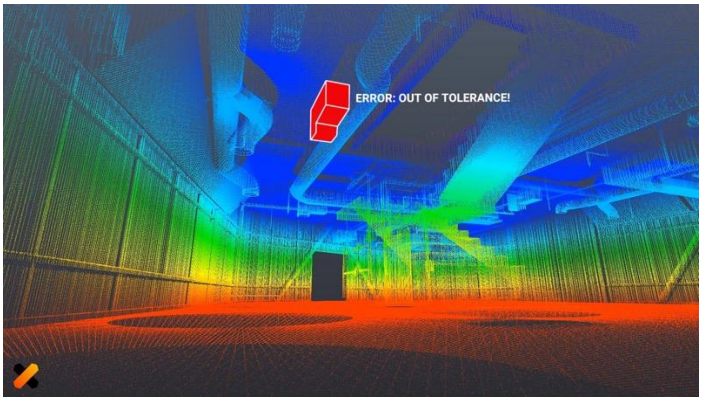


Figure 5: Visualization of On-site Construction Information by Doxel (Source: Doxel AI)

The key points and precautions of the occupational safety inspection of each working item are clearly-defined specifications, for example, the specifications for the categorization of hazardous areas and the key points of inspection. Hence, AI could serve as a 24/7 supervisor to check items such as whether safety equipment is worn and any staff is entering the hazardous areas by adopting the image recognition technology to conduct round-the-clock monitoring at the construction site. For example, the setting of geofencing allows personnel entering the hazardous areas to be adequately alerted and protected while notifications are also provided to have managers aware of the current presence of high-level risks of the staff (see Figure 6), and the real-time information could also be fed back to the Digital Twin's BIM environment integration platform and presented in a visualized manner. With

the integration of BIM and various AI techniques, the required time and cost for the construction site could be improved in terms of management efficiency, and the occupational safety and working environment for labors could also be assured, which would achieve a win-win situation[11], [12].



Figure 6: Construction Site Management Adopting AI Image Recognition Technology (Source: Smartvid.io)

The definition and application of BIM often evolve along with various technological advancements. And the pursuit of environmental friendliness and green living has always been its core value. Building O&M and management is a complex interface integration and multiple interdisciplinary coordination. However, when carrying out various intimate and highly customized O&M processes for future space users, a large number of people would be exposed to a high degree of on-site O&M risks for achieving the establishment of a smart city or its infrastructures. If the managers and participants could improve the building O&M mode in a safer and more efficient way, it will be more complete and environmentally friendly when establishing an overall smart city, which would also be a precious gift to the fields of building O&M.

3. The Planning and Design of Decision-Making BIM-Based Cloud Platform

This Section mainly explores the system adopting BIM in the auxiliary decision-making platform. It would discuss the information processing flow and architecture of the smart building O&M system in two parts respectively known as the information transmission architecture the system operates, and the information transmission loop of the system.

3.1. Information Transmission Architecture of the System

Figure 7 describe such information transmission architecture. The architecture itself adopts the IPO model as the application basis, and in the process of data transmission, a web portal is established between the endpoints in order to make the clearly defined division of labor of the system. The scalability and integration of the web portal also provide flexibility for the equipment and sensing devices installed in each period of the system's operation. And it is also helpful for the integration of information formats. For the transmission architecture specified in this Section, we will further describe it in detail in the following three subsections respectively titled (1) Input, (2) Process, and (3) Output.

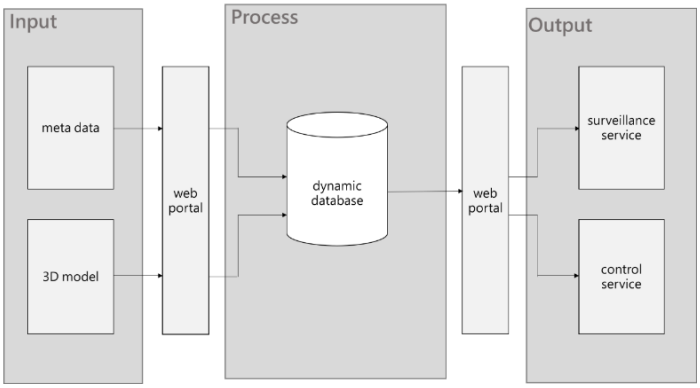


Figure 7: Diagram of Information Transmission Architecture for the Smart O&M System

- System Input: Information Collection

The information sources the smart building O&M requires could be summarized as three scales. One is the meta data that covers various real parameters and O&M mode scripts, which transmit the three-dimensional information model to the web portal and subsequently to the dynamic database for storage, integration, and further process. The meta data includes the real parameters, the scripts and modes related to various building O&M. If it is necessary to perform appropriate O&M for a certain building, the data monitoring operation would start the deployment independently without waiting for the completion of the system installation, and the scaling of input data monitoring could be performed even under the presence of an existing system which requires only the addition of new web portals. If the building usage data required by the system is transmitted to a specific web portal, such input would be designed to operate independently, and the system only retrieves the data from the web portal to achieve the division of labor and facilitate any subsequent scaling. Furthermore, upon the occurrence of equipment failure or update, the system would not be completely shut down but only some of the functions are temporarily suspended, which would also improve efficiency in troubleshooting for the system.

- System Process: Dynamic Database

Because the various O&M data and 3D models from the input would be updated continuously over time, a dynamic database that can be updated and accessed in real time shall be established for the process. Such a dynamic database for the process would retrieve the information of the web portal, then the system would analyze and compare the usage mode. Because the information carried by the web portal has been standardized, it is sufficient for the database to complete the data connection with the web portal while allowing the system to retain the scalability, which would be helpful for the maintenance and testing of some areas of the system or the addition of O&M areas.

- System Output: Provision of Services

The dynamic database compares and delivers the O&M data to the web portal of the output. Unlike the input web portal, the output web portal saves various service-related information which would be used for building O&M. Rather than being the big data, such

types of information required for O&M is integrated information, such as the parameter codes after script writing or the trigger values of execution, which are provided to the application services. The types of application services are divided into two types respectively known as "control type" and "service type". The output of the information transmission architecture would trigger related services according to different types of parameters.

3.2. Information Transmission Loop

In the adaptive O&M system, the transmission of information is not a straight-linear process but a continuous act of constant updating and modification. Therefore, a loop of information transmission shall be established to complete the architecture required for the O&M maintenance under the constant-changing relationship. When the data of the usage pattern is established in the dynamic database and compared with the analysis characteristics, these data can be used as a basis for system adaptation, learning, and reflective practice. Also, because both people and the built environment would continue to change and influence each other over time, the usage data will continue to be imported to the database, and then the updated results would be transmitted to the end sensors, which forms an information transmission loop accordingly. The three processes of this loop formation are respectively known as (1) Data Collection, (2) Database Comparison, and (3) Database Update. Figure 8 presents the data transmission loop formed by the system.

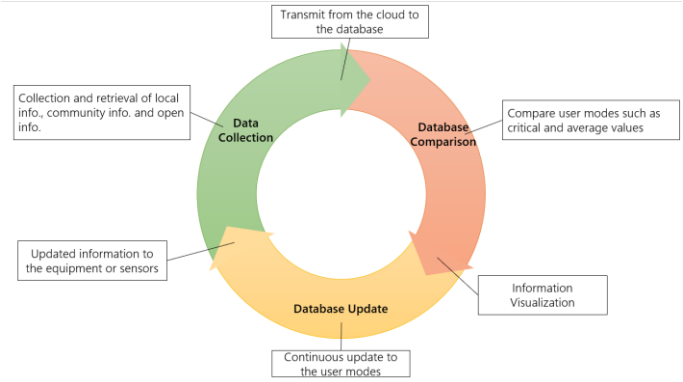


Figure 8: System Data Transmission Loop of SyncBIM

- Data Collection

After completing the layout of the sensor network for collecting building spatial data, the data via the cloud would start to be transmitted to the designated dynamic database. The types of data refer to the open information, local information, and community information mentioned in the previous paragraph. Such three types of information would be compared and analyzed within the database.

- Database Comparison

When the volume of data within the database is sufficient, it would be available for further analysis to find the average value, critical value, or usage habits of the particular space for the establishment of a spatial usage pattern. The usage pattern upon being determined established could be subject to continuous updates besides being compared with the data. By analyzing the dynamic pattern of the data, one could infer whether there is any device functioning abnormally or abnormal spatial activities, and if so, the information visualization would display the location and time where such abnormality takes place to allow the recording by the system and update to the user mode of the database.

- Database Update

With the increase of usage data and the modification of user mode, the operation of the system will keep all usage information updated along with the building's existence period and usage history, and update its information to various building equipment and sensors, which also means that the database would continue to be updated over time to continuously adjust the balance between people and the building environment.

4. SyncBIM that Integrates Computer Vision Applications

This Section presents the construction of a SyncBIM system based on the system architecture mentioned in the previous Section. It would discuss in two parts respectively titled "The System Interface Planning", and "The Process of Integrating Computer Vision Applications".

4.1. The System Interface planning of SyncBIM

The SyncBIM system is established based on the following five principles respectively known as (1) 3D Information Model Aiding Decision-making, (2) Object Orientation Design Delivering Service, (3) Dynamic Database based on Real-time Monitoring, (4) Mapping Data-driven Visualized Results, and (5) Hierarchical Integration of Contextualized Spatial Information. (see Figure 9).

- 3D Information Model Aiding Decision-making:

The spatial information presented in a three-dimensional manner allows viewers to view in an intuitive way, while also being able to efficiently be aware of the information and contextualized spatial data carried by the objects.

- Object Orientation Design Delivering Service:

The BIM model constructed with individual objects not only uses object-oriented design to guide information to the building objects, but also treats each object as a node that can provide the required services, and synchronously connect the functions and services required for the operation of the O&M system.

- Dynamic Database based on Real-time Monitoring:

When the spatial information from various usage data in real buildings is collected, the system could continuously transmit the parameterized data to the database. In addition to using the database as the source of review for system learning and reflection, such a database would also become a dynamic database for real-time display of building usage.

- Mapping Data-driven Visualized Results:

With the import of real-time data, users would be able to customize the colorized levels for visual alters and the threshold for event trigger with the system's logical decision-making function, which would allow the decision-making process more adaptable.

- Hierarchical Integration of Contextualized Spatial Information:

The system would be presented in an adaptive hierarchical manner for different scales and service targets. In addition to allowing precise deployment of various services, users could also perceive the contextualized spatial relationship.

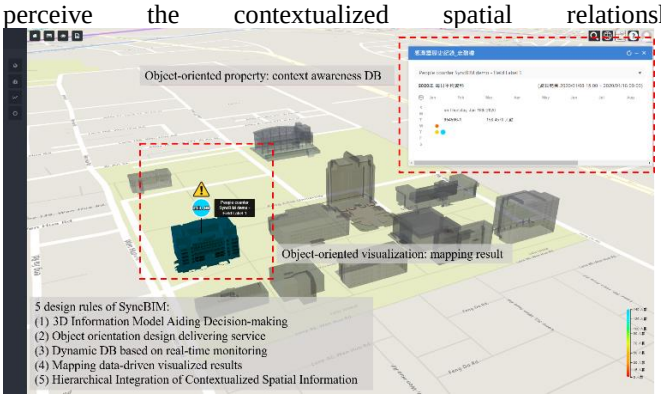


Figure 9: System Interface of SyncBIM

4.2. Importing SyncBIM into MATLAB for Computer Vision

MATLAB is software designed for currently commonly available algorithms, which is mainly used for data analysis and development of algorithms and also image processing and computer vision. The IoT platform "ThingSpeak" owned by MATLAB can transmit the analyzed results and parameters to the IoT platform, and further serve as a scalable modular web portal for the dynamic database of SyncBIM. The API could be used to access the data in the designated channel of the IoT platform (see Figure 10).

```
[initialNum,latestDate] = thingSpeakRead(954596,'ReadKey','QOX784947******');
if dateshift(latestDate,'start','day') < datetime('today','TimeZone','local')
    logThingSpeakData('reset')
end
```

Figure 10: ThingSpeak API

The designated IoT device would be set to ensure that the information is free of overlay (see Figure 11). This step would allow the continuous recording of the on-site data in the IoT platform and transmission to the SyncBIM database.

```
vidObj = webcam('STARTEC 1.3MP Webcam','Resolution','640x480');
faceDetector = vision.CascadeObjectDetector('MaxSize',[150 150]);
tracker = MultiObjectTrackerKLT;
if ~isempty(initialNum)
    tracker.NextId  = initialNum(1) + 1;
end
```

Figure 11: IoT Device and Object Tracker

The timestamp would be generated for the collected (Figure 12), and such data would be continuously uploaded to the IoT channel shown in Figure 10. The timestamp of the data will be able to provide a future database as a basis for the decision-making of spatial events and situations.

```
tim = timer;
tim.ExecutionMode = 'FixedRate';
tim.Period = 20;
tim.TimerFcn = @(x,y) logThingSpeakData(tracker);
tim.StartDelay = 5;
end
```

Figure 12: Timer for Logging Data

After running such a program in the IoT device, it can continue to perform image recognition and facial recognition (Figure 13). It would not only calculate the number of people, but also record the time point, upload the data to the designated IoT platform, and continue to be available for retrieval by the SyncBIM System database. Such a device has scalability, which allows it to be installed in different spatial architecture and scaled as required. For the system itself, only the specified channel data retrieval settings are required to allow the division of labor for O&M for achieving a higher level of efficiency.

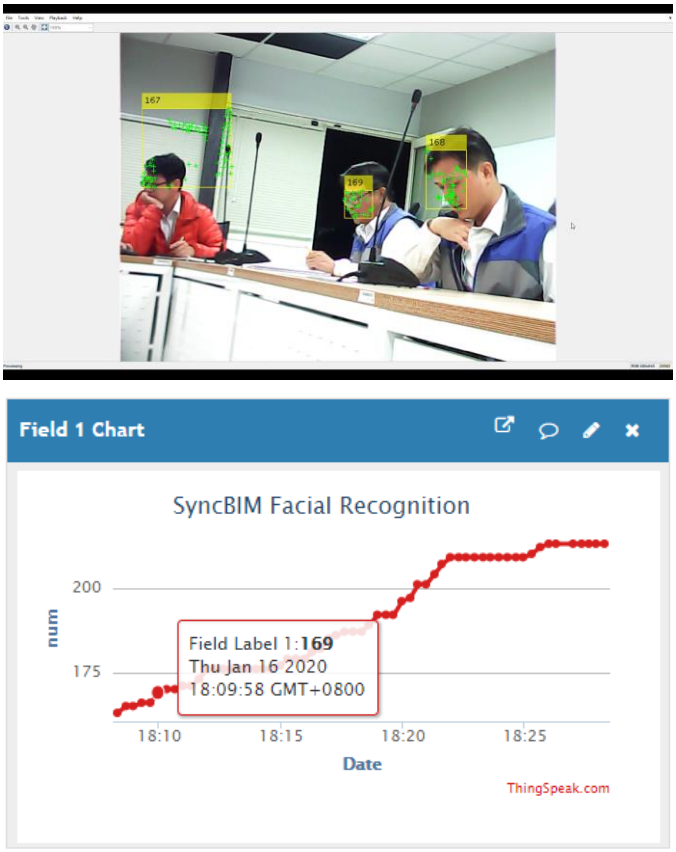


Figure 13: Top: Image Recognition Screen; Bottom: Data Collection Channel of IoT Platform

The ThingSpeak IoT platform would serve as a web portal to save the collected real-time data, while SyncBIM retrieves the data in the channel and continuously updates the dynamic database. The flexible architecture of the system allows the user preferences and spatial characteristics to be visualized for display (see Figure 14). The number and time data of user facial recognition collected in this study, combined with the building spatial information in BIM, could provide system operators with the required actions for interpreting the O&M situation, and even further adjusting the types of data collection such as various real-time information like air quality, temperature, and humidity. The adjustment for software and hardware would be feasible from simple video lenses to temperature detection cameras, and even computer vision algorithms could be modified according to events, such as detecting whether the person is wearing a mask (which is one of the major safety issues for spaces under the current global pandemic of COVID-19 and more application scenarios and services will become necessary). The spatial information and information integration in SyncBIM is the most appropriate

platform to provide these derivative services. In the future building environment, it can be foreseen that such type of cloud service platform will become a necessary asset, even as a part of infrastructures. For O&M, it can also effectively improve efficiency and further enhance the mutually friendly relationship between people and the environment.

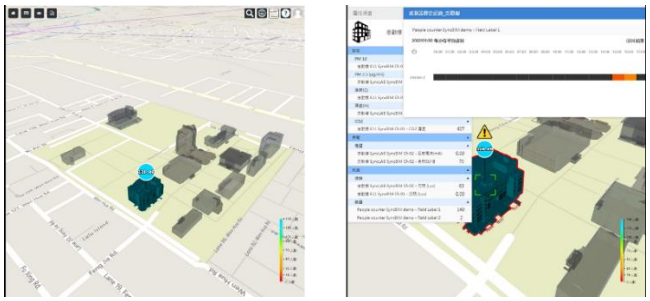


Figure 14: The system displays the facial recognition data screen. Left: The viewing panel of the 3D models. Right: The viewing panel of detailed object information, such as the time and current data of each sensor.

5. Conclusion

This study proposes a cloud platform that integrates computer vision and building information models, which is called SyncBIM system, it provides the monitoring and information visualization functions required for building O&M, and improves the passive and equipment troubleshooting-oriented building maintenance modes. In addition to being used for auxiliary O&M decision-making purposes, it can also immediately reflect the O&M state most suitable for the very moment and distinguish the relationship between the user and the spatial context, which is also a major feature of the SyncBIM system. By recording and reflecting the changes in the spatial context and usage situation to provide various instant feedback information required for O&M of buildings, it is expected that to use the SyncBIM system as the best driver for the common prosperity between people and the environment and the pursuit of well-being in the increasingly complex and multi-faceted building O&M context.

References

[1] Q.Qi, F.Tao, Y.Zuo, D.Z.-P.Cirp, undefined2018, "Digital twin service towards smart manufacturing," Elsevier, Sep. 2021.

[2] A.Matthews, B.Ta, "Applying a National BIM Model to Vietnam's National Implementation of BIM: Lessons Learned from the UK-Vietnam Collaboration for the Industry," Lecture Notes in Civil Engineering, **54**, 57–66, 2020, doi:10.1007/978-981-15-0802-8_6.

[3] Y.T.Shen, T.Y.Wu, "Sync-BIM: The interactive BIM-based platform for controlling data-driven kinetic façade," Communications in Computer and Information Science, **618**, 445–450, 2016, doi:10.1007/978-3-319-40542-1_72.

[4] C.E.Yang, Y.T.Shen, S.H.Liao, "Integration of Real-time Facial Recognition and Visualization in BIM," 2nd IEEE Eurasia Conference on IOT, Communication and Engineering 2020, ECICE 2020, 243–245, 2020, doi:10.1109/ECICE50847.2020.9301999.

[5] J.Kim, H.Kim, Y.Ham, "Mapping Local Vulnerabilities into a 3D City Model through Social Sensing and the CAVE System toward Digital Twin City," Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019, 451–458, 2019, doi:10.1061/9780784482445.058.

[6] F.Tao, M.Zhang, A.Nee, "Digital twin driven smart manufacturing," 2019.

[7] I.Din, M.Guizani, J.Rodrigues, …S.H.-F.G., undefined2019, "Machine learning in the Internet of Things: Designed techniques for smart cities," Elsevier, Sep. 2021.

[8] J.Chen, T.Tseng, …C.L.-2012 9th I., undefined2012, "An intelligent virtual fence security system for the detection of people invading," Ieeexplore.Ieee.Org, 2012, doi:10.1109/UIC-ATC.2012.64.

[9] M.Despotovic, D.Koch, S.Leiber, …M.D.-E. and, undefined2019, "Prediction and analysis of heating energy demand for detached houses by computer vision," Elsevier, Sep. 2021.

[10] E.I.Brunberg, K.E.Bøe, K.M.Sørheim, "Testing a new virtual fencing system on sheep," Acta Agriculturae Scandinavica A: Animal Sciences, **65**(3–4), 168–175, 2015, doi:10.1080/09064702.2015.1128478.

[11] B.Akinci, M.Fischer, J.Kunz, R.Levitt, "Representing Work Spaces Generically in Construction Method Models," Journal of Construction Engineering and Management, **128**(4), 296–305, 2002, doi:10.1061/(ASCE)0733-9364(2002)128:4(296).

[12] B.Akinci, M.Fischer, J.Kunz, "Automated Generation of Work Spaces Required by Construction Activities," Journal of Construction Engineering and Management, **128**(4), 306–315, 2002, doi:10.1061/(ASCE)0733-9364(2002)128:4(306).

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 4, 21-28 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Graph-based Clustering Algorithms – A Review on Novel Approaches

Mark Hloch[*,1], Mario Kubek[2], Herwig Unger[3]

[1]*Faculty of Electrical Engineering and Computer Science, University of Applied Sciences, Krefeld, 47805, Germany*

[2]*Central department I, FernUniversität Hagen, 58097, Germany*

[3]*Chair of Computer Engineering, FernUniversität Hagen, 58097, Germany*

ARTICLE INFO

Article history:
Received: 22 April, 2021
Accepted: 27 June, 2021
Online: 10 July, 2021

Keywords:
Graph-based clustering
Co-occurrence graph
SeqClu
DCSG

ABSTRACT

Classical clustering algorithms often require an a-priori number of expected clusters and the presence of all documents beforehand. From practical point of view, the use of these algorithms especially in more dynamic environments dealing with growing or shrinking corpora therefore is not applicable. Within the last years, graph-based representations of knowledge such as co-occurrence graphs of document corpora have gained attention from the scientific community. Accordingly, novel unsupervised and graph-based algorithms have been recently developed in order to group similar topics, represented by documents or terms, in clusters. The conducted work compares classical and novel graph-based algorithms, showing that classical clustering algorithms in general perform faster than graph-based clustering algorithms. Thus, the authors' focus is to show that the graph-based algorithms provide similar clustering results without requiring an hyperparamter k to be determined a-priori. It can be observed that the identified clusters exhibit an associative relationship reflecting the topical and sub-topical orientation. In addition, it is shown in a more in-depth investigation that the Seqclu (sequential clustering algorithm) can be optimized performance-wise without loss of clustering quality.

1 Introduction

Clustering is the process of grouping the most similar objects, e.g. images or text documents, in the same cluster in an unsupervised manner. In contrast to supervised classification, where the algorithm has been trained how to map its input to an according output, clustering only uses the provided input data and tries to find the best grouping of objects based on that information. Due to the sheer amount different data-types and according use-cases, dozens of architectural concepts, such as hierarchical, partitioning or graph-based clustering have been developed over the last decades[1]. Many of the classical, typically vector-based algorithms, such as the k-means[2], k-means++[3] or k-NN[4] algorithm come with the requirement of choosing the hyperparameter k, as the suggested number of expected clusters, a priori. This forces the user to estimate beforehand what number of result clusters are expected. This approach therefore softens the idea of an unsupervised algorithm providing the best possible result fully automatically, without user intervention. In addition, actual standard algorithms mentioned above, but also newer graph-based algorithms, e.g. Chinese Whispers[5] expose another weakness: they typically require a full set of documents beforehand and are not designed to adapt a growing or shrinking set of input data over time. In use-cases like building a web-engine clustering is known to improve the usability for the user: instead of having a long list of search results on a user's query the knowledge, that lies within the available document corpus, can be presented much more easily[6]. With the work of [7], [8] it is shown that the use of co-occurrence graphs is very useful for graph-based concepts on which novel clustering techniques can be built on. Recent work on graph-based clustering algorithms[9], [10] provide novel approaches in the field of clustering. This paper gives a comparison of these graph-based algorithms and shows their benefits over classical approaches, as well as requirements for further optimization.

2 Materials and methods

2.1 Classical clustering algorithms

2.1.1 K-means, K-means++, Mini Batch

Due to its simplicity the k-means algorithm and its variants are the most prominent representatives of vector based clustering algorithms used. As k-means assigns each input data point to exactly one cluster it is considered as a classical hard-clustering algorithm.

[*]Corresponding Author: M. Hloch, Reinarzstr. 49, 47805 Krefeld - Germany, mark.hloch@hs-niederrhein.de

The computational complexity, which is linearly proportional to the size of datasets, makes the k-means algorithm efficiently applicable even to larger datasets. Because each data point is represented numerically, the application field for k-means is wide: It can be used from document clustering up to other use cases, e.g. customer segmentation [11] or cyber-profiling criminals [12].

The general idea of the algorithm is partitioning the given data into k distinct clusters by iteratively updating the cluster centers and cluster associated data points. The k-means algorithm mainly performs two steps: Firstly, the user has to define manually the value k, which determines how many clusters should be a result of k-means. Secondly, a loop of two repeating steps of assigning the data set points to one of the clusters with lowest distance to the cluster's centroid and the calculation of a new centroid for each cluster is performed. Mathematically it can be said that the k-means is an optimization problem where the objective function that is employed is the Sum of Squared Errors (SSE). During the assignment and update steps the k-means algorithm tries to minimize the SSE score for the set of clusters. A more detailed overview on the mathematical implications and technical origins can be found at [13].

Due to its impact on the clustering result the distance measure has to be chosen carefully. In general, the most popular choice for estimating the closest centroid to each of the datapoints is the Euclidean distance. Other distance measures, such as the Manhattan distance or cosine similarity can also be used [13]. The k-means algorithm comes with two major disadvantages:

1. choosing the initial cluster centers
2. estimating the k-value.

The clustering result of k-means highly depends on the **initialization of the cluster centers** [14]. In order to optimize the results provided by k-means especially the initialization of the cluster centers has been subject of research over the years and lead to several approaches [15]–[16]. In contrast to the classical k-means algorithm as described in [2], the k-means++ more carefully determines the initial cluster centers and subsequently uses a weighted probability score to improve the finding of cluster centers over time.

Estimating the k-value and therefore the number of output clusters k a priori also is a big disadvantage in contrast to fully unsupervised clustering algorithms. From the practical point of view this would require the user to estimate a good k value before actual clustering can be performed. Especially in cases where the input data are very large or growing over time the manual selection of k is therefore not applicable.

Several approaches such as the Silhouette Coefficient [17] or Calinski–Harabasz Index [18] can be used to suggest the k-value automatically. In addition to the above-mentioned improvements, other algorithms like the Mini-batch k-means or fast k-means [19], [20] aim to improve the scalability and performance of k-means for large datasets such as web applications. In case of the mini-batch k-means algorithm small random batches of data are chosen and assign each of the sample points to a centroid. In a second step the cluster centroid is then updated based on the streaming average of all of the previous samples assigned to that centroid.

2.1.2 *Chinese Whispers*

The Chinese Whispers (CW) algorithm [5] is a randomized graph-based hard-clustering algorithm. Due to its simplicity and linear time properties, it performs very fast even for larger datasets. As it does not require a preliminary k-value it can be considered as an unsupervised clustering algorithm. The type of graph on which the CW algorithm is applied can be weighted, unweighted, undirected or directed. The CW is there-fore applicable to a wide range of use cases in natural language processing such as language separation or word sense disambiguation. The CW algorithm works in a bottom-up manner by first setting a random class label for each node of the graph and then merging class labels with those local neighborhood classes with the biggest sum of edge weights. In case of multiple winning classes, one is chosen randomly. Over time regions of the same class will stabilize and grow till they connect to another class region.

As shown in [5], CW scales are very well even for large datasets in linear time. It is also shown that the clustering quality is comparable to standard algorithms using the vector space model. As the CW algorithm has randomized properties its output changes on each run of the algorithm. This makes it hard for practical applications where the data change over time or in cases where the clustering process has to be repeated for the same data. Another disadvantage of the CW algorithm is its tendency of forming a large amount of often very small clusters. In order to extract the dominant clusters, e.g. to obtain an overview of related topics for each cluster, filtering is required which may result in unwanted information loss.

2.2 *Novel graph-based clustering algorithms*

2.2.1 *Dynamic clustering for segregation of co-occurrence graphs (DCSG)*

The DCSG [10] algorithm is a novel clustering algorithm inspired by the human's brain learning process. As the human brain develops from child to adulthood it continuously learns new words and categorizes them forming the entire knowledge of the human being. Transferring this method to the concept of co-occurrence graphs, clustering can be applied in order to identify topical related regions (clusters) formed by the terms within the graph. DCSG imitates the learning process by reading each document on sentence base while building up the co-occurrence graph. The nodes represent the terms, edges the relation of the words. While adding each term to the co-occurrence graph clustering is applied by measuring the distance between each new term and its distance to the existing cluster centers. The distance between two terms is determined by the inverse of the DICE-coefficient [21], [22]. The cluster center itself is represented by its centroid term as defined in [23], [24]. A cluster center is formed by the node with the shortest average distance Δd to every other node in the co-occurrence graph. In addition, the standard deviation μ is determined. Based on this information the cluster range $r_{cluster_i}$ is defined as

$$r_{cluster_i} = \Delta \mathrm{d} + 3\mu \tag{1}$$

and is used to determine which term belongs to which cluster. While processing new terms two major steps are performed

1. **Insertion of terms**

   New terms will be added to the graph as nodes and edges to the according co-occurrent terms. If a term is already existing the related connections will be updated.

2. **Clustering**

   Each term t_{new} will be merged into existing clusters if the distance $t_{new} \leq \Delta d + 3\mu$. In case that $t_{new} > \Delta d + 3\mu$ a new cluster will be created. In addition, the algorithm will check if due to the insertion any relations may have changed in order to update the cluster centers correctly.

Adding a new term that already exists in the graph will result in a change of the relation weights and cluster centers may change. In this case $\Delta d + 3\mu$ will also change. The membership of the according terms with that cluster therefore needs to be re-evaluated in case of exceeding the distance threshold might be moved to different clusters.

DCSG is a novel brain inspired algorithm that works in contrast to the classical approaches like k-means on dynamic data growing over time. As presented in the results on clustering text data show that this knowledge can be represented very well. Due to the high amount of recalculation steps, especially when the cluster centers move frequently, having a large amount of initial input data the DCSG algorithm will perform very slow. If the data is growing over time, just similar to the human brain's learning process, this disadvantage will be compensated as the graph will converge over time and recalculations will decrease.

2.2.2 *Sequential Clustering using Centroid Terms (SeqClu)*

The SeqClu-algorithm [9] is a co-occurrence graph-based hard-clustering algorithm that is capable of clustering documents in a sequential manner using the concept of centroid terms as described in [22]–[24].

In contrast to many of the standard clustering algorithms it does not require an a-priori definition of the number of clusters. As the algorithm works sequentially a set of feature vectors F of documents is processed incrementally at once or just as they appear over time. SeqClu is therefore applicable in cases where the input vector might change over time such as in a Web-Engine, where documents appear or disappear over time.

The general idea of the algorithm is to compare each new feature vector f (a document) against an existing set of clusters containing previously clustered documents. For each existing cluster the algorithm considers the membership of a new document by a distance determination process. If a closest cluster can be found without exceeding a certain threshold the document will join the according cluster. If not, a new cluster will be formed by the document extending the cluster model.

There are mainly three crucial parts that influence the behaviour of the algorithm:

- initialization,
- cluster membership value and
- threshold of membership (winning cluster determination).

The **initialization** can be performed in various ways and influences the quality of clustering result. If a set of one or more clusters are predefined it will be used as a reference for all further arriving documents to be clustered. The simplest initialization would be using a single cluster having a randomly chosen document as the first cluster. In combination with an inaccurately chosen membership threshold this approach most likely will tend to merge new documents, especially a low number of documents, into the initial cluster and imprecise the overall clustering result. Instead, it is suggested to form at least two initial clusters formed by the most distant documents (antipodean documents) existing. This requires an additional preprocessing step as for each of the initially existing documents the distance between each of the documents centroid must be determined. It is known that the co-occurrence graph will converge at about 100 documents and the process of distance determination can be very time consuming as the conducted experiments show. It is therefore suggested to consider only 2 to 100 randomly chosen documents and is also suspect of the experiments in this publication.

In order to determine a document's **cluster membership** each newly arriving document needs to be matched against the existing clusters. For this purpose, the average distance between the new documents centroid and all existing documents in each of the clusters is determined by using Dijkstra's shortest path algorithm. In order to speed-up the calculation process all previous path calculations are cached having a linear time complexly for reoccurring path determinations.

Regarding the judgement whether a document shall be merged into an existing or form a novel cluster, a **threshold** is required. As the threshold determination process directly influences the clustering result it must be chosen very carefully. To avoid having the user to choose this threshold manually, it is chosen dynamically for each of the clusters using the local connections between the cluster's centroid and its nearest neighbors. In order to reduce the amount of computation time, only neighbors within a certain radius are considered using a breadth-first search with limited depth.

As first experiments have shown the SeqClu-Algorithm is able to cluster documents unsupervised in a sequential manner providing good results without performing any optimization regarding the process of threshold determination. The initial step of finding the two initial clusters, especially having a large initial set of documents, is very time consuming. On the other hand, the calculation results of the initialization phase can be reused for the clustering process itself having a lookup table. In this case the disadvantage of slow initialization becomes an advantage for the clustering process itself and will speed-up this phase of the algorithm, which will be shown in the experiments conducted in this publication.

2.3 *Conceptual differences of tested algorithms*

Table 1 provides an overview over the six tested algorithms from conceptual point of view. All algorithms work with different requirements for initialization and are applicable to different use cases. In general, the classical algorithms are flexible regarding the type of input data, which can be numerical or textual data. From the user point of view the biggest disadvantage is the requirement to provide the number of output clusters (k-value) a-priori. Therefore, user

intervention is required in order to update the suggested number of output clusters or the use of additional algorithms that determine the k-value automatically. The classical algorithms are therefore especially inflexible in cases, where the corpus size changes over time. In contrast, the graph-based algorithms fully run unsupervised and therefore do not require any further change in initialization even if the corpus grows or shrinks. From clustering point of view the algorithm will adapt and provide automatically an according number of output clusters.

Table 1: Conceptual differences of tested algorithms

Algorithm	Type	Supervised/ unsupervised	A-priori no. of clusters	Growing corpora
K-means	vector-based	supervised	k-value	no
K-means++	vector-based	supervised	k-value	no
Minibatch	vector-based	supervised	k-value	no
CW	graph-based	unsupervised	none	no
SeqClu	graph-based	unsupervised	none	yes
DCSG	graph-based	unsupervised	none	yes

3 Results and discussion

3.1 Setup of experiments

For the conducted experiments, natural language preprocessing in form of sentence extraction, stop word removal and baseform reduction has been applied to all of the used documents. The according output was used as standardized input for all algorithms tested. In addition, the co-occurrence graph G has been built on sentence-based co-occurrences measuring the distance between terms by the reciprocal significance value. The significance value itself is determined using the DICE-coefficient. All small sub-graphs from the original co-occurrence graph have been removed finding the largest connected sub-graph within G in order to obtain a consistent single graph G. The graph is stored in an embedded graph database using Neo4j [25] where each node represents a term and the edges are annotated with their significance and respective distance value. As the vector-based algorithms don't use a co-occurrence graph the TF-IDF [26] matrix was created based on the documents term-vector, which was extracted during natural preprocessing. The implementation of k-means, k-means++ and mini-batch were realized by using the python-based machine learning programming library scikit-learn version 0.23. The Chinese Whispers algorithm was taken from the ASV Toolbox [27] from the university of Leipzig and is written in Java. SeqClu and DCSG algorithms were both implemented using the Java programming language. All experiments have been performed on ten workstations with identical hardware specs.: Intel Core i7-7700K CPU @ 4.20GHz and 16 GB Memory

3.2 Used corpora and parametrizations

The conducted experiments were performed over corpora consisting of 40,60,. . . ,260,280 documents and for each corpus the tested algorithms did run 100 times in order to obtain statistical relevant results. The corpora each contained a random number of equally distributed documents of the categories politics, cars, money and sports of the German newspaper "Der Spiegel". The documents themselves were each tagged by the author with their respective text category and therefore can be used as a gold standard in order to evaluate whether a document has been clustered correctly or not. Table 2 shows the parametrizations that were used for the individual algorithms.

Table 2: Parametrization of used algorithms

Algorithm	Parametrization
K-means, K-means++	Number of output clusters (k-value): 4 Initial cluster center: random Number of iterations: 100
Minibatch	Identical to K-means Batchsize: 100 Number of samples (init): (3 ∗ batchsize)
CW	Number of iterations: 100 Mutation rate: 0.0
DCSG	No initialisation required
SecClu	Initialisation: antipodean documents Dynamic threshold

3.3 Clustering Quality

Evaluating different clustering algorithms is difficult. Commonly used external evaluation measures, e.g. the f-measure or rand index, penalize false positive and false negative decisions. If the number of output clusters exceeds the number of classes it will result in a quality trade-off. Especially the tested graph-based clustering algorithms don't require a pre-defined k-value and break this condition by exceeding the number of output clusters. In order to limit the effect of quality trade-off the purity is preferred over other evaluation measures as it does not penalize if the number of output clusters is bigger than the number of class labels. The purity was calculated for each of the tested corpora and algorithms by

$$purity(C, M) = \frac{1}{N} \sum_{k} \max_{j} |c_k \cap m_j| \tag{2}$$

with N as the total number of documents, the set of clusters $C = \{c_1, c_2, \cdots, c_k\}$ and $M = \{m_1, m_2, \cdots, m_j\}$ as the set of classes. For each cluster c_k we determined the class m_j with the most members $n_{k,j}$ in c_k and then subsequently $n_{k,j}$ is summed up and divided by N.

Figure 1 shows the total average purity of all algorithms. The classical algorithms like k-means reside within a range of 0.7 to 0.8 exhibiting a good purity value. The graph-based algorithms, even with lower purity values, still are close having average values between 0.6 to 0.7.

It can be observed from Table 3 that for the standard algorithms the purity in average is almost constant and independent of the number of clustered documents.

The graph-based algorithms SeqClu and DCSG show, that their purity values increase in quality at about 100 to 160 documents per corpus. For SeqClu this behavior reflects that the co-occurrence graph converges at about 100 documents. Each additionally added document causes lesser changes to the co-occurrence graph resulting in lesser changing centroid terms, that are used by this clustering

algorithm for cluster determination. The DCSG also shows this tendency as the knowledge of the graph that is used for clustering also improves over time and stabilizes.

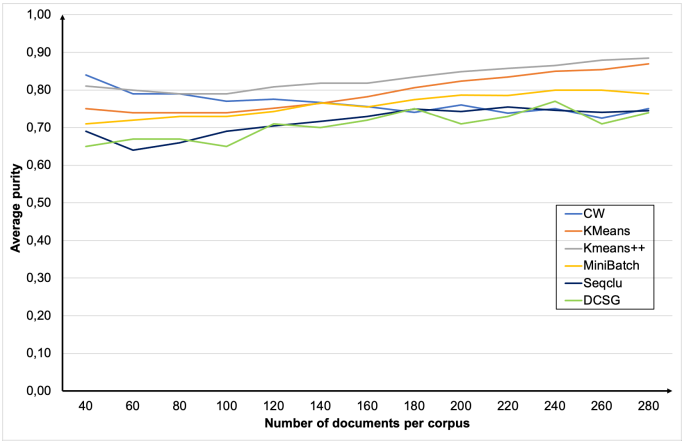


Figure 1: Comparison of average purity for different corpora sizes

Table 3: Average purity for each of the tested algorithms

# of doc.	K-Means	K-Means++	Minibatch	CW	SeqClu	DCSG
40	0.75	0.81	0.71	0.84	0.69	0.65
60	0.74	0.80	0.72	0.79	0.64	0.67
80	0.74	0.79	0.73	0.79	0.66	0.67
100	0.74	0.79	0.73	0.77	0.69	0.65
120	0.75	0.81	0.74	0.78	0.70	0.71
140	0.76	0.82	0.77	0.77	0.72	0.70
160	0.78	0.82	0.75	0.76	0.73	0.72
180	0.81	0.83	0.77	0.74	0.75	0.75
200	0.81	0.85	0.79	0.76	0.74	0.71
220	0.83	0.86	0.79	0.74	0.75	0.73
240	0.85	0.87	0.80	0.75	0.75	0.77
260	0.85	0.88	0.80	0.72	0.74	0.71
280	0.87	0.88	0.79	0.75	0.75	0.74
Avg. total	0.79	0.83	0.76	0.77	0.72	0.71

3.4 Cluster-sizes and number of clusters

As the purity can be influenced by having a large number of clusters - in worst case one cluster per document - it is required to take a closer look at the number of clusters and documents per cluster. This effect mainly has an effect to the graph-based algorithms as they expose in contrast to the standard algorithm ($k = 4$) a larger number of output clusters. As the SeqClu algorithm is in the authors focus of research the investigations were focused on this algorithm and resulted in the following two main observations:

1. The average cluster size in average is almost constant between 4 and 5 documents per cluster and

2. the number of clusters is slowly increasing from 20 to 50 for the tested cluster sizes from 40 to 280 documents.

In general, it can be concluded that the knowledge growth at this point seems to be not converged or the parametrization needs to be finer grained, as all the tests were using standard parametrization. In order to reveal possible indicators and starting-points for further optimization, the tested corpora were manually examined regarding the clustering result and the individual correctness of the clusters, even if the gold standard suggested number of output clusters is exceeded.

After examination the following commonly occurring observations have been made for the major number of tested corpora:

1. **Homogenous Clusters**

   Homogenous Clusters are found 4 to 5 times in average, having 10-20 documents with 100 percent accuracy and reflect an actual topic the documents focus on. Those clusters contain exactly one category and in addition, the topical relatedness within that category is identifiable. For example, a cluster of category “sports” can be clearly related to the topic formula 1 racing by the document centroids (number of occurrences in brackets): Season (3), Schumacher (8), Race (5)

2. **Heterogeneous Clusters**

   Heterogeneous Clusters occur in 1 to 2 clusters at a size of 10-20 documents of mainly two mixed topics. It could be observed that in some cases the centroid term that mainly influences the categorization of a document is generic. For example, terms like Euro, Zahl or Percent are very hard to map to a certain category. The according document could be related to Money, Politics or even Sports. In general, it seems difficult to identify the main scope of a cluster automatically for heterogeneous but also homogenous clusters. An according algorithm must be able to identify the topic and in case of having a heterogeneous cluster, differentiate between the mixed topics, which could occur on term level, which is focus on further research.

3. **Border Clusters**

   Border clusters occur at a varying number of times, depending on the topics within the corpus. These cluster types have an average size of 5-6 documents dominated by exactly one topic, but also having a single document of a different topic within. For example, the cluster may contain the centroid terms {War(2), Irak(2), USA(1)} of category politics, but one document with centroid {Beginning}. The politic topic (War, Irak, USA) can be clearly identified as the main topic for this cluster. In this example the document with the centroid Beginning cannot be accurately assigned to a single topic. The actual document refers to the beginning of a voting process regarding enrolment fees at the beginning of the semester. As the term {Beginning} also makes sense if related to the beginning of war between US and Irak the cluster has its border to a different topic marked by this term. Border clusters are interesting due to the fact that they mark the border between somehow topical related areas in the graph. This might be used e.g. within a search engine to lead the searching user to adjacent fields of knowledge.

4. **Single document clusters**

   Single document clusters are clusters formed by single documents that could not be assigned to any other cluster by exceeding the dynamic threshold used in SeqClu. In very early steps there might be not as many clusters with clearly defined topics that could lead into the creation of these new

clusters. As SeqClu is a single linkage algorithm it is prone to outliers. Further research in order to mitigate this problem has to be done in future work, e.g. by merging the single document cluster with a larger cluster within a short distance.

Putting things together it can be concluded that the graph-based clustering algorithm performs very well regarding purity. As the number of output clusters is greater than the gold standard suggests, the differentiation of different topics is much closer to real world applications. In contrast, the standard algorithms gain a high purity. But when manually evaluating the clustering results topical differences are not detected.

3.5 *Pure clustering performance*

Figure 2 shows the pure execution time of clustering different corpus sizes for all standard and graph-based clustering algorithms, except DCSG.

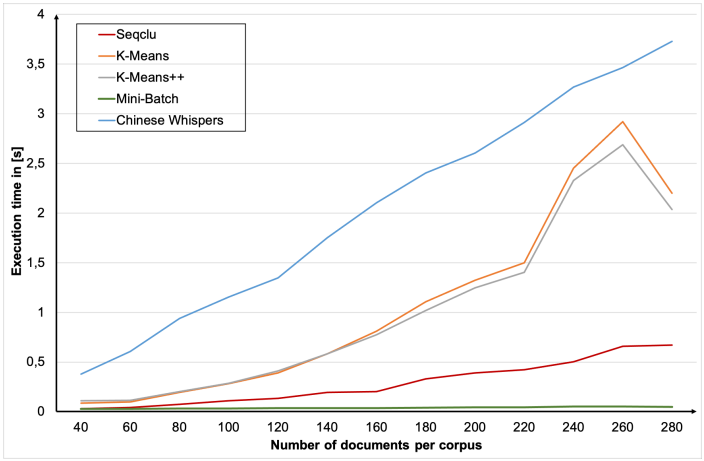


Figure 2: Pure execution time for clustering

In general, it can be concluded that the Minibatch algorithm performance is outstanding in comparison to the other algorithms even for larger corpora. As the Minibatch algorithm is designed especially for larger datasets, it is not surprising that it performs that well and faster than k-means and k-means++.

The graph-based SeqClu algorithm shows a linear increasing execution time for the pure clustering. It is also very fast in comparison of the other algorithms as it benefits from a caching mechanism of shortest path calculations that is used during the initial cluster initialization.

Figure 3 shows the overall execution time including all preliminary initializations for DCSG and SeqClu. In case of SeqClu and DCSG require, due to their focus on non-real-time applications and their heavy use of graph-based distance calculations, additional computations while clustering or during initialization. In case of the DCSG algorithm, that works on a sentence-based manner, a huge amount of distance determinations has to be performed which results in a high computation time of up to approx. 17 hours on 280 documents. As the scope of DCSG is not to be tremendously fast but trying to find an accurate representation of knowledge this performance drawback is not essential for the entire quality of the algorithm. Further investigations are beyond scope of the authors work and therefore not subject of this paper. SecClu's execution time in comparison to DCSG is about 2 hours in case of 280 documents. It is mainly influenced by the calculation of the nearest neighbors and number of documents used during the determination of initial clusters (antipodean documents). Especially the determination of antipodean documents leaves space for improvements.

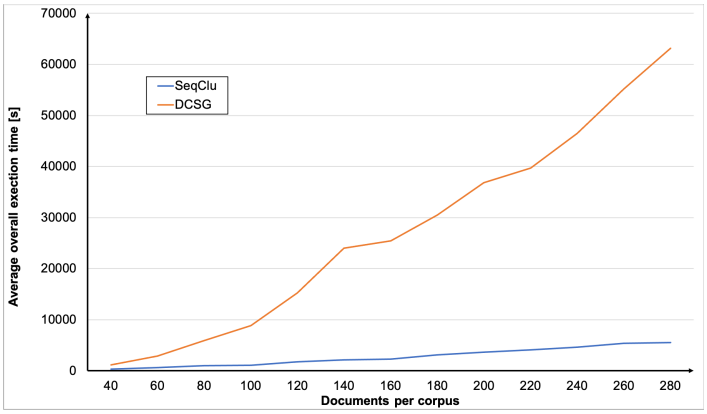


Figure 3: Execution time SeqClu and DCSG (initialization and clustering)

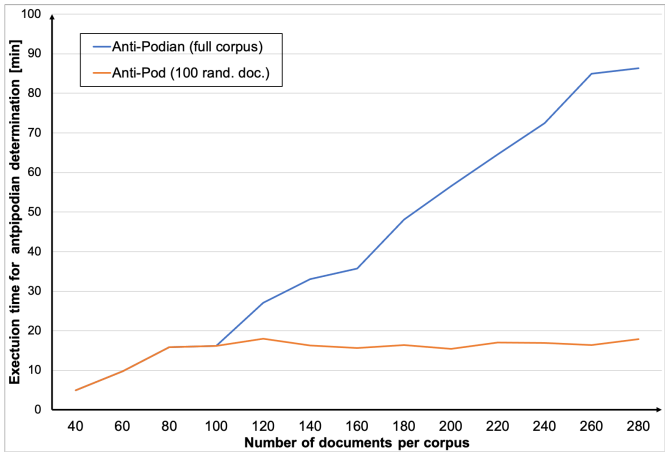


Figure 4: Execution times for initialization of SeqClu and SeqClu-100

The conducted experiments therefore have been performed by comparing the execution time for SeqClu using all available documents during initialization versus SeqClu-100, using only the first up to 100 randomly chosen available documents instead. It can be concluded (see Figure 4) that depending on the corpus size up to approx. 70 minutes of computational time could be saved for the tested corpora using SeqClu-100.

It is expected that co-occurrence graph converges at about 100 documents and can be considered as stable. When using 100 initial documents – even if randomly – chosen, the graph therefore is considered as stable and side-effects, e.g. on clustering quality should be negligible. Therefore, the purity has been also determined (Figure 5) concluding that there is no overall negative impact of using only 100 random documents for finding the two initial clusters on clustering quality.

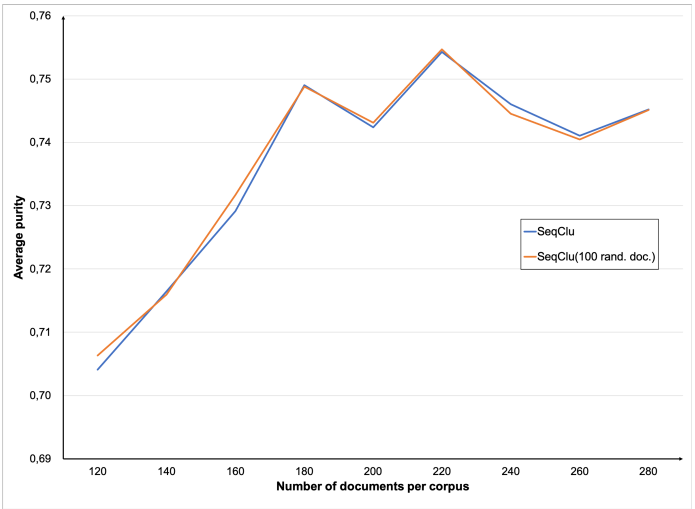


Figure 5: Comparison of purity for SeqClu and SeqClu-100

Further experiments also showed that the average cluster-sizes and number of output clusters do not change significantly and limiting the initialization of SeqClu to 100 documents is a performance enhancement which is reflected by having a better overall performance.

4 Conclusions

The provided work compares novel graph-based clustering algorithms against well-known vector-based algorithms. All algorithms were investigated regarding their clustering quality and general performance. The results show that the vector-based algorithms generally perform at a higher speed in contrast to the examined graph-based clustering algorithms. In contrast, the classical approaches such as the k-means algorithm forces the user's intervention a-priori, which limits the use cases where the user is able to investigate into the input data before actually clustering is applied. In contrast, the graph-based clustering algorithms show that a good categorization without the preliminary of a k-value requirement is possible. In addition, it can be concluded that graph-based clustering provides the property of having an associative representation - similar to the human brain – of the clustered data. A topical differentiation between individual topics subtopic is therefore much closer to the actual way of thinking of the user in contrast to classical approaches. The in-depth investigation into SeqClu's clustering structure and relations between the result clusters gained in further starting-points for optimization. Additionally, it was shown by limiting the number of documents during initialization that the performance of SeqClu can be significantly improved without negative impact on the overall clustering quality.

References

[1] V. Estivill-Castro, "Why So Many Clustering Algorithms: A Position Paper," SIGKDD Explor. Newsl., **4**(1), 65–75, 2002, doi:10.1145/568574.568575.

[2] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, 281–297, University of California Press, Berkeley, Calif., 1967.

[3] D. Arthur, S. Vassilvitskii, "K-means++: The Advantages of Careful Seeding," in Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07, 1027–1035, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2007.

[4] N. S. Altman, "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression," The American Statistician, **46**(3), 175–185, 1992, doi:10.1080/00031305.1992.10475879.

[5] C. Biemann, "Chinese Whispers: An Efficient Graph Clustering Algorithm and Its Application to Natural Language Processing Problems," in Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing, TextGraphs-1, 73–80, Association for Computational Linguistics, Stroudsburg, PA, USA, 2006.

[6] M. Christen, et al., YaCy: Dezentrale Websuche, Online Documentation on http://yacy.de/de/Philosophie.html, 2017.

[7] M. Kubek, Dezentrale, kontextbasierte Steuerung der Suche im Internet, Ph.D. thesis, Hagen, 2012.

[8] M. Kubek, Concepts and Methods for a Libarian of the Web, FernUniversität in Hagen, 2018.

[9] M. Hloch, M. Kubek, "Sequential Clustering using Centroid Terms," in Autonomous Systems 2019: An Almanac, 72–88, VDI, 2019.

[10] S. Simcharoen, H. Unger, "Dynamic Clustering for Segregation of Co-Occurrence graphs," in Autonomous Systems 2019: An Almanac, 53–71, VDI, 2019.

[11] M. Bacila, R. Adrian, M. Ioan, "Prepaid Telecom Customer Segmentation Using the K-Mean Algorithm," Analele Universitatii din Oradea, **XXI**, 1112–1118, 2012.

[12] M. Zulfadhilah, Y. Prayudi, I. Riadi, "Cyber Profiling using Log Analysis and K-Means Clustering A Case Study Higher Education in Indonesia," International Journal of Advanced Computer Science and Applications, **7**, 2016, doi:10.14569/IJACSA.2016.070759.

[13] C. Aggarwal, C. Reddy, DATA CLUSTERING Algorithms and Applications, 2013.

[14] J. Pena, J. Lozano, P. Larranaga, "An Empirical Comparison of Four Initialization Methods for the K-Means Algorithm," 1999.

[15] E. Forgy, "Cluster Analysis of Multivariate Data: Efficiency versus Interpretability of Classification," Biometrics, **21**(3), 768–769, 1965.

[16] S. Khan, A. Ahmad, "Cluster center initialization algorithm for K-means clustering," Pattern Recognition Letters, **25**, 1293–1302, 2004, doi:10.1016/j.patrec.2004.04.007.

[17] L. Kaufman, P. Rousseeuw, "Finding Groups in Data: An Introduction to Cluster Analysis," 1990.

[18] T. Caliński, H. JA, "A Dendrite Method for Cluster Analysis," Communications in Statistics - Theory and Methods, **3**, 1–27, 1974, doi:10.1080/03610927408827101.

[19] D. Sculley, "Web-Scale k-Means Clustering," in Proceedings of the 19th International Conference on World Wide Web, WWW '10, 1177–1178, Association for Computing Machinery, New York, NY, USA, 2010, doi:10.1145/1772690.1772862.

[20] A. Feizollah, N. Anuar, R. Salleh, F. Amalina, "Comparative Study of K-means and Mini Batch K- means Clustering Algorithms in Android Malware Detection Using Network Traffic Analysis," 2014, doi:10.1109/ISBAST.2014.7013120.

[21] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," Ecology, **26**(3), 297–302, 1945, doi:10.2307/1932409.

[22] M. Kubek, T. Böhme, H. Unger, "Empiric Experiments with Text Representing Centroids," in 6th International Conference on Software and Information Engineering (ICSIE 2017), 2017.

[23] M. Kubek, H. Unger, "Centroid Terms and their Use in Natural Language Processing," in Autonomous Systems 2016, VDI-Verlag Düsseldorf, 2016.

[24] M. Kubek, H. Unger, "Centroid Terms As Text Representatives," in Proceedings of the 2016 ACM Symposium on Document Engineering, DocEng '16, 99–102, ACM, New York, NY, USA, 2016, doi:10.1145/2960811.2967150.

[25] A. Vukotic, N. Watt, T. Abedrabbo, D. Fox, J. Partner, Neo4j in Action, Manning, 2015.

[26] G. Salton, A. Wong, C. S. Yang, "A Vector Space Model for Automatic Indexing," Commun. ACM, **18**(11), 613–620, 1975, doi:10.1145/361219.361220.

[27] C. Biemann, U. Quasthoff, G. Heyer, F. Holz, "ASV Toolbox: a Modular Collection of Language Exploration Tools," in Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), European Language Resources Association (ELRA), Marrakech, Morocco, 2008.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 22-28 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Service Lifetime Loss Evaluation Method for Transformers Exclusively Serving Solar Power Plants

Bonginkosi Allen Thango*, Jacobus Andries Jordaan, Agha Francis Nnachi

Department of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 29 May, 2021
Accepted: 07 September, 2021
Online: 13 November, 2021

Keywords:
Renewable Energy
Transformers
Service lifetime
Total Ownership Cost

ABSTRACT

In the last decade, South Africa has attracted and attained more investment by means of the Renewable Energy Independent Power Producer Procurement Programme (REI4P), which is, a structured invitation to Independent Power Producers (IPPs) to submit a bid to generate and supply power to the electrical grid. In spite of REI4P's undeniable triumph, a much-discussed drawback has been the related service lifetime cost of equipment that facilitate the renewable energy technologies. The description of the service lifetime loss method (SLLM) gets more complex in the new dawn of decarbonized electricity market. The SLLM must be modified for determining the total ownership cost of transformers facilitating Sustainable Energy Systems (SES's) in the decentralized energy market. The main focus of this work is to indicatively formulate a fundamental advancement upon the conventional service lifetime losses evaluation formula to contemplate the service lifetime loss evaluation method for transformers exclusively of service to solar power plants in South Africa. The distinct operational features of a solar plant have been embedded in the formulated service lifetime loss evaluation formula by way of the plants' Generation (GM) and Non-Generation Mode (NGM). Further, a levelized cost of energy supplied per unit of time by the solar plant is employed to determine the energy cost of the no-load and load losses that will be consumed by the studied transformers during their service life. Ultimately, the premier findings of this study indicate that the annual solar potential has an effect on the transformer service lifetime loss factors and the conventional method is not suitable thereof for this application. This is a characteristic that should be precisely considered, as it may influence the tender adjudication process to purchase a transformer based on the total ownership cost offers of various transformer manufacturers.

1. Introduction

The rapidly escalating cost of energy from coal power generation in South Africa has aroused a substantial demand for feasible recourse energy sources. In spite of the fact that at present it contributes to only a small proportion of South Africa's power generation, deployment of solar power plants (SPPs) is rapidly increasing for both utility-scale and distribute generation power generation systems (DPGS). Reduction of the energy tariffs are impelled by the technological advances, proportionate saving in costs gained by an increased level of production in manufacturing, and novelties in finance has introduced solar power at arm's length of grid parity in an increasing renewable energy markets. Ongoing advances and additional cost savings will broaden these window of opportunities especially in South Africa where opportune solar conditions prevail. The energy tariffs for renewable energy technologies has been incessantly plummeting since the inception of the Integrated Resource Plan (IRP) [1] and [2]. Figure 1 shows the annualized cost of energy (ACOE) for wind and solar [3]. Wind and solar photovoltaic (PV) are observed to cost about 0.91 ZAR/kW and 0.755 ZAR/kW while coal-based energy cost about 1.03 ZAR/kW [3].

Meanwhile the rapid increase of the South African energy mix has been attended by an observed uneven proficiency revealed the new dawn of renewable energy market. For instance, the energy mix in South Africa by the vertically-integrated utility and Independent Power Producers (IPP's) exist together at the same time but have distinctive methodologies of evaluating their capital expenses, network expenditures and load curves. Therefore, the processes for the transformer loss capitalization must be distinctive. The methods available in the literature straggle the proficiency to appropriately unbundle the loading and energy components of the cost transformer losses in the concerned

*Corresponding Author: Bonginkosi Thango, thangotech@gmail.com

https://dx.doi.org/10.25046/aj060604

entities. This will significantly guarantee that each loss component is apportioned to suitable stakeholders in decentralized energy markets with regards to who is accountable to insure the losses of the transformer.

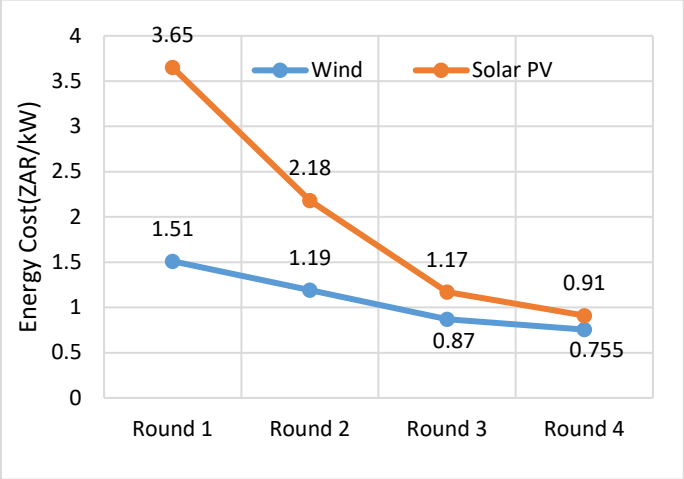


Figure 1: Energy tariffs for Wind and Solar in South Africa

Transformers are often anticipated to be in service between 25-30 years [4]. The jurisdiction procedure by utility owners for purchasing a transformer should therefore not only be based on the initial purchase cost but also the operating and maintenance costs. The total ownership cost (TOC) method [5] take into consideration the purchase cost, operational cost and maintenance costs over the transformer designed service lifetime. It has become possible for utility owners to compute the most technical and economic choice between competing transformer designs over its designed service lifetime through the TOC method. The TOC method also incorporates other cost related parameters such as annualized cost of electricity (ACOE), discount rates, tax rates et cetera. Considering that the TOC method comprises of the cost of service losses, in which will materialize in the future, there is a need to discount these losses to equalize them to current day South African rand (ZAR) value. The capitalization values cost of no-load loss (CNLL) and cost of load losses (CLL) in ZAR/kW studied by authors in [6 - 10] derive the cost of service losses. Transformer manufacturers submit a technical and commercial offer in a competitive bid and if the capitalization values are provided, then the manufactures should be grounded on these values in all cases. The service no-load losses are reliant on the core steel grade and are unvarying for a transformer during service [11] and [12]. The service load losses on the other side are dependent on the loading profile of the transformer [13] and [14]. Significant service losses of a transformer during service can potentially culminate in excessive heating and consequently degradation of insulation of materials due to overloading and presence of harmonics especially in renewable energy technologies suchlike wind and solar [15] - [20]. As a result, a low-loss transformer would practically be in service for a longer duration with minimum potential for premature failure. Additionally, a low-loss transformer discounts the amount of electrical energy necessitated to supply the service losses.

Firstly, this work presents classical SLLM for transformers that are intended to be in service for 8760 hours annually and from a single utility energy source. This method is based on transformer loss data supplied by the manufacturers and the utility characteristics.

Nevertheless, the understanding of the SLLM becomes knotty within the bounds of transformers intended to be of service to renewable energy application. A SLLM for solar power plants is then proposed in this work. This method corrected the classical method intended for transformers intended to be in service for 8760 hours annually and from a single utility energy source. In the case of solar power plants, during the non-generating period, the solar plant will be feeding-off the electric grid with coal power generation

2. Transformer cost analysis

In order to bestow the surged energy demand, the excessive loading of transformers during service and related remaining service lifetime on account of the hotspot temperature rise is structural components must be cautiously estimated in opposition to the replacement of transformers or supplementing existing unit with new transformer. IPP's generally ascertain definitely the size of suitable units for service based on current loading cycle, forecasted load growth and other sound engineering acumen. Additionally, IPP's determine the economic effect of service losses of the transformer over its planned service lifetime and the initial purchase price prior to procuring new transformers. The classical Total Ownership Cost (TOC) procedure provides a comprehensive analysis for the economic planning aspects of investing in purchasing and operating a transformer

2.1. Service lifetime loss method: Classical Method

The purchasing of transformers by SLLM is predominantly employed based on the premise of a lowest TOC amongst competing offers of various transformer manufacturers. This method provide details on the cost of transformer service losses and the initial purchasing price of the unit. The service lifetime of the unit is presumed to be even irrespective of the MVA rating. Various transformer loading guides suggest a normal operating service lifetime of 30 years [4]. The aspect of investing in purchasing a transformer is determined to reflect an annual cost of the service no-load and load losses inclusive of the demand rate. The total revenue needed is treated by the ratio of the annual cost of these losses to the fixed charge rate (FCR). The FCR illustrates all the incurred charges in the utility independent of energy (kWh) traded, suchlike depreciation, tax, insurance and service maintenance.

For each unit of kilowatt (kW) of the service no-load and load losses, there is an inclusion of the demand cost (DC) portion derived from capital cost of the transformer. Consequently, there is an inclusion of the component for the energy cost (EC) in kilowatt-hour (kWh) by the utility. The cost of service losses (CSL) are hence computed as follows in (1).

$$CSL(ZAR/kW) = DC + EC \quad (1)$$

here:

$DC = \text{System Investment}(ZAR/kW)$

$EC = \frac{8760 \times ACOE}{FCR}$

System investment (SI) illustrates the acquisition in the bulk movement of electrical energy from a generating site to distribution facilities essential to deliver supplementary energy demand arising from transformer service losses at the system peak. There are essentially two approaches to assessing the SI rate in ZAR/kW. The first one is classical and is based on the established costs of constructing a generating site and related distribution facilities. The second approach is based on a scenario where the power utility does not self-generate power and instead purchase power. The SI rate can be ascertained by ratio of the DC (ZAR/kW-yr.) and the FCR.

Consequently, the ACOE is evaluated by forecasting and weighting the annual inflated cost of energy by means of capital recovery factor (CRF). Analytically, the latter can be expressed as shown in (2).

$$ACOE = Current\ Energy\ Cost \times \mathrm{CRF} \tag{2}$$

here,

$$\mathrm{CRF} = \left(i \times \left[\frac{(1+i)^n}{(1+i)^n - 1}\right]\right)\left(X \times \left[\frac{1-X^n}{1-X}\right]\right)$$

where,

i – Discount rate

n - Designed transformer service lifetime

e – Energy escalation rate

$$\mathrm{X} = \frac{1+e}{1+i}$$

It follows that the cost of losses can be expressed as follows in (3). This equation is handily feasible for power utilities with their own electricity generation.

$$CSL(ZAR/kW) = \frac{DC(ZAR/kW-Year) + 8760 \times \mathrm{ACOE}}{CRF} \tag{3}$$

Additionally, (3) segregate apiece class of service losses in ZAR/kW as shown in (4), (5) and (6). The cost of no-load losses ($CNLL$) are expressed as shown in (3).

$$CNLL(ZAR/kW) = CGTD + \frac{8760 \times ACOE}{CRF} \tag{4}$$

Here, $CGTD$ is the utility costs for the generation, transmission and generation. The cost of load losses (CLL) are expressed as shown in (4).

$$CLL = CGTD \times PRF^2 \times EPR^2 + \frac{8760 \times ACOE \times TLF \times EPR^2}{CRF} \tag{5}$$

Here, PRF and EPR is the peak responsibility factor and the equivalent peak ratio. The PRF is the ratio of the load's demand to peak power demand at the course of the system peak as expressed in (7).

$$PRF = \frac{Transformer\ loading\ during\ system\ peak}{Transformer\ peak\ load} \tag{7}$$

The PRF can be implemented to unique customer classes. Needless to say that a loading with a PRF of one peaks concurrently with the overall system. The electrical loading of particular customer classes have a propensity to fluctuate at a similar fashion. For instance in South Africa, commercial electrical load are at peak from about 08:00AM in the morning to about 18:00PM in the evenings while residential electrical loads peak from around 18:00PM in the evenings. Inevitably, the weather has a direct impact on the loading extent. In the winter for instance, electrical heaters increase the demand and lower the variety among loads. At the distribution transformer level, the loading factors are largely ranging from about 0.4 to 0.9 with 0.75 being common [6 - 8] and [10]. The standard values of PRF can vary from 1.0 to about 0.35. Given that PRF is a ratio based on the loading (kVA), the services losses are a function of PRF^2based on transformer manufacturers empirical survey. The most common values of the PRF for various transformers is tabulated in Table 1.

Table 1: PRF values for various transformer classes

Power System Level	**PRF**
Generator Step-Up	1
Transmission Substation	0.9
Distribution Substation	0.8
Distribution	0.35

The interrelation between the service losses and the rated transformer loading is described by the peak ratio (PR). The latter can be expressed analytically as follows in (8).

$$PR = \frac{Annualized\ peak\ transformer\ loading}{Maximum\ rated\ transformer\ load} \tag{8}$$

Supposing the load increase by a specific percentage annually, then from (5), the EPR may be employed as expressed as follows in (9).

$$EPR = PR \times \sqrt{\frac{(1+g)^{2t} - 1}{In(1+g)^{2t}}} \tag{9}$$

where,

g – Rate of annual load growth

t – Time in years

Lastly from (5), the annualized TLF can be described by the ratio of the average service load loss to the maximum load losses. The latter can be expressed as follows in (10).

$$TLF = \frac{Average\ transformer\ load\ loss}{Maximum\ transformer\ service\ load\ loss} \tag{10}$$

If the power utility provides the annual loading cycle data, then the services load losses can be evaluated from this data. Nevertheless, the utility owners endeavour to link up the transformer service loss factor with a more easily accessible information viz. the load factor (LF). The determination of the TLF is dependable on some empirical data and can then be evaluated as expressed in (11) [11].

$$TLF = a \times LF^2 + b \times LF \quad (11)$$

here,

a – varies between 0.8-0.89

b – Varies between 0.2-0.11

In practice, the recommended a and b constants are 0.84 and 0.16 respectively. The total ownership cost is then expressed as shown in (12). The CSL are them multiplied by the corresponding no-load (NLL) and load losses (LL) in kW.

$$TOC(ZAR) = TP + CNLL \times NLL + CLL \times LL + CAL \times AL \quad (12)$$

Here, TP is the transformer price offered by the most competitive transformer manufacturer and CL is the cost of auxiliary losses. Prevalently, (4) and (5) are referred to as the A and B factors. Subsequently, (13) can be introduced as new factor C to the classical TOC method, in which ignores this factor.

$$TOC(ZAR) = PP + A \times NLL + B \times LL + C \times AL \quad (13)$$

2.2. Service lifetime loss method: Proposed Method

A majority of SPP's in South Africa are owned by the Independent Power Producers (IPP's). The plant consists of an array of solar photovoltaics modules connected in series and thereupon connected to an inverter that carries out the direct current (DC) to alternating current (AC) conversion. Furthermore, a step-up transformer is connected to the inverter output to increase the voltage to the desired transmission voltage level. Transformers in SPP's are connected on a permanent basis to the electric grid to guarantee the plant is fed with power to fulfil its auxiliary losses during non-generating mode (NGM). The energy cost of the coal power generation (EC_{CPG}) will be affected in this regard. Conversely, during SPPs generation mode (GM) the solar irradiation is available and the solar modules can harvest power. Only, the energy cost of the solar PV (EC_{PV}) in kWh will be affected.

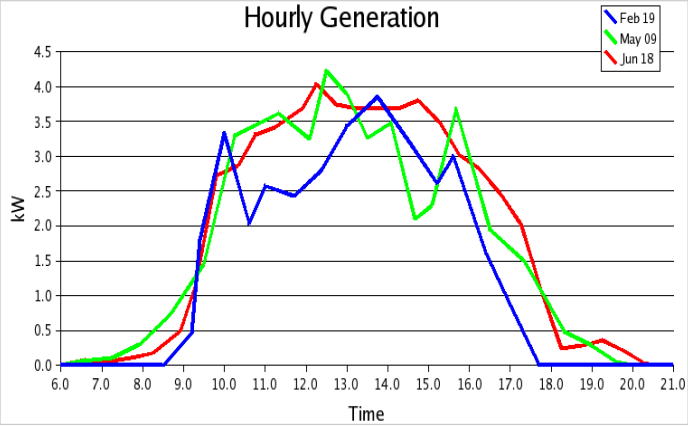


Figure 2: SPP generation profile in South Africa

Figure 2 demonstrate a solar PV generation profile of a SPP as acquired by field measurements and the data is collected over the period of five months. It should be noted that the SPP profile is reliant on the effective average daily solar radiation profile on the surface area of the solar PV modules and corresponding solar PV technology. As demonstrated in Figure 2, the GM of the SPP can be described by the affective area under the curve between the period 08:30AM to about 18:30PM

The generation modes of the SPP generation profile is then tabulated as shown in Table 2. A 24-hours generation state is considered in the analysis and about 14 hours the solar radiation will be available for the solar modules to harvest power (GM) while bout 10hours of that time the sun will be down (NGM). The ratios of GM and NGM against the 24-hours daily generation state result in the generating (FGM) and non-generation factors (FNGM)

It is reaffirmed that the proposed SLLM proposed in this work is relevant to Independent Photovoltaic Power Producers

Table 2: Generation modes: solar power plant

Generating State	**24 hours**
GM (Hours)	14
NGM (Hours)	10
FGM	0,583
FNGM	0,417

(I3P) supplying electricity to the electric grid by means of step-up transformer. The critical element in capitalizing the service losses in these transformers is the correct definition of the energy component of the cost of service no-load (ZAR/kW) and load losses (ZAR/kW). The energy component is the current energy tariff that will be utilised by a unit kilowatt (kW) of each service loss over the course of designed service lifetime of the SPP. In (14), the proposed method to evaluate the $CNLL$ for transformer intended to operate in SPP. Over the course of the day, the SPP has a greater likelihood to operate GM and NGM. Whenever, the SPP is operating in GM, then it is solely accountable to generate, solely supply of energy and service loss demands as well as transmission of energy to the electric grid. Conversely, during the course of NGM, the SPP purchase energy in a South African context from a coal power generation utility in order to supply its auxiliary requirements and losses. The latter is expressed by (14).

$$CNLL(ZAR/kW) = CGTD + \frac{8760 \times \text{NGM} \times EC_{CPG}}{FCR} \quad (14)$$

In this case, the energy tariff of the coal power generation utility (EC_{CPG}) applies. Moreover, the annual operating hours are weighted in accordance with the availability of the solar radiation. In (15), the case of the evaluating for the cost of service load losses for transformers intended for SPP is also demonstrated by improving (5) from the previous section. The energy tariff of the SPP (EC_{PV}) and corresponding generating factor is considered when the plant is generating its power.

$$CLL(ZAR/kW) = CGTD \times PRF^2 \times EPR^2 + \frac{8760 \times \text{GM} \times EC_{PV}}{FCR} + \frac{8760 \times \text{NGM} \times EC_{CPG}}{FCR} \quad (15)$$

Additionally, NGM component is added into the formulation to take into account for the when the transformer is operated in reverse as a step-down transformer to supply the auxiliary and loss requirements of the SSP from the grid supply.

3. Case Study: 1500KVA oil-immersed transformer

In this section, the data in Table 3 is considered to conduct the service lifetime loss evaluation. The data is intended to make a substantiated decision to purchase a 1500kVA oil-immersed transformer with 65 ℃ mean winding temperature. Transformer designs from three distinct manufactures will be considered in this case study. Each design attempt to attain coveted specifications with cost-effective materials, minimum dimensions and lowest overall cost of transformer ownership. In South Africa, increasingly in South Africa, utility owners now purchasing transformers on the basis of the transformer TOC over its intended service lifetime, than just the purchasing price.

Table 3: Service lifetime parameters

Parameter	**Value**
Designed Service Lifetime	30 years
No. of service hours per year	8760
Annualized Energy Cost	0,61 ZAR/kW
Fixed Charge Rate (FCR)	0,192
System Investment (SI)	23000 ZAR/kW
Discount Rate (*i*)	10,00%
Energy Escalation rate (*i*)	2,00%
Load Factor (LF)	0,6
TLF	0,3984
EPR	0,97
Peak Responsibility Factor (PRF)	0,6
Peak Ratio (PR)	0,7
Load Growth Rate (LGR)	2,00%
Coal	1,03 ZAR/kW

In Table 4, a design comparison (A, B and C) of the service losses, selling prices and respective cost of service losses and TOC is presented for three different competing transformer manufacturers. This information is useful in the jurisdiction of the more technically and economically suitable transformer. Design A is the highest loss design of the three offers. Design B and C are lower loss designs. Manufacturer B and C on that account have used better grades of core steel, optimized conductor dimensions and improved cooling medium and procedures

Table 4: Service lifetime loss analysis: classical method

Mn	NLL	LL	SP	CLL	TOC
A	6	25	R460,839	R970 966	R1,431,806
B	5	20	R510,215	R789 677	R1,299,892
C	5	16	R559,591	R696 260	R1,255,851

In Table 4, the classical TOC method has been employed to evaluate the SLLM. This method makes the assumption that the transformer will be in service for 8760 days annually over its designed service lifetime. This premise does not however hold true for renewable energy applications considering the spasmodic nature of renewable energy sources. In Table 5, the proposed solar power plant procedure for evaluating the SLLM is presented. This method takes into consideration the GM and NGM of the plant over the designed transformer service lifetime as described in the previous section. Design C is observed to have lower service losses in comparison to manufacturer B and C even though the selling price of the manufacturer is higher. It follows that the TOC of the unit is also lower over the normal operating service lifetime

Table 5: Service lifetime loss analysis: solar PV method

Mn	NLL	LL	SP	CLL	TOC
A	6	25	R460,839	R737,986	R1,198,825
B	5	20	R510,215	R597,308	R1,107,523
C	5	16	R559,591	R512,443	R1,072,033

The case studies in Table 4 and Table 5 illustrate that cost of the transformer service losses over the transformer designed lifetime are critical to evaluate over just the purchase price.

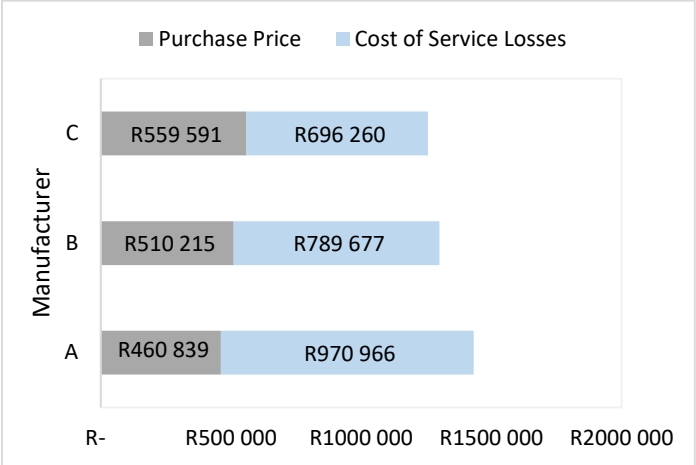


Figure 3: TOC of the transformer manufacturers: Classical method

The results in Figure 3 indicate that manufacturer C has the highest purchase price in comparison to manufacturer B and C by 9% and 18 % respectively. Additionally, the CSL for manufacturer C is lower by 13% and 39% respectively over the designed service lifetime. By employing the classical method, this results indicate that over the designed service lifetime of the transformer under study the IPP can purchase the most expensive transformer while keeping a relatively lower cost of service losses over the intended transformer lifetime. Alternatively, the IPP can choose the cheapest transformer offer with significantly higher service losses over the designed lifetime. However, from a maintenance perspective of the IPPP the later will not be economical.

Results of the proposed SLLM for the three transformer offers are presented in Figure 4. As indicated earlier, throughout the day the SPP will undoubtedly operate in two modes. When the plant is operating in GM, the SSP is accountable to supply its own energy and losses in addition to supplying power to the electric grid. When the SSP is operating on the NGM, then auxiliary requirements of the SSP will be supplied by the electric grid, that is to say that the SPP will purchase power from the coal power generation utility and their energy tariffs will apply.

Figure 4, the results indicate that manufacturer A has the cheapest purchase price and highest CSL. Manufacturer B is the second cheapest offer while manufacturer C is more expensive. The CSL is cheaper by 44% and 17% in comparison to manufacturer A and B respectively. In this regard, the most techno-economical purchase decision for the IPP will be that of manufacturer C.

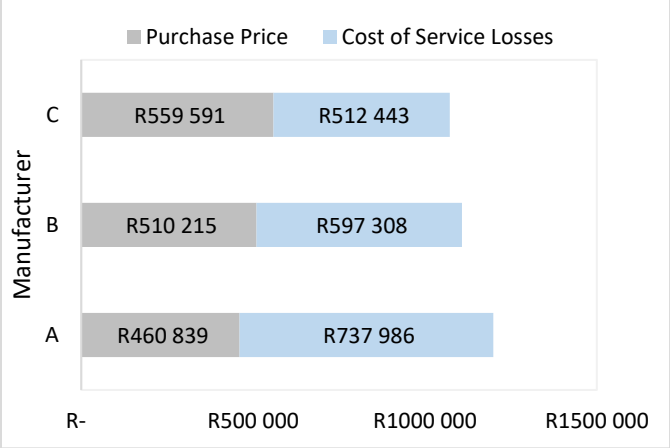


Figure 4: TOC of the transformer manufacturers: Solar power plants

In order to demonstrate the differences between the classical and proposed SPP SLLMs, Figure 5 is presented. It can be observed from Figure 3 and Figure 4 that the transformer offers from each respective method that manufacturer C has a significantly higher selling price but is more economical over the 30 years designed service lifetime.

Manufacturer B and C have managed to optimise their no-load losses by 17% in comparison to manufacturer A. in practice, this loss is minimised by four methods:

- Enhancement of the core steel construction
- Enhanced cutting of the core laminations and employing step-lapped joints as they interleave the core laminations that connect the joint gap.
- Enhancement of the core laminations stacking; and
- Computational modelling of the no-load losses.

The use of amorphous steel in the transformer design also plays a critical role and can minimise the no-load loss by 70-80% as opposed to silicon graded core steel.

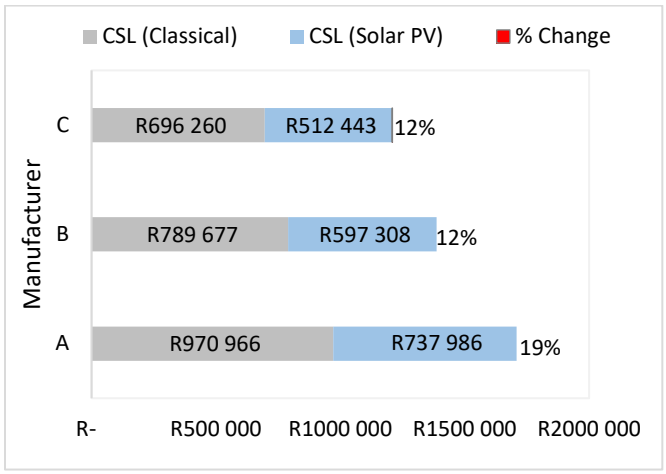


Figure 5: TOC of the transformer manufacturers: Comparison

Overall, manufacturer C has managed to optimise the service load losses by 56% and 25% in comparison to manufacturer A and B respectively. In practice, the service load losses is instrumental to the efficiency of the transformer during service. It therefore critical to reduce these losses for the purpose of enhancing and achieving the most optimal service performance. Reduction of the load loss is treated by the selection of large conductor dimensions in the interest of minimizing the resistance per unit length of winding conductors. A comparison of the CSL between manufacturer A and B is 12% while between manufacturer A and manufacturer C it is 19%.

The results presented in Figure 5, show some benefits of using the proposed method over the classical method. Taking into consideration the generation modes of the SPP and the ACOE has some techno-economic benefits for the IPP.

4. Conclusion

This work proposes a holistic SLLM to compute the TOC of transformers designed to be of service to solar power plants. These transformers are predominantly owned by IPP's in South Africa. More notably, the proposed method acknowledge precisely how the cost of the service no-load and load losses should be calculated, taking into consideration the TOC of the transformer in relation to the operational requirements of solar power plants they operate in. The distinct operational features of a SPP have been incorporated in the classical SLLM by means of the two generating modes (GM and NGM). Additionally, the proposed SLLM for the transformer under study is influenced by the reality that this transformer will exclusively facilitate the SSP, as distinguished from another generation utility. As a result the ACOE for SPP in South Africa is employed to evaluate the CSL that will be consumed by the transformer under study during service. At large, it is shown that the solar irradiation plays a pivotal role on the calculation of the CNLL and CLL. This is a property of the SPP that should be cautiously covered, as it may influence the tender jurisdiction criterion to choose the lowest TOC of a transformer over its designed service lifetime.

The insights gained in this work indicate that most economical transformer purchase choice for SPP owners is a high purchase price and low-cost transformer over low-purchase price and high-loss transformer.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Department of Energy, "Executive Summary of the Draft Integrated Electricity Resource Plan for South Africa - 2010 to 2030 IRP 2010," in Department of Energy, 2010, doi: executive-summary-draft-integrated-electricity-resource-plan-south-africa-2010-2030-irp

[2] Department of Energy, "Integrated Resource Plan 2019, in Department of Mineral Resources and Energy, 2019, doi: IRP/2019/IRP-2019.pdf.

[3] J. Calitz, C. Mushwana and T. Bischof-Niemz, "Statistics of utility-scale solar PV, wind and CSP in South Africa in 2016," in CSIR Energy Centre, 2017.

[4] K.S. Takayuki, K.T. Kobayashi, "How Transformers Age," in T&D World, 2019, doi: substations/article/20972255/how-transformers-age.

[5] A. Indarto, I. Garniwa, R. Setiabudy and C. Hudaya, "Total cost of ownership analysis of 60 MVA 150/120 kV power transformer," 2017 15th International Conference on Quality in Research (QiR) : International Symposium on Electrical and Computer Engineering, Nusa Dua, 291-295, 2017, doi: 10.1109/QIR.2017.8168499.

[6] A. L. Lazari and C. A. Charalambous, "Life-cycle loss evaluation of power transformers serving large photovoltaic plants in vertically integrated and decentralised systems," in IET Generation, Transmission & Distribution, **9** (8), 759-766, 2015, doi: 10.1049/iet-gtd.2014.0465.

[7] T. Knutson, "Conducting Distribution Transformer Evaluations Using the Total Ownership Cost Method," 2015 IEEE Rural Electric Power Conference, Asheville, NC, 97-101, 2015, doi: 10.1109/REPC.2015.14.

[8] C. A. Charalambous, A. Milidonis, S. Hirodontis and A. Lazari, "Loss Evaluation and Total Ownership Cost of Power Transformers—Part II: Application of Method and Numerical Results," in IEEE Transactions on Power Delivery, **28** (3), 1881-1889, 2013, doi: 10.1109/TPWRD.2013.2262507.

[9] A. Eliasson, H. Elvfing and V. R. Ramanan, "Amorphous Metal core material shows economic and environmental benefits when pre-existing transformers are to be replaced within Vattenfall Group´s distribution network," 2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), Gothenberg, 1-7, 2010, doi: 10.1109/ISGTEUROPE.2010.5638963.

[10] A. L. Lazari and C. A. Charalambous, "Probabilistic Total Ownership Cost of Power Transformers Serving Large-Scale Wind Plants in Liberalized Electricity Markets," in IEEE Transactions on Power Delivery, **30** (4), 1923-1930, 2015, doi: 10.1109/TPWRD.2014.2365832.

[11] IEEE Guide for Transformer Loss Measurement - Redline," in IEEE Std. C57.123-2019 (Revision of IEEE Std C57.123-2010) - Redline, 1-86, 2020, doi: 10.1109/IEEESTD.2010.5954117.

[12] J. Wojtkun, B. Bródka and D. Stachowiak, "The influence of core geometry on no-load losses of medium power transformers," 2018 International Interdisciplinary PhD Workshop (IIPhDW), Swinoujście, 123-127, 2018, doi: 10.1109/IIPHDW.2018.8388339.

[13] P. Tandeaw, C. Manop and N. Jirasuwankul, "Design and Testing of Low Loss Distribution Transformer for Non-Linear Loading Current," 2018 International Electrical Engineering Congress (iEECON), Krabi, Thailand, 1-4, 2018, doi: 10.1109/IEECON.2018.8712270.

[14] S. Taheri, H. Taheri, I. Fofana, H. Hemmatjou and A. Gholami, "Effect of power system harmonics on transformer loading capability and hot spot temperature," 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Montreal, QC, 1-4, 2012, doi: 10.1109/CCECE.2012.6334834.

[15] B.A Thango, D.B Nyandeni, P.M Molepo J.A Jordaan, A.F Nnachi, “Solar Power Plant Transformer Loss Computation under Harmonic Currents using Finite Element Method”, 9th CIGRE Southern Africa Regional Conference, 2019.

[16] DB Nyandeni, M Phoshoko (Pr. Eng.), R Murray, BA Thango, “Transformer Oil Degradation on PV Plants – A Case Study”, 8th South Africa regional conference, 14-17, 2017,

[17] Bonginkosi A. Thango, Jacobus A. Jordaan, Agha F. Nnachi, “Effects of Current Harmonics on Maximum Loading Capability for Solar Power Plant Transformers”, 2020 International SAUPEC/RobMech/PRASA Conference, 1-5, 2020, doi: 10.1109/SAUPEC/RobMech/PRASA48453.2020.9041101.

[18] A. Elmoudi, M. Lehtonen, H. Nordman, "Effect of harmonic on transformers loss of life", in: Conference Record of IEEE International Symposium on Electrical Insulation, **2**, 408-411, 2006.

[19] R. Singh and A. Singh, "Aging of distribution transformers due to harmonics," Proceedings of 14th International Conference on Harmonics and Quality of Power - ICHQP 2010, Bergamo, 1-8, 2010.

[20] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Selection and Rating of the Step-up Transformer for Renewable Energy Application," in SAIEE Africa Research Journal, **111**(2), 50-55, 2020, doi: 10.23919/SAIEE.2020.9099492.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 23-31 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Cyber Incident Handling and the Perceptions of Learners on Cyber Incidents in South African Schools

Naume Sonhera[*,1], Elmarie Kritzinger[2], Marianne Loock[2]

[1]*Information and Communication Technology, Vaal University of Technology, Vanderbijlpark, 1911, South Africa*

[2]*School of computing, College of Science, Engineering & Technology, University of South Africa, Florida, 1709, South Africa*

ARTICLE INFO

Article history:
Received: 21 April, 2021
Accepted: 27 August, 2021
Online: 10 September, 2021

Keywords:
Aggressor
Cyber Incident
Learner
Role Player
Schools
South Africa

ABSTRACT

With increases in technological usage, cyber incidents are also on the rise and have become a major concern in schools across the globe. What is of significant concern is that cyber incidents in South African schools are also on the rise. Existing evidence suggests that, in South Africa there are no clear procedures that are consistently followed by schools on how to report cyber incidents. The aim of this research is therefore to propose cyber incident handling procedures to enhance the effectiveness in handling cyber incidents as well as ensuring that each role player has an important contribution in the intervention process that is designed to reduce cyber incidents in South African schools. The study also assessed the perception of learners on cyber incidents in South Africa. Using the literature review approach and thematic analysis of the data collected from learners the study highlighted the procedures and roles of role players that can assist in cyber incident handling in South African schools. The study also came up with a detailed analysis of the views of learners on cyber incidents in South Africa. The results presented can help to provide a framework that will act as a guide on reporting cyber incidents and directing school management, and all within the school, towards appropriate reporting procedures and intervention processes. The study also found out that the rise in cyber incidents in South African schools, if left unaddressed, can have a devastating effect on learners. Therefore, the government of South Africa, through the Department of Basic Education, must prioritize the handling of cyber incidents in schools as cyber incidents are now a threat to the efficient and effective execution of the mandate of the department.

1. Introduction

The 21st century came with changes in almost all the sectors of the world economies. In the education sector, both the academics and practitioners of the 21st century have applauded the paradigm shift in schools of encouraging learners to be computer literate [1-4]. The technological paradigm shift has been complemented by the efforts of schools through promoting access to technology in education [5-6]. The socialization patterns among learners have also changed drastically because of the growth and proliferation of electronic communication devices. Learners access social network sites to communicate with their friends at any time of the day or night [7]. However, access to new technology has led to an increase in misuse and abuse of technology by learners and this has brought about incidents of threatening, harassing, embarrassing, and humiliating behaviour and actions online [8]. Cyber incidents have consequently become prevalent in schools worldwide, and it is evident that they present certain unique problems for schools in the regulation thereof [9-12]. What is of significant concern is the fact that cyber incidents in South African schools are also on the rise [13].

South African schools are adopting Information and Communication Technologies (ICTs) as part of an educational and social approach to prepare learners for the future. However, there is a lack of supporting cyber incident handling procedures to empower schools to deal with cyber incidents [14]. Cyber incidents are increasing among learners and yet there is very little guidance for schools on how to deal with cyber incidents [15]. With the potential for cyber incidents and their negative effects, schools have the responsibility to intervene in the occurrences [16]. Cyber incidents do not have an impact only on the physical and mental health of learners but also on the academic performance which in turn influences learners' learning outcomes negatively.

[*]Corresponding Author: Naume Sonhera, Email: nqume@vut.ac.za

Therefore, there is a need to ensure that schools are provided with cyber incident handling procedures that can contribute to a safe and caring environment for learners and all involved role players. Far from being limited to cyber victims, the effects of cyber incidents extend to the learners collectively, the school environment, and the entire school system [17]. To address the elements of cyber incidents in South African schools and to narrow the focus of this research to an educational environment, a cyber incident has been defined as: An act that violates explicit or implied cyber safety, computer safety or acceptable use of ICTs policies of South African schools or the Department of Basic Education (DBE) [18]. It is an event that involves the intentional or negligent use of digital platforms or electronic media to cause harm to other learners/self, or to impact on their confidentiality and integrity.

Some studies were done to try and understand the critical issue of cyber safety. In [14], the author investigated the maturity levels of cyber safety in South African schools. The study was intended to assess whether schools are prepared to help learners and educators to come up with cyber safety culture within the school environment. The study involved 24 schools in South Africa, and the investigation was through a critical evaluation of the elements that are important in improving cyber safety in schools. The elements in [14] which were evaluated include leadership and policies, infrastructure, education and standards, and inspection. After measuring cyber safety successfully using one of the UK approved measurement tool, the 360safe tool, the study discovered that all the schools had low levels of cyber safety maturity and compliance. In [14], the author further discovered that there is a clear lack of cyber safety awareness policies, practices, and procedures in the schools in South Africa. The study went on to propose a step-by-step cyber safety plan to empower schools to create their own cyber safety culture.

In [19], the authors stated that cyber safety is one of the critical aspects in the world today. Their study highlighted that learners require education on how to operate safely in cyber space and to protect themselves. In [19], the authors developed a cyber safety curriculum through open educational resources. The authors believed that this curriculum can empower educators in schools when educating learners about cyber safety. In [20], the author argued that business schools can meet the multidisciplinary needs of training cyber safety professionals. Motivated by these findings the study aims at proposing cyber incident handling procedures to enhance the effectiveness in handling cyber incidents as well as ensuring that each role player has an important contribution in the intervention process that is designed to reduce cyber incidents in South African schools. The study also assesses the views of learners on cyber incidents in schools.

2. The Education Sector in South Africa

The education sector in South Africa is divided into two departments, the Department of Basic Education (DBE) and the Department of Higher Education and Training (DHET) [21-22]. The DBE is responsible for primary and secondary education while the DHET is responsible directly for tertiary education and vocational training [23]. These two departments were once represented as a single department of education before 2009, then there was a split. The DBE looks after the public schools, the private schools, Early Childhood Development centres (ECD) and the special needs schools. Private schools are also known as independent schools and these private schools and public schools are known collectively as the ordinary schools [22]. On the other hand, the DHET deals with tertiary education where we have further education and training, popularly known as the Technical and Vocational Education Training colleges (TVET) [21]. The DHET is also responsible for Adult Basic Education and Training (ABET) Centres and Higher Education Institutions (HEI) [24]. In South Africa, each province has its education department which is responsible for the implementation of policies of the national department. These departments are also there to deal with local issues that are related to the execution of the mandate of the department at national level.

As of 2016, the DHE reported that there were 29 749 public and registered independent schools in South Africa [23]. The number of ordinary schools was 25 574 while 4 175 were other educational institutions such as special schools and Early Childhood Development (ECD) centres. The DBE also reported that the total number of educators were 425 000 in 2012 and 440 151 in 2016 who were reporting to the national department and the provincial department in the 9 provinces and the 86 districts of South Africa [22]. District and provincial DBE offices in nine provinces and 86 districts administer all the schools and have considerable influence over the implementation of policy. The goals of the DBE and its district and provincial offices are to improve the quality of teaching, undertake regular assessments, and improve on the Early Childhood Development, to ensure a system of outcomes-focused accountability [23].

3. Literature Review

Schools worldwide are being charged to take action to reduce cyber incidents because of the potential of cyber incidents disrupting the educational processes, creating a hostile environment, and threatening learners' feelings of safety and mental well-being [8, 12, 25]. Due to the increase of cyber incidents that cross over into both home and school environments, coupled with confusion as to who should issue consequences and how, role players must work together to create effective cyber aggression strategies [16]. The intervention during cyber incidents takes group effort from all role players and learners themselves. While it may take a village to raise a child, the virtual village inclusive of social media platforms, Internet Service Providers, phone companies, non-profit youth organisations, non-governmental organizations, educational institutions, parents or guardians, attorneys, and media need to join forces to have dialogue nationally and globally on ways to protect learners from cyber incidents [26].

The 21st century has brought with it a different revolution - learners who are on the cutting edge of technological proficiency [7]. While awareness of the use of the Internet is growing among learners in South African schools, there is no increase in awareness of safe practices in the use of ICTs [27]. Learners receive mixed messages with regard to online behaviour as they strive for technology literacy, sometimes without appropriate support [28]. Learners seem unaware of the risks of inappropriate behaviour online, viewing them as trivial. Cyber incidents are causing major challenges for school officials who are called upon to respond to

cyber incidents involving learners [29]. Online threats take place off the radar screen of educators and parents; this makes it difficult to detect cyber incidents in schools and more impossible to monitor off school premises [10]. The language that the learners use on the Internet has evolved and created a generation gap between learners and adults around them. This gap enables learners to participate in cyber incidents without fear of being discovered by adults [15, 30].

There are no clear procedures that are followed consistently by South African schools, governing boards, and educators on how to report or handle cyber incidents [10, 31]. At present, the principals, educators, and parents are not sure what procedures to follow when learners are being harassed online [32-33]. Research in South Africa (SA) that examines the procedures for schools on how to report cyber incidents are limited [10]. The lack of clear procedures for reporting cyber incidents in schools, makes educators feel unsupported. Educators will rather ignore these unethical violations than follow ill-defined and unenforceable policies [34]. In [34], the authors investigated the reason learners should be educated about risks in cyber space and the ways in which the schools can intervene to assist. The study discovered that despite the usefulness of the Internet there are still negative issues that are related directly to the Internet use. Users of the Internet across the world are being subjected to various forms of abuse such as cyber incidents, cyber bullying, online fraud, racial abuse, pornography, and gambling. In [34], the authors believe that issues related to cyber incidents are on the rise, especially among learners, because of a lack of anti-cyber incident awareness. The other argument put forward by them was that the level of awareness among the Internet users is low or moderate. As a result, in [34], the authors stated that it is important that the anti-cyber incidents awareness be cultivated among learners through various initiatives like education so that learners are able to operate safely in cyber space.

In another study, in [35], the authors verified the effectiveness of gamification in reducing cyber incidents and offer a constructive application of a framework that can help to improve cyber safety skills and capabilities for learners in schools. The study also discovered that the rise in cyber incidents shows that the traditional methodologies used to train and bring awareness are insufficient to build the necessary cyber safety skills and capabilities. The study also discovered that gamification can be a promising methodology that can bring some changes to the behaviour of learners in cyber space at early stages. In [35], the authors came up with a proposal for a cyber hero framework for information security awareness and training programs that can fight human error in cyber space. In [35], the authors discovered massive progress in the skills and capabilities of cyber safety for learners using the cyber hero framework.

In [36], the authors discussed the applicability of the safety standards and frameworks used in various industries to avoid threats of cyber incidents due to the rise in the use of technology in schools. In [36], the authors discovered that it is very critical to have strong cyber safety controls in schools due to the growth in cyber incidents. The other notable finding in [36] was that it is very important to have a cyber safety framework that accommodates safety standards and frameworks in an education environment. In [15], the authors investigated the perceptions of cyberbullying among student teachers in South African schools, in Eastern Cape. In [15], the authors found out that in South Africa cyber incidents are some of the issues affecting learners. The study went on to discover that the Department of Basic Education in South Africa is providing little guidance for schools on issues related to cyber incidents. Using a quantitative survey approach, the study discovered that cyber incidents are some of the critical issues in South African schools, but the topic has not been incorporated in the policy or even in the curriculum. In [15], the authors recommended that the Department of Basic Education should come up with a standardised policy that can be used to implement and enforce cyber safety behaviour in schools.

In [37], the authors did an investigation to assess the policies, procedures and measures which schools need to have in place to ensure that there is cyber safety awareness for learners in South African schools. They discovered that Information Communication Technology is now becoming part of the daily lives of learners, opening various opportunities and cyber risks as well. Therefore, just like in [15], [37], the authors believe that it is important for cyber users to become aware of the cyber safety issues so that they are able to protect themselves. It was also highlighted that in South Africa, cyber safety awareness and cyber education is still lacking, especially in schools. In [37], the authors also believe that it is of utmost importance for South Africa to begin the creation and implementation of cyber safety and cyber incident awareness policies, procedures, measures, and initiatives for the education sector, especially for schools and learners.

4. Research Methodology and Design

The research methodology is explaining how the literature review was done and how data was collected from learners to get their experiences and perceptions on cyber incidents. This research consists of two research methods. The first one is the literature review. The researcher conducted a systematic search on literature that reported the prevalence and impact of cyber incidents among learners in South African schools and the extent to which role players responded to cyber incidents in schools. The main idea behind the reviews was to identify the roles and responsibilities of role players in cyber incident handling for the Schools in South Africa. The systematic process included the electronic search in databases, title and abstract review and hand searches in Google scholar and article reference lists. Online databases contained bibliographic references to academic and peer-reviewed journal articles, as well as references to theses, books, and chapters. Secondly, the research used a qualitative approach and purposive sampling to collect data from the learners to get their experiences and perceptions on cyber incidents [38-40]. The rationale for selecting learners was based on reviews in [41-44], with the authors who indicated that adolescence is a peak period for involvement in cyber incidents. A total of 85 learners participated in three focus groups that ranged in size: group one had twenty participants, group two had thirty participants and group three had thirty-five participants as shown in the table1.

Forty-eight percent were male, and fifty-two percent were female. The interview questions enabled learners to explain their perspectives in depth. The questions were related to the themes that emerged from each group. All the conversations were recorded and transcribed. The data was then coded using ATLAS.ti. 7; themes

emerged, and the frequencies of these themes were noted by the researcher. Several ethical guidelines were followed to ensure that the research study was conducted ethically. In an endeavor to ensure the reliability and validity of data, the verbatim transcribed interviews were presented to the respondents to verify and sign off. The researcher reviewed all the transcripts from focus group interviews to check for accuracy, ensuring that no obvious mistakes appeared. The researcher also ensured that there was no drift in the definition of codes or shift in the meaning of codes during the coding process.

Table 1: Participants in the Three Focus Groups.

Focus Groups	Participates in the Groups
Group 1	20
Group 2	30
Group 3	35
Total	85

Source: Author's Analysis

Table 1 above shows the number of students who participated in each focus group.

5. Discussion of the Roles and Responsibilities of Role Players in Cyber Incident handling for the Schools in South Africa

The study came up with a theoretical framework that outlines the roles and responsibilities of role players in addressing the cyber incident handling needs. The responsibilities of role players within schools and relevant role players outside schools for specific and specialized interventions and support were identified during the literature review. To elaborate on the "Needs for Addressing Cyber Incidents", each need is linked to role players and recommended responsibilities. The aim of the theoretical framework and the roles of role players for schools in South Africa is to ensure that relevant role players are drawn in to create a supportive structure that can contribute to the development of the cyber incident handling procedures. To achieve this, the researcher highlighted the roles and responsibilities of role players within and outside the school when responding to cyber incidents. The framework is designed to complement normal schooling duties and activities and to integrate cyber school safety into the daily activities of the school. The framework also gives a guide on how relevant role players can assist in responding to cyber incidents. The framework may not make provision for all cyber incidents in schools or cyber safety concerns for schools but is a starting point to address cyber incidents in South African schools. The linking of "needs for addressing cyber incident" to "the role players" assisted the researcher to present the complex ideas and the theoretical concepts, from the literature review to a new structure which will assist in developing reporting procedures [41]. Table 2 is outlining the roles and responsibilities of role players in cyber handling in South Africa as discovered from the reviewed literature.

Table 2: Summary of the Roles and Responsibilities of Role Players in Cyber handling in South Africa as discovered from the Reviewed Literature.

Role Players	Responsibilities
School Principals, Learners, Parents, Guardians, school staff members, and School Governing Body	Role players should be competent in handling cyber incidents.
	All role players should be informed of the cyber incident reporting procedures.
	School procedures should be readily accessible to learners, parents, guardians, and school staff members.
	Need clear and easily understandable school procedures to address cyber incidents.
	Consistent, universal, and effective ICT policy aimed directly at regulating cyber incidents in schools, should be developed.
	Codes of conduct or code of behaviour should be formulated and adopted.
	Need a referral system to appropriate services and community partnerships to support learners and build a cyber school safe environment.

Source: Author's Analysis

Full explanation of the roles and responsibilities of role players in cyber incident handling in South Africa as discovered from the reviewed literature is given in the following section.

5.1. Role Players should be Competent in Handling Cyber Incidents.

The principal should ensure that school personnel and parents are trained in handling cyber incidents. The advisory Team should implement an effective approach in intervening and responding to all cyber incidents. Educators should participate in school cyber safety trainings. Educators should be aware of the steps to take and advice to give if learners notify them of cyber incidents. Parents should report their children's cyber aggression. This was in line with the findings that were discovered in [15] and [37], the authors discovered that it is important for cyber users to become aware of the cyber safety issues so that they will be able to protect themselves from the various forms of cyber incidents. It was also highlighted that in South Africa, cyber safety awareness and cyber education is still lacking, especially in the education sector [15]. In [36], the authors argued that it is very important to have a cyber safety framework that accommodates safety standards and frameworks in an educational environment.

"All Role Players should be informed of the Cyber Incident Reporting Procedures"

The school community should know who to approach if they become aware of or suspect any cyber incident taking place. The advisory team should establish a process that communicates the cyber incident, reporting procedures to all parents, learners, and all staff members. The School Management Board (SMB), School Governing Body (SGB) and Counsellors should explain the cyber incident reporting procedures to learners and their parents or

guardians. Educators should establish consultation and training for parents and other staff members on cyber incident reporting procedures. Social Media Service Providers (SMSP) should make their educational and safety materials easy to find and promote the materials to users regularly. Social Media Service Providers should improve the visibility, consistency, and accessibility of reporting tools on their platforms. Several studies support this. In [15] and [31], the authors pointed out that in South Africa there are no clear procedures that are followed consistently by schools, governing boards, and educators, on how to report or handle cyber incidents. The authors believe that it is very important for role players to be aware of the reporting procedures because at the moment, the principals, educators, and parents are not sure of what procedures to follow when learners are being harassed online [10, 32]. In [35], the authors emphasized that role players should be informed of the cyber incident reporting procedures especially if these frameworks are to be usable.

5.2. School Procedures should be readily Accessible to Learners, Parents, Guardians, and School Staff Members.

Provincial education departments must ensure that all schools are trained about procedures to address cyber incidents. The principal, Information Technology (IT) Unit, SGB, SMB and the Advisory Team should develop, support, and evaluate procedures to address cyber incidents. Officials from the National Department of Education, in collaboration with provincial, regional and district officials responsible for ICT school safety, should monitor the implementation of procedures to ensure that learners are safe from cyber incidents. Educators should assist in the development and implementation of anti-cyber incident procedures according to delegated roles and responsibilities. This has been supported in [37] by the authors who argued that for South Africa to have meaningful progress towards addressing cyber incidents in schools, the departments of education must be equipped with the appropriate policies, procedures, and measures to ensure that there is cyber safety awareness for leaners in South African schools. In [37], the authors discovered that Information Communication Technology is now becoming part of the daily lives of learners, with various cyber opportunities and cyber risks as well, hence the need for proper policies, education, and training for those responsible for cybersecurity.

5.3. Need Clear and Easily Understandable school procedures to address cyber incidents.

The other important aspect is that procedures outlining the school response to cyber incidents and how to report cyber incidents, should be in place and these procedures should be easily understandable. Provincial Departments of Education should ensure that school cyber incident reporting procedures are in place and should be adhered to strictly in schools. School cyber incident reports must be submitted to the district after which reports must be consolidated and forwarded to the Provincial School Cyber Safety Coordinator. A consolidated report from provinces must be submitted every quarter to the national Department of Basic Education (DBE). This was also supported in [14], [15], and [37] by the authors who agreed with many other scholars. These authors believe that it is important for learners to be aware of the cyber safety issues and the cyber incident reporting procedures so that they can get help whenever they need it. It was also highlighted that in South Africa, cyber safety awareness and cyber incident handling education is still lacking, especially in schools [14, 15, 37]. Scholars also believe that it is of utmost importance for South Africa to begin the creation and implementation of cyber safety and anti-cyber incident awareness policies, procedures, measures, and initiatives for the education sector, schools and learners in the country [14, 15, 37].

5.4. Consistent, Universal, and Effective ICT Policy aimed at Regulating Cyber Aggression directly in schools, should be developed.

The other critical aspect is that the Department of Basic Education should take a lead to ensure that ICT policies for provinces have been developed. Provincial officials must provide direction to their regional or district counterparts to ensure that the development, enforcing and reviewing of ICT policies is being implemented in schools. Public policy should also be informed by ongoing research on the prevalence of cyber incidents. A school ICT policy that includes an ICT Acceptable Use Policy (AUP), ethical use of ICTs or digital technologies and anti-cyber incidents should be developed, implemented, and enforced. All role players should be aware of the contents of these policies. The principal, IT Unit, SGB, SMB and the Advisory Team should ensure that they have an appropriate ICT school policy in place to safeguard the overall cyber safety and wellbeing of learners. The IT Unit should manage the ICT policy through recording and monitoring Internet use for compliance with ICT policies. The SMB and the SGB should convene regular meetings to review and update the ICT Policy. Parents should ensure that their children adhere to the ICT policy.

5.5. Codes of Conduct or Code of Behaviour should be Formulated and Adopted.

The information gathered also reveals that the provincial education departments should ensure that all schools have a Cyber Safety Code of Conduct for learners. The principal, SGB, SMB and the Advisory Team should develop, support, and evaluate the Cyber Safety Codes of Conduct for learners. Learners should adhere to the Code of Conduct and to ensure appropriate behaviour always. All schools should be obliged to comply with Child Protection legislation, which includes Internet safety for learners. DBE needs to develop and integrate cyber safety programs into existing curricula to support human rights and child protection. The Child Protection and Abuse Organization should be contacted for help if a cyber incident is a suspected child protection issue.

5.6. Need a Referral System for appropriate Services and Community Partnerships to support Learners and build a Cyber School Safe Environment.

The South African Police Service (SAPS) should be obliged to investigate any unlawful cyber incident in schools. SAPS should advise and educate learners, parents, and school personnel about legal responses to serious cyber incidents. Law enforcement authorities must ensure that the all-role players have a clear awareness and understanding of how existing criminal offences can be applied to cyber aggression behaviour. Learners should contact SAPS in case of any cyber incident. Counsellors should give immediate support to all affected learners. Parents should

partner with schools to ensure that their children are following appropriate guidelines for online behaviour. Parents of the victims should be able to work with the Internet Service Provider, Cell Phone Service Provider, or Content Provider to investigate the cyber incident or to remove the offending material. Internet Service Providers should track instant messaging which could have been used as evidence in a court of law. SA Depression and Anxiety Group (SADAG) should assist learners suffering emotionally because of the cyber incident ordeal. Educational Organizations should continue to hold workshops, do research, and give presentations on cyber incidents and cyber safety topics to help the school communities. The attorneys should be able to provide parents with sound legal advice on how to open a possible criminal case and how to get restraining orders against the accused.

5.7. Findings of the Focus Group Interview

The interviews revealed learners' perspectives about cyber incidents and were specific to cyber incidents as they manifested within the schools. The results indicated the negative cyber incident effects on victims, the no or ineffective consequences for aggressors, and the inability to stop or control cyber incidents once they have started. The themes which emerged are Learners' Perceptions about Victims, Aggressors and Bystanders; Learners' Perceptions about Technology; Learners' Perceptions about Cyber Incidents, and Learners' Perceptions about Adults. A collection of the conceptual valid and reliable evidence contributed to a better understanding of learners' interactions in cyber space and how much help they need when they are in trouble. The results of the focus group interview also showed that learners are acutely aware of cyber incidents occurring in their schools and how they are reluctant to report the experiences of cyber incidents when they occur. The researcher recognized that cyber aggression, especially if left unaddressed, can have a devastating effect on learners; it can create a barrier to learning and have serious consequences for mental wellbeing. The views of learners are summarized in the table below.

Table 3: The Findings from the Focus Groups on the Views of Learners about Cyber Incidents

Findings From Focus Groups	**Description**
	Learners' Perceptions about Victims, Aggressors and Bystander
	Learners' Perceptions about Technology
	Learners' Perceptions about Cyber incidents
	Learners' Perceptions about Adults

Source: Author's Analysis

The table above is summarizing the views of learners on issues related to cyber incidents and the full description of these views is given below.

5.8. Learners' Perceptions about Victims, Aggressors and Bystanders

The themes generated from the focus group interviews were linked to the suggestions from learners and from the literature reviewed. Pillars for reporting procedures were derived from the whole integration process. Themes from Learners' Focus Group Interviews indicated the Challenges for Victims, Aggressors' Behaviour, and Learners' Perceptions about Bystanders Supporting Aggressors. The results revealed that the challenges for cyber victims are that cyber victims are humiliated and harassed online. Also, learners were of the view that victims fear aggressors and are affected psychologically which will lead them to suffer in silence. On the behaviour of aggressors, the results revealed that aggressors gain power by controlling victims and most of them are mean; they seek attention, are insecure, and they take out their pain or anger on other learners. The results also revealed that technology allows the aggressors to hide behind screens and in many circumstances, they are perceived as mentally unstable. The study went on to check the learner's perception about bystanders supporting aggressors. The study found out that usually bystanders are afraid to become the next victims and to suffer rejection by peers and sometimes uncertainty about whom to tell.

In summary, the results pointed to the fact that cyber victims feel like outcasts, are hurt, offended, embarrassed and feel inferior. These victims fear aggressors because aggressors take out their pain or anger on other learners. The results also pointed out that bystanders support aggressors because of fear of becoming the next victims, being rejected by peers or uncertain of whom to report to. The identified pillars for reporting procedures derived from the discussions are Legislation, Referral System, Cyber Incident Reporting Platform and Cyber Awareness. Learners suggested that all schools should be compelled to comply with Child Protection Legislation. There is a need for a referral system to appropriate services, and need for community partnerships to support learners. Learners should report cyber incidents to the principal or ICT personnel, school staff members should provide support and counselling, school psychologists should help affected learners, school assemblies and public service announcements should be used for awareness. Learners need to be educated about cyber incident problems, there should be a hotline for reporting cyber incidents and finally there must be consequences for the aggressors.

5.9. Learners' Perceptions about Technology

Leaners are of the view that technology aggravates cyber incidents, and, in many ways, technology has decreased empathy among learners. The results indicated that aggressors use profiles of innocent people to threaten other learners. The anonymity online makes threatening behaviour more readily achievable and protects aggressors. The other aspect that was highlighted is that technology enables aggressors to have the biggest audience, making it difficult for those who want to champion anti-cyber incident awareness. Learners also indicated that due to technology some learners put themselves at risk by taking naked videos of themselves and these videos end up going viral on social media. In a nutshell, the results from the interviews have revealed that technology has decreased empathy for victims due to space differences, instant gratification and communication with the biggest audience. As a result, learners are of the view that schools need to have a clear ICT policy, procedures, and codes of conduct in place to regulate cyber incidents among learners. Parents and learners should be given the sets of ICTs rules so as to guide learners against cyber incidents. Learners are also of the view that

schools should introduce education against cyber incidents and independent service providers should assist to solve cyber incident problems, for example, the network companies should be able to identify phones that could have sent the harassing messages. It was also highlighted that learners should stop taking photos which could later be used against them, they should avoid instigating cyber incidents. Lastly, it was discovered that the ICT Policy, School Codes of Conduct, Identification of Aggressors and Anti-Cyber Aggression Awareness were the pillars contributing to the development of reporting procedures for cyber incidents in schools.

5.10. Learners' Perceptions about Cyber Incidents

The results indicated that cyber incidents among peers could be perceived as harmless or humorous and with no ways of stopping cyber incidents once they are viral. Learners are of the view that cyber incidents are not reported as often as they should. At times learners do not have the courage to talk about the incidents; in most cases victims are only willing to report to someone they feel comfortable with or someone they can trust. The other important observation is that in most cases it is not known how these issues are solved or what the outcomes are. Aggressors perpetrate incidents to seek revenge and sometimes a victim may become an aggressor due to revenge. The other way learners manage cyber incidents is physical fighting. The results also pointed out that to facilitate learner confidentiality and trust towards adults, learners must be able to report cyber incidents anonymously and their confidentiality must be respected. Learners' voices are essential, and their positive suggestions should be included in the design and implementation of the cyber incident handling procedures. However, without a confidential, voluntary, and anonymous reporting system, it is unlikely that learners will feel free to report cyber incidents. Learners fear being singled out; they fear that someone might retaliate against them. The results indicated that the Reporting System and Management of Cyber Incidents were the pillars for reporting procedures.

5.11. Learners' Perceptions about Adults

The study also assessed the leaner's perception about adults. The results indicated that in most cases learners are scared that adults will worsen the situation because of them being too judgmental about the behaviour of learners. Many parents are seen to always remind their children of the issues which they would have done wrong. Leaners also believe that adults do not have any role in the intervention of cyber incidents and are ineffective when they attempt to intervene. Leaners are also of the view that educators are not helping when it comes to cyber incidents to an extent that some learners are not comfortable telling them about their cyber incident problems. This is why some learners end up being absent from classes, avoiding being embarrassed. All these results point to the fact that there is no open and good relationship between learners and the adults around them. The results indicated that learners wish their parents could be more involved in their online lives from an early age. Most learners want parents to have a good relationship with them, to communicate with them, to spend more time with them, to support them and even to join them in cyber space. Learners believe that the most important issues that many parents are interested in are about schoolwork. As a result, many parents do not know what is happening in their children's daily lives. The other critical factor is the huge generation gap between children and their parents; there is a great difference between how children are being raised and how their parents were raised. As a result, this gives learners the idea that adults are not confident to help and intervene during cyber incidents. Some learners pointed out that it is not known how principals are handling cyber incident cases and the outcomes are not known. This is due to adults being absent from cyber space, which compromises their ability to assist. Leaners believe that the fact that most parents are not using social media, makes it very difficult for them to identify cyber aggressors.

In summary, learners indicated that they do not have confidence in adults, especially when it comes to their online problems. Role players need to be competent in handling cyber incidents and learners need to be assured of their safety when reporting cyber incidents, and their reports should be taken seriously and acted upon. All role players should be informed of the cyber incident reporting procedures and should be aware of how important it is to report cases as early as possible. The school community should know whom to approach if they become aware of or suspect any cyber incident taking place. Parents should check their children's phones if they notice any behaviour change. Parents should block the aggressor, notify their service provider, and report cyber incidents to the school and law enforcement. Parents should tell their children not to go about taking naked pictures of themselves and posting them online. All school role players should be educated about cyber incidents. Parents need to learn the legal issues associated with cyber aggression. Severe cyber incidents should be considered as crime. Law enforcement should have a specialised task team to deal with cyber incidents. SAPS should have a cyber division. The pillars for reporting procedures derived from the themes are Competence in Handling Cyber Incidents, Awareness of The Reporting Procedures, Learner Assistance from Parents, Cyber Aggression Awareness, Legislation, And Cyber Division for Law Enforcement.

6. Summary and Conclusion

Access to new technology in schools has led to an increase in misuse and abuse of technology by learners and this has led to a rise in cyber incidents related to threats, harassment, embarrassments and humiliating behaviour, and various bad activities online. The study, therefore, proposes that cyber incident handling procedures, and the roles and responsibilities of role players should be in place for cyber incidents to be addressed in South African schools. Using the literature review approach and thematic analysis of the data collected, the results indicated that it is important that role players should be competent in handling cyber incidents, and they should be informed of the cyber incident reporting procedures. The other important finding was that the cyber incident handling procedures should be readily accessible to learners, parents, guardians, and school staff members and they should be clear and easily understandable. The other issue that was highlighted was the need for a consistent, universal, and effective ICT policy directly aimed at regulating cyber incidents in schools. The codes of conduct or a code of behaviour should be formulated and adopted. The study also discovered that there should be a need for a referral system to appropriate services, and community partnerships to support learners and build a cyber school safe environment.

The study also come up with the views of learners on various aspects of cyber incidents in schools which include their perceptions on victims, aggressors and bystanders; learners' perceptions about technology; learners' perceptions about cyber incidents and learners' perceptions about adults. For instance, learners' perceptions on cyber victims, aggressors and bystanders were that cyber victims feel like outcasts, hurt, offended, embarrassed and inferior. These victims fear aggressors because aggressors take out their pain or anger on other learners. The results also pointed out that bystanders support aggressors because of fear of becoming the next victims, being rejected by peers or uncertain of whom to report to. The results presented helps to provide a framework that will act as a guide on reporting cyber incidents and directing school management, and all within the school, towards appropriate reporting procedures and intervention processes. The study also found out that the rise in cyber incidents in South Africa if they are left unattended, can have a devastating effect on learners. Therefore, the government of South Africa through the department of basic education must prioritize the handling of cyber incidents in schools as cyber incidents are now a threat to the efficient and effective execution of the mandate of the department.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] E. B. Meier, "Designing and using digital platforms for 21st century learning," Educational Technology Research and Development, **2021**(1), 1–4, 2021, doi.org/10.1007/s11423-020-09880-4.

[2] M. Finn-Stevenson, Schools of the 21st century: Linking child care and education, Routledge, 2018.

[3] S. McCoy and S. Lyons, Digital Technologies and Student Learning. Ireland's yearbook of education 2018-2019, Education Matters, 2018.

[4] J. Remón, V. Sebastián, E. Romero, and J. Arauzo, "Effect of using smartphones as clickers and tablets as digital whiteboards on students' engagement and learning.," Active Learning in Higher Education, **18**(2), 173–187, 2017, doi.org/10.1177/1469787417707618.

[5] L. K. Butola, "E-learning-A New Trend of Learning in 21st Century During Covid-19 Pandemic.," Indian Journal of Forensic Medicine & Toxicology, **15**(1), 422–426, 2021.

[6] J. Byrne, D. Kardefelt-Winther, S. Livingstone, and M. Stoilova, Global Kids Online Research Synthesis, 2015-2016, London, 2016.

[7] I. . Ogaji, P. Okoyeukwu, I. Wanjiku, E. A. Osiro, and D. Ogutu, "Pattern of use of social media networking by Pharmacy students of Kenyatta university, Nairobi, Kenya," Computers in Human Behavior, **100**(66), 211–216, 2017, doi.org/10.1016/j.chb.2016.09.035.

[8] Queensland Government, "Adjust our settings: A community approach to address cyberbullying among children and young people in Queensland," Queensland, 2018.

[9] J. D. Shapka, H. Z. Onditi, R. J. Collie, and N. Lapidot-Lefler, "Cyberbullying and cybervictimization within a cross-cultural context: A study of Canadian and Tanzanian adolescents.," Child Development, **89**(1), 89–99, 2018, doi.org/10.1111/cdev.12829.

[10] C. A. Hills, Developing a law and policy framework to regulate cyber bullying in South African schools, PhD Thesis, University of South Africa, 2017.

[11] J. M. Lee, J. S. Hong, J. Yoon, A. A. Perguero, and H. J. Seok, "Correlates of adolescent cyberbullying in South Korea in multiple contexts: A review of the literature and implications for research and school practice.," Deviant Behavior, **39**(3), 293–308, 2017, doi.org/10.1080/01639625.2016.1269568.

[12] D. Cross, T. Shaw, K. Hadwen, P. Cardoso, P. Slee, C. Roberts, L. Thomas, and A. Barnes, "Longitudinal impact of the cyber friendly schools program on adolescents' cyberbullying behavior," Aggressive Beviour, **42**(2), 166–180, 2016, doi.org/10.1002/ab.21609.

[13] P. Burton, L. Leoschut, and J. Phyfer, "South African Kids Online: A glimpse into children's internet use and online activities, Cape Town," http://www.cjcp.org.za/uploads/2/7/8/4/27845461/south_africa_kids_online_full_report.pdf, Accessed: 17-Nov-2017.

[14] E. Kritzinger, "Improving cybersafety maturity of South African schools," Information, **11**(10), 471–488, 2020, doi.org/10.3390/info11100471.

[15] L. Cilliers and W. Chinyamurindi, "Perceptions of cyber bullying in primary and secondary schools among student teachers in the Eastern Cape Province of South Africa.," Electronic Journal of Information Systems in Developing Countries, **86**(4), 1–10, 2020, doi.org/10.1002/isd2.12131.

[16] S. G. Denby, Cyberbullying: School Administrators' Perceptions of Law and Prevalence, and Roles in Prevention, Intervention, and Discpline., PhD Thesis, Virginia Commonwealth University, 2020, doi.org/10.25772/NQZM-FM42.

[17] M. Astatke, C. Weng, and S. Chen, "A literature review of the effects of social networking sites on secondary school students' academic achievement.," Interactive learning Environments, **2021**, 1–17, 2021.

[18] L. M. Hellsten, An Introduction to Cyberbullying, Macerata, 2017.

[19] R. Von Solms and S. Von Solms, "Cyber safety education in developing countries," Journal of Systemics, Cybermetics Informatics, **13**(2), 14–19, 2015.

[20] S. C. Yang, "A meta-model of cybersecurity curriculums: Assessing cybersecurity curricular frameworks for business schools.," Journal of Education for Business, **96**(2), 99–110, 2021, doi.org/10.1080/08832323.2020.1757594.

[21] D. Mhlanga, "Industry 4.0: The Challenges Associated with The Digital Transformation of Education in South Africa.," https://www.researchgate.net/publication/344230555_Industry_40_The_Challenges_Associated_with_The_Digital_Transformation_of_Education_in_South_Africa, Accessed: 02-Aug-2021.

[22] South African Market Access, "Education Statistics - South African Market Insights," https://www.southafricanmi.com/education-statistics.html, Accessed: 02-Aug-2021.

[23] Department of Basic Education, Education Statistics in South Africa 2016. Department of Basic Education, 2018.

[24] D. Mhlanga and T. Moloi, "COVID-19 and the Digital Transformation of Education: What are we learning on 4IR in South Africa?," Educatin Sciences, **10**(7), 180, 2020, doi.org/10.3390/educsci10070180.

[25] A. DeSmet, N. Aelterman, S. Bastiaensens, K. Van Cleemput, K. Poels, H. Vandebosch, G. Cardon, and I. De Bourdeaudhuij, "Secondary school educators' perceptions and practices in handling cyberbullying among adolescents : a cluster analysis," Computers and Education, **88**, 192–201, 2015, doi.org/10.1016/j.compedu.2015.05.006.

[26] J. L. Tinstman Jones, L. O. Campbell, J. Stickl Haugen, and C. C. Sutter, "Cyberbullying considerations for school counselors: A social media content analysis," Professional School Counseling, **23**(1), 2020, doi.org/10.1177/2156759X20919365.

[27] C. A. Myers and H. Cowie, "Cyberbullying across the lifespan of education: Issues and interventions from school to university," International Journal of Environmental Research and Public Health, **16**(7), 1–14, 2019, doi.org/10.3390/ijerph16071217.

[28] P. Redmond, J. V. Lock, and V. Smart, "Developing a cyberbullying conceptual framework for educators.," Technology in Society, **60**, 101223, 2020, doi.org/10.1016/j.techsoc.2019.101223.

[29] C. Angus, "Cyberbullying of Children: E-Brief," Australia, 2016.

[30] D. H. Tustin, G. N. Zulu, and A. Basson, "Bullying among secondary school learners in South Africa with specific emphasis on cyber bullying; South African Professional Society on the Abuse of Children: ISSN 1562-1383 Child Abuse Research," A South African Journal 2014, **15**(2), 13–25, 2014.

[31] M. Bulger, P. Burton, B. O'Neill, and E. Staksrud, "Where policy and practice collide: Comparing United States, South African and European Union approaches to protecting children online," new media & society, http://www.cjcp.org.za/uploads/2/7/8/4/27845461/protecting_children_online.pdf, Accessed: 03-Mar-2018.

[32] P. Burton, L. Leoschut, and J. Phyfer, "South African kids online: barriers, opportunities, and risks. A glimpse into South African children's internet use and online activities. Techical Report.," Cape Town, 2016.

[33] J. Phyfer, P. Burton, and L. Leoschut, "South African Kids Online.," Cape Town, 2017.

[34] N. A. A. Rahman, I. H. Sairi, N. A. M. Zizi, and F. Khalid, "The importance of cybersecurity education in school," International Journal of Information and Education Technology, **10**(5), 378–382, 2020, doi:10.18178/ijiet.2020.10.5.1393.

[35] H. Qusa and J. Tarazi, "Cyber-Hero: A Gamification Framework for Cyber Security Awareness for High Schools Students," in 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), 2021, pp. 0677–0682, doi:10.1109/CCWC51732.2021.9375847.

[36] M. Torres and N. Thompson, "Toward a Cyber Security Adoption Framework for Primary and Secondary Education Providers," in 31st Australasian Conference on Information Systems, 2020.
[37] D. Scholtz, E. Kritzinger, and A. Botha, "Cyber Safety Awareness Framework for South African Schools to Enhance Cyber Safety Awareness," in Computer Science Online Conference, 216–223, 2020, doi:10.1007/978-3-030-51974-2_19.
[38] A. S. De Vos, H. Strydom, and C. B. Fouche, Research at Grassroots for the Social Sciences andHuman Service Professions, Van Schaiks, 2009.
[39] J. H. McMillan and S. Schumacher, Research in education: A conceptual introduction, Longman, 2001.
[40] S. Hymel and S. M. Swearer, "Four decades of research on school bullying: An introduction.," American Psychologist, **70**(4), 293–299, 2015, doi.org/10.1037/a0038928.
[41] D. Adom, E. K. Hussein, and J. A. Agyem, "Theoretical and conceptual framework: mandatory ingredients of a quality research," International Journal of Scientific Research, **7**(1), 158–172, 2018.
[42] P. K. Smith, Cyberbullying and cyber aggression, Routledge, 2012.
[43] E. Maramwidze-Merrison, "Innovative methodologies in qualitative research: Social media window for accessing organisational elites for interviews," The Electronic Journal of Business Research Methods, **12**(2), 157–167, 2016.
[44] R. S. Tokunaga, "Following you home from school: A critical review and synthesis of research on cyberbullying victimization," Computers in Human Behavior, **26**(3), 277–287, 2010, doi.org/10.1016/j.chb.2009.11.014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 29-33 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Assessment of Transformer Cellulose Insulation Life Expectancy Based on Oil Furan Analysis (Case Study: South African Transformers)

Bonginkosi Allen Thango*, Jacobus Andries Jordaan , Agha Francis Nnachi

Department of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 04 June, 2021
Accepted: 21 August, 2021
Online: 13 November, 2021

Keywords:
Ageing
Oil
Transformer
Dissolved Gas Analysis (DGA)
Degree of polymerization (DP)

ABSTRACT

The ageing of oil-immersed power transformers triggers several defects and damages in the insulating materials, particularly in the cellulose insulation. The decomposition of the cellulose paper produces dissolved gases into the insulating oil, in which the Dissolved Gas Analysis (DGA) of the oil samples can provide insights to incipient faults sustained by the transformer as well as the status of the insulation. The condition of the cellulose insulation can be established by the measurement of the Degree of Polymerization (DP). In some instances, when the DP is intricate to measure, the estimation can be realized by collating information with respect to oil Furan analysis (2-furfuraldehyde (2FAL)). In this work, regression analysis based on 120 transformer oil samples is developed to establish new formulae by generating an analogy between the DP and 2FAL to indicate the status of the cellulose insulation and percentage of remnant life expectancy. In present study, five different transformers are tested using the proposed formulae to compute the DP based on the 2FAL concentrations. The results, are compared with measurement results and four other existing models. The results indicate an error of estimate of less than 2% and 1% against the measured DP.

1. Introduction

Oil-immersed power transformers form an integral part in the continuous operation of the South African power grid. Retaining the continuous operation of power transformer is thereby a critical issue for the power utility for the reason that any operational malfunction leads to loss of revenue. On that account, scheduled maintenance contributes to preventing transformer failures and power outages. A large percentage of transformer failures is contributed by the damage of the insulation materials, which is composed of insulating oil and cellulose paper [1-5].

The decomposition process of transformer insulating materials amplify at higher hotspot temperature, presence of oxygen and production of moisture. The examination of the transformer insulation condition can be established by diagnostic methods among which includes DGA [5, 6], DP by measuring furan compounds [7, 8], or Gel permeation chromatography [9, 10].

Furfural analysis has gained immense popularity in the South African transformer manufacturing industry in the evaluation of the strength of insulating cellulose paper and corresponding remnant service lifetime. Furans are the derived products of the decomposition of the cellulose paper that disseminate into the transformer oil. In measuring the furan concentration on the transformer oil, then the cellulose paper conditions can be ascertained. A rigorous method to assess the tensile strength in the cellulose paper that is intricately connected to cellulose insulation strength of a transformer is unearthing the Degree of Polymerization (DP) of the insulation. However, it is infeasible to obtain the paper samples from a transformer that is in-service, particularly paper samples in winding hotspot areas that are typically localized in the centre of the transformer steel tank. A number of researchers [11–14] have concluded that there is a remarkable connection between the DP and furan concentration, specifically 2-Furaldehyde (2FAL). The benefit of analysing the 2FAL concentration correlated to the DP is that the transformer oil samples can be appropriately extracted from the unit in-service without engendering any impairment to the transformer.

The most common procedure in measuring Furanic compounds in the transformer oil is recommended by the standard published in [15]. In view of the fact that this procedure necessitates particular expertise and high-end equipment, utility owners and manufacturers subcontract the Furan test to certified service enterprises. The utilisation optical spectroscopy techniques is also

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com

more and more notable in the Southern African transformer condition monitoring.

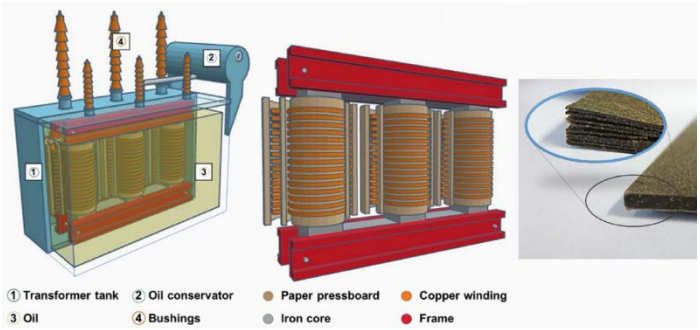


Figure 1:Transformer cellulose paper

The current work proposes an expedient alternative to evaluate the DP in the transformer cellulose paper using a statistical survey. A set of tests were carried out in the laboratory based upon oil and cellulose paper samples of transformers that have aged and also units that were collected from site as a results of various faults. Regression models have been developed to estimate the DP in the cellulose paper using the measured 2FAL concentrations. The results indicate an error of estimate of less than 2% and 1% of the proposed logarithmic and polynomial fit DP models against the measured DP respectively.

2. Furan compounds mechanism

The Furan compounds present in the transformer oil that can be measured by the High-Performance Liquid Chromatography (HPLC) are namely the 2FAL, 2-Furfurol (2FOL), 5-Hyroxy methyl-2-furfural (5HMF), 5Methyl-2-Furfural (5MEF), and 2-acetylfuran (2ACF) as demonstrated in Figure 2 [16].

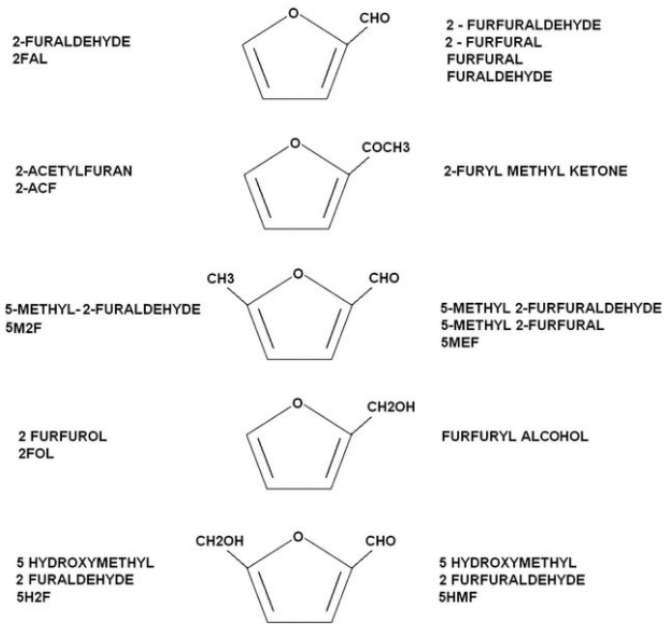


Figure 2: Furan compounds present in the transformer oil.

It is widely adopted the Furans are triggered from the decomposition of cellulose insulating material during the course of the transformer operational lifetime.

In several researches from the literature, case studies of transformer diagnostic status have been presented where the measured concentration of the furan in oil is employed as a key performance indicator of the chemical decomposition of the insulating cellulose paper. Evaluation of the results of furan analysis may be implemented with other assessments such as the monitoring of in-service transformer insulating oil and winding temperatures. As a result of intricacy of the cellulose decomposition mechanics and plethora of factors that can influence the process, the correlation between the 2FAL concentration and the service life of the transformer is not fully unravelled. Based on field experience, the factors that have a material impact on the composition of 2FAL in the transformer insulation oil are summarized below:

- The commencing 2FAL concentration in the transformer oil is characterised by the ratio of the generation/decomposition rate to a particular temperature;
- The occupancy of moisture considerably affect the dissemination of furan compounds between insulating cellulose and oil. The foregoing espouse the 2FAL composition;
- The winding temperature allures the formation / decomposition of 2FAL. The variability of the operating temperature as a result of miscellaneous loading conditions substantially affect the formation of furans;
- Types of transformer oil are also critical on the formulation of 2FAL. Oxidation inhibitors have a particular safeguarding effect on cellulose thermoset and triggers the deceleration of furan compounds in comparison with non-additive oils;
- Thermally upgraded paper generate minimal quantity of furan compounds than standard Kraft paper in oil;
- Types of transformer (cellulose/oil ratio) and cooling state is of paramount importance. Accruing the quantity of cellulose add up to the composition of furans;
- Oil reclamation bring about a lower furan content throughout a lapse of time. However, as a chemical equilibrium, the initial 2FAL concentration will be attained once more.

Several studies have been published in modelling the interrelation between the oil 2FAL concentration and the DP of cellulose material. A universal correlation that is endorsed by the transformer manufacturing industry is presently not available. Nevertheless, there is a common characteristic on the models proposed by authors in [17–21], which concludes that there is a correlation between the 2FAL concentration and the DP of cellulose. Other furan compounds such as 2FOL, 5HMF, 5MEF, and 2ACF are more erratic than 2FAL and tend to decompose rapidly.

Table 1 depicts the DP models that are employed in the comparison approach to substantiate the proposed models. The proposed Eq. (6) and Eq. (7) above are formulated based on logarithmic and polynomial fit regression models using furan analysis data collected from a fleet of mineral oil-filled transformers, particularly in the South African network. These models show excellent correlation coefficient (R^2) between the 2FAL concentration and the DP of cellulose with the polynomial fit model having an improved R^2. A graphical of these models and existing models is then demonstrated as shown in Figure 3.

Table 1: Various DP models

Eq.	Author	DP experiential formula	Ref.
1	Chendong	$DP = \frac{\log_{10} 2FAL - 1,51}{-0,0035}$	[17]
2	Dong	$DP = 402,47 - 220,87 \times \log_{10} 2FAL$	[18]
3	Vaurchex	$DP = \frac{2,6 - \log_{10} 2FAL}{0,0049}$	[19]
4	Heisler	$DP = 325 \times \left(\left(\frac{19}{13}\right) - \log_{10} 2FAL\right)$	[20]
5	De Pablo	$DP = \frac{1850}{2FAL + 2,3}$	[21]
6	Thango (Eq. 6)	$DP = -123,6 \times \ln 2FAL + 456,38$ $R^2 = 0.9809$	
7	Thango (Eq. 7)	$DP = 0,2358 \times 2FAL^6 - 5,4819 \times 2FAL^5 + 49,471 \times 2FAL^4 - 220,64 \times 2FAL^3 + 520,55 \times 2FAL^2 - 692,51 \times 2FAL + 811,15$ $R^2 = 0,9913$	

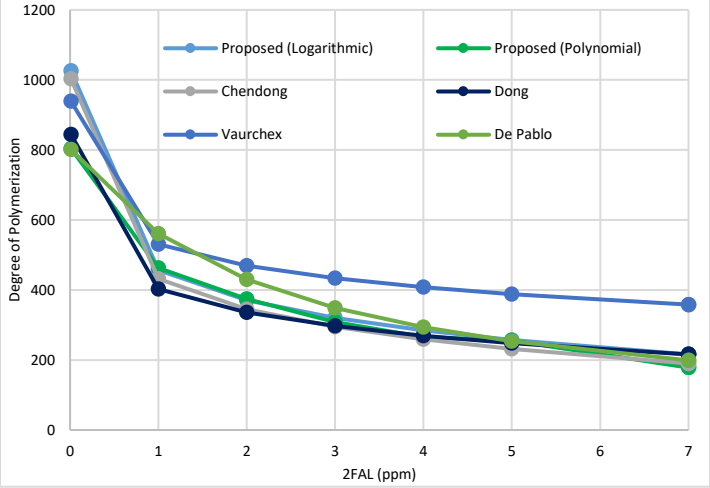


Figure 3: Comparison of DP models

2.1. Transformer remaining service lifetime

Consequently, the remaining life of the transformer can determined using Eq. (8), where the DP values range from 1200 ppm to 1000 ppm, with 200 ppm for degraded paper insulation [22].

$$RSL = 20,5 \times ln\left(\frac{1100}{2FAL}\right) \quad (8)$$

The percentage of the decomposed cellulose life for a unit in-service can be ascertained the Eq. (9) and Eq. (10) below.

Table 2: percentage of the decomposed cellulose

Eq.	Author	%RSL experiential formula	Ref.
9	Kanumuri	$\%\ RSL = \frac{LOG10(DP) - 2,903}{-0,006021}$	[23]
10	Doble	$\%\ RSL = 166,1 \times \log_{10}(2FAL) - 382,2$	[24]

Eq. (9) gives the decomposed cellulose life as a percentage, but to measure transformer paper life in years, Eq. (5) is proposed in [22] may be employed. The parameters used assume initial DP (DP_0) as 1100 and a 35-year transformer lifespan.

3. Case Study: South African oil-immersed transformers

In this section, the sample results of furan analysis of five transformers are presented. The DP has been measured for each of the case studies. At glance, transformer #5 has the lowest 2FAL concentration with a slightly higher measured DP, which suggest this unit is older than the other units. Additionally, transformer #2 has a higher 2FAL value and lowest measured DP.

Table 3: Service lifetime parameters

#	2FAL	Measured DP
1	1,563	408
2	0,231	648
3	1,601	405
4	1,778	393
5	2,377	351

In Table 4, the proposed models in Eq. (6) and Eq. (7) to evaluate the amount of DP in the transformer insulating cellulose paper are compared with measured data and other models in the literature. The comparison will be carried out based upon the error of estimated of the estimated and measured samples.

Table 5 presents the error of estimate of some case studies using Eq. (6) and Eq. (7) and other DP models from the literature. The 5 cases in Table 3 are employed to examine the accuracy of Eq. (6) and Eq. (7). Only one case (Case 2) shows a slightly higher error of estimate against the measured data, but all the cases for Eq. (6) and Eq. (7) are less than 2% and 1 % respectively. The error of estimates for Chendong are less than 8%, Dong less than 12% except for higher case (case 2), Vaurchex has less than 29% and De Pablo less than 17%.

Table 4: Degree of Polymerization

#	2FAL	Measured DP	Thango (Eq. 6)	Thango (Eq. 7)	Chendong	Dong	Vaurchex	De Pablo
1	1,563	408	401	406	376	360	491	479
2	0,231	648	637	676	613	543	660	731
3	1,601	405	398	403	373	357	489	474
4	1,778	393	385	390	360	347	480	454
5	2,377	351	349	349	324	319	454	396

Table 5: Error of estimate (Degree of Polymerization)

#	Thango (Eq. 6)	Thango (Eq. 7)	Chendong	Dong	Vaurchex	De Pablo
1	2%	1%	8%	12%	20%	17%
2	2%	4%	5%	16%	2%	13%
3	2%	1%	8%	12%	21%	17%
4	2%	1%	8%	12%	22%	15%
5	0%	1%	8%	9%	29%	13%

Table 6: Remnant Service Lifetime

#	2FAL	Measured DP	Thango (Eq. 6)	Thango (Eq. 7)	Chendong	Dong	Vaurchex	De Pablo
1	1,563	20,33	20,68	20,46	22,01	22,92	16,53	17,05
2	0,231	10,85	11,18	9,97	11,98	14,47	10,46	8,38
3	1,601	20,48	20,83	20,60	22,17	23,05	16,62	17,25
4	1,778	21,10	21,51	21,27	22,90	23,64	17,02	18,16
5	2,377	23,42	23,51	23,54	25,06	25,35	18,15	20,97

Table 7: Percentage Remnant Service Lifetime (Eq. 9)

#	2FAL	Measured DP	Thango (Eq. 6)	Thango (Eq. 7)	Chendong	Dong	Vaurchex	De Pablo
1	1,563	48,55	49,77	48,99	54,44	57,66	35,19	37,00
2	0,231	15,18	16,36	12,09	19,16	27,93	13,81	6,50
3	1,601	49,09	50,31	49,51	55,02	58,12	35,51	37,70
4	1,778	51,26	52,69	51,85	57,58	60,18	36,89	40,90
5	2,377	59,41	59,74	59,86	65,18	66,21	40,87	50,79

Table 8: Percentage Remnant Service Lifetime (Eq. 10)

#	2FAL	Measured DP	Thango (Eq. 6)	Thango (Eq. 7)	Chendong	Dong	Vaurchex	De Pablo
1	1,563	51,43	50,21	50,99	45,54	42,33	64,79	62,99
2	0,231	84,80	83,62	87,89	80,83	72,05	86,18	93,49
3	1,601	50,90	49,68	50,47	44,97	41,86	64,48	62,28
4	1,778	48,73	47,29	48,13	42,41	39,80	63,09	59,08
5	2,377	40,58	40,24	40,12	34,80	33,77	59,12	49,20

After examining the veracity of Eq. (6) and Eq. (7) to evaluate the DO, the remnant service lifetime of the insulating cellulose paper was established, at which point, the studied transformers was determined by applying Eq. (8). Table 6 shows the corresponding remnant service lifetime of the transformer.

Table 7 and Table 8 demonstrate the percentage remnant service lifetime of the 5 cases in Table 6. The results indicate that there is an inversely proportionate relationship between the percentage remnant service lifetime and the 2FAL concentration while a directly proportional relationship is observed against the degree of polymerization. The results of the percentage remnant service lifetime by applying Eq. (9) is tabulated below.

An increase in the 2FAL concentration results in the decrease of the degree of polymerization and consequently the transformer's remnant in-service lifetime. The degree of polymerization of insulating cellulose paper is high for newly manufactured transformers ranging above 800 to 1200 and the cellulose paper with decomposition activity due to the operational conditions will decrease radically until it reaches a DP of about 2000 where the unit must be removed from service as it indicates end of lifespan. Subsequently, the results of the percentage remnant service lifetime by applying Eq. (10) is tabulated.

Based on the results presented in Table 7 and Table 8, it can be observed and concluded that the proposed Eq. (6) and Eq. (7) may be adopted by local manufactures to reliably evaluate the remnant in-service lifetime of transformers connected to the South African grid. Results of Chendong, Dong, Vaurchex and De Pablo are erratic in evaluating the transformer service life.

4. Conclusion

Examination of cellulose paper in transformers is highly complex when the unit is in-service, consequently chemical tests were conducted to survey the condition of the insulating cellulose paper. The degree of polymerization was wielded as a key benchmark for evaluating the condition of the cellulose paper by measuring the Furan (2FAL) concentration. A statistical survey through logarithmic and polynomial fit regression models was realized to develop the prognosis equations allying the degree of polymerization and the measured 2FAL concentration. The statistical survey results demonstrated that there is a high correlation between the cellulose DP and the 2FAL concentration,

yielding correlation coefficients of 0.9809 and 0.9913 for the proposed models.

The results of the studied cases inveterate that the 2FAL concentration is a suitable index for predicting the degree of polymerization. By comparing the proposed models in Eq. (6) and Eq. (7) with the DP models in the literature, the results revealed that the proposed DP models based on South African transformers to be an excellent measure of the degree of polymerization at which point the remnant and percentage service lifetime of the units can be evaluated. The error of estimate of the proposed Eq. (6) and measured data is less than 2 % while Eq. (7) is less than 1% except for an anomaly observed in case number 2. Chendong, Dong, Vaurchex and De Pablo yielded and error of estimate of no more than 8%, 16%, 29% and 17% in comparison to the measured data. Similar trends in the estimation of the remnant and percentage service lifetime are observed.

The continuation of this work in the future will comprise the application of least squares support vector machine for classifying the condition of the insulating cellulose paper and then evaluate the remnant service lifetime of a transformer.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. K. Pradhan and T. S. Ramu, "Diagnostic testing of oil-impregnated paper insulation in prorated power transformers under accelerated stress," Conference Record of the 2004 IEEE International Symposium on Electrical Insulation, 66-69, , 2004, doi: 10.1109/ELINSL.2004.1380452.

[2] C. Chakradhar Reddy, E. Johnson and T. S. Ramu, "On the life estimation of oil impregnated paper insulated transformer bushings," 2007 Electrical Insulation Conference and Electrical Manufacturing Expo, 13-16, 2007, doi: 10.1109/EEIC.2007.4562578.

[3] Y. Tu, W. Sun, C. Yue, G. Chen and W. Wang, "Research of insulation properties of polymer materials using in oil-filled transformers under high temperature," 2011 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 583-586, 2011, doi: 10.1109/CEIDP.2011.6232724.

[4] S. Missas, M. G. Danikas and I. Liapis, "Factors affecting the ageing of transformer oil in 150 / 20 kV transformers," 2011 IEEE International Conference on Dielectric Liquids, 1-4, 2011, doi: 10.1109/ICDL.2011.6015409.

[5] T. K. Saha, M. Darveniza, Z. T. Yao, D. J. T. Hill and G. Yeung, "Investigating the effects of oxidation and thermal degradation on electrical and chemical properties of power transformers insulation," in IEEE Transactions on Power Delivery, **14**(4), 1359-1367, 1999, doi: 10.1109/61.796229.

[6] D. M. Mehta, P. Kundu, A. Chowdhury and V. K. Lakhiani, "DGA diagnostics save transformers - case studies," 2015 International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), 116-120, 2015, doi: 10.1109/CATCON.2015.7449519.

[7] S. Apte, R. Somalwar and A. Wajirabadkar, ""Incipient Fault Diagnosis of Transformer by DGA Using Fuzzy Logic"," 2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), 1-5, 2018, doi: 10.1109/PEDES.2018.8707928.

[8] K. Baburao, N. M. Bhangre, A. M. Wagle, A. Venkatasami and S. E. Chaudhari, "The experience of DP and furan in remnant life assessment of power transformer," 2008 International Conference on Condition Monitoring and Diagnosis, 555-558, 2008, doi: 10.1109/CMD.2008.4580348.

[9] N. Das, A. Abu-Siada and S. Islam, "New approach to estimate furan contents in transformer oil using spectroscopic analysis," 2012 22nd Australasian Universities Power Engineering Conference (AUPEC), pp. 1-4, 2012.

[10] S. P. Lai, A. Abu-Siada, S. M. Islam and G. Lenco, "Correlation between UV-Vis spectral response and furan measurement of transformer oil," 2008 International Conference on Condition Monitoring and Diagnosis, 659-662, 2008, doi: 10.1109/CMD.2008.4580372.

[11] D. Feng, Z. Wang and P. Jarman, "Transmission power transformer assessment using furan measurement with the aid of thermal model," 2012 IEEE International Conference on Condition Monitoring and Diagnosis, 521-524, 2012, doi: 10.1109/CMD.2012.6416194.

[12] M. Meissner, M. Darmann, S. Schober, M. Mittelbach and C. Sumereder, "Reliability Study of Furan Level Analysis for Transformer Health Prediction," 2019 IEEE 20th International Conference on Dielectric Liquids (ICDL), 1-4, 2019, doi: 10.1109/ICDL.2019.8796785.

[13] K. B. Shaban, A. H. El-Hag and K. Benhmed, "Prediction of Transformer Furan Levels," in IEEE Transactions on Power Delivery, **31**(4), 1778-1779, 2016, doi: 10.1109/TPWRD.2016.2521320.

[14] M. A. Talib, M. A. A. Aziz, Y. Balasubramaniam, Y. Z. Y. Ghazali, M. R. Z. Abidin and M. F. M. Yousof, "Transformer Ageing Investigation and Correlation of Furanic Compound and Degree of Polymerization for Power Transformer Life Assessment," 2020 IEEE International Conference on Power and Energy (PECon), 240-244, 2020, doi: 10.1109/PECon48942.2020.9314543.

[15] "IEEE Guide for Evaluation and reconditioning of Liquid Immersed Power Transformers," in IEEE Std. C57.140-2006, 1-73, 2007, doi: 10.1109/IEEESTD.2007.353650.

[16] I. Hohlein and A. J. Kachler, "Aging of cellulose at transformer service temperatures. Part 2. Influence of moisture and temperature on degree of polymerization and formation of furanic compounds in free-breathing systems," in IEEE Electrical Insulation Magazine, **21**(5), 20-24, 2005, doi: 10.1109/MEI.2005.1513426.

[17] X. Chendong, F. Qiming, X. Shiheng, "To Estimate the Ageing Status of Transformers by Furfural Concentration in the Oil", In Proceedings of the CIGRE Study Committee 33 Colloquium, Leningrad, Moscow; 1991.

[18] M. Dong, Z. Yan and G. J. Zhang, "Comprehensive diagnostic and aging assessment method of solid insulation in transformer," 2003 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 137-140, 2003, doi: 10.1109/CEIDP.2003.1254813.

[19] H. Vaurchex, I. Höhleim and A. J. Kachler, "Transformer aging research on furanic compounds dissolved in insulating oil", CIGRE, 2002, doi: publication/15-302_2002-transformer-ageing-research-on-furanic-compounds-dissolved-in-insulating-oil

[20] H. Andreas et al. "Zustandsbeurteilung von Transformatoren mit Furfurol-Bestimmung." 2003.

[21] A. de Pablo, "Furfural and ageing: how are they related," IEE Colloquium on Insulating Liquids (Ref. No. 1999/119), 5/1-5/4, 1999, doi: 10.1049/ic:19990667.

[22] M. K. Pradhan and T. S. Ramu, "On the estimation of elapsed life of oil-immersed power transformers," in IEEE Transactions on Power Delivery, **20** (3), 1962-1969, 2005, doi: 10.1109/TPWRD.2005.848663.

[23] D. Kanumuri, V. Sharma, O.P. Rahi, O P, "Analysis Using Various Approaches for Residual Life Estimation of Power Transformers", International Journal on Electrical Engineering and Informatics, Bandung **11**(2), 389-407, 2019, doi:10.15676/ijeei.2019.11.2.11.

[24] Doble," Testing & Sampling", doi: services/laboratory-services/testing-and-sampling/.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 32-42 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

The Design and Implementation of Intelligent English Learning Chabot based on Transfer Learning Technology

Nuobei Shi, Qin Zeng, Raymond Shu Tak Lee*

Division of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, 519000, China

ARTICLE INFO

Article history:
Received: 21 April, 2021
Accepted: 27 August, 2021
Online: 10 September, 2021

Keywords:
NLP-based Chatbot
Transfer learning
OpenAI GPT-2
English Learning Chatbot
Artificial Intelligence

ABSTRACT

Chatbot operates task-oriented customer services in special and open domains at different mobile devices. Its related products such as knowledge base Question-Answer System also benefit daily activities. Chatbot functions generally include automatic speech recognition (ASR), natural language understanding (NLU), dialogue management (DM), natural language generation (NLG) and speech synthesis (SS). In this paper, we proposed a Transfer-based English Language learning chatbot with three learning system levels for real-world application, which integrate recognition service from Google and GPT-2 Open AI with dialogue tasks in NLU and NLG at a WeChat mini-program. From operational perspective, three levels for learning languages systematically were devised: phonetics, semantic and "free-style conversation" simulation in English. First level is to correct pronunciation in voice recognition and learning sentence syntactic. Second is a converse special-domain and the highest third level is a language chatbot communication as free-style conversation agent. From implementation perspective, the Language Learning agent integrates into a WeChat mini-program to devise three user interface levels and to fine-tune transfer learning GPT-2 as back-end language model to generate responses for users. With the combination of the two parts about operation and implementation, based on the Neural Network model of transfer learning technology, different users test the system with open-domain topic acquiring good communication experience and proved it ready to be the industrial application to be used. All of our source codes had uploaded to GitHub: https://github.com/p930203110/EnglishLanguageRobot.

1. Introduction

Chatbot operates similar to *virtual personal assistant (VPA)* and *question-answer (QA) system* in its development. In general, a practical chatbot consists of two categories, one has no *artificial intelligence (AI)* technologies i.e. rule-based and pattern recognition that collects tremendous high quality artificial corpus as database for *question-answer (QA)* matching [1]. Another has AI that uses current models and algorithms enabling chatbot to learn necessary articles before use. A typical example is speech recognition with text-to-speech interconversion to listen and converse at real-world human-machine interaction with database as machine brain or corpus with various meanings used as *natural language understanding (NLU), natural language generation (NLG), neural network (NN) models* with sufficient target language corpus training to equip the brain with thoughts and communication. This new human-machine interaction method provides convenience but it generates interior questions on how the chatbot can perform intelligently. It has been an ongoing research topic on user interface and expected functions by many internet organizations. However, back-end support require both AI and/or non-AI technologies to operate effectively. A literature review on human-like chatbot with more than one algorithm to select the best response that similar to Microsoft's Xiao Ice, a software assistant with emotions will be presented. A chatbot's performance is dependent on core components enriched with sufficient algorithms. Thus, a transfer-based technology with different language learning models would be proposed in this paper.

Turing test progress enticed researches on rule-based human-machine interaction system using pattern recognition in 1950s. Corpus as response matching outperformed Turing test showed

*Corresponding Author: Raymond Lee, Email: raymondshtlee@uic.edu.cn

that database can be enlarged by sufficient human dialogue directing that human syntactic simulation is not necessary as semantic matched natural human dialogue for response. Since then, when natural language is generated by trained machine with big data, the rule-based chatbot is updated to an *information-retrieval (IR)* or a *neural network (NN)* based system automatically and reduced manpower considerably. IR based system is mostly applied in a search engine optimizer with constructed knowledge base whereas NN system is densely data-driven to perform high-level prediction and classification. It can be used as a transfer learning method to pre-train and fine-tune the basic concepts of a chatbot. As shown on Figure 1, machine linguistic is depended on its data learned such as the basic level with words, phrases and sentences etc. For *natural language processing (NLP)*, an initial step is to transform natural language into a word vector followed by mathematical computation to illustrate the basic meanings equivalent to real-world. Thus, NN has the ability to simulate human brain and self- generate natural language as compared with other methods.

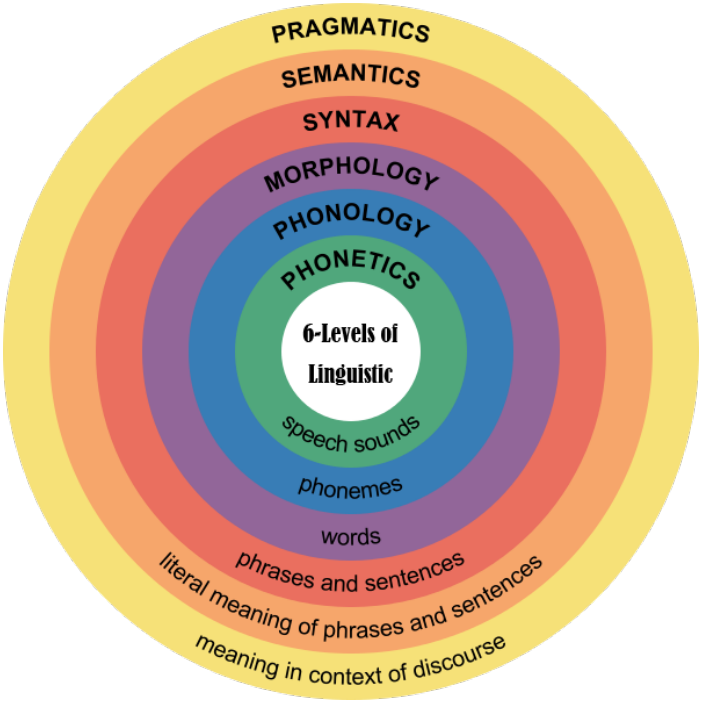


Figure 1: Levels of Linguistic in human languages [2]

There are numerous chatbot models equipped with different algorithms. Transfer learning is amongst performed better than various unsupervised learning sequence-to-sequence (seq2seq) models i.e. *recurrent neural network (RNN)* and its related bi-direction RNN and *long-short term memory (LSTM)* models. BERT from Google and GPT-2 from Open-AI are the most famous and pre-trained transformer models to generate natural language and classification. In this paper, GPT-2 is used as English language learning back-end model to converse daily dialogues because of GPT-2 inherent priority of natural language generation. For pre-trained transfer learning language model, a small dataset is used to fine-tune the transformer model to outperform traditional counterparts. Also, a self-attention mechanism is included. It is the main theme of transfer learning to find the relationship between sequences. Further, a language mini-program based on GPT-2 with a fine-tuned model will be included to provide fluent and syntactic responses. Lastly, a mini-program equipped with a fine-tuned GPT-2 language model and speech recognition service from Google to deliver a three systematical English learning level functions at WeChat for implementation.

2. Literature Review

2.1. Overview of industrial English Learning apps

Duolingo is an English language learning application with growing popularity since coronovirus. It is accredited by universities in UK, US and Australia as English language proficiency qualifications similar to IELTS and TOEFL. It contains spoken tests' pronunciation and oral presentation on text questions or images understanding with scores given depend on voice recognition. Test contents are preinstalled without human participation and gradings are given by statistic algorithm that are different from IELTS by physical examiner. It covered more than 30 mainstream languages transmitted in English and more than 6 languages relayed to Chinese language. Prior course commencement, the application would match users' demands to set step by step learning schedules and provided tests to detect grammar and pronunciation proficiciency. Level detection is an important component to improve language skills because it depended on applications or chatbots functions and users' willingness to select appropriate learning levels at different courses. [3] It shows that the online English learning or human-machine interaction English test will be the mainstream in these days. However, the technologies it used are not clear or transparent for users, the statistical algorithm or non-AI technologies also cannot substitute the official recognized IELTS which is communicate with real human or the machine recoreds judge by human. On the contrary, our model are natural based on AI technologies and growing with the interaction between human and machine.

The model we proposed is based on IELTS Part ABC speaking module examination syllabus covering pronounciation, topic discussion and free chat. They are named as Level 1-3 where the highest level 3 is a human-machine interaction dialogue. As responses from reverse party generated by IELTS are in random so the same format is followed where responses are generated by a back-end transfer learning model similar to a proficient English Tutor. Corpus of big data used for transfer learning model training covers topic knowledge discussion.

Other reference is based on a different English language learning application called LiuliShuo. It contains 7 levels to improve verbal skills. Level 4 refers to proficiency equivalent to College English Test (CET)-4 or 6, where the highest level 7 is equivalent to IELTS 7.5 and TOEFL 105. LiulShuo settings, compared with Duolingo in user interface and functions design are fitted to the model design. However, they are sufficient for grades evaluation but are lack of AI advanced intelligent technologies for response according to user's learning progress.

Thus, the design of the proposed model began with daliy dialogue for model training to improve its learning ability, then a database was added to track every users' learning history and levels improvement. For this segment, daily chitchat data in natural langauge are transferred into json format to improve machine learning. Those unstructed raw data or called heterogeneous data

required cleaning and sorting automatically in accordance with AI ecosystem.

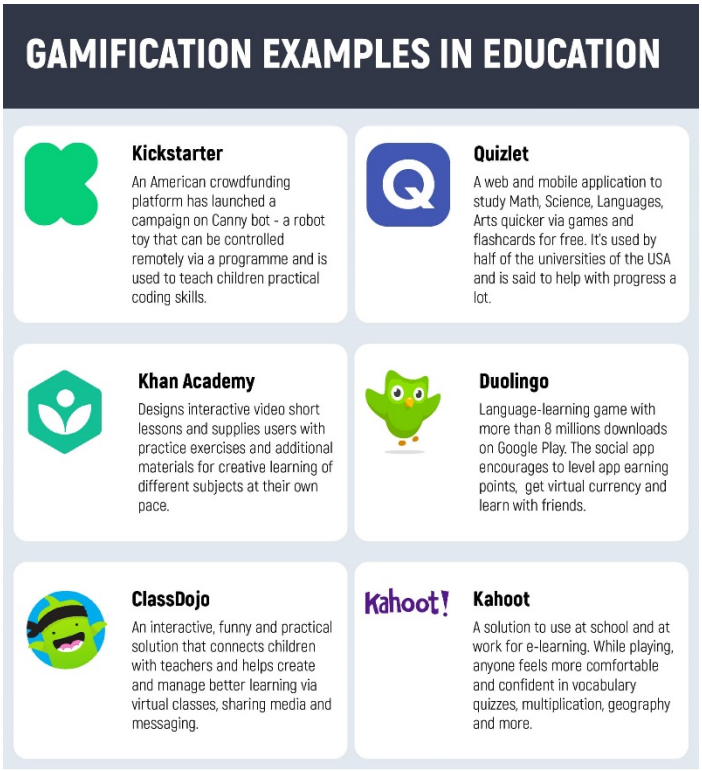


Figure 2: Gamification examples in Education [4]

Figure 2 shows a set of gamification examples in education. It is a game module to calculate grade levels to motivate users compete within app for improvement.

2.2. Academic research about Chatbot.

2.2.1. AliMeChat: A Sequence to Sequence and Rerank based Chatbot Engine (rule-based IR+seq2seq) [5]

Recurrent neural network (RNN) is capable to generate responses from end to end leading to sequence to sequence (seq2seq) model becomes chatbot mainstream generation prior 2018. Since then, NLP tasks were greatly improved using transfer learning model.

TaoBao App uses AliMe to substitute online human customer service for most of the predicted questions. Statistical data showed that 5% questions were chats from a commercial-domain Question-Answer system to an open-domain chatbot. This process resembled to English language learning agent similar to level 2 scenario dialogue to level 3 open-domain free style conversation of this paper. AliMe integrated a rule-based IR and pre-trained seq2seq application model to real-world industry and performed better than both IR and public chatbot. However, its pre-trained seq2seq model is used twice for response generation to re-rank a set of answer from matching input questions with knowledge base paired dialogues. It suggested that free style format can be based on usual conversation scenario exist in daily activities.

The proposed model has two similar AliMe settings in generation for information retrieval but in different formats. IR-based model uses natural language word to match knowledge base, the generation seq2seq model and model evaluation to re-rank output response in generation are embedded words as vector. IR-based dataset is approximately 9,164,834 natural questions and answers paired dialogues from users and staff collected from commercial domain. Researchers use inverted index based on search engine concept to match these conversation with input sentence containing identical words and used BM25 algorithm to measure input sentence similarity to select questions and respond to input questions. Traditional chatbot back-end also avoided commonsense questions that cannot be responded by chatbot and generation model used seq2seq model.

AliMe also selected GRU, a type of RNN units to reduce computation time for response generation. Further, an optimizer SoftMax is used in calculate to time control sample words set covering 512 random words. Beam Search in decoder assisted to find the highest probability in Conditional Probability and obtain the optimized response sentence within parameters. Results performance showed that *IR+generation+rerank* shown on Figure 3, by seq2seq model and mean probability scoring function evaluation approach achieved the highest score compared with other separated methods.

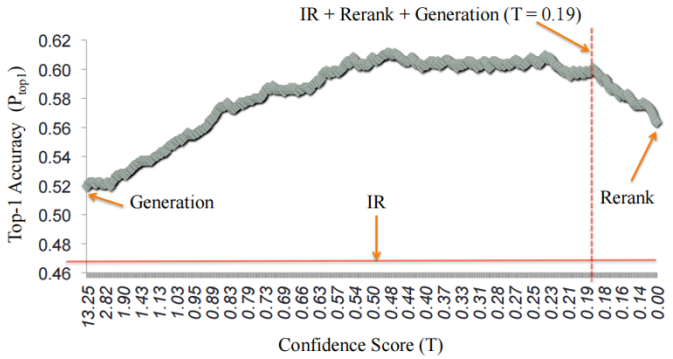


Figure 3: Top-1 accuracy of candidate approaches [5]

Due to the characteristic of NN and chatbot usage or a QA-based system, special-domain system in contents apart from the task-oriented or goal-oriented system for special usage than open-domain system applications are essential for model design. Microsoft chatbot is open-domain with emotions and AliMe specialized in business as Taobao customer service restricted in business corpus. Also, grounded knowledge contents becomes another important trend for chatbot developments[6]. In summary, rule-based dialogue with seq2seq RNN model are used frequently than others because most daily conversation cannot be generated by RNN model. Since attention mechanism proposed in 2017 in Attention is all you need [7], language model in natural language processing changed to *seq2seq+attention* and self-attention in transfer performed better than seq2seq.

2.2.2. The Design and Implementation of Xiao Ice, an Empathetic Social Chatbot（IR+seq2seq+KG）[8]

Microsoft released an empathetic social chatbot based on two segments at various platforms. One is intelligence quotient (IQ) that used for thinking and answer human questions. Another is emotional quotient (EQ) to understand and analysis the emotional meaning of natural language to provide 24 hours user services, research team use three back-end to equip the machine with human-like brain to realize these two segments. First is the traditional IR with corpus for matching. Second is a RNN of sequence to sequence model to generate response and the third is a knowledge graph for entities extraction to extract related entities

and sorting for responses. After that, an optimizer would rank potential responses and select the most suitable response to users.

Each Microsoft operation system has a unique Cortana for its computer owner [8]. Xiao Ice has different Cortana as a personal assistant resembled to Siri. It preferred to chat with IQ and EQ system design similar to human interaction. It is by far the most dialogue-oriented AI robot with emotions to provide 24 hours interaction. It has over 660 million users since its inception with *conversation-turns per session* (CPS) reached 23 which is higher than other chatbots. It is also an open-domain chatbot aimed to optimize content responses with emotions to understand users' thoughts. It has expanded to more than 5 countries with different names. Users willing to use not only depended on IQ but also dialogue contents provided a user-friendly first perception. Thus, IQ part would also be used in the proposed model.

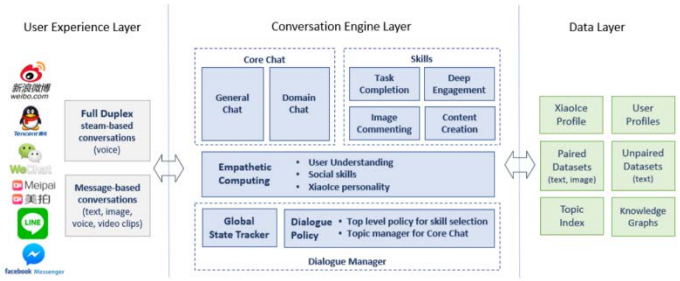


Figure 4: XiaoIce system architecture [8]

In Figure 4, for users implementation, Xiao Ice exists in 11 mainstream platforms including WeChat and Tencent QQ and at mini blog, Weibo and Facebook. It has the highest recognition function that include text, image voice and message-based conversation. Its conversation engine layer was developed after four years and its Core Chat can distinguish and change topics automatically. Thus, the proposed model is set with a second level at topic dialogues for special domain scenario and a third level for free style conversation to improve users' dialogue proficiency. A dialogue manager is to define dialogue management policy so that it can monitor the state of dialogue consisting Global State Tracker and Dialogue Policy to select a state action. Global State Tracker is a vector consisting of Xiao Ice responses and analyses text strings to extract entities and empathy. Dialogue policy is designed for long-term users to improve interactive engagement optimization. Initially, topic manager is the first to greet users and awake system with natural language understanding (NLU). It will then obtain users' responses' to adjust the "Core Chat" into open-domain or special scenario with users' interests. If system always repeats or response information is bland and within three words, the topic will be changed accordingly. Secondly, is the skill selection. Once user input is a particular format, the skills would activate to process different input. Its images can be classified into different task-oriented scenario e.g. if the image is food, a restaurant will be displayed. Lastly, a personal assistant would provide information such as weather, space availability and reservation for task completion.

In other words, it contained certain knowledge graphs in data layer where dataset is originated from popular forum such as Instagram in English or douban.com in Chinese. These datasets listed as several topics with relative small knowledge base as candidate answer and are sufficient for questions. When a new topic appears, system would obey the policy to refresh its knowledge base by machine-learning. However, not all new entity or topic will be collected. It is similar to a ranking list unless the entity is contextual relevance or the topic has higher popularity or news latest update and users' personal interests will be considered.

The core of Xiao Ice is language model, using the RNN shown in Figure 5 to create the open and special domain knowledge.

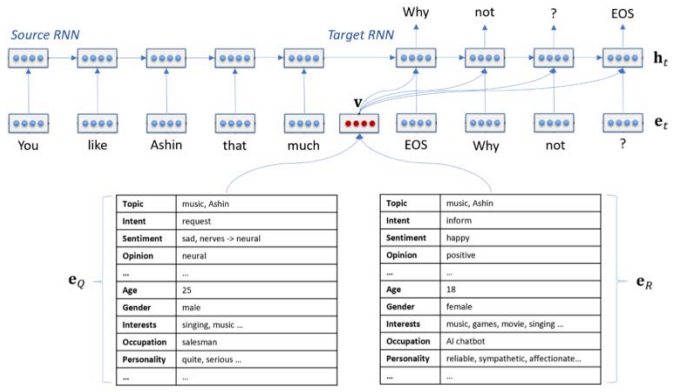


Figure 5: RNN-based neural response generator [8]

[Human] How old are you?
[S2S-Bot] 16 and you?
[Human] What's your age?
[S2S-Bot] 18.

[Human] How old are you?
[Xiaoice] I am 18, of course.
[Human] You age?
[Xiaolce] 18. Why?
[Human] Were you 18 last year?
[Xiaoice] I made a wish to stay 18 forever. Oh, my wish has come true.

Figure 6: Examples of inconsistent responses generated using a seq2seq model [8]

Figure 6 shows a chatbot with emotional consciousness. It enhanced interaction to improve system EQ performance.

Xiao Ice, in general, has retrieval and generated-based back-end models. The former is a mature search engine used for decades. Its knowledge graph is a database for graph matching and entities extraction where the latter is trained by big data with related dialogue based on entities relations to supply contents and explain output sentences from NN. These two directions using neural model to generate responses and rank sentences to match relations and optimization. Once entity recognition is linked with neural generator conversation and dynamic knowledge graph and the knowledge-based will be refreshed. If user's feedbacks match with English syntax, the knowledge-based would always be collected together with users' interests. In addition, a rule-based chatbot with *artificial intelligence markup language (AIML)* as backup will be included to avoid senseless or repeat answers.

Thus, cognition, reinforcement learning of history dialogue[9] and empathetic computing are essential segments in the proposed model for competitiveness.

2.2.3. Attention Is All You Need [7]

Although chatbots mentioned above are acceptable for usage, its training back-end testing about sequence to sequence model showed good performance for text generation. Considering neural network relied on big data, data-driven model for better performance. Thus, transfer learning model shown on Figure 7 with magnitude training and fine-tune data has been the architecture mainstream. With the recognition of attention mechanism, the updated self-attention created the priority of text generation to be proved by BERT and GPT and GPT has natural advantage for text generation within sequence.

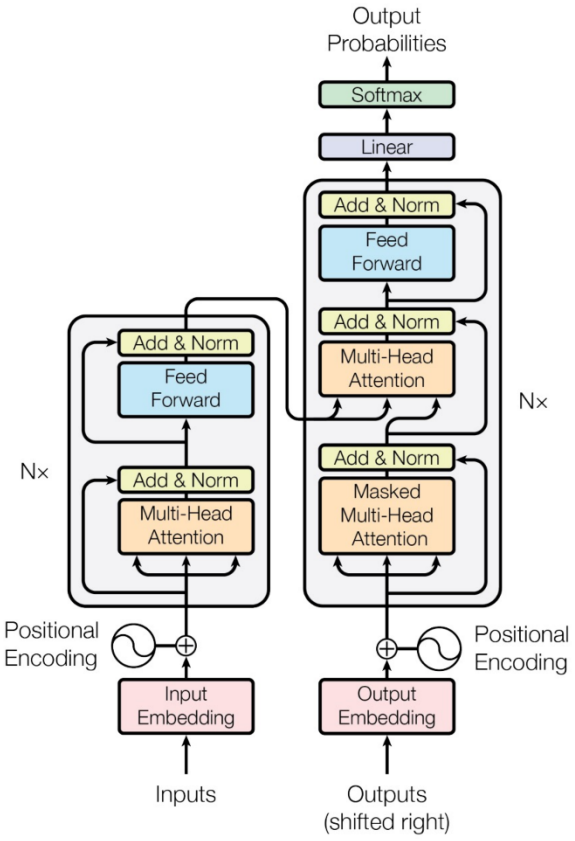


Figure 7: Transfer learning-model architecture [7].

The proposed model used as back-end is *Transfer Transfo*[1], as a daily dialogue system with common sense which integrate transfer learning training architecture and high-quality dataset. By fine-tuned persona-chat, the original utterance of GPT-2 with text generation based on context changed to format dialogue with fine-tune content dataset and English skills with original magnitude GPT-2 model. Also, the fine-tuned dataset also added the contents of English Tutor with users' personality and background resemble to native IELTS examiner.

Attention is human preference as a weight added to other words computation. It means that the encoding related to context, input and output. Transformer use self-attention that is search relations within sequence finding dependency among words in one sentence.

BERT is the first system proposed in transfer learning and fine-tuning. It has widely used in especially encoder part of pre-trained language model. However, it cannot predict decoder outputs. Open-AI GPT-2 has better application for generation with corpus and decoder limitation. The most famous performance of GPT-2 is text generation, since it releases larger dataset, more users are willing to make text prediction and dialogue. With the same transfer learning mechanism, GPT-2 not only can build dictionary use tokenizer like T, also play a role of language teacher with pre-trained grammar in natural language processing. Also, the GPT-2 original training source is 40G text from Internet, it has better adaption for text generation.

When the first edition of GPT-2 open source, researchers would like to test its performance with larger data. The 40G text in open-domain enabled GPT-2 equip with sufficient common sense similar to AIML(Artificial Intelligence Markup Language) which consist of human daily dialogues, but with more contents like grounded knowledge of Microsoft chatbot. For researchers, fine-tuning is the process to make GPT-2 transfer as own language model with special tasks. Text generation, text classification, automatic summarization, music modeling even writing code, all are functions of fine-tuned GPT-2 by users.

Thomas Wolf and his partners applied GPT-2 into dialogue generations called *Transfer Transfo* [1] to develop conversational system. Besides rule-based and seq2seq dialogue system, Transfo is another data-driven dialogue agent based on GPT-2, it combines transfer learning method to training and the transfer learning model with abundant contents with high-capacity, which shows big progress compared with end to end like seq2seq and information retrieval models. For traditional language models, neural network cannot track dialogue history within the model, even it learns the conversation with users. It is only big data to learn which cannot be tracked like rule-based system with database to store history dialogues. For training and generation algorithm, Transfo has priority to improve in relevance, coherence also grammar and fluency for output. For the model, multi-layers with 12 decoders and masked self-attention heads that token only range from left context. The large transfer learning model used in several practical task pays more attention to decoder part for the down-stream NLP tasks [10]. Pre-trained is an important part for down-stream performance to avoid discrete sentence to construct systematical learning in grammar and continuous content, document level corpus is better than sentence level. So, an Open AI release different size corpus for researchers to select suitable pre-trained and fine-tuning model afterwards. Fine-tuning decides the down-stream task text format in this paper. Thomas Wolf use persona-chat dataset as fine-tuning data to train input and output utterance change from long-text to dialogue format. Persona-chat in real-world helped to shape the speakers' backgrounds to make system further define topics and better understand users' input sentences .

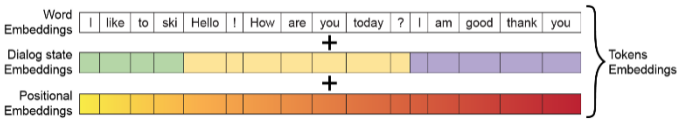


Figure 8: TranferTransfo's input representation [1]

In Figure 8, tokens embedding in one sequence is consist of 3 embedding of word, dialogue state and positional as shown in Fig 8. Rather than simple end to end dialogue with only one sentence, it collects usually 4-6 sentence and 3-5 of history dialogue belongs to same user in one sequence for training helping system to judge , state and position representation of every token to predict the next sentence in real application.

A dialogue system based on GPT-2 is included in the proposed model. In [11], the author made dialogue task more specific to goal-oriented personal assistant to help users complete special tasks. For task-oriented training, it has higher quality for corpus to fine-tune model and track conversation state of various situations. Not only it can interact with external system rather than pipelined system but also explain why the system generate such response.

End-to-end neural architecture for dialogue systems built on the top of GPT-2 with dialogue format. So, the dataset for fine-tuning should be dialogue as shown in Figure 9, which include external database of restaurants, hotels and others scenarios traveling in UK. The dialogues link to database contents to predict the system dialogue state and system action. Even more, following

the traditional end to end pipeline plus the database to be more explainable in output sentence that show why the system provides direction or operation advice as responds to users. Although it is a simple gradient descent, GPT-2 pre-trained model already equipped competitive dialogue as response.

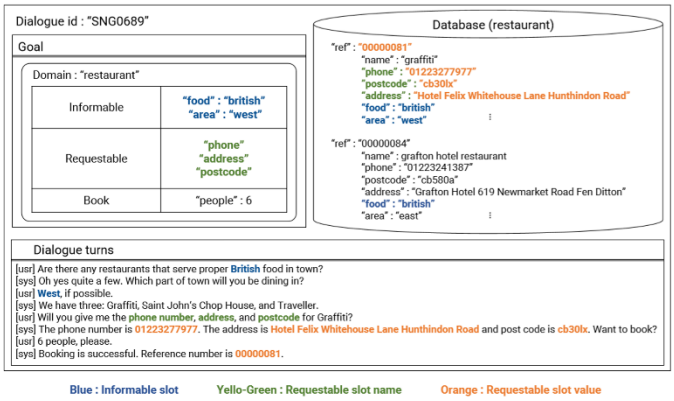


Figure 9: A single-domain example in MultiWOZ dataset [11]

Figure 10 shows the architecture of model similar to the model of grounded knowledge by Microsoft. Because the external system also a supplement for normal case which is a sequence to sequence model to generate substantial knowledge. When transfer learning-based system requires contents to engage specific task, knowledge base to fill the slot is another method to accomplish same object.

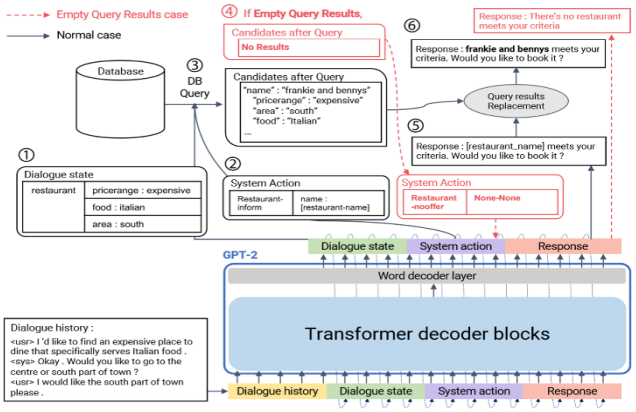


Figure 10: Architecture of end-to-end neural dialogue model[11]

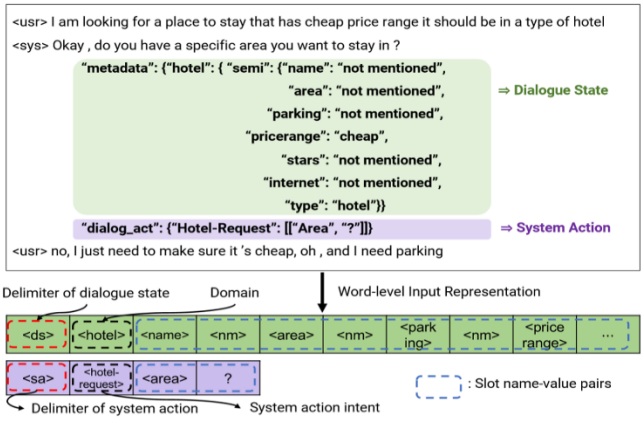


Figure 11: MultiWOZ dataset for test [11]

Dialogue state and system action, shown on Figure 11, is fundamental standard for the procedure of architecture. With personal assistant setting, this model also requires decoding strategy to optimize output sentence conditional on probability distribution. In this paper, top-k and top-p sampling is better than general beam search which are not suitable for high-entropy NLG and greedy decoding get highest probability over all probability such as for machine translation. To some extent, even though the database requires query to get real-world information tokens, there are not all sentence need external system to fill the slot, GPT-2 model will consider the situation to handle empty query for common sense responses and completion [12].

So far, progress in dialogue system in recent years due to the development in language model of natural language processing. In fact, most of techniques are for machine translation, sequence to sequence model and attention mechanism also the ones for machine translation. Until GPT-2, which natural has ability to generate text when the model is tested in below section. So, GPT-2 with transfer learning changeability for downstream natural language tasks with fine-tuning. Thus, we will use GPT-2 will be used as language model in the proposed model to obtain better response than seq2seq models.

After investigation and test of industrial products and some research projects about chatbot, with the development of AI technology, it was decided to use frameworks and transfer learning and fine-tune chatbot for English language learning. Chatbot in the proposed model will be a response generation chatbot with own common sense - the highest level is *free-style conversation agent* of a language system superior than other popular applications.

3. Methodology

In view of system framework, the design is firstly complied with the chatbot general flow at Figure 13, to decide whether it is functional. A top layer would be added followed by the application layer for systematically learning English which include three levels at User Interface designing scheme. As shown on Figure 12, the lower layer of architecture is the back-end model and the usage of language model using GPT-2 as language model responsible for the natural language understanding and natural language generation.

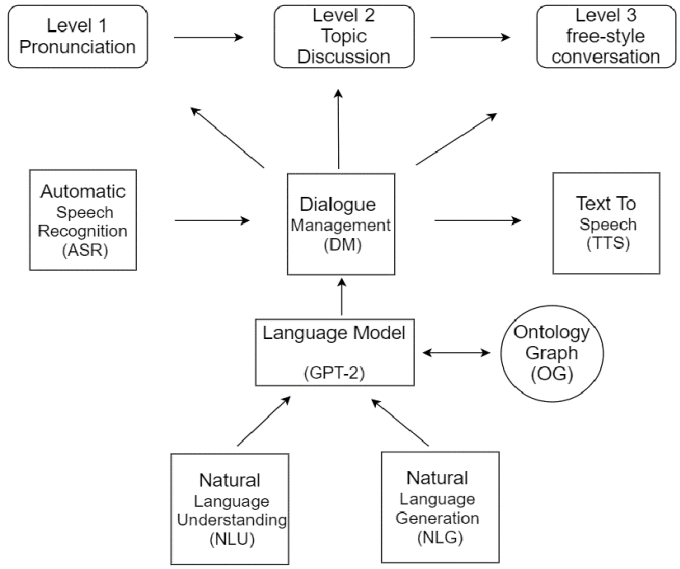


Figure 12: English learning chatbot architecture.

3.1. Chatbot system flow

In general, chatbot requires five parts to make communication with users and responses. The first step is voice recognition then change to text version natural language via ASR, the Automatic Speech Recognition is the same from text to voice in last step. For the central part, chatbot should handle natural language understanding and natural language generation using language model, it requires plenty of dataset as corpus to pre-train and fine-tune back-end model. About the optimization for daily dialogue and language model, with MySql database support the track of history dialogue distributed the user individually, the chatbot will become more intelligent and easier to modify with the long-term users.



Figure 13: System Flow of chatbot [13]

With the extension of system flow, in our architecture, the application layer are all English learning levels, which detail functions shown at implementation parts with mini-program User Interface and Develop Tool.

3.2. Open AI GPT-2 language model

Before we talk about Open-AI GPT-2, we should illustrate what is *language model (LM)*. For example, input editor is a language model which can predict the next word by language habits for behavior. From that, we can consider GPT-2 as a connection function of input editor. So far, GPT-2 already enlarge it language model with different size covering different needs of industrial NLP tasks.

Compared with the sequence to sequence we mentioned at section 2, the *transfer learning model* could provide better response that are more related to dialogue questions. The pre-trained dataset of Open AI GPT-2 are 40G text, it is no doubt that Open AI GPT-2 perform better than sequence to sequence model as it is data-driven neural network. The encoder part of GPT-2 will helps in the generation of natural language. Figure 14 shows the different size open source GPT-2 model for download and training.

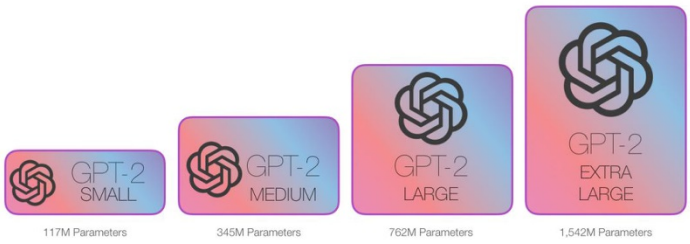


Figure 14: Size of Open AI GPT-2 pre-trained model [14]

Original GPT-2 always predicts the next token based on the information only at left hand of whole sentence. In summary, on probability condition, it uses history information with sufficient vocabulary to predict next token. After one prediction, it will provides a list for select, that is algorithm of top-k. In BERT, GPT-2 is also an architecture of transfer learning. Transfer learning consists of encoder and decoder, BERT focuses on the encoder and GPT-2 deletes encoder but keep decoder. Without encoder, decoder layers will be accumulated and magnified data. The idea of self-regression used by GPT-2 resemble RNN where it puts the token generated at the back of sequence, which the sequence will be the new input for the condition of next token generated.

In Figure 15, the encoders and decoders have 6 encoder and decoder. It removed traditional RNN and the pre-trained model fine-tuned by researchers will become mainstream in the future.

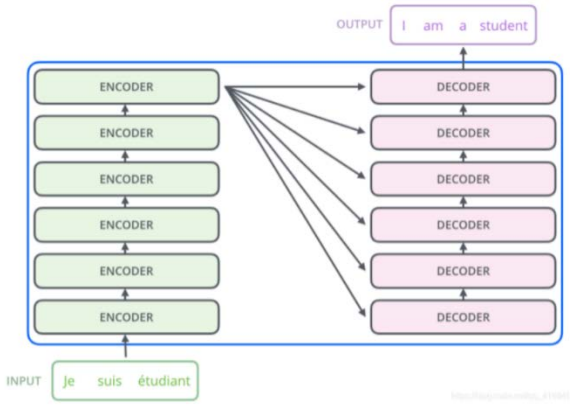


Figure 15: 6 layers encoder and decoder [14]

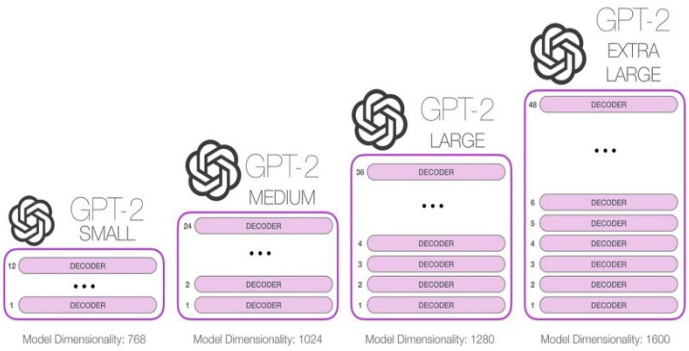


Figure 16: Decoder layers of Open AI GPT-2 [14]

The transfer learning change from BERT, shown on Figure 17, to GPT-2, the self-attention changes to masked self-attention shown on Figure 18. In Figure 16, due to GPT-2 have no encoder, input sentence as reminder is incomplete. The empty part of BERT covered with <eos> and other characters but GPT-2 with self-attention masked. So, it is masked self-attention. The impact for computation is only calculated the left words of token. The illustration of two different self-attention mechanism is shown in Figure 19.

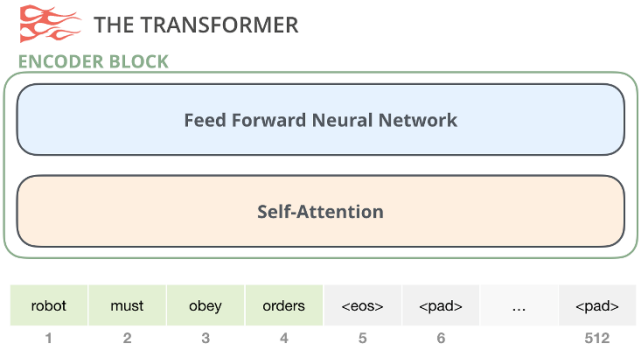


Figure 17: Encoder block of Google BERT [14]

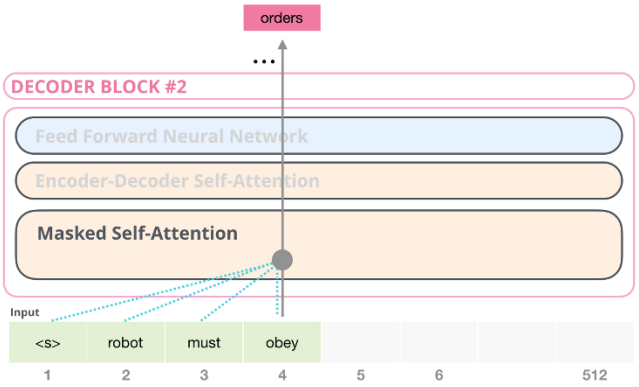


Figure 18: Decoder block of Open AI GPT-2 [14]

The masked self-attention GPT-2 used shows different with self-attention is that the masked one only know the left part of sentence, which include all information on left to predict next word to form the complete sentence even with articles. In addition, the token position encoding and embedding also considered into computation. From the start point of masked self-attention to the neural network to calculate the next word, the process will retain all the weight for the score of final values.

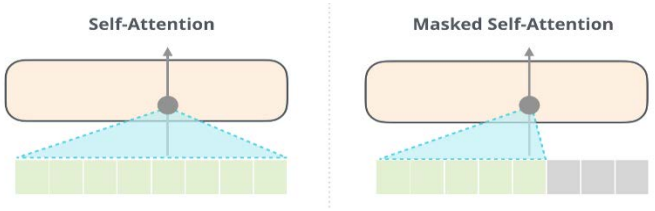


Figure 19: self-attention mechanism[14]

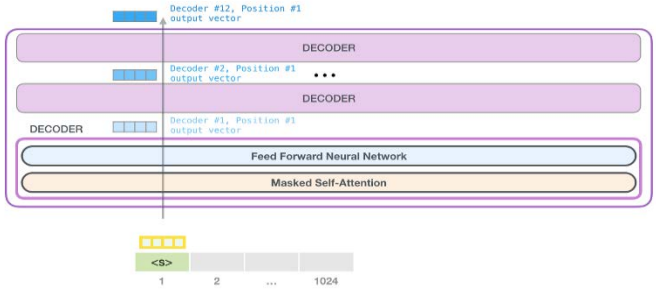


Figure 20: stack in Open AI GPT-2[14]

Figure 20 and 21 shows the process of prediction words. Sending a word to the first transfer learning block means looking up its embedding and adding up the positional encoding vector for position.

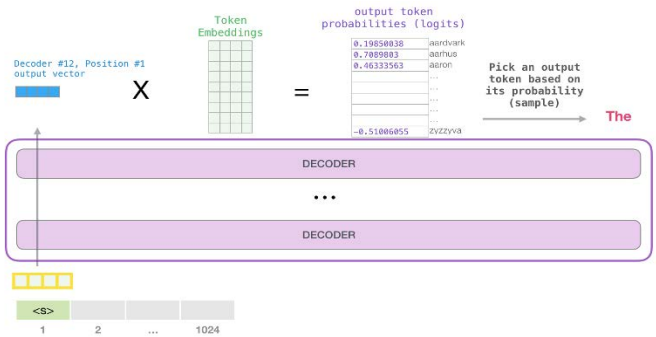


Figure 21: Open AI GPT-2 how to predict word in computation [14]

Initially, end-to-end and transfer learning are designed for machine translation with promising results. Since translation and dialogue are all sequence, more and more users used the same method to dialogue system. For research, only fine-tune with magnitude pre-trained model can obtain ideal language model for special tasks such as text classification, dialogue system etc.

4. System Implementation and Experimental Test

For the implementation of three levels English language learning, for users to access the system anytime, the platform of mini-program at WeChat for the front-end as User Interface to integrate our fine-tuned GPT-2 model based on the Google cloud server will be used in the proposed model. To realize the top layer applications of global architecture, Google speech recognition server is used as the recognizer of pronunciation due to Google has mature technology in speech recognition as compared with others. Level 1 includes the sentence distance comparison of recognized voice and sample sentence. Levels 2 and 3 use generate based language model to provide response to user. Level 2 is topic discussion from daily scenario with 8 topics is the representative of practical English. The highest level is called *free-style conversation* without any limitation for themes. In addition, the Open AI GPT-2 model also could process response to avoid meaningless response with common sense of English syntactic and semantic.

In Figure 22, our *user interface (UI)*, English learning agent has three levels for spoken English training. In this section, system usage and the different interaction performance levels with human users will be presented in this section.

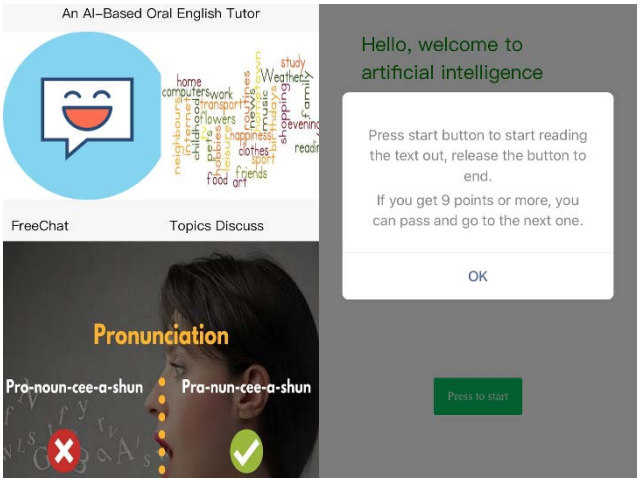


Figure 22: UI of English Learning Chatbot and Level 1 Pronunciation.

Pronunciation is the most important in English learning. This module is designed to practice. Once chatbot can recognize your voice to text and match the text read by user meaning that pronunciation is correct and can be recognized by a native speaker.

Firstly, the rule of pronunciation is the evaluation score shown on system more than 9 where the score comes from a native speaker's recognition set by Google speech recognition server. If user pass the pronunciation meaning that the words from users can be clearly recognized by voice recognition and match with the sample sentence. With the computation of voice recognized words and preinstalled sample sentence by *sequence matcher algorithm*, the essence of distance between two texts can be also calculated to distinguish the difference between user and the native speaker.

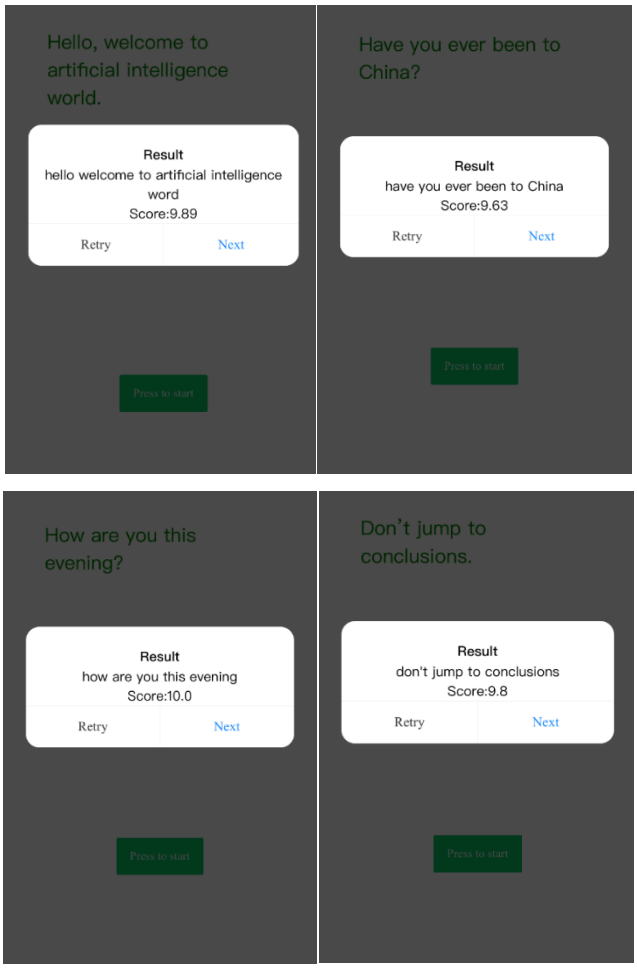


Figure 23: Test of Level 1 Pronunciation

During test in Figure 23, the first level only set 4 sample sentences as demonstration. After several rounds of test, the system gives higher score when user's voice is slow enough to distinguish every phonetic. Also, the testing or practice environment should be silent to avoid some acute noises. In addition, when different words with approximate phonetics such as 'word' and 'world', it will reduce the score pronunciation test in level 1, but the system may not transform voice to text accurately at level 2s and 3 which will influence the feedback from chatbot.

Second part is the process of a transfer learning model based on an Open-AI GPT-2 architecture so that the pre-trained model is able to generate high-quality text. The fine-tuned model used is *Transfer Transfo* where fine-tuned dataset is an open-domain daily dialogue. During the test, the UI settings are designed for dialogue scenario restricted in daily dialogue such as weather, travel, movies and other daily related activities.

For data training, the character always talks about movie, however, the range of movie contents are limited in US, with the influence of voice recognition. It is difficult to converse about the of movies contents. The model proposed has GPT-2t characteristic which according to the several reminders to write an article. Even the dialogue format limits the sentence length, it also follows the words from users to provide response and take no account of occasionally wrong input and fantastic output. Thus, the whole dialogue is worthy for further research. As mentioned, chatbot only focus on chatter. It has little knowledge base to generate meaningful sentences. So, the conversation contents cannot be limited within the topic as shown on Figure 24.

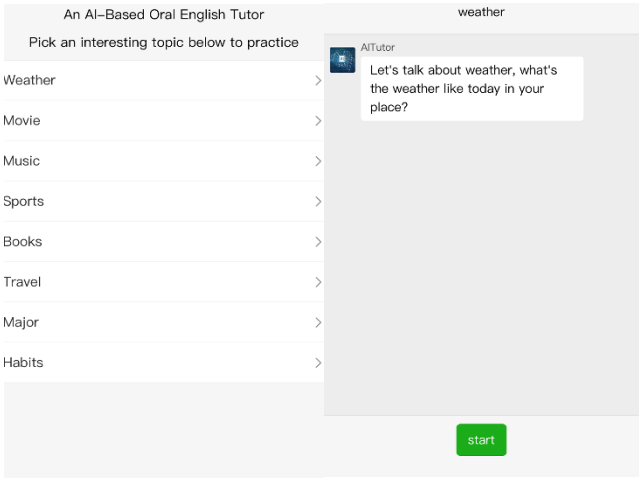


Figure 24: User Interface designed for our Chatbot Level 2 Topic Discussion.

In Figure 25, the level 2 test shows topics without limitation, in general, when the chatbot greets users, the topic are pre-established and dialogue are also limited in this topic. However, since same language model for levels 2 and 3 are used, the discussion after several rounds would exceed the original one and can shows other key-words related to tutor personality in random. At shown in the movie module, the topic would be changed from movie to music at second one.

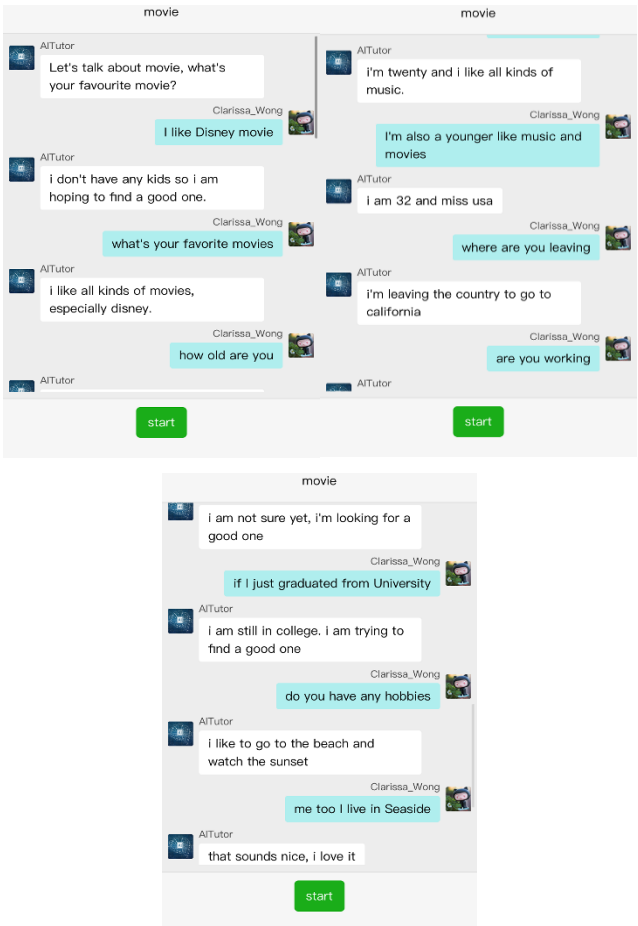


Figure 25: Test of Level 2 Topic Discussion.

At level 3 the same model is used for free chat. To be specific, the system cannot provide sustainable response about the query such as what, why or how. From the visualization of training dataset, the dialogue stayed at information level of the knowledge pyramid. Because the fine-tuned dataset is persona-chat dataset, it cannot provide response like specific *question-answer (QA) system* but prefer to chatbot direction with daily dialogue.

The highest level is free conversation. User and system can have several rounds of dialogue without any restriction . It was found that that the most response is related to fine-tuned dataset contents. Due to the transfer learning model and sequence to sequence model are all data-driven models, with more data used for training, the model has higher language level in syntax in English. Assuming that the text book of an *AI language tutor* is an English teacher, the fine-tuned data will be learned by the machine. After several rounds of test, the tutor uses text book contents to construct sentence content from 200M with high-quality grammar in English by data-driven of 40G text pre-trained.

According to the author of *Transfer Transfo*, the model is equipped with personality of responses generation. During test of Figure 26, every launch of system are different users with different personalities and personal information to chat with users.

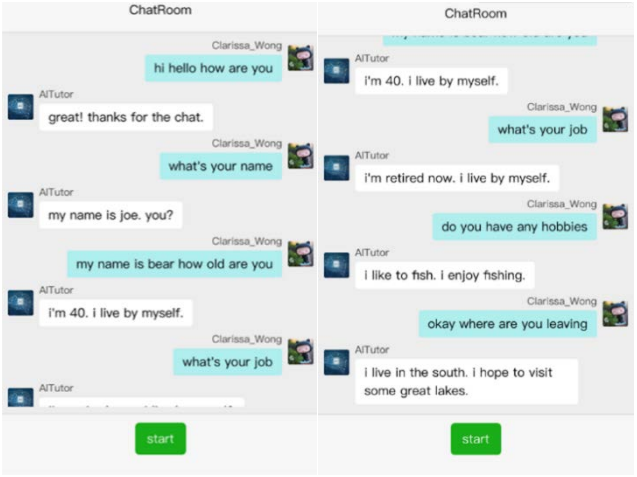


Figure 26: Test of Level 3 Free-Chat with different personality

The daily greetings initially similar to the chat with English teachers in some websites are face to face remote spoken English training. Users cannot guess the next sentence when it is unrelated to any practical topic. Every entry to chatroom has different personality, it is easier to follow the response to continue the subsequent rounds of dialogues.

Different with Level 2 *topic discussion*, Level 3 shows nothing is required to follow in UI. Users could converse anything with chatbot. However, in Figure 27, we can see the response is also limited in the corpus of fine-tune data even though the language model has 40G pre-trained corpus. Level 3 dialogue test results as above showed that English speaking practice with system, the contents cover normally personality with hobbies like music and reading at beginning . When the *AI tutor* ask the key-word 'work', the content will correspond to change and tie to some special words which is related to the key words. Also the contents will be changed from open domain chitchat to another topic proposed by the machine. In this test, "*animal*" and its related words from tutor can also be found at the fine-tune dataset. It was noted that when output consists of semantic of relations amount among different words.

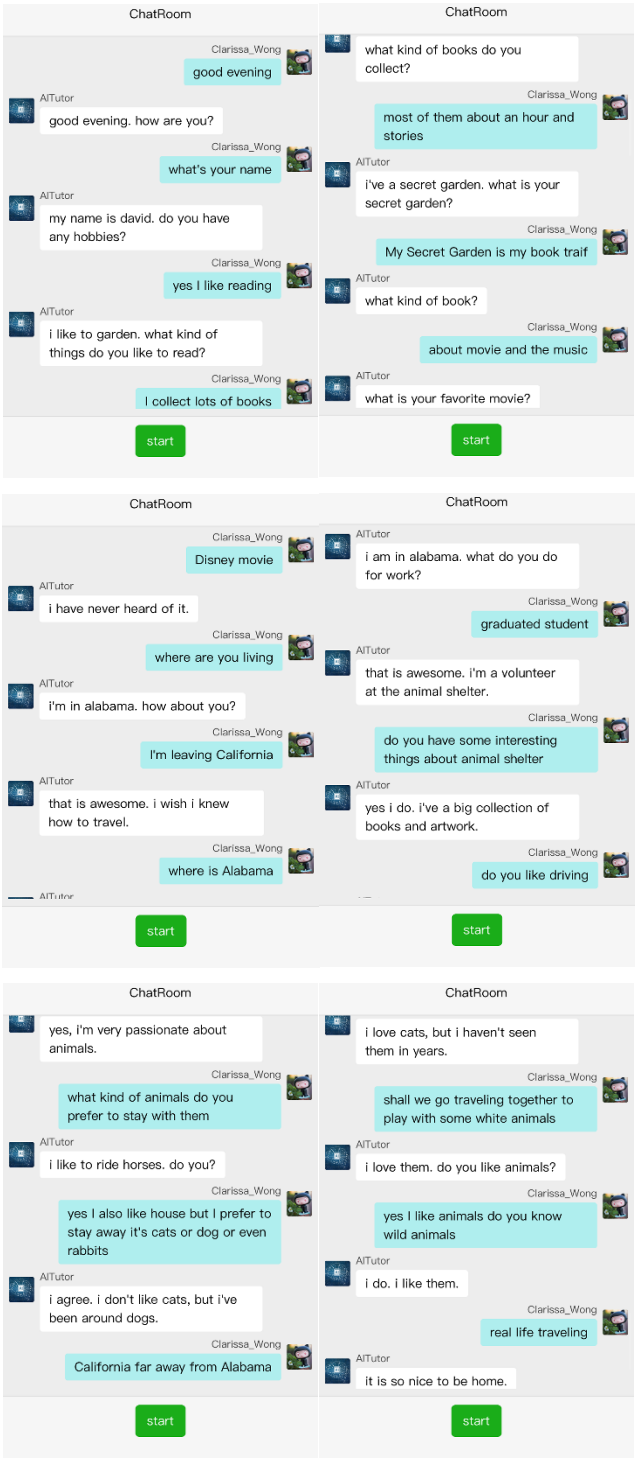


Figure 27: Test of Level 3 Free-Chat

5. Conclusion

In this paper, we used *transfer learning*, an advanced AI technology focusing on technology implementation and system construction to design and implement an *English language chatbot* which is cross-discipline project being proposed the first time and used by real-world university students for learning English round the clock. After around one year improvement, the stable back-end

supports the test with different users 24 hours. Our contribution and the system function shows:

- AI technologies in chatbot field can be well used and combine with the real-world learning needs, not limited in the usage only for personal assistant.
- With magnitude dataset and fine-tune corpus for machine against the AI model, it can generate the meaningful and good syntax response to imitate human response. It overcomes the problem with rule-based or knowledge search engine type chatbot relies on human to prepare the preliminary work, such as label or key-words matching construction. These two methods also range used by ongoing industrial chatbot.

For business aspects, especially in English learning market, there are many applications with splendid *user interface (UI)* and plenty of users declared that their back-end technology uses AI technology. However, during investigation and test commercial apps, the non-mature but with high-expectation neural network model required and relied on dataset quality. For users and system design, the advantages with learning language systematically combines three levels of phonetic, syntactic and semantic which are in both linguistic and the application of *natural language processing* of AI ecosystem.

Based on Turing test theory and *strong AI*, human-machine interaction agent is the main theme of the system. It is proficient in both grammar and knowledge grounds. An Open GPT-2 has already solved the problem to correct error character generated by RNN with pre-trained big data in grammar [3].

6. Future Work

Due to lack of hardware and server, our system performance and capability requires further improvement. At present, 8 GPU with GTX1080Ti on Tencent cloud is used to fine-tune GPT-2 with 345M objective data to make the language model with more meaningful entities to build a QA-based system. However, the computing capacity cannot reach the required magnitude. Also, the official GPT-3 is unavailable for download modification on our GPT-2 language model is continuous until the release of GPT-3. For further research, system optimization for system display is as follows:

- Find a higher capacity server for project implementation
- Optimize UI with better users' experiences and extend the three levels exercises with more options.
- Use 8 V100 GPU to train the largest GPT-2 with existing dataset to compare performance.

Acknowledgment

The authors would like to thank for UIC DST for the provision of computer equipment and facilities. This paper was supported by Research Grant R202008 of Beijing Normal University-Hong Kong Baptist University United International College (UIC) and Key Laboratory for Artificial Intelligence and Multi-Model Data Processing of Department of Education of Guangdong Province.

References

[1] T. Wolf, V. Sanh, J. Chaumond, C. Delangue, "TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents," (ii), 2019.

[2] R.S.T. Lee, Artificial Intelligence in Daily Life, 2020, doi:10.1007/978-981-15-7695-9.

[3] N. Shi, Q. Zeng, R. Lee, "Language Chatbot-The Design and Implementation of English Language Transfer Learning Agent Apps," in 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering, AUTEEE 2020, 403–407, 2020, doi:10.1109/AUTEEE50969.2020.9315567.

[4] Nazar Kvartalnyi, Gamification and Simulation in Education and Corporate Learning, Https://Inoxoft.Com/Gamification-and-Simulation-in-Education-and-Corporate-Learning/, 1, 2020.

[5] M. Qiu, F.L. Li, S. Wang, X. Gao, Y. Chen, W. Zhao, H. Chen, J. Huang, W. Chu, "AliMe chat: A sequence to sequence and rerank based chatbot engine," ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), **2**, 498–503, 2017, doi:10.18653/v1/P17-2079.

[6] M. Ghazvininejad, C. Brockett, M.W. Chang, B. Dolan, J. Gao, W.T. Yih, M. Galley, "A knowledge-grounded neural conversation model," 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, 5110–5117, 2018.

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, "Attention is all you need," Advances in Neural Information Processing Systems, **2017-Decem**(Nips), 5999–6009, 2017.

[8] L. Zhou, J. Gao, D. Li, H.Y. Shum, "The design and implementation of xiaoice, an empathetic social chatbot," Computational Linguistics, **46**(1), 53–93, 2020, doi:10.1162/COLI_a_00368.

[9] C. Agents, "applied sciences Human Annotated Dialogues Dataset for Natural Conversational Agents," 2020.

[10] S. Hoppe, M. Toussaint, "Qgraph-bounded Q-learning: Stabilizing Model-Free Off-Policy Deep Reinforcement Learning," 2020.

[11] D. Ham, J.-G. Lee, Y. Jang, K.-E. Kim, "End-to-End Neural Pipeline for Goal-Oriented Dialogue Systems using GPT-2," **2**(1), 583–592, 2020, doi:10.18653/v1/2020.acl-main.54.

[12] J.B. Tenenbaum, C. Kemp, T.L. Griffiths, N.D. Goodman, "How to grow a mind: Statistics, structure, and abstraction," Science, **331**(6022), 1279–1285, 2011, doi:10.1126/science.1192788.

[13] N. SHI, Q. Zeng, R. Lee, "The Design and Implementation of Language Learning Chatbot with XAI using Ontology and Transfer Learning," (Dm), 305–323, 2020, doi:10.5121/csit.2020.101124.

[14] Jay Alammar, The Illustrated GPT-2 (Visualizing Transformer Langua ge Models), Https://Jalammar.Github.Io/Illustrated-Gpt2/, 1, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 34-42 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Service Robot Management System for Business Improvement and Service Extension

Hideya Yoshiuchi*, Tomohiro Matsuda

Hitachi, Ltd., Research & Development Group, Center for Technology Innovation – Digital Platform, 1-280, Higashi-koigakubo, Tokyo,185-8601, Japan

ARTICLE INFO

Article history:
Received: 14 July, 2021
Accepted: 26 October, 2021
Online: 13 November, 2021

Keywords:
Service Robot System
Data Analysis
External Equipment Association

ABSTRACT

Service robots are applied in many kinds of business fields and its various kinds of data can be collected with service robot in business scenes. Therefore, it becomes more important to utilize such data for business improvement. Additionally, due to limitation of design such as body size and battery life, a service robot cannot prepare much peripheral equipment on its body. In this paper, we will show two research results as for service robot. One is data analysis technology for business improvement as a function of operation and management system for service robot. Another is service extension of service robot by association of external equipment with proper design of utilization condition of external equipment. Through evaluation experiment, we confirmed potential effect of business improvement is 8.1 % by modifying service scenario of robot.

1. Introduction

This paper is an extension of work originally presented in 5th International Conference on Robotics and Artificial Intelligence [1]. Due to the several social trends such as labor power and human resource issue, service robots technologies are being developed for realizing not only simple operation automation but also business improvement [1]. Recently, more people highly expect service robot system [2], [3] can solve social issues such as lack of labor power [4], improvement of productivity, human care [5] and so on. In manufacturing field, industrial robots were deployed and applied in 1970s for factory automation, and Internet of Things (IoT) technologies are utilized for productivity improvement and reduction of operation cost [6] with safe environment [7]. In service business field, service robots are expected to realize business and service quality improvement [8], [9] in the future aging society with fewer children. Based on these social and business background, we develop service robot and its application with intellectual media processing technology such as voice recognition and image analysis.

As for the robot deployment, people have higher attention to how to utilize various kinds of robots in actual fields for the purpose of collaboration with people. In industry field, Industry 4.0 [10] aims effective manufacturing with various technologies. Industrial robots are one of the most important technology in this fields. Industrial robots are designed to realize both production efficiency and safety for work with human in the same fields [11]. On the other hand, service robots are assumed to be applicable to many kinds of service and business scene through verbal and multimodal communication [12] not only for replacement of human labor load by automation but also for giving helpful function to people, therefore, applications for service robot have many variations for their purposes. For example, in a large facility such as an airport or a shopping mall, customers sometimes have difficulties for finding the way to their destination due to the scale and complexity of the facility. In such case, a service robot can give customers facility guidance information and support customers. In another case, a service robot can provide detail explanation of sold goods in a large shopping mall.

The value provided by service robots is considered in Business-to-Business-to-Consumer (B to B to C) scheme, and there are differences between the value of service provider (business side) expected and the value of consumer expected. Service providers apply service robots to their facilities for not only labor power but also some special skills such as multi-language guidance that normal employees cannot learn in short time. Consumers thinks service robots give them dedicated guidance. For example, when customers visit a large shopping mall for the first time and does not have enough knowledge for the mall, they ask a service robot where the shop that they are looking for. The service robot can give the location information by facility map, or the service robot can bring the customers to their destination by physical movement in the mall.

*Corresponding Author: Hideya Yoshiuchi, hideya.yoshiuchi.fq@hitachi.com

https://dx.doi.org/10.25046/aj060606

People consider that service robots have much potential for social innovation. On the other hand, especially service providers, people would like to know how effective service robots are. Therefore, it is important to evaluate and visualize the effect of service robot for service providers and consumers. To evaluate service robots' effect by a kind of quantum key performance indicator (KPI), we have to clarify what service provider expect to service robots for their business improvement in each service field. After specifying KPI for business improvement measurement, we consider service scenario for robots and verify the effect of service robots operated by the service scenario with monitoring the specified KPI.

In addition to business improvement, it is important to consider how to realize applications by service robots with limited capability. Because service robot has several limitations such as body size and battery duration, it is difficult to prepare enough devices for applications to be realized by service robot. In such case, we need to consider how to extend capability of service robot for extended service.

In this paper, we explain two technologies. One is the service robot system and data analysis technology for business improvement. The other is external device association technology for service extension with service robot system. For both proposed technology, we will show evaluation results. The structure of this paper is as follows: In chapter 2, we will show basic architecture of service robot system. In chapter 3, we will explain data analysis technology of service robot system for business improvement and its evaluation result. In chapter 4, We show external device association technology for service extension. Chapter 5 conclude this research.

2. Service Robot System

2.1. Summary of Service Robot System Architecture

Figure 1 shows general system architecture of service robot system. The service robot system consists of service robots and application servers that provides media processing function such as voice recognition, image processing and language processing. Additionally, the service robot system includes operation and management server that gives us management function for system operation such as machine condition monitoring and remote operation for emergency case. Because the main function of a service robot is service application through interaction with people, service robots are deployed in facilities of service providers. The methods of interaction between a service robot and people are sound and image. Therefore, a service robot has microphone and video camera for voice and image processing. A service robot system gathers voice and video data through microphone and video camera on a service robot, processes such media data, and provides basic interaction function to end users such as replying questions and giving useful information. Media processing function like voice recognition and video analysis consume much computational resources such as Central Processing Unit (CPU), sometimes media processing function is realized by application servers located out of a service robot.

Implementing media processing function by external application servers makes the size of a service robot smaller one. However, the data size of voice, image and video is relatively large and sending such media data from a service robot to application servers requires much larger network resources. Therefore, it is better to put application servers in the same service facilities as service robot's location. However, if it is not possible to put application servers in service facilities for some reasons such as facility size or no server room, application servers can be set in a management center.

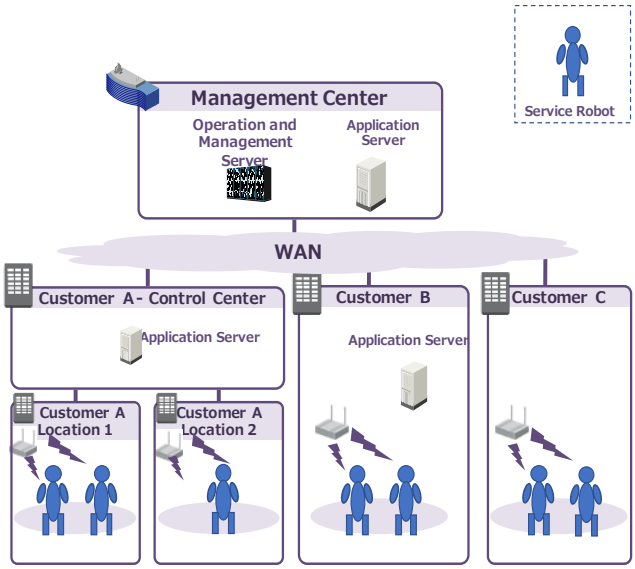


Figure 1: General architecture of service robot system

2.2. Research Task of Service Robot System

In service robot system, media processing is primitive function for interaction with end users. The combination of media processing realizes voice communication, and questions and answers. The one of the main purposes of developing service robot system is to organize service scenario by the set of communication and Q&A list for various kinds of business scene such as guidance in shopping mall or large transportation facilities (ex. airports and railway stations). Service scenario consists of timing information of voice communication, contents of communication, gestures and movement of a service robot. Service scenarios are designed for satisfying business and service requirement of service provider. It is difficult to evaluate the effect of service robot system before service robot system start its operation. Therefore, one of the research challenges of service robot system is to establish the evaluation method of the application of service robot system for assumed KPI suitable for business and service requirement, and proposal scheme of service scenario modification to realize business improvement. Besides KPI issues, we have to consider how to realize service application with limited resource of service robot. If we set many devices on service robot, its size is getting bigger, and weight is getting heavier. Because service robots work among peoples, sometimes in busy crowdy area, it needs to keep human safety [13], [14]. If robots are bigger and heavier, service robots easily violate human safety. Therefore, robot developer has to consider not only functionality realized by various devices, but also risk assessment for safer robot operation [15], [16].

3. Data Analysis Technology for Service Robot System

In this section, data analysis technology for service robot system that realizes business improve is explained.

3.1. System Architecture

Figure 2 shows the architecture of service robot system with data analysis technology for business improvement. This system consists of the following components.

- Service robot: Communicate with end users by interaction with voice and video, gesture and movement and provides various kinds of service to the end users
- External equipment and system: Realize additional function that the service robot cannot provide such as GUI by digital signage.
- Application server: Providing media processing primitive function such as voice, video and language processing.
- Integrated service system: Store and manage all the data that can be collected in ser\vice robot system. For the detail of collected later, we will explain in later section
- Business analysis server: Provide data analysis function for business improvement with all the data stored in the integrated service system.
- Customer business system: Various external system for customers' business purpose such as marketing, accounting and so on.

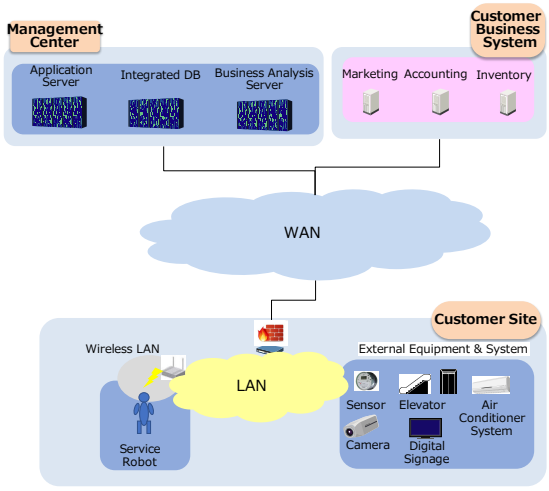


Figure 2 Architecture of Service Robot System with Data Analysis Technology

3.2. Data Analysis for Business Improvement

In service robot system, service robot provides various functions to end user by service scenario. The possible actions of a service robot consist of communication method by voice, image and language processing, Service scenario combines these actions. From the viewpoint of service provider, they expect service robots to do some business operations instead of their staff. The typical business operation realized by service scenario is guidance activity in commercial facilities or transportation facilities.

To realize business improvement, KPI related to business operation is measured and monitored. KPIs are indicators that evaluates business activities. For example, ISO 22400 defines important KPIs for manufacturing field mainly focusing on Manufacturing Execution System (MES) [17]. Commonly, KPIs focus on the following items:

- Cost: Cost KPI allows you to evaluate the activities that utilize economic resources in different processes. This can include energy consumption, cost of raw materials, cost of human resources and overhead costs.
- Time: Time KPI measures the time spent on an activity or a process. This can be the time from the beginning to the end of a process or a step of a process. Ways to measure time can be by the percentage of products or services delivered on time or the percentage of late deliveries.
- Productivity: Productivity KPI allows your organization to measure the efficiency of the operation. This can include measuring the percentage of downtime your equipment or machine experience. It also can be an output of operations such as product items or the number of served customers.
- Quality: Quality KPI help calculates the effectiveness of processes. These numbers may be based on the number of errors occurring and the number of completed processes. Other way to measure is the percentage of products that need corrective action and the number of satisfied customers.

Service robot applications have different tendency from industrial robot use case. Service robots are mainly used for support of end user's activities in service facilities such as commercial mall, railway station, airport, hospital and so on. Service robots mainly provide intellectual works to those customers; therefore, it is not always suitable to measure KPIs related to manufacturing productivity or automation activities. Suitable KPI candidates are service-related value such as service quality, the number of users that communicate with service robots and so on. After we decided to utilize service robot into our service environment, we have to choose KPIs that have relationship with service robot application. To analyze how the behavior of a service robot affects the KPIs, we need to monitor and store the data which describe service robots' activities. As for KPI evaluation, one or more service scenario affects the KPI value change. For example, daily sales information of commercial complex facilities is a sum of daily sales of each tenant. Some shop items are very popular and have strong effect to the daily sales of the tenant. In another viewpoint, daily sales information has relationship with likability of a facility.

To realize business improvement measured by KPIs, it is effective to figure out the service scenario that has strong relationship with business KPI and try to apply the service scenario as much as possible to service robots. For this purpose, we need to gather the enough data from service robot system for analysis of relationship between business KPI and service scenario. Then, we consider how to realize KPI improvement by modifying service scenario that much affects the KPI. The modification ways of the service scenarios are just candidate method and there is no assured result. Therefore, we must apply

the modified service scenario to service robots and evaluate its efficiency by checking KPI increase or decrease. If the KPI is improved, we decide to apply scenario modification permanently. If the KPI decreases, we discard scenario modification and try another modification. The data type to be collected in service robot system for business improvement is as follows:

- KPIs to be improved as time-series data
- Service scenarios that are applied to service robots
- The type of service scenario
- Relationship between the KPI and service scenarios
- Execution result of service scenario
- Action history of service robots as time-series data
- Motion history of servicer robots: raising a hand, nod, showing bye-bye, etc.
- Image processing results
- Voicer recognition results
- Statements of service robots corresponding to the end users

Here we show the detail of data analysis process of service robot system for business improvement.

- Select one KPI for business improvement that a service provider applying service robot system would like to improve.
- Extract service scenario for service robot that has strong relationship with the KPI to be improved. The service scenario that gives very good effect to the KPI to be solved is selected the scenario that has possibilities for further business improvement.
- Analyze the data that relates to the extracted service scenario and decide whether the execution result of service scenario gives good or bad effect. By analyzing the result of service scenario execution and KPI change, we can understand tendency of service scenario and KPI behavior.
- Consider how to modify the extracted service scenario for business improvement: One of the ways to realize business improvement is to increase service scenario execution for good result.
- Modify the service scenario according to the way of step 4 and deploy the modified service scenario. After that, do evaluation test for a pre-determined term.
- Evaluating the KPI to be improved. If the KPI is improved, the service provider permanently applies the modification of the service scenario. If KPI does not change or show bad result, the service provider discards the modification and try to consider another modification way.

3.3. Evaluation

We applied service robot system to the exhibition. The task of a service robot is guidance of our company booth in the exhibition hall.

Figure 3 shows the layout of exhibition hall. The exhibition was held in 2 days and around 40,000 visitors came. In the exhibition, there were 171 topics are shown within 7 business categories. The service robot was set only in one category booth. It has microphone, speaker and camera. By utilizing these devices, the service robot can communicate with people through verbal way. The service robot also has wheel and it can move around the exhibition hall, however, due to safety issue, we do not use moving function of the service robot. The contents of guidance are concierge service. The service robot waits for the visitors at the entrance of the booth and give the information related to the exhibition contents or facilities when visitors get close to the service robot. Additionally, stage show is periodically done and the service robot make announcement when the stage show begins soon. The basic behavior of the service robot for the exhibition is as follows:

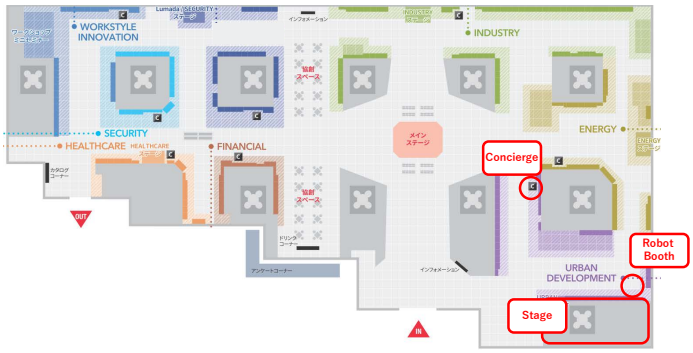


Figure 3: Architecture of Service Robot System with Data Analysis Technology

- The service robot waits for visitors' arrival at the entrance of the booth
- If there is no visitor, the service robot make announcement for its own service and try to get attention of visitors: "Hello, here we show our company's technology and service. Please come and visit our booth!"
- If the stage demo begins soon, the service robot makes announcement for the stage demo: "Our stage show starts soon. Please come and see the show!"
- When a visitor comes in front of the service robot, the service robot starts to talk and ask the visitor whether they have some questions or not.
- When the visitor asks a question, the service robot try to understand the question and reply answer. The assumed questions are questions for the exhibition contents, facility guidance and general question such as daily conversation.
- After answering a question, the service robot asks the visitor whether the visitor still have question or not.
- If the visitor has more question, the service robot continues to answer the question. If the visitor has no more question, the service robot finishes the service for the current visitor and wait for another visitor's arrival.

The service scerio of the service robot for the exhibition includes the following items:

- Announcement for the visitors
- Physical Motions that attract visitors' attention and promote natural communication
- Understanding the question by voice recognition and language processing
- Answering the question

Service robot has two kind of tasks. The important task is guidance of exhibitions and it has higher priority than other task. When the service robot detects somebody's talking, it start guidance service to the person in front of the service robot. Service robots are recently deployed in actual fields. It has basic communication function through microphone and voice recognition technology. However, if people has no experience to communicate with service robot, they might have difficulty to let them understand what people say. The communication of service robots are controlled by computer program and it has detection function of human utterance section. Because not all the people understand this detection mechanism, sometimes communication gap happens between service robot and human. Semiotics[18] is an approach for smooth mutual communication. It sometimes applies sign or non-verbal communication method. In our system design, we applied sign of service robot to let the people understand the service robot is ready for communicating with the customers. For example, a service robot aim his face to the customer or raise its hand to start communication in natural way. The data related to the previous listed items are recorded as time-series data and we used the data for business improvement analysis. For the business improvement analysis, there are two important tasks.

- Clarifying the KPI to be improved and related service scenario
- Making data formatting for data analysis of service scenario related data

Several KPI is suitable for evaluating the effect of exhibition. In this case, we focused on the understanding contents of exhibition by visitors. We set up one service robot in exhibition booth. It is difficult to measure how much the visitor understands exhibition contents, we chose the number of question and contents of question as a KPI that reflects visitors' interest to the exhibition. We classified the contents of question by judging how much the question relates to the exhibitions and monitored how the service robot behaves according to the service scenario. we recorded the data of service robot action for each visitor.



Figure 4: Structure of Concierge Service Data for Exhibition

Figure 4 shows the architecture of service robot action data for exhibitions. The data is classified into two categories. One is behavior with visitors, and the other is behavior without visitors. To separate time-series data of the service robot for each visitor, we used voice recognition result with specific words to indicate service start and end for a visitor. For example, when the service robot starts to ask question, the service robot says, "Are there any question?". In the exhibition, the service robot serves to visitors for two hours and 25 minutes for 37 visitors. After we prepare data for each visitor, we extract the feature values that represents the data. In this evaluation, we chose the following items:

- The number of speaking of the service robot
- The number of voice recognition of the service robot (that means the number of speaking of a visitor)
- The number of motion that the service robot executes.
- The number of repeated questions
- Service time for each visitor

When the service robot cannot understand the meaning of the visitor's question, the service robot asks the visitor to repeat the same question once more. We calculated the feature values shown above, then we classified the result of service scenario for each visitor. We considered the best result of service scenario is directly related to the exhibition content because the purpose of service robot system for the exhibition is to let visitors understand the exhibition contents (that is, our company's technology and service). However, service robots are just now in the beginning stage and start to be deployed in actual world. Therefore, some visitors do not know what to ask to the service robot. In such case, visitors make the question for facilities such as "Where is a rest room?" or start daily life conversation. We inspected all the data for 37 visitors and classified the result of service scenario into the following four categories, then we give the label for the data for each visitor to represent the category.

- Answer the question for the exhibition contents
- Answer the question for facility information (such as restroom or exit)
- Start daily conversation
- Cannot understand the visitor's question and does not answer anything

Smaller number means visitors have higher interest for the exhibition contents. After data classification, we did cluster analysis with K-means method. We analyzed the clustering result with changing the number of clusters. When we set the number of clusters three, we confirmed the following tendencies.

- Cluster 1 (data of 22 visitors belongs to): In this class, the service time is relatively short, few visitors repeated their questions. The service robot answers a few questions.
- Cluster 2(data of three visitors belongs to): Service time is quite long. The visitor repeated question and asked so many questions, and the service robot answered all the related questions.
- Cluster 3 (data of 12 visitors belongs to): The service time is middle. In this class, visitors keep silent and the service robot try to make the visitor do some questions. However, visitors seem to have no idea and do not know what to be asked. Finally, the service robot regarded as the visitor has no question and ends the service for this visitor.

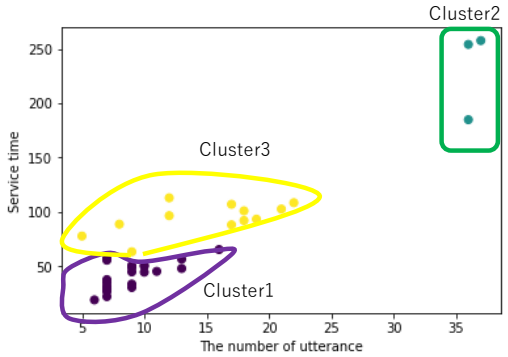


Figure 5: Clustering Result of Exhibition Data

Figure 5 shows the result of clustering. In Figure 4, service time and the number of utterance of the service robot are chosen as the

evaluation axis that most reflects the clustering result. In cluster 1, visitors and the service robot make smooth conversation. In some case, a visitor repeated the question, however, in most case, a visitor asked the exhibition contents and the service robot properly answered the question. In cluster 2, visitor seems to have some knowledge for service robots, and try to figure out what kind of questions. In this case, the service robot recognizes the visitors' question with high accuracy and the visitor and the service robot can talk long time. In cluster 3, the service robot cannot understand the questions, or the visitor does not know what to say.

We made detail analysis of the data classified into cluster 3 and figure out two major cases.

- Case1: The service robot failed to understand what the visitor says. The visitor quitted to communicate with the service robot.
- Case2: The visitor does not know what to ask to the service robot and there exists long silence time. Even though the service robot talks to the visitors and try to make the visitor ask some questions, the visitor has no idea and conversation is terminated.

As for the number of cluster, we have tested several cases. When we set the number of cluster as 4, we confirmed the samples included in cluster 3 and cluster 4 show the similar features. Therefore, we chose 3 as the number of the cluster. In case 1, main reason is the accuracy of voice recognition and technology improvement is needed. In case 2, the issue is scenario design problem. In the evaluation experiment, the service robot just asked visitors whether the visitor has question or not. However, if we designed the service robot gives more detail information such as what kind of question is assumed, we can expect the service robots behaves to visitors more effective way. For case 2, three visitors are classified in this case. We defined service quality as the ratio of number of the visitors that the service robot properly gives service for exhibition demo. The number of whole visitor for service robot is 37, that is, the sum of the data in all the clusters. If we can modify the service scenario and give better guidance to these three visitors, we can expect 8.1% (3/37) business improvement.

This is our first trial of service scenario analysis. Through this field test, we found the following issue to be solved in the future.

- In some cases, the service robot cannot correctly understand human speaking contents. In such case, we should provide alternative communication method by remote operation by human.
- Information on voice is sometimes not correctly understood by listener. In such case, visualizing method is useful for proper information providing. Since our service robot used in this field test does not have monitor, we should consider how to associate external device and the service robot in this application.

For further service scenario improvement, it is important to set target value of selected KPI, and making evaluation with Plan-Do-Check-Action cycle. Through continuous improvement, we can figure out root causes of the problem by utilizing analytic tools such as fishbone diagrams [19].

4. Service Extension Technology with External Equipment Association

In this chapter, we explain external device association technology for service robot system to realize service extension.

4.1. Functions for External Equipment Association

Figure 6 shows function blocks of external device association technology for the service robot system. Application server provides the following functions to realize external device association for service robots.

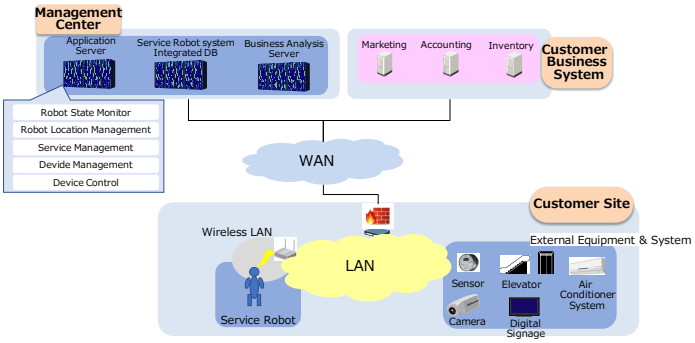


Figure 6: Functions for External Equipment Association

- Robot State Monitor: Collecting information on the status of various parts (motors, cameras, etc.) of service robot, then monitoring and visualizing the operating status of service robots. When some error occurs, this function notifies an alert.
- Robot Location Management: Managing location information of service robots based on the map information of a facility where service robots are working.
- Service Management: Managing services provided in facilities where service robots are installed. It includes the list of external equipment required for each service.
- Equipment Management: Managing equipment information about the installation location of external equipment in a facility and the conditions of use of external devices.
- Device Control: Provides an interface to control external equipment to realize various operation.

4.2. Usage of External Equipment with Association Condition

Service robot has several limitations on its design. Typical one is body size. If we set many peripheral equipment on a service robot, its size gets bigger and bigger. Even though we satisfy size limitation with great effort, many peripherals consume much electric power, and it requires larger battery. Therefore, service robot cannot prepare all the equipment for service realization and some equipment should be prepared as external devices. To solve this issue, we propose service management function that manages a list of devices required for services. The contents of service information are shown in Table 1.

For example, if a service robot does not have a display, an external display must be used to display visual information. The

facility guidance service using a map is a typical example, and the facility guidance service can be provided only when the external display is available for a service robot.

Table 1: Service Information

Item name	Description
Service Identifier	Unique identifier of a service
Device List	Device list that is essential for a service providing (ex. camera, display, etc.)

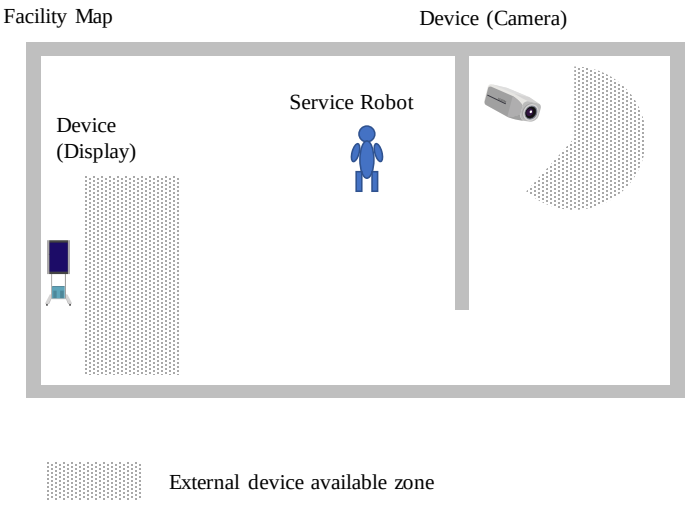


Figure 7: Example of external device availability

When a service robot starts to provide service, the service robot must check whether all the external devices necessary for the service are available or not in advance. The availability of external device depends on the position of external devices and a service robot. Also, we should consider whether the aimed devices are free and not used by other service robot. Figure 7 illustrates the positional relationship between a service robot and external devices. In this example, two external devices, a display and a surveillance camera, are installed in a facility. Taking a display as an example, for a service robot to use the display for facility information guidance service to end users such as facility users, the service robot must be close to the display. In addition, since the display has a display surface, the service robot needs to take care so that the display surface of the display can be seen by the end user. Based on this assumption, the available zone of the display is shown as the shaded area in Figure 7. The available zone of the camera differs depending on the service used, but if the recording area of the camera needs to include EMIEW, the available zone of the camera is the fan-shaped shaded area on the upper right of Figure 7. In either case, it is not enough to simply consider the physical distance between the service robot and the external devices.

Table 2 is a list of information managed by equipment management function. The external device has X and Y coordinates on the map created by using a laser range finder or the like as the position information of the device in the facility. In addition, information of available zone of external devices is stored as a list of parameters that express the range of area. Using these parameters, a service robot realizes extended service with external devices by the following procedure shown in Figure 8.

Table 2: Item List of Device Management

Item name	Description
Device Identifier	Unique identifier of a device
Device Location	Device location in a facility: X and Y coordination in a map
Device Availability Parameter List	Device list that is essential for a service providing (ex. camera, display, etc.)

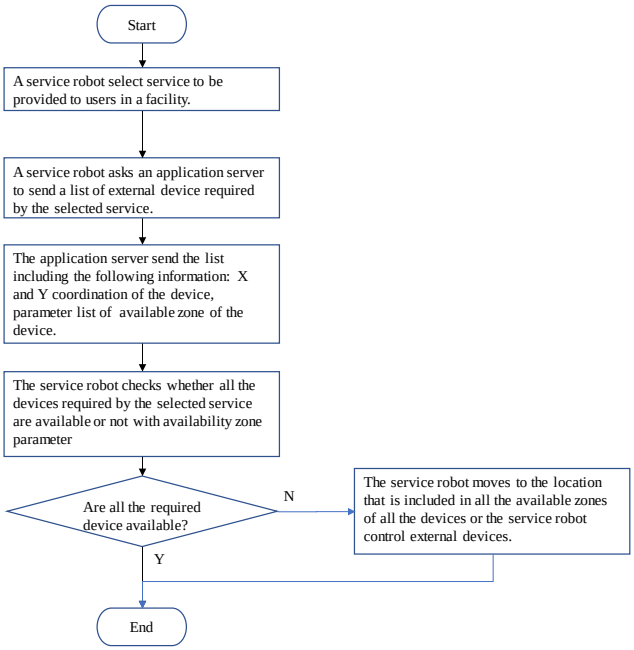


Figure 8: Flow chart of device availability check

4.3. Evaluation

In order to measure the effect of the proposed method, consider a concrete application example in a business field. As one of the typical application examples, let us consider a product and facility guidance service in a commercial facility. A typical commercial facility in Japan opens at 10 am and closes at 9:00 pm. A service robot provides shoppers with on-site facility guidance and product explanation services for 11 hours when commercial facilities are open. Figure 9 shows the timetable for commercial facilities.

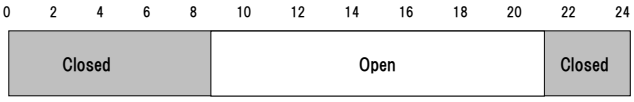


Figure 9: Timetable of the Shopping Mall

Figure 10 shows the layout of the shopping mall. A shopping mall consists of an information center that guides visitors to the facility and a shopping area that sells products. A service robot shall provide facility guidance using digital signage at the information center and product explanations at the shopping area.

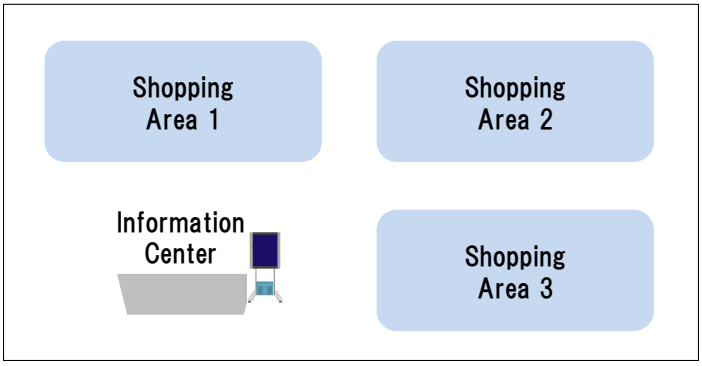


Figure 10: Layout of the shopping mall

Without the proposed method, the service robot cannot determine available services by comparing its location with the location of external devices. For this reason, shopping mall owner must set service robots at each location that provides different services. In this case, the arrangement of service robots is shown in Figure 11. In this case, 4 service robots are required, one in the information center and one in each shopping area.

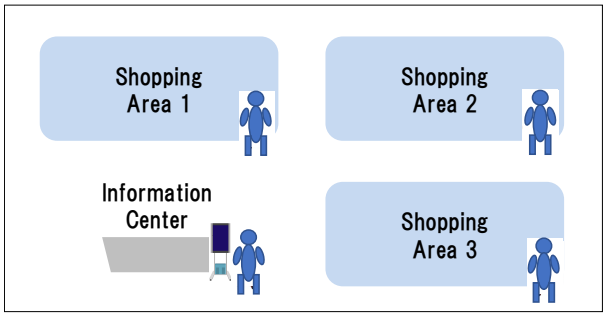


Figure 11: Arrangement of Service Robots with Conventional Technology

When applying the external device association technology, service robots can check whether its position is the information center and near the signage, or the location is the shopping center and away from the signage. Therefore, after providing facility information at the information center, it is possible to move to the shopping area and provide product information if necessary. We show the arrangement of service robots in this case in Figure 12. As for the association among service robots and signage, Simultaneous Localization And Mapping (SLAM) [20] technology is applied. The utilization area of signage is indicated in robot map managed by application server. The robot map has both signage utilization area and robot current location and application server can check whether signage is available by robot or not based on the robot location and signage condition. If robot is not at near signage, robot cannot use signage. Also, if the signage is used by another robot, signage is not available, too.

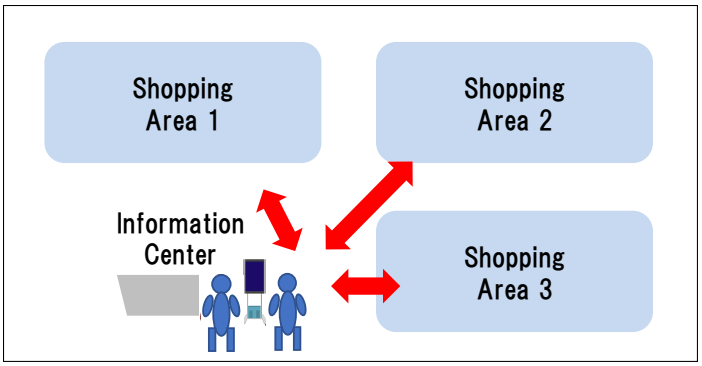


Figure 12: Arrangement of EMIEWs with Proposed Technology

The following conditions are assumed in the shopping mall.

- Time required for commercial facility guidance: 3 minutes
- Travel time between shopping area and information center: 1 minute
- Time required for product description: 5 minutes
- Visitor frequency: Once every 5 minutes

It takes 10 minutes for a service robot to move to the shopping area, explain the product, and return to the original position after guiding the commercial facility at the information center. If the customer's visit frequency is once every 5 minutes, it is possible to respond to customers without waiting for the customer with two service robots. In this case, the number of service robots installed in one commercial facility can be reduced by 50% compared to the conventional technology.

5. Conclusions

We developed data analysis technology of service robot system for business improvement and service extension with external device association with service robots. Through the evaluation by exhibition case, we found out the reason that the service robot failed to serve to the visitors, and how to modify the service scenario. 8.1% business improvement can be expected for exhibition case. In another case, we consider facility and product guidance service in a commercial facility and confirmed our proposed external device association technology can reduce the number of service robots set in a facility 50% compared with conventional technology.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] H.Yoshiuchi, T. Matsuda, and J. Dai, " Data Analysis Technology of Service Robot System for Business Improvement," Proceedings of the 2019 5th International Conference on Robotics and Artificial Intelligence (ICRAI), 7-11, 2019, doi:10.1145/3373724.3373733

[2] Z. Zhaohui, M. Xuesong, B. Xu, C. Hanghang and T. Jian, "Development of an intelligent interaction service robot using ROS," Proceedings of 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, IMCEC 2016, 2017, doi:10.1109/IMCEC.2016.7867516

[3] C. Park, S. Kang, J. Kim and J. Oh, "A study on service robot system for elder care," 2012 9th International Conference on Ubiquitous Robots and

Ambient Intelligence (URAI), IEEE, 2012, doi;10.1109/URAI.2012.6463070
[4] T. Jyh-Hwa and L. Su Kuo, "The Development of the Restaurant Service Mobile Robot with a Laser Positioning System," Proceedings of the 27th Chinese Control Conference, IEEE, 2008, doi:10.1109/CHICC.2008.4605839
[5] H. S. Ahn, J. Choi, H. Moon, M. Jang, S. S. Kwak and Y. Lim, "Social Human-Robot Interaction of Human-Care Service Robots," 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019, doi:10.1109/HRI.2019.8673282
[6] D. J. Ahn, J. Jeong and S. Lee, "A Novel Cloud-based Fog Computing Network Architecture for Smart Factory Big data Applications," 2018 South-Eastern European Design Automation, Computer Engineering, Computer Networks and Society Media Conference (SEEDA_CECNSM), 2018, doi:10.23919/SEEDA-CECNSM.2018.8544946
[7] M. Dohi, K. Okada, I. Maeda, S. Fujitani and T. Fujita, "Proposal of Collaboration Safety in a Coexistence Environment of Human and Robots," 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, doi:10.1109/ICRA.2018.8460869
[8] T. Savickas and O. Vasilecas, "Business Process Event Log Use for Activity Sequence Analysis," 2015 Open Conference of Electrical, Electronic and Information Sciences (eStream), 2015, doi:10.1109/eStream.2015.7119500
[9] G. Park, L. Chung, L. Khan and S. Park, "A Modeling Framework for Business Process Reengineering Using Big Data Analytics and a Goal-orientation," 2017 11th International Conference on Research Challenges in Information Science (RCIS), 2017, doi: 10.1109/RCIS.2017.7956514
[10] K. Zhou, T. Liu and L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges," 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2147-2152, 2015, doi: 10.1109/FSKD.2015.7382284.
[11] J. A. Marvel, J. Falco and I. Marstio, "Characterizing Task-Based Human–Robot Collaboration Safety in Manufacturing," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **45**(2), 260-275, 2015, doi: 10.1109/TSMC.2014.2337275.
[12] Y. Kim and W. C. Yoon, "Generating Task-Oriented Interactions of Service Robots," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **44**(8) , 2014, doi: 10.1109/TSMC.2014.2298214
[13] G. Yu, S. Hong, K. Sung and J. Seo, "A study on the risk investigation and safety of personal care robots," 2017 17th International Conference on Control, Automation and Systems (ICCAS), 904-908, 2017, doi:10.23919/ICCAS.2017.8204353
[14] A. J. Kim, B. B. Kang, C. J. Shin, D. H. Kim and E. S. Jang, "A Study of Safety Requirements of Service Robot Appliances for Household and Similar Use," 2019 19th International Conference on Control, Automation and Systems (ICCAS), 1410-1411, 2019, doi:10.23919/ICCAS47443.2019.8971666
[15] W. Wang, Z. Wang, F. Gao and Y. Ma, "Design of Wireless Remote Control Service Robot," 2020 IEEE International Conference on Information Technology,Big Data and Artificial Intelligence (ICIBA), 651-654, 2020, doi: 10.1109/ICIBA50161.2020.9276882
[16] S. B. Lee and S. Hun Yoo, "Design of the companion robot interaction for supporting major tasks of the elderly," 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), 655-659, 2017,doi: 10.1109/URAI.2017.7992695
[17] K. Kumagai, M. Fujishima, H. Yoneda, S. Chino, S. Ueda, A. Ito, T. Ono, H. Yoshida and H. Machida, "KPI element information model (KEI Model) for ISO22400 using OPC UA, FDT, PLCopen and AutomationML," 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 602-604, 2017, doi: 10.23919/SICE.2017.8105550.
[18] W. Takano and Y. Nakamura, "Statistically integrated semiotics that enables mutual inference between linguistic and behavioral symbols for humanoid robots," 2009 IEEE International Conference on Robotics and Automation, 646-652, 2009, doi: 10.1109/ROBOT.2009.5152574.
[19] O. A. Makinde, T. Munyai and B. I. Ramatsetse, "Establishing suitable process improvement methodologies for optimizing servicing operations in the banking industries," 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 860-864, 2017, doi: 10.1109/IEEM.2017.8290014.
[20] F. Hashikawa and K. Morioka, "Mobile robot navigation based on interactive SLAM with an intelligent space," 2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), 788-789, 2011, doi: 10.1109/URAI.2011.6146017.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 40-44 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Plummeting Makespan by Proficient Workflow Scheduling in Cloud Environment

Juhi Singh, Shalini Agarwal*

Department of Computer Science, Shri Ramswaroop Memorial University, Lucknow, 225003, India

ARTICLE INFO

Article history:
Received: 21 January, 2021
Accepted: 12 April, 2021
Online: 05 May, 2021

Keywords:
Cloud Computing
Deadline
Decisive Path
Makespan
Scheduling
Workflow

ABSTRACT

Cloud is an Internet-based computing technology in which on-demand shared resources such as software, platforms, repositories, and information are delivered to customers. In the emerging era of computing cloud environment provide the use of resources with the concept of virtualization. Workflow of the tasks has vital role for the improvement of computing performance which leads to improved quality of service. As per the demand of user's number of tasks are scheduled in such a way so that better performance is computed using partial deadline of the workflow. In this paper we have introduced with the workflow concepts, further we aim to diminish makespan for the proposed workflow scheduling algorithm. Here makespan refers to overall time duration taken for the sequence of tasks, by the resources so as to complete the execution of each and every task.

1. Introduction

Every day, everyone is connected in one way or another to this digital world, and this is because the field of information technology is escalating. User-friendly environment is the main factor of internet, so that diverse groups of individuals like students, researchers and business people complete their work by providing numerous opportunities to accomplish their goals. Many users connect to the Internet and use their IT infrastructure to meet their daily needs. As the demand for the internet increases, the services provided over the internet such as software, platforms, database services, storage services, etc. are also escalating. Here the imperative term cloud computing comes into existence. It provides countless diverse services to its users over a network. Because of concept "Pay as you Go", end user can get the most out of using this service at a lower cost.

1.1. Workflow

The increasing demand and heterogeneity of cloud computing is gaining recognition among the scientific commune to leverage cloud services to implement large-scale electronic applications. These applications are in the form of a set of tasks representing workflow. Computations performed considering task dependencies are regarded as Workflows. It allows users to straightforwardly elucidate multifaceted multi-step calculation tasks. Workflow tasks are associated to the mechanization of procedures where tasks, information or documents are delivered between partakers in accordance with a specific policy set and allow the formation of different applications in Directed Acyclic Graph. Each and every in DAG is represented as task and each edge symbolizes the dependencies [1], [2].

A mechanism to manage various workflows on cloud, is known as Workflow Management System (WMS). Figure 1 depicts some significant constituents of WMS. Workflow can be categorized as:

- Single workflow - It includes one or more instances using the same architecture
- Multiple workflow – It can take into account manifold cases of diverse structures of workflow.



Figure 1: Components of Workflow Management System

1.2. Workflow Design

Figure 2 portrays the workflow design mechanism. It is concerned with how to delineate and configure constituent of a

*Corresponding Author: Juhi Singh, Department of Computer Science, SRMU, Lucknow, India, 225003, juhisingh.srmcem@gmail.com

workflow. The workflow portrays the relationship and dependencies between workflows. DAG and Non-DAG [3] are two categories of workflow design. Directed Acyclic Graph is further characterized as Selection, Sequence and Parallel, while Non-DAG can be classified as Repeat, Parallel and Selection. In sequence architecture, the tasks are executed in a sequence, whereas in a parallel architecture, the workflows can be executed synchronously.

In a selection structure, workflows can run in sequence or in parallel. The recurrence pattern structure performs tasks iteratively [4]. Another is the workflow model, which is a constituent of workflow design, delineates the workflow at both the task and structure levels in both abstract and concrete workflows. Abstract workflows are characterized as a nonfigurative template with no commit to cloud resources to carry out the tasks, while concrete workflows are called actionable workflows. The workflow configuration permits users to coalesce diverse components through user-oriented and stand-alone systems.

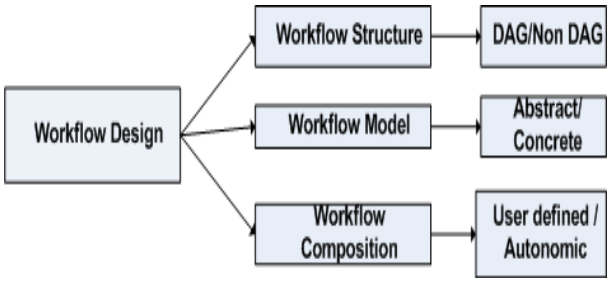


Figure 2: Workflow Design Components

1 . 3 . Workflow Scheduling

The workflow scheduler is necessity for the arrangement of workflows task on the cloud resources that are utilized to implement the workflow. The components that must be delineated to schedule a workflow are revealed below in Figure 3:

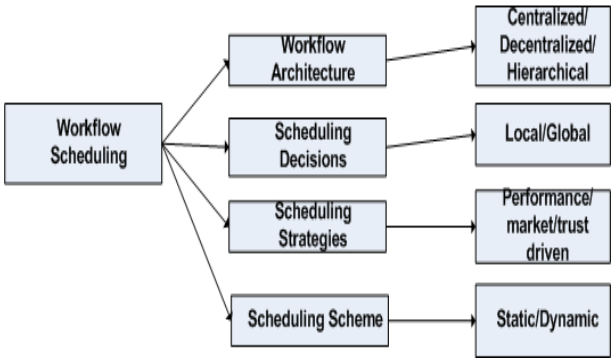


Figure 3: Workflow Scheduling Components

1.4. Fault Tolerance in Workflow

Fault tolerance is linked to tackle errors that can take place during the scheduling and execution phase of workflow tasks for various reasons such as unavailability of resources, resource breakdown, task malfunction, resource overload, network collapse, out of memory, etc

2. Literature Review

Arrangement of tasks can affect cloud system performance, so numerous workflow techniques as well as scheduling algorithms for scientific workflows have been studied and discussed below:

- Enhanced Scheduling of Resources: In [5], the author introduced a scheduling algorithm to attain optimization or more precise to sub-optimization for scheduling tasks. The authors exploited IGA (Improved Genetic Automated Scheduling Policy) to produce better results.
- Transaction Exhaustive Cost Restraint: In [6], author introduced an algorithm for scheduling tasks that took into account time and cost. Their simulation showed that this algorithm diminishes costs while adhering to deadlines.
- In [7], a superior algorithm based on cost was proposed by author. This algorithm capably assigns tasks to available resources in the cloud. Resource cost, computing performance, convalescing computing connectivity ratio is evaluated in this algorithm.
- In [8], author introduced a new cost scheduling algorithm based on deadlines. It took into account the features of cloud computing to hold cost-concentrated and limited workflows. It diminishes execution time and cost while enabling consumer input instantly.
- Inferences based on PSO for programming workflow applications: In [9], author introduced computation and data communication for applications that include the cost for both and provided a guideline based on Particle Clustering Optimization (PSO). This algorithm can be utilised for workflow applications that have dissimilar computing and communication overheads. Experiential results showed that PSO can attain cost savings and well distribute workloads to cloud resources. Also in [10], authors expanded PSO to provide deadline-based resource scheduling and provisioning. However, these authors did not explain resource failures or extreme dependence on essential tasks.
- Market-oriented hierarchical scheduling technique was proposed by author in [11], which includes both the levels of scheduling that is task-level scheduling and service-level scheduling, where task-level scheduling concept copes with optimizing device allocation from a task to a VM on cloud data centers and service-level scheduling concept copes with task assigned to service.
- Stretchy workflow scheduling: In [12], author proposed SHEFT workflow scheduling algorithm which is a stretchy workflow scheduling in cloud environment. Investigational upshots showed that this algorithm performs better than various other workflow scheduling algorithms. This algorithm perks up workflow uptime, as well as it also facilitates resources to flexibly measure uptime.
- Multi-workflow Multi-QoS (MQMW) constrained scheduling strategy: In [13], research work author introduced a scheme for workflows with multiple QoS. Authors boosted the rate of access to scheduling and also diminished the duration and outlay of workflows for the cloud platform, so improved Quality of Service was proposed for multiple workflows using constrained scheduling strategy.
- In [14], research work author suggested SA to schedule all the tasks on platform with the aspire of plummeting execution time, but they did not analyse malfunctioning.

3. Proposed Workflow Scheduling Approach

The literature survey on various workflow algorithms is done one different aspect. The proposed workflow scheduling

algorithm generates planned scheme of tasks which reduces the entire cost of implementing a workflow which meets a deadline defined by user. The algorithm consists of two chief stages:

- Deadline allocation
- Scheduling.

Deadline allocation doles out the time limit for the overall workflow between individual tasks. Scheduling arranges every task with the economical service that is able to run the task earlier than its deadline.

The proposed algorithm focuses on the two concepts of task start times, which is Earliest Begin Time (EBT) and Actual Begin Time (ABT). EBT is calculated prior to schedule of workflow while the ABT is calculated when the tasks are arranged. The EBT of each non-arranged task t_i, EBT (t_i), is delineated as follows:

$$EBT(t_{entry})=0 \quad (1)$$

$$EBT(t_i)= \max_{t_p \in t_i s \ parents} \{EBT(t_p)+LET(t_p)+TT(e_{p,i})\} \quad (2)$$

where

LET(t_i): Least Execution Time of a task t_i, on service $s_j \in S$ with the least ET (t_i, s_j) between all available services.

TT($e_{p,i}$) is the transmission time from parent node to task t_i.

$$LET(t_{entry}) = LET(t_{exit}) = 0. \quad (3)$$

As a result, the Earliest Completion Time (ECT) of an unscheduled task t_i, ECT (t_i), can be delineated as follows:

$$ECT(t_i) = EBT(t_i) + LET(t_i) \quad (4)$$

In addition, we delineate Latest Completion Time (LCT) of task that are unscheduled t_i, LCT (t_i), refers the latest time for t_i which can terminate its execution such that the deadline D of the entire workflow is reduced.

It can be calculated as follows:

$$LCT(t_{exit})=D \quad (5)$$

$$LCT(t_i) = \min_{t_c \in t_i s \ children} \{LCT(t_c) + LET(t_c) + TT(e_{p,c})\} \quad (6)$$

The Service that is selected for every arranged task t_i, SS (t_i) = $s_{j,k}$, is delineated as the selected service for executing t_i while scheduling process, where $s_{j,k}$ is the k_{th} occurrence of service s_j.

The Decisive Parent of a node t_i refers to the not assigned parent of t_i that has the newest arrival time of data at t_i. It is taken as the parent t_p of t_i for which ECT (t_p)+ TT ($e_{p,i}$) which has maximal value.

The partial Decisive Path (PDP) of a node t_i is:

- Blank if t_i does not have some unassigned parents.
- Composed of Decisive Parent t_p of t_i and the PDP of t_p if has any unassigned parents.

Algorithm 1 : Scheduling Algorithm

```
procedure Schedule_Workflow (G (T, E), D)
{
Step 1: verify existing computation services

Step 2: add t_entry , t_exit and their corresponding dependencies to G

Step 3: calculate EBT (t_i), ECT (t_i) and LCT(t_i) for each task in G

Step 4: ABT(t_entry) ← 0, ABT(t_exit) ← D

Step 5: mark t_entry and t_exit as assigned

Step 6: call Assign_Parents(t_exit )
}
```

Algorithm 2 : Algorithm for Parents Assignment (Step 6)

```
procedure Assign_Parents (t)
{
  while (t has any unassigned parent)
  {
    PDP = null , t_i = t

    While (t_i has any unassigned parent)
    {
      add Decisive_Parent (t_i) at the start of PDP

      t_i = Decisive_Parent (t_i)
    }

    Step 7: call Assign_Path(PDP)

    for each t_i ∈ PDP
    {
    update EBT and ECT for every successor of t_i

    update LCT for every predecessor of t_i

    Step 6: call Assign_Parents(t_i)
    }
  }
}
```

Algorithm 3 : Algorithm for Path Assignment (Step 7)

```
procedure Assign_Path(P)
{
  s_i,j = the cheapest applicable existing instance for P
    if (s_i,j = null)
    {
     launch a new instance s_i,j of the cheapest service s_i
     which can finish each task of P before its LCT
     }

   Schedule P on s_i,j and set SS (t_i ), ABT (t_i ) for each t_i ∈ P
   Set all tasks of P as assigned
}
```

4. Performance Evaluation

To appraise the proposed algorithm, we need to measure its performance in some workflow models. One of the structures

we used in this article is the Montage. Figure 4 depicts the approximate structure format of this workflow:

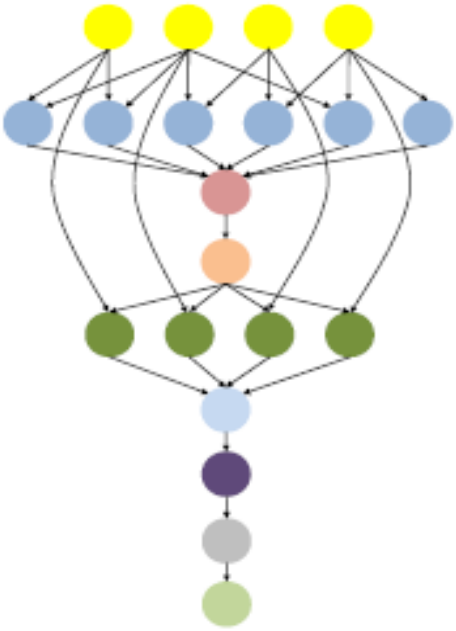


Figure 4: Structure of Montage workflow

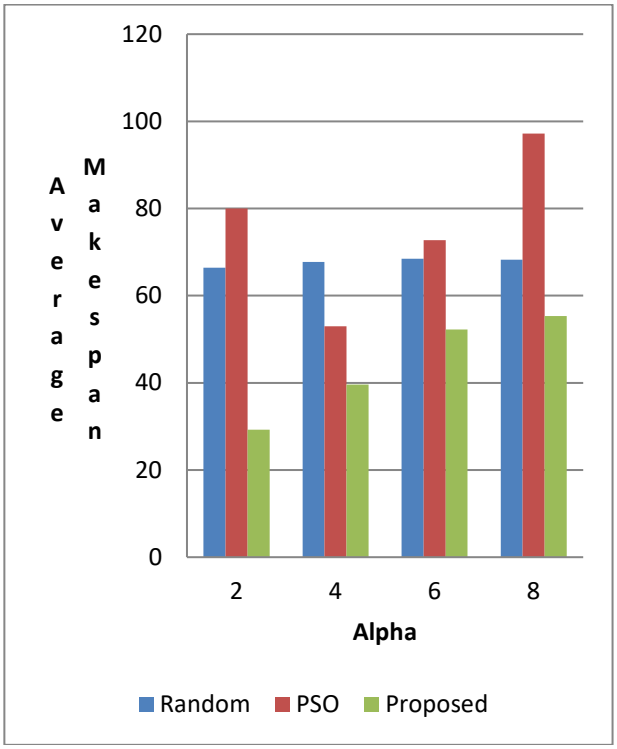


Figure 5: Comparison of average makespan for proposed algorithm

For calculation of the effectiveness of the proposed algorithm, we use makespan as a parameter. Makespan is the total time taken for resources to complete the overall execution of the tasks.

For evaluation of our proposed algorithm, we set a deadline for each workflow. We first delineate the fastest schedule. MF indicates the makespan that is fastest program in a workflow. It is only a least duration for that workflow to execute, so to establish time limit for every workflow, we delineate the deadline with factor α and delineate the new deadline of the workflow with its arrival plus α * MF. Since there is no elucidation for α = 1, we take the range of α from 2 to 8 in our experiment.

Below is a comparison of the proposed approach with the stochastic approach and the PSO [9]. The graph is plotted for the parameter value of average makespan and alpha factor as deadline of the workflow.

Table 1: Comparison of average makespan for proposed algorithm

Alpha Value	Algorithm	Average Makespan
Alpha Factor =2	Random	68
	PSO	80
	Proposed	30
Alpha Factor =4	Random	68
	PSO	52
	Proposed	40
Alpha Factor =6	Random	68
	PSO	72
	Proposed	54
Alpha Factor =8	Random	68
	PSO	97
	Proposed	56

5. Conclusion

In this research paper, we aimed to design an algorithm that takes an approach to schedule workflow in one phase. The approach is intended to schedule each and every partial decisive path into a one occurrence of a computation service. The time complexity of algorithms is $O(n^2)$, where n indicates the number of tasks in the workflow. The complexity of polynomial time makes it appropriate for bulky workflows. We assess our algorithm by measuring its recital in Montage synthetic workflow. The results show that the proposed algorithm works better than other approaches.

References

[1] X. Zhou, G. Zhang, J. Sun, J. Zhou, T. Wei, S. Hu, "Minimizing cost and makespan for workflow scheduling in cloud using fuzzy dominance sort based HEFT," Future Generation Computer Systems, **93**, 278–289, 2019, doi:10.1016/j.future.2018.10.046.

[2] J. Yu, R. Buyya, "Scheduling Scientific Workflow Applications with Deadline and Budget Constraints Using Genetic Algorithms," Scientific Programming, **14**, 217–230, 2006, doi:10.1155/2006/271608.

[3] J. Yu, R. Buyya, "A Budget Constrained Scheduling of Workflow Applications on Utility Grids using Genetic Algorithms," Workshop on Workflows in Support of Large-Scale Science, 2006, doi:10.1109/WORKS.2006.5282330.

[4] J. Yu, R. Buyya, "A Taxonomy of Workflow Management Systems for Grid Computing," Journal of Grid Computing, **3**(3), 171–200, 2005, doi:10.1007/s10723-005-9010-8.

[5] H. Zhong, K. Tao, X. Zhang, An Approach to Optimized Resource Scheduling Algorithm for Open-Source Cloud Systems, 2010, doi:10.1109/ChinaGrid.2010.37.

[6] Y. Yang, K. Liu, J. Chen, X. Liu, D. Yuan, H. Jin, An Algorithm in SwinDeW-C for Scheduling Transaction-Intensive Cost-Constrained Cloud Workflows, 2008, doi:10.1109/eScience.2008.93.

[7] S. Selvarani, S. Sadhasivam, "Improved cost-based algorithm for task scheduling in cloud computing," Computational Intelligence and Computing Research (ICCIC), 2010, doi:10.1109/ICCIC.2010.5705847.

[8] K. Liu, H. Jin, J. Chen, X. Liu, D. Yuan, Y. Yang, "A Compromised-Time-Cost Scheduling Algorithm in SwinDeW-C for Instance-Intensive Cost-Constrained Workflows on a Cloud Computing Platform," IJHPCA, **24**, 445–456, 2010, doi:10.1177/1094342010369114.

[9] S. Pandey, L. Wu, S. Guru, R. Buyya, A Particle Swarm Optimization-Based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments, 2010, doi:10.1109/AINA.2010.31.

[10] M.A. Rodriguez, R. Buyya, "Deadline Based Resource Provisioningand Scheduling Algorithm for Scientific Workflows on Clouds," IEEE

Transactions on Cloud Computing, **2**(2), 222–235, 2014, doi:10.1109/TCC.2014.2314655.

[11] Z. Wu, X. Liu, Z. Ni, D. Yuan, Y. Yang, "A market-oriented hierarchical scheduling strategy in cloud workflow systems," The Journal of Supercomputing, **63**(1), 256–293, 2013, doi:10.1007/s11227-011-0578-4.

[12] C. Lin, S. Lu, Scheduling Scientific Workflows Elastically for Cloud Computing, 2011, doi:10.1109/CLOUD.2011.110.

[13] M. Xu, L. Cui, H. Wang, Y. Bi, "A Multiple QoS Constrained Scheduling Strategy of Multiple Workflows for Cloud Computing," in 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, 629–634, 2009, doi:10.1109/ISPA.2009.95.

[14] M. Kaid, M. Othman, "Simulated Annealing Approach To Cost-Based Multi-QoS Job Scheduling in Cloud Computing Enviroment," American Journal of Applied Sciences, **11**, 872–877, 2014, doi:10.3844/ajassp.2014.872.877.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 5, 43-52 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Analysis of Grid Events Influenced by Different Levels of Renewable Integration on Extra-large Power Systems

Christoph Rüeger*, Jean Dobrowolski, Petr Korba, Felix Rafael Segundo Sevilla

School of Engineering, Zurich University of Applied Sciences, IEFE, Winterthur, 8401, Switzerland

ARTICLE INFO

Article history:
Received: 14 July, 2021
Accepted: 23 August, 2021
Online: 10 September, 2021

Keywords:
Power System Dynamics
Stability
Control
European Grid
Lyapunov Exponent
Inter-Area Oscillation

ABSTRACT

In this work, the impact of implementing a large amount of decentralized renewable energy sources (RES) of different scales on an extra-large power grid is investigated. Three scenarios are created, substituting 10%, 20%, and 30% of the conventional energy production by RES. For this purpose, the initial dynamic model of Continental Europe in combination with the industrial power system application DIgSILENT PowerFactory was used. In order to compare the behavior of different applied scenarios, a performance index was developed to evaluate and rank the effects of network disturbances by means of time-domain simulations. The performance index was designed based on three different criteria that analyze the oscillatory content and thus, the severity of a given event. The initial power flow of the dynamic model was identified as a limiting factor for the integration of RES, therefore two additional power flows were developed following an innovative procedure. Through the methodologies mentioned above, it was found that Turkey is the most sensitive to such changes, which are amplified by increasing implementation of RES and often lead to inter-area oscillation.

1 Introduction

The release of Al Gore's film "An Inconvenient Truth" in 2006 as a campaign to educate people about environmental awareness in the USA or the protest and notable criticism of the Swedish environmental activist Greta Thunberg to world leaders, that started in 2015, are good examples of events that have helped to draw people's attention on global warming problems, in particular among the young population. Pollution, climate change, and damage to the earth's ozone layer are some of the issues on the public agenda of governments around the world. Confinement, the considerable decrease in land and air mobility forced by the ongoing pandemic caused by the COVID-19 virus [1] has demonstrated how a sudden change in human behavior can have a positive impact on our environment. With lock-downs all over the world, major industrial processes at a standstill, lack of massive events, and a home-office policy, the daily routine of millions of people have been affected and might never be the same again. In world capitals such as New York, the measured carbon emissions dropped up to 50%, and satellite images have also shown a decrease in NO_2 (nitrogen dioxide) all over Europe, particularly in northern Italy, Spain, and England [2]. Even if it could be proved that the current pandemic situation was a short-term phenomenon, environmental effects have shown that extreme measures provide positive results to reverse global warming. Moreover, it has been also shown that these extreme measures have a direct impact on the way of generating electricity in the upcoming years. In Switzerland, for instance, after a parliament vote and approval of a new referendum, a revision in the legislation entered into force in 2018 in the form of an amendment to the current Energy Strategy 2050, which is a document known as the New Energy Act [3]. The new reform considers measures to increase energy efficiency, measures to increase the use of renewable energy sources (RES) and withdrawal from nuclear energy. Furthermore, the implementation of further and more ambitious measures to the energy sector have not been limited only within the Swiss borders and similar measures have been implemented in parallel throughout the European Union. In December 2019, the European Green Deal [4] was released, containing 50 measures that, similar to those in Switzerland, intent to push the integration of RES in the electrical system and to promote an emissions trading system. An important point to highlight from this document, is the definition of nuclear power plants as an unsustainable source of energy, leading to nuclear phase-out across Europe in the future. As a result of the aforementioned measures, two main challenges for Transmission System Operators (TSOs) in Europe have envisaged namely 1) the decentralization of RES and 2) the exclusion of large

*Corresponding Author: Christoph Rüeger, Technikumstrasse 9, CH-8401 Switzerland, +41 (0) 58 934 69 06, christoph.rueeger@zhaw.ch

rotating masses from the system, presenting a loss of inertia into the system. Regarding the first challenge, a widespread decentralization represents a reshuffle of the energy flow in the transmission system as a result of the installation of RES in remote locations with weak interconnections. With respect to the second problem, existing rotating masses play an important role in the stability of the system as a source of inertia. Switzerland has been privileged with its geography, allowing it to produce more than half of his annual consumption from alpine hydro-power, which also contribute with a non-trivial amount of inertia [5]. However, not every country can benefit from this particular type of energy generation. Nevertheless, it is a fact that load and demand profiles on power systems are changing and thus, its impact must be assessed in order to manage the security of the supply [6, 7]. Contrarily, blackouts like in 2003 in Italy will happen more often [8]. Motivated by these questions, the current conditions of the European transmission power grid and the expected effects on the system as a result of the integration of the aforementioned measures are analyzed in this paper.

Table 1: Number of elements in the dynamic model

Code	Lines	Buses	Loads	Generators
Albania (AL)	193	339	110	77
Austria (AT)	123	104	40	31
Bosnia & Herz.(BA)	312	294	164	37
Belgium (BE)	178	140	36	52
Bulgaria (BG)	787	798	419	77
Switz. (CH)	244	193	82	88
C. Rep. (CZ)	106	288	76	113
Germany (DE)	3378	3939	859	898
Denmark (DK)	250	397	66	71
Spain (ES)	1338	1385	646	496
France (FR)	2599	2665	991	1564
Greece (GR)	1133	1312	367	128
Croatia (HR)	334	329	171	73
Hungary (HU)	94	120	37	28
Italy (IT)	817	1264	341	458
Luxembourg (LU)	41	38	11	12
Montenegro (ME)	80	94	35	18
Macedonia (MK)	146	163	85	25
Nether. (NL)	905	1031	244	178
Poland (PL)	988	648	199	140
Portugal (PT)	365	506	87	158
Romania (RO)	1194	1171	654	185
Serbia (RS)	619	553	303	62
Slovenia (SI)	111	230	65	73
SK Slovakia (SK)	52	48	18	82
Turkey (TR)	1943	4888	1245	1022
Europe (EU)*	9	316	26	1
Total	18339	23253	7377	6147

N_i = Elements not labeled for a specific country.i

The presented studies are based on the initial dynamic model of Continental Europe [9] developed by the European Network of Transmission System Operators (ENTSO-E), which is the more comprehensive representation of the European transmission power grid available today and that has been extensively studied in previous projects [10].

This paper is organized as follows, as background, it is worth mentioning that this document describes an extension of [11], which is based on two additional results in order to give more insight to the main work. In the first section, the Lyapunov exponent is presented [12] and is used as preliminary stability evaluation of the most severe grid events already reported in [13]. Additionally, the characteristics of the dynamic model are provided in detail, as well as a ranking of events in terms of severity. In the subsequent section, three variations of the power system model are fully described. Considering different future scenarios, different levels of RES integration were evaluated [14]. Subsequently, the effects of faults (introduced on Section 2) for each model variation are investigated. Then, the results are analyzed for all simulated scenarios and ranked using the algorithm based on the Lyapunov exponent. From the time series resulting from Root Mean Square (RMS) simulations, special attention is paid to low-frequency fluctuations (in the range of 0.1-1Hz) also known as inter-area oscillations, which are the main problem to solve in extra-large power grids [15, 16]. One disadvantage of the dynamic model used for the analysis in this document is the fact that there is only a single initial power flow available. This constraint plays an important role in the investigation of the stability of the system. To overcome this limitation, new power flow scenarios were developed using an algorithm reported in Section 3.3. Finally, a possible solution will be discussed. The progressive replacement of the rotating masses and the consequent decrease of the grid inertia[17] makes the power grid more vulnerable to faults especially in the eastern part of the grid. Thus, three different approaches are presented using wide-area monitoring to control the inter-area oscillations. New in these methods is the architectural design of the controllers.

2 Background

2.1 The Dynamic Model of Continental Europe

The dimensions of the model, with more than 55,000 interconnected elements distributed over 27 grids (see table 1), is one of the main characteristics and made the selected system very challenging to work with. It is the most comprehensive representation of the continental European power grid, which was developed by ENTSO-E and that we have used to demonstrate the effectiveness of the proposed algorithms. In the following section, multiple variations of the topology of the system will be introduced but first, the base case is presented so it later can be used as reference for the comparison. To do so, different types of contingencies or faults were simulated and the response of the system was analyzed by means of frequency measurements on randomly selected buses at the high voltage level in all countries [9]. Given the size of the model, the number of contingencies to be investigated and thus the number of simulations to be performed, was narrowed down to 184 based on the seven criteria displayed in figure 1.

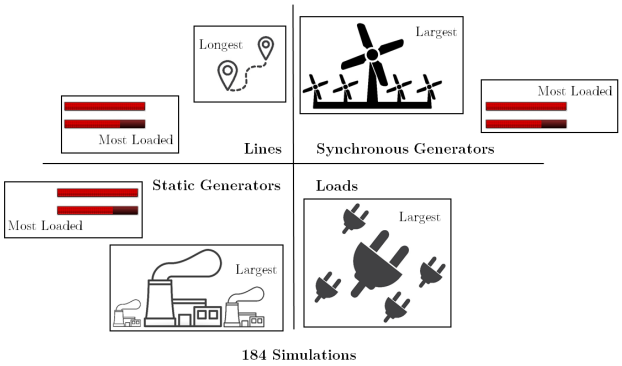


Figure 1: Event selection based on the seven criteria.

2.2 *Severity evaluation and ranking of most dangerous events*

For the base case, 184 simulations were performed representing one fault of an asset at the time, e.g. 3-phase short circuit on a line or the sudden disconnection of a generator or load. Following the event, the frequency was measured for a period of 20 seconds at 27 busbars randomly selected (one bus of each country). In order to evaluate the impact on the network and to rank the severity, the measurement results of each simulation were compared. To do this, a performance index that analyzed three aspects of the time series was first developed as follows:

1. Measuring the damping of the oscillations based on the Lyapunov Exponent [18] with f_j^i as the frequency in the country i at the time j (1),

$$\lambda_k^i = \frac{1}{N\Delta t} \times \sum_{m=1}^{N} \log\left(\frac{\left|f_{(k+m)\Delta t}^i - f_{(k+m-1)\Delta t}^i\right|}{\left|f_{(m)\Delta t}^i - f_{(m-1)\Delta t}^i\right|}\right) \quad (1)$$

$$\alpha_{le} = \frac{\sum_{i=1}^{27} \alpha_{le}^i}{27} \quad (2)$$

and where α_{le} in Equ. (2) is the average value of each simulation.

2. The next aspect was the analysis of the average amplitude of the oscillation in all countries [19] (3)

$$\alpha_{am} = \frac{\sum_{i=1}^{27} max\left(|f_0 - min(f_i)|, |f_0 - max(f_i)|\right)}{27} \quad (3)$$

with the nominal frequency f_0 before the event and f_i being the frequency in the i-th country.

3. Finally, the rate of change of frequency or RoCoF for each country was used as last indicator [20]

$$RoCof = \frac{\partial f(t)}{\partial t} = \frac{f_m - f_{m-1}}{\Delta t} \quad (4)$$

$$\alpha_{de} = \frac{\sum_{i=1}^{27} max\left(RoCoF_i\right)}{27} \quad (5)$$

where in equation (4) f_m is the frequency at time t and similarly to the Lyapunov Exponent, α_{de} in Eq. (5) is the average value across the 27 countries.

In order to avoid that an individual criteria had more impact on the final index value than others, the individual indicators (α_{le}), (α_{am}) and (α_{de}) were multiplied by a weighting coefficient as shown in Equ. (6).

$$\alpha = \alpha_{le} * \omega_{le} + \alpha_{am} * \omega_{am} + \alpha_{de} * \omega_{de} \quad (6)$$

The final outcome, resulting from equ. (6) is a scalar that can be used to evaluate the severity of a given contingency.

A partial result is depicted in Figure 3, which displays the severity index for the base case, where the normalized contribution of the criterions defined on Eq. (1-3) to the final index are depicted in blue, red, and green, respectively. The worst event has a magnitude of +1 and as the magnitude decreases its value, the lower the relevance of the subsequent event. Fig. 3 displays the worst events (ranked 175 - 184) and their geographical location is depicted on Figure 2. The inter-area oscillations caused by the sudden disconnection of a large synchronous machine in Turkey, are displayed in Figure 4 for a selected number of countries across Europe, where the severe effects of this fault can be clearly observed. Table 2 summarized the events displayed in Fig. 2 and it can be observed that 9 out of 10 of the most critical faults are related to the sudden loss of a large synchronous generation, in contrast to only one event that was caused by the trip of an important transmission line. It is worth noticing that all machines listed in Table 2 are located in Eastern Europe.

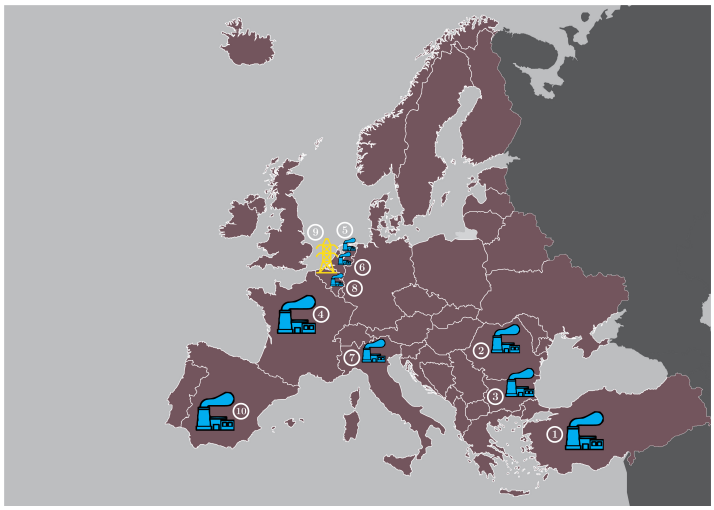


Figure 2: Location of the 10 worst events according to performance value.

Table 2: Performance Value of the ten worst events

Ranking	Country	Type	Perfomrance Value
184	TR	Sync. Machine	1
183	RO	Sync. Machine	0.903
182	BG	Sync. Machine	0.886
181	FR	Sync. Machine	0.699
180	NL	Sync. Machine	0.571
179	NL	Sync. Machine	0.569
178	IT	Sync. Machine	0.559
177	BE	Sync. Machine	0.518
176	BE	Line	0.517
175	ES	Sync. Machine	0.512

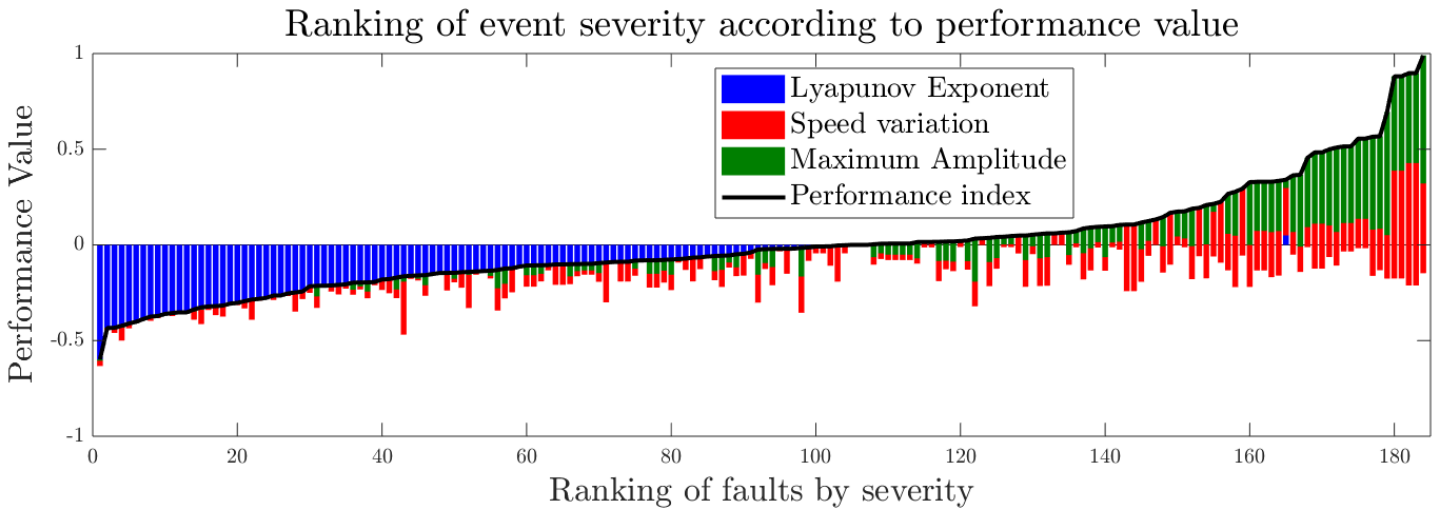


Figure 3: Ranking of events by severity.

3 Stability effects of RES integration

An insight of the results, obtained after the analysis of the base case, confirms the accuracy of the model in comparison with the real system. In this case, the base case corresponds to the conditions of the European power grid in 2016, approximately. In the next step, a modified grid is investigated, where conventional synchronous machines were replaced by decentralized renewable energy sources. In this form, the original dynamic model was topologically changed three times in order to represent different future scenarios.

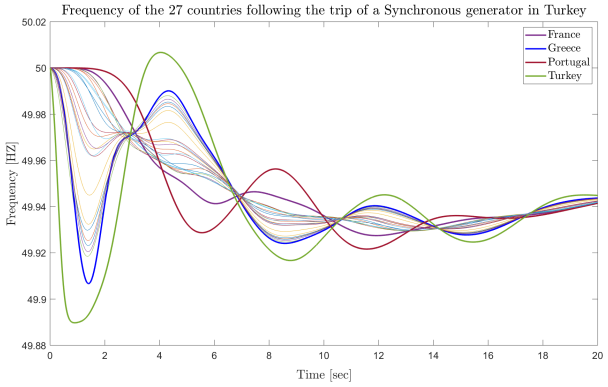


Figure 4: Inter-Area oscillations (0.125Hz) affirmed by the results of a tripped synchronous generator in Turkey.

3.1 Implementation of different levels of RES

Three future scenarios were developed with a systematic increase of RES. First, 10% of the country's conventional energy production was replaced with RES, this scenario is referred to as *S1* in the subsequent sections. Then, a 20% RES penetration was developed and labelled as scenario 2 or *S2* and finally a third scenario *S3*, with 30% of RES integration. It is important to underline that the procedure to achieve an increase in the different levels of RES is not trivial nor simple. In DIgSILENT PowerFactory, the power system software application in which the model was built, is possible to combine renewable energy with existing models in different ways. However, since the focus of this work was to analyze stability phenomena, an accurate representation of renewable devices, namely one to one component representation, was omitted. In other words, there is no distinction between the replacement of conventional generation between Wind, PV or any other technology. The key idea was to substitute rotating masses with static elements and thus, indirectly affect the inertia of the system. Therefore, renewable assets were modelled as static generators connected via static converters [21].

Furthermore, a decentralized renewable generation is assumed [14], which has a direct influence on the topology of the different variants. In this form, synchronous generators representing nuclear power plants, coal-fired power plants and other conventional production, were replaced by several smaller static generators located on multiple locations. The voltage level at which these large quantities of RES were connected played also an important role. In [22] it is recommended to connect new elements at busbars where a load is also attached. In our case, the model has more than 23'000 buses with more than 7'000 loads. The new static generators integrated into the system were selected to operate in the 15kV voltage level.

From the countries considered in the model, not all of them are represented with the same level of detail or do not have a pronounced infrastructure as in Switzerland, Germany or Turkey. As a result, some countries were modelled with fewer amount of synchronous machines. This fact was taken into consideration in the development of scenarios 1, 2 and 3, respectively. Only countries modelled with less than 20 synchronous machines were excluded from the modification. Consequently, only 13 countries were modified and their names are listed in Table 3. The topology of these countries was changed individually to create the new system variations, leading to 39 model variations (13 countries times 3 different variations). In addition, Table 3 displays the amount of sync. machines (N_k) existing in each country and the total amount of active power in MW that these machines contribute. Similarly, this table indicates the static generators (N_i) required in order to replace the energy produced by the conventional machines N_k.

Technically, the 39 variants mentioned above were created and the share of active power to be replaced was calculated for the

respective country using an optimization procedure. Next, all synchronous machines available in a given country were analyzed in order to find the fewer amount of synchronous generators to be substituted by RES. Then, the selected machines were replaced by randomly distributed static generators to compensate the same amount of active power. Given that this task involved the modification of hundreds of elements, the procedure was automatized using python. In the same way, all simulations performed (7'000 resulting from analyzing 184 events at the time for 39 variations), were analyzed and compared, using the python interface of DIgSILENT PowerFactory and the results are presented in the following subsection.

Table 3: Transmutations in the selected countries

Country	***RES***	***Scenario 1 10%***		***Scenario 2 20%***		***Scenario 3 30%***	
	N_i	N_k	*MW*	N_k	*MW*	N_k	*MW*
Austria (AT)	40	4	409	4	1092	4	1640
Switz. (CH)	80	4	712	4	1458	4	2186
C. Rep. (CZ)	76	4	653	4	1304	4	1955
Germany (DE)	111	6	5301	13	8012	21	14046
Spain (ES)	646	4	2620	6	5241	9	7861
France (FR)	956	4	6600	11	13185	16	19784
Greece (GR)	363	4	726	4	1069	6	1604
Italy (IT)	319	4	2025	9	3211	16	4030
Nether. (NL)	166	4	1538	4	3076	5	4630
Poland (PL)	185	4	1125	4	2252	5	3376
Belgium (BE)	36	4	981	4	1959	4	2940
Slovenia (SI)	59	4	155	4	517	4	613
Turkey (TR)	316	5	3641	14	7278	23	10918

N_i = number of static generators, N_k = number of synchronous generators.

3.2 Simulation results using the initial Power Flow

In this section, the 7'000 simulations resulting from the 184 events on the different levels of RES integration are compared against the results of the 184 events on the original model [13] and presented in three parts as follow:

i. *Scenario 1 (S1) vs. Base case (B)*. The 13 variations implementing 10% of RES in different countries are compared against the original model.

ii. *Scenario 2 (S2) vs. Base case (B)*. The 13 variations implementing 20% of RES in different countries are compared against the original model.

iii. *Scenario 3 (S3) vs. Base case (B)*. The 13 variations implementing 30% of RES in different countries are compared against the original model.

The comparison was performed as follows:

a) The measure used to compare scenarios is the absolute value of the difference between the frequencies of each variation and the frequencies of the original model, more details are provided next.

b) The main results are presented in the form of 3D-plots where:

- The $x-$ axis depicts the country where the RES were deployed. For simplicity, only the country codes are presented and the country names are listed in Table 3.
- The $y-$ axis provides the simulation number from 1 to 184, which is refers to different types of events (discussed in Section 2) and is summarized below:
  - 1-27 severe synchronous machine events
  - 28-57 most loaded synchronous generators
  - 58-84 most loaded lines
  - 85-111 events related to loads
  - 112-138 longest lines
  - 139-161 largest static generators
  - 162-184 most loaded static generators events
- the $z-$ axis corresponds to the sum of the frequency differences, which are given by:

$$\begin{pmatrix} \sum_{i=1}^{N_G} \left| f^i_{B_1} - f^i_{AT_1} \right| & \cdots & \sum_{i=1}^{N_G} \left| f^i_{B_{N_E}} - f^i_{AT_{N_E}} \right| \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N_G} \left| f^i_{B_1} - f^i_{TR_1} \right| & \cdots & \sum_{i=1}^{N_G} \left| f^i_{B_{N_E}} - f^i_{TR_{N_E}} \right| \end{pmatrix} \tag{7}$$

where N_G is the number of measurements per simulation (27), N_E is the number of events (184), $f^i_{B_j}$ represents the frequency measured in the i-th country after the j-th event for the base case model [13] and $f^i_{AT_j}$ is the frequency measured in the i-th country (Austria in this case) after the j-th event for the n-th scenario (S_n) of renewable penetration.

c) After the frequency differences have been estimated, the stability indexes introduced on equation (6) are calculated for every model variation and then compared against the base case [13] using a heat map. In these graphical representation, the $y-$ axis represents the country where the RES were implemented and the $x-$ axis provides the simulation number related to the type of event as in the 3D plot. The colors of the heat map represent the relation among stability indexes, which are given as:

$$\begin{pmatrix} \alpha_{AT_1} - \alpha_{B_1} & \cdots & \alpha_{AT_{N_E}} - \alpha_{B_{N_E}} \\ \vdots & \ddots & \vdots \\ \alpha_{TR_1} - \alpha_{B_1} & \cdots & \alpha_{TR_{N_E}} - \alpha_{B_{N_E}} \end{pmatrix} \tag{8}$$

where α_{B_i} is the stability index of the base case for the i-th event and α_{AT_i} is the stability index in the j-th country (Austria in this case) after the i-th event for the n-th scenario (S_n) of renewable penetration. It is important to note that:

- An increase in stability is represented by negative values in matrix (8) as a consequence of RES penetration.
- Positive values imply a larger deviation from the base case, leading to a negative impact in terms of stability as a result of the RES implementation.

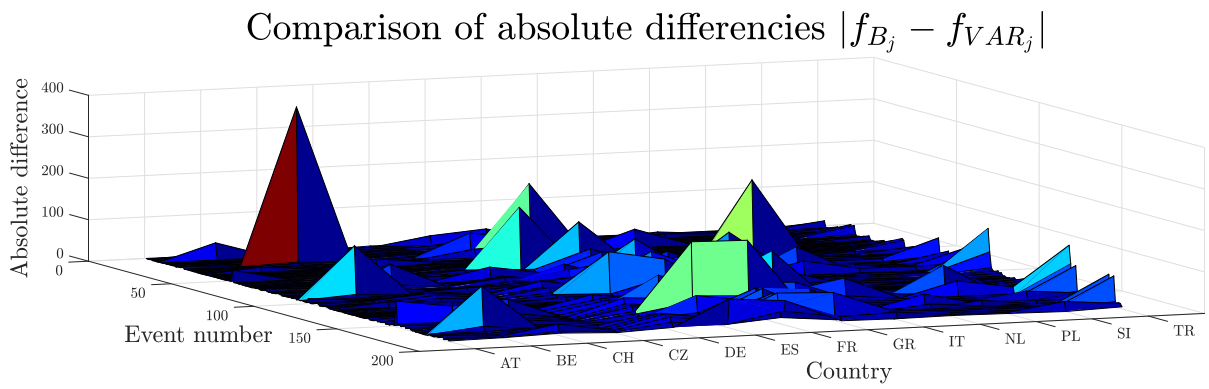


Figure 5: Comparison between scenario 1 (implementation of 10% RES) and the original model by the absolute difference of the frequency.

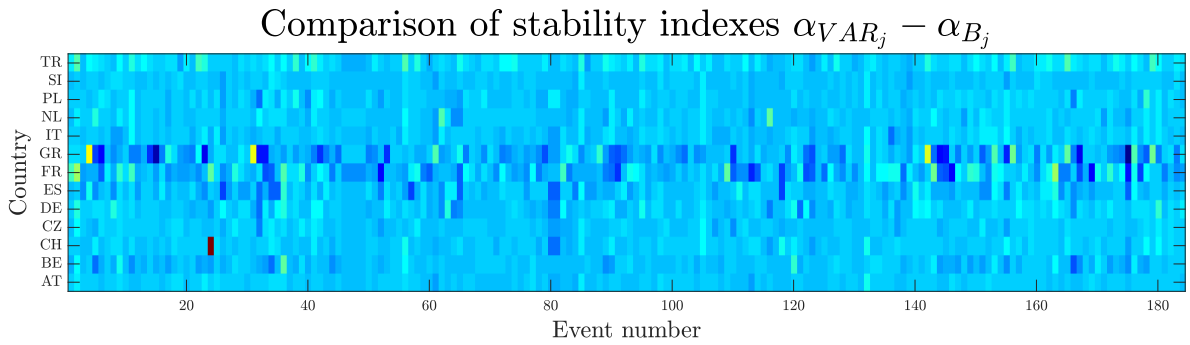


Figure 6: Heatmap visualizing the change in the stability index between scenario 1 (implementation of 10% RES) and the original model.

3.2.1 Scenario 1 (10%) vs. Base Case

Figure 5 shows how the compensation of 4-6 synchronous generators by RES affects the frequency. For this purpose, the results from the base case are compared with those from scenario 1. This corresponds to a 10% increase in energy generated by renewables, which means up to 163 Mvar of reactive power and up to 6.6 GW of active power, depending on the model variation. As mentioned in equation (7), the absolute difference of the frequencies serves as a metric for the comparison. By spanning a surface and plotting the frequency on the z-axis in Figure 5, unevenness indicates deviations from the base case. Apart from a few outliers and small peaks, no major differences can be detected. This observation is backed up by the heat map in Figure 6. Here, changes in the stability index are shown in color. Coloring in the dark blue area represents a negative value and thus indicates a more stable grid. The positive effect of an increment of 10% of RES in some countries, particularly in Greece (GR), is explained by the stability index on equation (6). As predicted on [15], reducing inertia produces an improvement of the rate of change of frequency (RoCoF) and at the same time produces an increase of the damping and the natural frequency. For most of the studied events in Greece, the increase in the RoCoF is lower than the increase on the damping, leading to a smaller value of α.

At this point, an event in the model variation Switzerland shall be briefly discussed. This event can be seen in Figures 5 triggering the largest outlier, as well as in Figure 6 as a dark red colouration. This is not a fault, but a disconnection of a synchronous generator, which is directly connected to the interconnection between Switzerland and Italy. In the ENTSO-E model, this element is subject to market restrictions, so the active power flowing through this line must remain constant. Additionally, one synchronous generator decommissioned in Switzerland was geographically close to this line. As result, the reallocation of the power was changed causing a large variation in the frequency.

3.2.2 Scenario 2 (20%) vs. Base Case

In this comparison, the same pattern is found in relation to the previous case. However, Figure 7 displays more frequency variations, particularly in Germany, Spain, France and Turkey. In the corresponding heat map depicted in Figure 8, it can be observed that Turkey happens to be the most sensible country to the RES variation and this is reflected by the number of times where the magnitude α exceed the value of 5, which is denoted with a turquoise/green color on the heat map. Similarly, it can be seen that 20% of renewable penetration has a positive impact in some events, especially in Greece and Italy because also, in this case, reducing the inertia has a lower negative influence on the RoCoF, than a positive one on the damping and the natural frequency in the system. Finally, it is important to note that similarly as in the previous case, some random negative events in Greece and Italy occur as a result of reallocation of the power after an event with respect to a trip of a static generator and a highly loaded line.

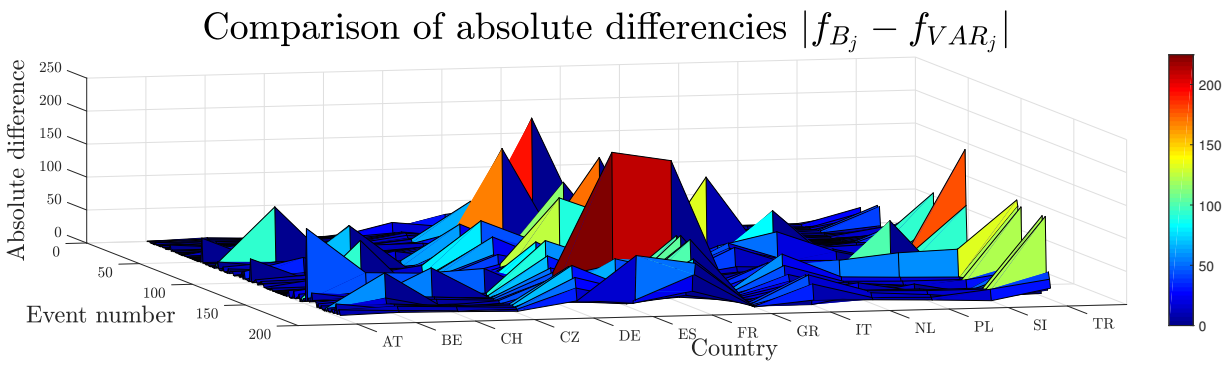


Figure 7: Comparison between scenario 2 (implementation of 20% RES) and the original model by the absolute difference of the frequency.

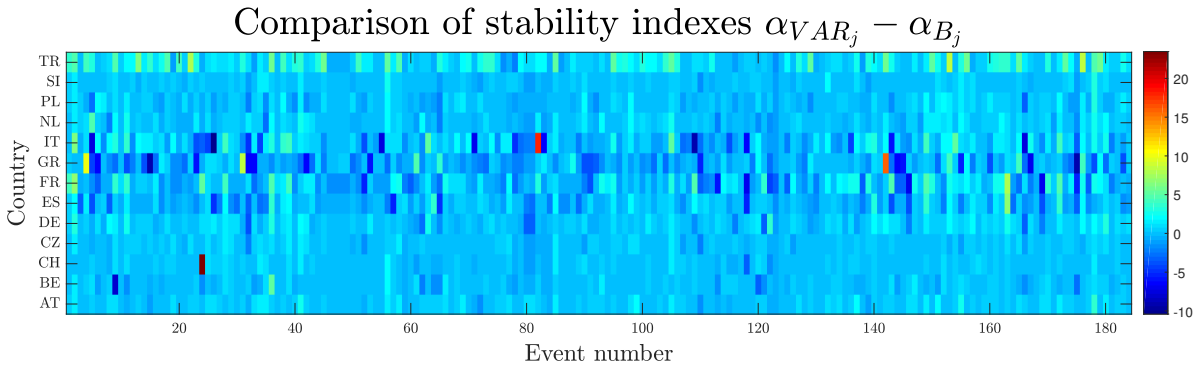


Figure 8: Heat map visualizing the change in the stability index between scenario 2 (implementation of 20% RES) and the original model.

3.2.3 Scenario 3 (30%) vs. Base Case

The largest topological changes were made in the last scenario and the effects on the models can be seen very well in Figure 9 and 10. When looking at the absolute frequency difference in Figure 9, there are clearly more irregularities than in subsection 3.2.1 and 3.2.2. What was only slightly apparent in scenario 2 can now be confirmed in this comparison. Events in countries resembling Germany, France, Spain, Italy and Turkey are most affected by the aggressive implementation of RES and show clear differences to the base case. It is noticeable that in all cases Turkey reacts most sensitively to the network changes. However, it is a fact that Turkey is the country with the highest change in share of renewable energy. The problems in Turkey can be explained by the results shown in Figure 12, where 27 measurements resulting after a severe event in one of the most loaded static generators in Turkey, are displayed (in blue) and the results are compared against the base case (in red). Through these results, it is evident that although the damping of the oscillations is improved for the 30% RES implementation (lines in red), there is a significant degradation of the amplitude and RoCoF compared to the base case (lines in blue). This results in electromechanical oscillations caused by groups of generators oscillating against each other (west versus east), which were already identified using the stability index developed in (6).

3.3 Development of additional power flows

It can be concluded from the results in the previous subsection that massive integration of RES can change the stability of the European power grid. In particular, Turkey is more prone to present stability constraints following RES integration and after analyzing the initial power flow, it can be seen that around 13% of the production in Central Europe comes from Turkey and that 70% of the global production is produced in the western part of Europe namely in France, Germany, Spain, Italy, and the Netherlands. The current generation distribution in the system anticipates the results obtained and opens the following question: does the redistribution of the current ratio play an important role in the stability of the system? The answer is not trivial and one potential solution is to redistribute the current generation and repeat a similar procedure presented in Section 3.2. This solution represented to create new power flow scenarios and to do so, an iterative algorithm to automatically randomly calculate a new set of initial conditions was developed and is explained in the next steps:

1. Calculation of the total active power in the model contributed from all participating countries. It is worth mentioning that in this initial approach the reactive power was not change.
2. Random reallocation of the active power among the countries of the model, respecting the maximum active power provide from each country.
3. Perform a load flow calculation in PowerFactory with the new conditions
4. Check if power flow converge (a solution to the new operating conditions exist). If solution exist: algorithm is finish, otherwise return to step 2.

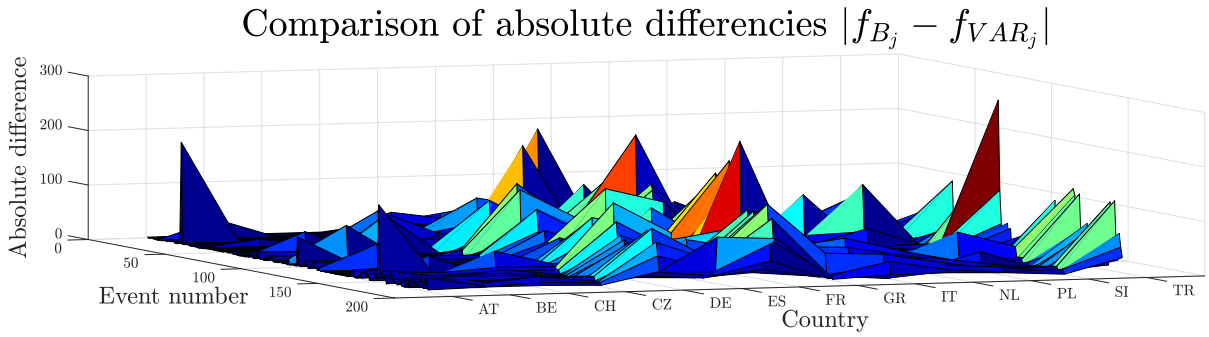


Figure 9: Comparison between scenario 3 (implementation of 30% RES) and the original model by the absolute difference of the frequency.

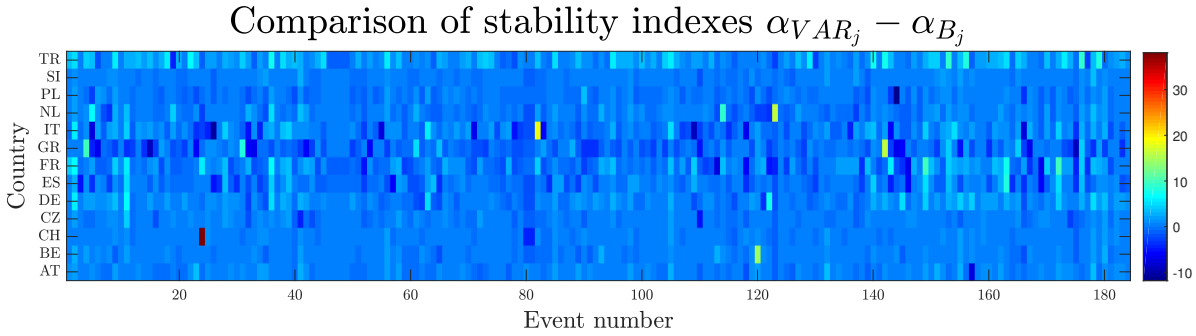


Figure 10: Heatmap visualizing the change in the stability index between scenario 3 (implementation of 30% RES) and the original model.

Given the complexity of the model, more than 500 iterations were required in order to obtain a new power flow. In this document, we present two additional initial power flows and the difference between the base case and the new operating points are depicted in Figure 11. In the first developed power flow (represented in orange bars), the shares of Turkey have been significantly reduced almost by half and the shares of Italy and Spain (Western-Europe) have been increased. In contrast, in the second developed power flow (green bars) the shares of Turkey have significantly increased and the share of Western-Europe has been reduced. After the two power flows were defined, the 184 simulations were performed again. The simulation results have been compared against the simulations performed using the original power flow and the results are summarized in Figure 13. In this figure, the evolution of the stability index (α) between the base case (presented in Section 2), and the RES integration for the new power flows is presented. For each simulation j, the performance index of the base case (α_{B_j}) and of the variation (α_{VAR_j}) are calculated and the results are compared in Figure 13, with the x-axis representing $\sum(\alpha_{VAR_j} - \alpha_{B_j})$. To get a better idea of these results, we can refer to Figure 3. In the same way, this ranking is created for each power flow and the performance value is added up.

The results displayed in Figure 13 show that the power flow plays an important role in the stability of the system and that incrementing the balance difference between the western and the eastern part of Europe is not realistic from a stability point of view. Particularly, it was noticed that the most critical event identified in Section 3 (a trip of a synchronous generator in Turkey) is even worst in the new power flow. On the contrary, the second power flow shows that the system becomes more robust or in other words, produces a positive impact on the stability of the system. From the simulations results, it can be concluded that reducing the gap between the East and West seems to be a reasonable solution to improve the stability of the system.

4 Discussion & Future Work

The methodologies and results presented in this document represent a substantial extension to the works presented before [11]. The main objective was to investigate the dynamic behavior of the European power grid before and after the implementation of decentralized renewable energy sources. Thus, in a first part, a performance index was developed with which the measurement results from the numerous simulations could be compared. The final results are covered in the second part of this work. Three phases of RES penetration were investigated; 10%, 20%, and 30% increase respectively, resulting in 39 variations of the initial dynamic model. The 30% variation case caused the greatest amount of stability issues in the system in comparison with the 10% and 20% variation case. In general, the results of 10% and 20% more RES differ very marginally. However, what was already identified as a potential problem in the original model has been confirmed by the three tested and analyzed scenarios; Turkey is one of the most sensitive countries. As the percentage of implemented RES increases, the effects of the outage of a certain sync. machine (see Table 2) are getting worse. These negative effects are namely inter-area oscillations and are not rare, yet to be avoided in such super grids. With the increase of interconnections over longer distances, the risk of groups of generators in different geographical locations starting to oscillate against each other is

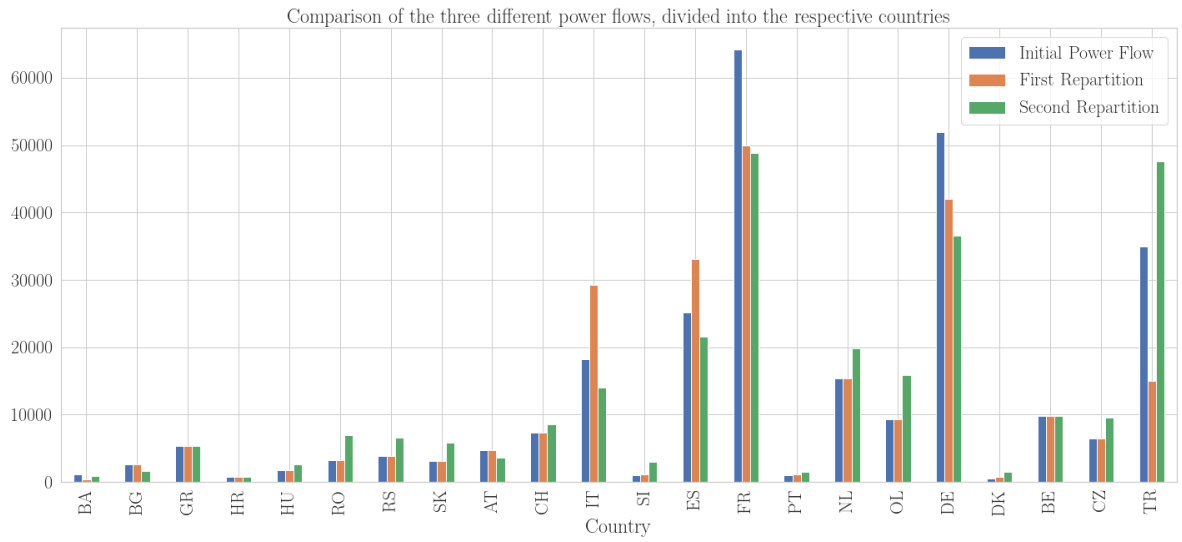


Figure 11: Representation of the three working power flows, divided into the different countries.

increasing respectively, creating electromechanical oscillations over the entire grid (in our case west to east). These events can limit the power transfer capacity and can even lead to a system collapse and blackouts if left undamped [23, 24].

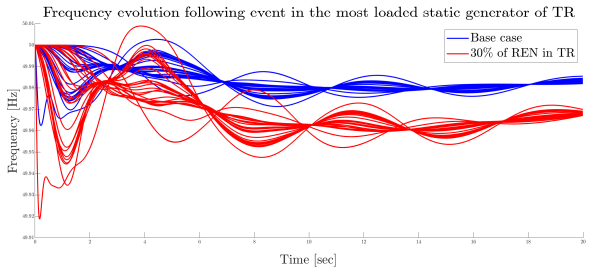


Figure 12: Frequency comparison of grid response to the trip of the most loaded static generator located in Turkey.

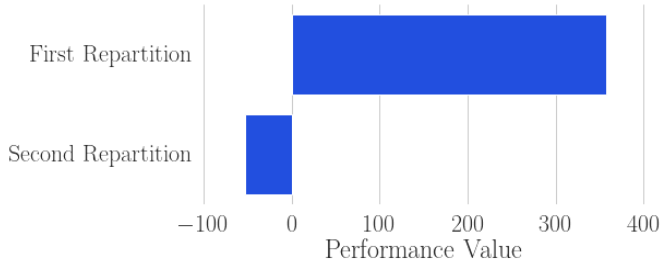


Figure 13: Stability index evolution depending on the initial power flows.

Power system stability problems are a topic that has been discussed for many years. Most approaches refer to the adjustment of Power System Stabilizers (PSS). However, considering the size of the European grid, dynamically adjusting all necessary PSS seems rather unrealistic. With the help of Wide Area Monitoring (WAM) and Wide Area Control (WAC), a different approach could be taken. By synchronously measuring the phase angle of the node voltages and currents via satellite-based Phasor Measurement Units (PMUs) distributed over a large area, a large system can be easily monitored and inter-area oscillations can be quickly detected. Wide Area Daming Controllers (WADC) play a key role in WAMs and help to damp such oscillations faster. The design of the controller is crucial to how efficiently such a WADC works.

In a next step, different methodologies for the design of such WADCs will be investigated and tested based on benchmark models. In a future step, these controllers will be implemented and tested in a real laboratory experiment. An adapted variant of Kundur's two area system has already been successfully developed in the smart grid laboratory in Winterthur, Switzerland [25]. Furthermore, electromechanical oscillations have been reproduced as well. This laboratory test will then be extended accordingly to enable the use of WACD. But even a test laboratory will eventually reach its limits in terms of size and complexity of a power & transmission grid. Therefore, these test trials are to be complemented by implementing the initial European model into Opal-RTs real-time simulator, coupled with control-in-the-loop to validate the effectiveness of a WADC and address main challenges like measurement noise & errors and control noise during the communication.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors acknowledge the Swiss National Science Foundation (SNSF) under the project number PZENP2 173628 of the program Ambizione Energy Grant (AEG). This research is also part of the activities of SCCER FURIES, which is financially supported by the Swiss Innovation Agency (Innosuisse-SCCER program).

References

[1] M. E. E. Zowalaty, S. G. Young, J. D. Järhult, "Environmental impact of the COVID-19 pandemic – a lesson for the future," Infection Ecology & Epidemiology, **10**(1), 1768023, 2020, doi:10.1080/20008686.2020.1768023.

[2] S. Saadat, D. Rawtani, C. M. Hussain, "Environmental perspective of COVID-19," Science of The Total Environment, **728**, 138870, 2020, doi: https://doi.org/10.1016/j.scitotenv.2020.138870.

[3] Swiss Federal Office of Energy SFOE, "Energy Strategy 2050 once the new Energy Act is in Force," PowerPoint Presentation, 2011. [Online].

[4] European Comission, "The European Green Deal," PDF Document, 2019. [Online].

[5] B. Schaefli, P. Manso, M. Fischer, M. Huss, D. Farinotti, "The role of glacier retreat for Swiss hydropower production," Renewable Energy, **132**, 615–627, 2019, doi:https://doi.org/10.1016/j.renene.2018.07.104.

[6] European Network of Transmission System Operator for Electricity (ENTSO-E), "EC Green paper: A 2030 Framework for climate and energy policies, an ENTSO-E consuiltation response," PDF Document, 2014. [Online].

[7] Swissgrid, "Overview of Ancillary Services," PDF Document, 2019. [Online].

[8] A. Berizzi, "The Italian 2003 blackout," in IEEE Power Engineering Society General Meeting, 2004., 1673–1679 Vol.2, 2004, doi:10.1109/PES.2004.1373159.

[9] European Network of Transmission System Operator for Electricity (ENTSO-E), "ENTSO-E Dynamic Model of Continental Europe," Operation Report, 2019. [Online].

[10] F. R. Segundo Sevilla, P. Korba, K. Uhlen, E. Hillberg, G. Lindahl, W. Sattinger, "Evaluation of the ENTSO-E initial dynamic model of continental Europe subject to parameter variations," in 2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), 1–2, 2017, doi:10.1109/ISGT.2017.8085986.

[11] C. Rüeger, J. Dobrowolski, P. Korba, F. R. Segundo Sevilla, "Stability Effects After Massive Integration of Renewable Energy Sources on Extra-Large Power Systems," in 2020 IEEE PES Transmission Distribution Conference and Exhibition - Latin America (T D LA), 1–6, 2020, doi:10.1109/TDLA47668.2020.9326181.

[12] S. Dasgupta, M. Paramasivam, U. Vaidya, V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," in 2014 IEEE PES General Meeting — Conference Exposition, 1–1, 2014, doi:10.1109/PESGM.2014.6939198.

[13] C. Rüeger, J. Dobrowolski, P. Korba, F. R. Segundo Sevilla, "Lyapunov Exponent for Evaluation and Ranking of the Severity of Grid Events on Extra-Large Power Systems," in 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 1–5, 2019, doi:10.1109/ISGTEurope.2019.8905564.

[14] European Network of Transmission System Operator for Electricity (ENTSO-E), "TYNDP 2016 Scenario Development Report," PDF Document, 2015. [Online].

[15] P. Kundur, Investigation of low frequency inter-area oscillation problems in large interconnected power systems, Canadian Electrical Association, Montreal, Canada, 1993.

[16] V. Vittal, N. Bhatia, A. Fouad, "Analysis of the inter-area mode phenomenon in power systems following large disturbances," in 1991., IEEE International Sympoisum on Circuits and Systems, 982–985 vol.2, 1991, doi: 10.1109/ISCAS.1991.176528.

[17] P. H. J. Nardelli, M. S. Baptista, C. A. Pomalaza-Raez, P. Cardieri, M. Latva-aho, "Network models for the structures of modern electric power grids," CoRR, **abs/1401.0260**, 2014.

[18] S. Dasgupta, M. Paramasivam, U. Vaidya, V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," in 2014 IEEE PES General Meeting — Conference Exposition, 1–1, 2014, doi:10.1109/PESGM.2014.6939198.

[19] C. Sourkounis, P. Tourou, "Grid Code Requirements for Wind Power Integration in Europe," Conference Papers in Science, **1**, 9, 2013, doi:https://doi.org/10.1155/2013/437674.

[20] M. Bollen, M. Häger, "Power quality: interactions between distributed energy resources,the grid, and other customers," Leonardo Energy, **1**, 2005.

[21] DIgSILENT, Gomaringen, Germany, DIgSILENT PowerFactory 2018 User Manual, 2018.

[22] M. Reza, "Stability analysis of transmission system with high penetration of distributed generation," 2006.

[23] A. P. G. A. (APG), "Almost Blackout on January 8, APG Publishes Video-Simulation," Website, 2021. [Online].

[24] E. N. of Transmission System Operator for Electricity (ENTSO-E), "Analysis of ce inter-area oscillations of 1st december 2016," PDF Document, 2017. [Online].

[25] D. Baltensperger, J. Dobrowolski, A. Obushevs, F. R. Segundo Sevilla, P. Korba, "Scaling Version of Kundur's Two-Areas System for Electromechanical Oscillations Representation," in 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 243–249, 2020, doi:10.1109/SPEEDAM48782.2020.9161960.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 43-53 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Antecedents to Learners' Satisfaction with Serious Games: An Investigation Using Partial Least Square

Ruben Chambilla[1], Daniel Tomiuk[2,*], Cataldo Zuccaro[2], Michel Plaisent[2], Prosper Bernard[2]

[1]*Faculté des Sciences, Université du Québec à Montréal, Montréal, H2X 3Y7, Canada*

[2]*Faculté des Sciences de la Gestion (ESG), Université du Québec à Montréal, Montréal, H3C 3P8, Canada*

ARTICLE INFO

Article history:
Received: 16 July, 2021
Accepted: 11 October, 2021
Online: 13 November, 2021

Keywords:
Serious Games
Learner's Satisfaction
Partial Least Square

ABSTRACT

Serious games are at the forefront of discussions about the future of learning. Research shows that they can help improve student motivation and knowledge transfer by making learning experiences more pleasurable. Teaching practices are increasingly enhanced or complemented by serious games; skills and knowledge are developed by recreating realistic situations allowing learners to enhance their procedural knowledge, all while having fun. Literature is scarce when it comes to identifying factors that influence learners' satisfaction when using serious games. In this paper, we present empirical results from an ongoing research project. We developed a questionnaire using factors identified in the literature and collected the responses of n = 110 business students from classes using serious games as a tool. Analyses were performed using partial least squares structural equation modeling (PLS-SEM). We used the following predictor variables for our model: performance and status feedback and tracking, ease of use, reliability, perceived control, instructor support, aesthetics, realism entertainment, goal clarity, immersion, and progressive challenge. Results show that sense of control, entertainment and effectiveness have a direct positive influence on learners' satisfaction while other factors influence satisfaction by mediation or as components of theoretically justified higher-order constructs. The sample size and its composition limit the generalization of results. Further studies are needed.

1. Introduction

With today's many new applications and technological innovations, learning using serious games appears as an increasingly viable alternative to more traditional teaching environments. Serious games combine both playful aspects (e.g., rules of control, scoring, challenge, and artificial characters with utilitarian contents (e.g., learning, teaching, information). [1]. Serious games are often thought of as educational games and these terms are often used synonymously [2]. Although they are games, they do not pursue entertainment, enjoyment, or fun as their primary purpose. Instead, the gaming component is used as a motivator to engage and captivate players to encourage and facilitate learning. Serious games, particularly those that simulate reality follow a learning model often qualified as *experimental* [3]. This category has two characteristics: (1) it is linked to a learning cycle based on experimentation and reflection and (2) it is based on anchoring learning in a simulated reality, building on practical experiences [3] while learning from one's mistakes without the inherent risks associated with failure in the real-world.

Virtually recreating real-world situations allows learners to develop procedural knowledge and skills by performing complex tasks without any of the real risks typically associated with bad decisions [4]. Learning objectives are all achieved while having fun. The ludic aspect is, in fact, what separates serious games from other learning environments that may also make use of information technology. Industry predictions attest to the popularity of game-based learning. According to a recent Global Game-Based Learning Market report, serious games will probably reach $17 billion by the year 2023, surpassing 37% in terms of growth rate. This is being driven by improvements and innovation in virtual and augmented reality and artificial intelligence [5].

Nevertheless, the literature on serious games also acknowledges limits and difficulties. Primarily, good games are not easy to design. Games developed by the entertainment industry, set up exclusively for amusement, often fail. Arguably, designing successful serious games is even more difficult. Serious games must not only be entertaining; they must also help

*Corresponding Author: Daniel Tomiuk, tomiuk.daniel@uqam.ca

https://dx.doi.org/10.25046/aj060607

students achieve learning goals. Presented by this dual challenge, many serious game projects continue to fail [6].

When a learning environment is well-designed, it generates a more positive impact on the player's attitude and learning experience [7]. However, relatively few studies have tried to identify the antecedents of serious game success in the eyes of learners. Although their popularity continues to steadily increase, few studies have empirically examined the factors that affect learners' satisfaction with serious games. The purpose of our research was to identify factors affecting learner satisfaction in a serious gaming environment and to assess the relative importance of these factors.

We collected data collection from 110 university business students enrolled in courses using serious games to simulate real life managerial situations. We used SmartPLS3 software to conduct statistical analysis. Results from our analysis show that among the twelve antecedents of satisfaction identified through our literature review, sense of control, fidelity, effectiveness, and entertainment influence directly students' satisfaction with serious games. However, further analyses revealed the presence of multicollinearity among the independent variables of the model making some factors appear statistically unimportant. Treatment of our initial model was attempted to remedy the situation by considering the existence of mediating variables and theoretically justified superordinate constructs depicted as second-order factors. Our results show that while some factors affect learners' satisfaction with serious games directly, others do so indirectly through mediating factors or as part of higher-order factors and, as such, should not be neglected when designing successful games.

2. Literature Review

2.1. Benefits of Serious Games

Serious games are increasingly used in academia [8]; [9]. Several reasons explain the growing interest in serious games. First, an increasing number of educational institutions are transitioning from traditional pedagogical and didactic to more learner-centered approaches. The latter emphasizes the more active and participative role of learners [10] and serious games can create authentic (ecologically valid) and immersive environments [11]. Their interactivity helps learners in better understanding complex tasks, knowledge acquisition, and retention. Serious games also promote better concentration more than passive learning environments [12]. This represents a major transition from a traditional listening/observing only learning model to one founded on live experiences and practice. Compared to traditional classrooms where knowledge transmission is typically passive, serious games are more effective because of the interactivity they allow [13]. Moreover, the playfulness of games promotes greater interest in learning activities, engagement, focus, and enthusiasm. Serious games provide several benefits to learners [14]. Well-designed games are fun, enjoyable and increase student motivation. Linked to motivation is the concept of satisfaction. Studies have evidenced the importance of this concept in education [15], and yet, few studies have examined the antecedents of learners' satisfaction while interacting with serious games [16]; [17].

2.2. Satisfaction with Serious Games

In business, the concept of customer satisfaction is a vital component of a company's success. Research shows that employee satisfaction enhances productivity and organizational performance [18]. Customer satisfaction is positively linked to perceived service quality. Satisfaction also impacts positively on customer loyalty [19]. In the area of education, studies have demonstrated a positive relationship between learner satisfaction and both motivation and achievement [20]. Gameplay can generate strong, positive feelings in learners. These, in turn, motivate learners to continue playing (i.e., continue using the technology) and learning. Learners' satisfaction with serious games is more pronounced when positive emotions are high [21].

Satisfaction is one of the most researched concepts in the area of Information Systems (IS) [22]. When compared to other measures, user satisfaction is among the strongest of indicators employed by researchers to measure IS success [23]. According to [24], user satisfaction is "*the affective attitude towards a specific computer application by someone who interacts with the application directly*". Through a literature review, our paper first seeks to identify antecedents of serious game satisfaction from the perspective of learners. We then propose and empirically test a model to investigate the degree to which these antecedents relate to satisfaction.

2.3. Antecedents of Satisfaction with Serious Games

Twelve main factors were identified from an intensive review of the literature [25] as potential predictors of satisfaction, defined below :

Ease of Use is borrowed from the Technology Acceptance Model [26], it is defined as the degree to which a particular technology is perceived to be intuitive and simple to use.

Reliability is the degree to which the serious game worked smoothly and without problems. In Marketing, reliability has been identified as a crucial dimension of perceived service quality (i.e., the literature on SERVQUAL), while in Information Systems (IS) /Information Technology (IT) research, it refers to whether a system functions properly and can be defined as the degree to which a user possesses inherent confidence in the system [27].

Aesthetics reflects a subjective experience that is evoked by players during the process of interaction with the game interface [28]; [29]; [30], namely using sound, sight, and motion [31].

Fidelity/Realism describes the psychological and physical similarities that exist between a particular game and the environment it seeks to represent [32], the realism of a simulation [33].

Instructor assistance and support. Several studies show that the teacher plays an essential role in the success of IT integration in classrooms [34]. We define instructor assistance as the degree to which learners perceive that the instructor provided encouragement, help, and information on the use of the serious game. Instructors provide assistance at two levels, theoretical and technical [35].

Perceived control is defined as the degree to which learners feel capable of influencing and bringing about change to the world of play [36]. This can be achieved via multiple customization options such as, the pace of the game, the choice of the type of feedback the learner believes would be useful to progress more successfully through the game and even the capability of manipulating the appearance of one's playable character [37]. Research shows that when the learner has more control over their learning, they become more involved in the game and even try more complex strategies [38] and a resulting increase in positive attitudes [39].

Progressive challenge is the degree to which the game challenges the learner with increasingly difficult problems and increasingly complex challenges. Successful game progression is built on previously acquired knowledge, abilities, and skills developed and learned while playing [40]. Successful learning ensues when players progressively adapt to and succeed in performing tasks that are increasingly challenging and complex.

Effectiveness is defined as the capacity/power of the serious game to improve the knowledge of learners [41]. Whereas conventional games seek to boost the playfulness aspect of the game, serious games seek to create learning environments that support improvements in skills training (e.g., professionals). These skills are other than those associated purely with play. Indeed, research shows that serious games can be very effective in improving professional and academic skills [42].

Status/Performance tracking and feedback is the degree to which the simulation's progress information and feedback is perceived as constructive and helpful. Assessment and feedback are pillars of teaching effectiveness [43]. They help learners achieve the game's objectives [44]. Performance feedback must be specific and presented at the right time [45]. Feedback includes information on the state of the player, instructor help/assistance during play, tips and clues [41]. It can take the form of visual, auditory, tactile prompts [46].

Entertainment is the degree to which the learner finds the game play pleasurable. Research defines *pleasure* as an attitude or an emotion that can satisfy intrinsic needs and produce a positive reaction [47]. The entertainment value derived from gameplay is largely determined by the quality of design of the "educational ludo" environment [48]. Entertainment value impacts positively on learners' satisfaction.

Immersion is the ability of a game to create an immersive experience in the learner [49], the feeling of "being there" [50]. It occurs when the learner temporarily accepts the reality of a game and focuses on this reality through his involvement in the game. It reflects the degree of player engagement in the game environment [49].

Goal clarity refers to the need to clearly define what objectives must be achieved during game play. It is a fundamental principle of instructional design [51]; [52]. Studies (e.g., [53]) demonstrate that serious game environments can fail to promote concentration or rouse curiosity in learners mainly because their goals are ill-defined, ambiguous, or misunderstood. Clarity of objectives is crucial in effective game design because it serves to increase commitment and concentration [54].

3. Hypotheses

In this section, we present the hypothesized relationship between each factor and the learner's satisfaction with the serious game (see Table 1). The questionnaire items for each measure/factor are presented in the Appendix. For the sake of parsimony, only questionnaire items that successfully passed the card-sorting exercise (item validation) are presented.

Table 1: Hypotheses of Original Model

H1	A positive and significant relationship exists between ease of use and learners' satisfaction
H2	A positive and significant relationship exists between reliability and learners' satisfaction
H3	A positive and significant relationship exists between aesthetics and learners' satisfaction
H4	A positive and significant relationship exists between fidelity and learners' satisfaction
H5	A positive and significant relationship exists between instructor assistance and support and learners' satisfaction
H6	A positive and significant relationship exists between feeling of control and learners' satisfaction
H7	A positive and significant relationship exists between progressive challenge and learners' satisfaction
H8	A positive and significant relationship exists between and learners' satisfaction
H9	A positive and significant relationship exists between performance/status tracking and feedback and learners' satisfaction
H10	A positive and significant relationship exists between entertainment and learners' satisfaction
H11	A positive and significant relationship exists between immersion and learners' satisfaction
H12	A positive and significant relationship exists between goal clarity and learners' satisfaction

4. Methodology

PLS-SEM (Partial Least Square - Structural Equation Modeling) is a multivariate statistical modelling technique increasingly being used by researchers in IS/IT, education and in other areas because of several advantages. It is often favoured when analyzing models using smaller sample sizes or data with non-normal distributions. Specifically, we used the statistical software package called SmartPLS.

4.1. Item Creation and Refinement

We generated an initial set of items and then refined and validated them before including them in a questionnaire. Steps were as follows.

- Creation or adaptation of items based on the definition of concepts [55].
- Validation of the questionnaire items using the card-sorting method [56].

The card sorting method seeks to verify, whether the items initially chosen by the researcher to be included in the questionnaire relate accurately to the concepts that (s)he is trying to measure. For this, four (4) judges were recruited: two professors from the IS/IT area and two students enrolled in a master's program in Information Systems/Information Technology. The validation process required each judge to sort the initial questionnaire items by properly combining each one with the actual construct that item was intended to measure. When at least 3 out of 4 judges (or 75% of judges) would successfully associate a specific item to its expected construct, the item in question was deemed to accurately reflect the concept it was intended to measure and was therefore included in the questionnaire.

4.2. Data Collection and Sample Size

After validating the initial set of items using the pretest and the sort-by-card method, the items were included in an online questionnaire created using the Google Forms application. Three (3) lecturers from the business school of our university were contacted to help us recruit students. These lecturers were chosen because, in their classes, they used a serious gaming platform that simulated the typical challenges and problems encountered by managers of small- and medium-sized businesses. Over a period of approximately one (1) week, we visited the classrooms of these lecturers to explain to students the purpose of our study and invite them to participate. We printed information sheets containing our contact details and the URL for the online questionnaire. During our visits to classrooms, these sheets were provided to students. All questionnaire items were of the Likert type ranging from 1 (Strongly disagree) to 5 (Strongly agree).

Data collection was done over a period of three (3) weeks. Two emails were sent to the lecturers to remind their students to participate in the study. Participation was completely voluntary. All in all, there were 345 students enrolled in all the classes we had visited. A total response of n = 110 was received across all students enrolled in the courses, a response rate of about 33%.

Demographic information was as follows: As for age, 23- to 27-year-old students represented the largest group of participants (45% of respondents). As for gender, there was just about the same number of women as there were of men: 55 and 49 respectively while six (6) respondents preferred not to specify their gender or chose "other". In terms of education, 37% of participants hadn't finished their undergraduate studies. The remaining students specified that they had, at least, obtained a bachelor (or equivalent degree).

4.3. Original Model

The Partial Least Square - Structural Equation Modeling (PLS-SEM) approach allows the modelling and testing of complex cause-and-effect relationships. Two of the strengths of PLS-SEM is that the technique allows the modeling of latent variables and works well with smaller samples [57]. We used SmartPLS version 3.3.0.

Measurement Model - In SmartPLS, the measurement model was solved using the *factor* weighing scheme. Evaluating the measurement model includes an analysis of (1) reliability, (2) convergent validity, and (3) divergent validity [58]. We used Cronbach's alpha to analyze reliability and refine our scales. The cut-off value was 0.7 as recommended by [59]. Convergent validity was achieved by retaining items with loadings above or equal to 0.7. We also tentatively retained items with loadings between 0.4 and 0.7. These were later scrutinized further by examining their *composite reliability* (CR) and *average variance extracted* values (AVE). If CR was greater or equal to 0.7 and AVE was greater or equal to 0.50, the item was considered acceptable and was retained in the model [60]. All items with loadings that did not attain a minimum threshold of 0.4 were removed, one by one, starting with the item with the smallest loading. After each item removal, the model was recalculated and re-evaluated. This process continued until only acceptable items remained. The final loadings of indicators from the measurement model are presented in the Appendix.

Next, we turned towards establishing discriminant validity. For this, the *Heterotrait-Monotrait Ratios* (HTMT) were examined. A value of 0.85 or less for HTMT indicate good discriminant validity [61]. The two highest HTMT values in our model were between *effectiveness* and *entertainment* and between *entertainment* and *immersion*, with values of 0.724 and 0.703, respectively. Acceptable Cronbach alpha, CR, and AVE values presented in Table 2 all point to a good measurement model.

Table 2 - Construct Validity and Reliability

Construct	Cronbach alpha	Composite Reliability	Average Variance Extracted
Assistance and Support	0.922	0.944	0.809
Fidelity	0.922	0.945	0.810
Effectiveness	0.902	0.928	0.720
Aesthetics	0.874	0.911	0.641
Entertainment	0.859	0.914	0.780
Goal Clarity	0.839	0.886	0.611
Immersion	0.834	0.881	0.599
Satisfaction	0.832	0.900	0.749
Reliability	0.829	0.883	0.654
Progressive Challenge	0.804	0.872	0.633
Performance and Feedback	0.775	0.869	0.689
Ease of Use	0.767	0.852	0.660
Control	0.700	0.869	0.768

Structural Model - The structural model was generated using a path weighing scheme. A bootstrap of 5000 subsamples was used to generate the p-values. Path coefficients and t-values using SmartPLS3 are presented in Table 3 below which also shows whether the associated hypothesis is supported or not.

Out of all the paths specified in our original model, only four (4) were significant (at the 0.05 level). The highest path value is between entertainment and satisfaction (0.258), followed by effectiveness and satisfaction (0.205), control and satisfaction (0.187), and fidelity and satisfaction (0.164), in decreasing order. As such, only hypotheses H4, H6, H8, and H10 were supported. The paths not showing up as significant were perplexing. For instance, in our model, ease of use has no significant effect on

learners' satisfaction. This result contradicts numerous studies which evidence the importance of perceived ease of use on users' technology acceptance. We suspected the presence of multicollinearity in our model.

Table 3: Path Coefficients and p-values (Original Model)

Supported Hypotheses	**Path Coeff.**	**p-value**
H4 (Fidelity → Satisfaction)	0.164	0.049*
H6 (Control → Satisfaction)	0.187	0.049*
H8 (Effectiveness → Satisfaction)	0.205	0.038*
H10 (Entertainment → Satisfaction)	0.258	0.014*
Note: * = p < 0.05		
Unsupported Hypotheses	**Path Coef.**	**p-value**
H1 (Ease of Use → Satisfaction)	0.006	0.937
H2 (Reliability → Satisfaction)	0.088	0.250
H3 (Aesthetics → Satisfaction)	-0.021	0.788
H5 (Assistance and Support → Satisfaction)	0.035	0.656
H7 (Progressive Challenge → Satisfaction)	0.098	0.250
H9 (Performance and Feedback → Satisfaction)	-0.027	0.702
H11 (Immersion → Satisfaction)	0.096	0.267
H12 (Goal Clarity → Satisfaction)	0.082	0.334

4.4. Multicollinearity, Higher-Order Constructs, and Mediation

Statistical models that contain several independent variables that are highly correlated to one another and not just with the dependent variable can cause a serious problem called multicollinearity. It makes accurate interpretation of statistical results difficult because it increases the standard errors of the coefficients. By overinflating the standard errors, multicollinearity may cause coefficients of some independent variables to appear to be non-significant. In other words, multicollinearity makes it look like the relationships between some independent variables and the dependent variables are statistically insignificant.

To test for multicollinearity, we first ran a SmartPLS model which contained only two (2) constructs: ease of use and satisfaction. As expected, and supported by literature, the path was now highly significant (path=0.408; p=0.000) which suggested the presence of multicollinearity [62] and/or mediating variables [63] in the original model (i.e., misspecified paths).

According to [62], the presence of multicollinearity among the independent factors of a model may indicate that higher-order structures are present and should be considered. Respecifying a model with these higher-order structures can effectively reduce the effects of multicollinearity but should be theoretically argued and empirically tested. In other words, higher-order constructs should not be extracted through a purely data-driven approach. They should be grounded in theoretical logic extracted from the existing literature.

In a model, a lack of statistically significant paths when these are supported by theory (or strong conjectural evidence), can suggest the presence of mediating variables that are not being considered or are incorrectly specified in that model. When mediation occurs (figure 1), an independent variable (X) has little or no statistically discernable direct effect on a dependent variable (Y). Instead, X acts on Y via an intervening variable (M). As such, the independent variable (X) affects the mediating variable (M) which, in turn, affects the dependent variable (Y). In other words, an independent variable (X) influences the dependent variable (Y), but this influence is not direct. Instead, variable X influences variable Y through some other variable (M). Mediation can be complete or partial.

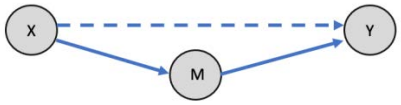


Figure 1: Mediation

4.5. Alternate Model

Following established guidelines to assess your measurement model and testing that discriminant validity falls within acceptable boundaries are very important steps in Structural Equation Modelling (SEM) but it does not guarantee that your structural model will yield significant paths [61]. Indeed, many of our initial hypotheses were not supported in our original model. Given that our analyses suggest the presence of higher-order constructs and mediation, we reworked our original model by respecifying alternative paths, paired with theoretical justification. These modifications to our original model are currently specified as propositions given that, in the coming year, we intend to explore in more detail existing literature to better substantiate these proposed alternate paths and the addition of higher-order structures.

Higher-Order Construct 1 → Platform Transparency

During our initial analyses, a significant correlation was noted between reliability and ease of use (0.434). Also, among all constructs considered in our study, these two were the only factors that were "technological" in nature.

In a now-famous article entitled "*The Computer for the 21st Century*" published in Scientific American in 1991 (pp 66-75), the author Mark Weiser writes that "*(t)he most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it*". The article argues that the best technologies are made or eventually become transparent to users. Technology should never be cumbersome to use. Learners' attention should be drawn to the contents of the communication rather than to the equipment used to communicate it, and players/learners should not be concerned by technological issues. Arguably, difficult to use, unintuitive interfaces, and unreliable IS/IT platforms riddled with recurrent bugs and glitches, are likely to take away from the learner's overall serious game experience and diminish one's enjoyment of it, and thus, should reflect negatively on the user's satisfaction with the serious game.

To represent this in our model, we re-specified *reliability* and *ease of use* as dimensions of a second-order construct which we named "*Platform transparency*". We define this second-order construct as *the perception that the game's underlying technological platform does not distract from the user's overall gaming experience and learning objectives because of inadequate design and technical problems*.

When technological platforms supporting serious gaming are inherently unstable, unintuitive, and/or difficult to use, it is likely to impede on the overall learning experience, reducing gaming enjoyment and negatively impacting one's sense of control over the entire experience. Previous empirical studies in IS support our assertions. There is a negative relationship between *ease of use* and one's perceptions of control and enjoyment while using information technologies (e.g., [64]).

- P1: The more the game's technological platform is perceived to be transparent during gaming (easy to use and reliable), the greater the learner's entertainment while playing the serious game.
- P2: The more the game's technological platform is perceived to be transparent during gaming (easy to use and reliable), the greater the learner's sense of control while playing the serious game.

Higher-Order Construct 2 → Guidance

In the original model, our analysis showed non-significant relationships between satisfaction and three (3) independent variables, namely *goal clarity, performance and feedback,* and *assistance and support.* Rather than modeling three (3) direct paths to satisfaction, we considered the possibility that mediation was present. It could be argued that, rather than impacting on satisfaction directly, these three (3) antecedents may be factors influencing learners' perceived effectiveness of the serious game, which would then impact satisfaction. In other words, perceived effectiveness could be a moderating variable. However, the path coefficients from *assistance and suppo*rt (path=0.132; p=0.102), from *performance and feedback* (path=0.016; p=0.843), and from *goal clarity* (path=0.130; p=0.112) to perceived *effectiveness* of the serious game were all non-significant.

Next, we reviewed the definitions of these three constructs collectively while re-examined their individual items to see whether an over-arching concept could connect them together. This led us to re-specify all three (3) concepts as dimensions of a common, second-order construct. We labeled this higher-order structure as "*Perceived guidance*". We define this as the *degree to which the user perceives that the serious gaming experience provides a clear understanding of goals and sufficient guidance and orientation on how to achieve them by adequately informing learners of their progress and by providing information about what they are doing right/wrong.* We propose that the more effective serious games are those characterized by clear, achievable goals where learners' progress is guided by regular, useful feedback reinforcing successful actions and results all while encouraging users to take corrective measures or try alternate strategies following unsuccessful ones. As such, we reason that *guidance* will impact satisfaction, but indirectly. The relationship will likely be mediated by perceived *effectiveness*. The following proposition is then stipulated.

- P3: The greater the *perceived guidance* received during the serious gaming experience, the more the learner will perceive the serious game as effective.

Mediation

Based on our literature review, our original model hypothesized a significant path between the serious game's *progressive challenge* and learner's *satisfaction*. Statistically, this path was not significant. We then reasoned that the learner's perceived *effectiveness* of the serious game could be mediating this relationship, making it look like *progressive challenge* has no effect on learners' satisfaction. Arguably, games designed to be progressively challenging are also likely to be perceived as being more effective by learners given that such games would allow learners to progress in their knowledge and skill acquisition gradually. We postulate the following proposition:

- P4: The greater the perceived progressive challenge of the serious game, the more the learner will perceive the game as being effective for learning.

In our original model, the path between *immersion* and *satisfaction* was statistically non-significant. This path was dropped and, instead, we specified a path between the concepts of *immersion* and *entertainment.* Studies have recognized that games' ability to create immersive experiences is among the most pleasurable aspects of video gameplay [65]. This brings us to the following proposition:

- P5: The greater the perceived immersion during serious gameplay, the greater the entertainment.

A path was also drawn linking *aesthetics* to *immersion.* Aesthetically superior games help in generating an immersive experience while playing suggesting that both concepts are linked. This relationship is expressed in the following proposition:

- P6: The better the aesthetics of a serious game, the greater the immersive experience perceived by the learner.

Research has also identified a link between the *entertainment* value of games and their perceived *effectiveness* [66]. Well-designed, fun games can put learners in a better mood which then positively impacts their perceived effectiveness of the serious game in supporting learning [67]. As such, we propose that a positive relationship exists between the entertainment value of a serious game and the perceived effectiveness of the game as a learning platform.

- P7: The greater the perceived entertainment value of a serious game, the greater the perceived effectiveness of that serious game.

In our original model, the path between *fidelity* and *satisfaction* was found to be positive and significant. However, the literature that pertains to the realism of games also suggests that there may be mediating effects between these two factors. It is likely that positive relationships exist between fidelity (how closely the game simulates reality) and the training effectiveness

of the serious game and immersion as well [68]. This brings us to the following propositions:

- P8: The greater the perceived fidelity of the serious game, the more the learner will perceive it as effective.
- P9: The greater the perceived fidelity of the serious game, the more the learner will perceive it as immersive.

With these propositions, we specified a new, alternative model which included the higher-order constructs described previously and added paths for mediation (Figure 2). We ran it using bootstrapping with 5000 subsamples.

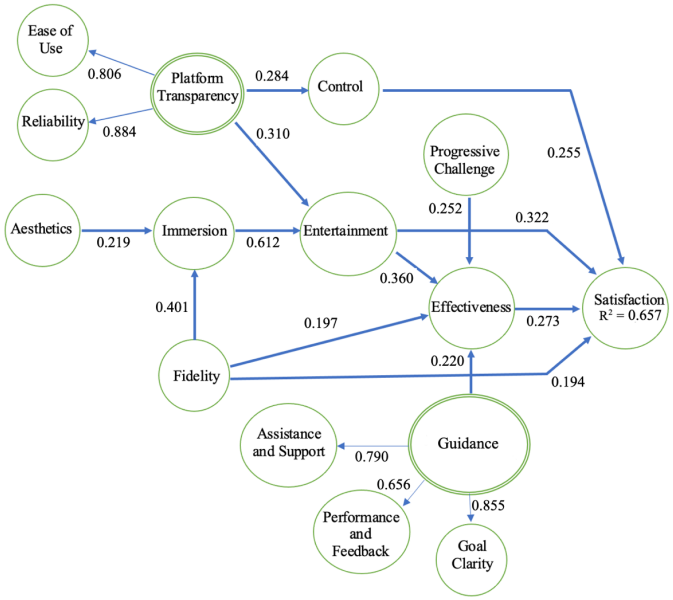


Figure 2: Structural Model (alternate model)

Tables 4 (hypotheses) and 5 (propositions) present the results of the structural model. In this respecified/alternate model, all paths are significant. This alternate model explains $R^2 = 65.7\%$ of the variance in our data on satisfaction.

Table 4: Hypotheses Path Coefficients and p-values (Alternate Model)

Hypotheses	**Path Coeff.**	**p-value**
H4 (Fidelity → Satisfaction)	0.194	0.009**
H6 (Control → Satisfaction)	0.255	0.001**
H8 (Effectiveness → Satisfaction)	0.273	0.002**
H10 (Entertainment → Satisfaction)	0.322	0.000***
Note: ** = p < 0.01; *** = p < 0.001		

Table 5: Proposition Path Coefficients and p-values (Alternate Model)

Propositions	**Path Coeff.**	**p-value**
P1 (Platform Transparency → Entertainment)	0.310	0.000***
P2 (Platform Transparency → Control)	0.284	0.017*
P3 (Guidance → Effectiveness)	0.220	0.001**
P4 (Progressive Challenge → Effectiveness)	0.252	0.002**
P5 (Immersion → Entertainment)	0.612	0.000***
P6 (Aesthetics → Immersion)	0.219	0.011*
P7 (Entertainment → Effectiveness)	0.360	0.000***
P8 (Fidelity → Effectiveness)	0.197	0.017*
P9 (Fidelity → Immersion)	0.401	0.000***
Note: * = p < 0.05; ** = p < 0.01; *** = p < 0.001		

5. Conclusion

Through a literature review on serious games, we identified several factors which seemingly impact learners' satisfaction with serious games. Using these factors, we tested an original model using PLS-SEM. This original model specified paths going directly from each factor directly into satisfaction suggesting direct effects.

However, several paths from our original model were statistically non-significant. We respecified an alternate model supported by theoretical foundations and previous research. It included both mediation and second-order factors which helped provide a richer, more complete picture of how factors identified through our literature review relate with one another to positively impact learners' satisfaction with serious games. Every factor considered in this study was found to have a positive effect on learners' satisfaction, although not necessarily a direct effect. Several relationships were mediated through other factors while some factors are best combined into higher-order structures. Results suggest that none should be neglected when designing serious games.

Overall, the ludic component plays a central role in serious games. Entertainment value derived from playing serious games has a strong, positive impact on learner satisfaction. Although they are primarily learning tools, people perceive them as games, and as such, expect them to be designed to be fun to play. The importance of entertainment value is followed by whether the learner perceives the serious game as effective, the learners' sense of control during play, and last, by the fidelity/realism of the game. These findings suggest that to design successful serious games that satisfy learners, it is important to create a ludic environment that is fun, simulates reality well, all while making learners feel like they are in control of this environment. Perceived effectiveness of the serious game holds an important role in generating satisfaction in learners. Perceived effectiveness is a direct antecedent to satisfaction and acts as a mediating variable for several of the other variables identified in our literature review. Although other variables gained mediator status in our alternate model, none mediated a greater number of variables than did effectiveness. Our results also show that some variables do not impact satisfaction directly. Their effects on satisfaction are completely mediated through other factors. This includes aesthetics, immersion, and progressive challenge.

An important limitation of this study is the sample size. Although PLS-SEM handles small sample sizes well, small sample sizes raise questions about the generalizability of results. Another limitation of this study is that it is cross-sectional. This means that results do not take into account how time influences the relationship among variables. A factor can be more influential on satisfaction from the onset of play. Others may become more important in explaining satisfaction later on, as the learner amasses more experience. For instance, the importance of providing increasingly challenging and more varied learning

situations may become more and more important in explaining whether players continue to derive satisfaction from a serious game as their experiences and skill set grow. As such, the temporal stability of factors as antecedents of satisfaction should be investigated in future studies. Moreover, our surveys use self-reporting. These can cause respondents to inflate their opinions and responses making it hard to capture the true relationships among constructs. Interviews could be conducted in the future to validate our results. Since the respondents are all business students a generalization can be challenged for other disciplines. Finally, research is further needed to identify additional factors that may impact satisfaction with serious games and to explore the nature of the relationships that may exist between these factors.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] J. Alvarez, D. Djaouti, O. Ramouroux, Apprendre avec les serious games? Editions Canopé, Poitiers, France, 2016.

[2] B. Cowan, B. Kapralos, "An overview of serious game engines and frameworks," Intelligent Systems Reference Library, **119**, 15-38, 2017, doi:10.1007/978-3-319-49879-9_2.

[3] P. Cohard, "L'apprentissage dans les serious games : proposition d'une typologie," @GRH, **16**(3), 11-40, 2015, doi : 10.3917/grh.153.0011.

[4] M.S. Calabora, A. Moraa, S. Moyab, "The future of serious games in accounting education: a delphi study," Journal of Accounting Education, **46**, 43-52, 2019, doi:/10.1016/j.jaccedu.2018.12.004.

[5] R. Valentine, "Metaari: game-based learning market will reach $17 billion by 2023," Gamesindustry.biz, 8th August 2018, https://www.gamesindustry.biz/articles/2018-08-08-metaari-game-based-learning-market-will-reach-usd17-billion-by-2023.

[6] A. Usoltsev, "Serious games: 5 gamification pitfalls to avoid," August 8 2018, https://zenodo.org/record/2847262

[7] A. Messaoud, "Conception d'un jeu sérieux actif pour l'apprentissage intergénérationnel du français," Québec, Canada: Université de Laval, 2016.

[8] S. De Freitas, "Are games effective learning tools? a review of educational games," Educational Technology & Society, **21**(2), 74–84, 2017, http://www.jstor.org/stable/26388380.

[9] C.J. Stettina, T. Offerman, I. Sidhu, B. De Mooij, "Gaming for agility: using serious games to enable agile project & portfolio management capabilities in practice," Engineering, Technology and Innovation (ICE) & IEEE International Technology Management Conference, Stuttgart, 2018, doi.org/10.1109/ice.2018.8436384

[10] R. Rodela, A. Ligtenberg, R. Bosma, "Conceptualizing serious games as a learning-based intervention in the context of natural resources and environmental governance," Water, **11**(245), 1-15, 2019, doi.org/10.3390/w11020245.

[11] J. Lumsden, A. Skinner, A.T. Woods, N.S. Lawrence, M. Munafò, "The effects of gamelike features and test location on cognitive test performance and participant enjoyment," PeerJ., **4**, 2-19, 2016, doi.org/10.7717/peerj.2184.

[12] I. Bououd, R. Fafi. "Using serious games to correct french dictations: proposal for a new unity3d/nooj connector," Association for Computational Linguistics, 49–52, Septembre 4 2017, doi.org/10.18653/v1/W17-3808.

[13] A. Duguet, S. Morlaix, "Le numérique à l'université : facteur explicatif des méthodes pédagogiques?," Revue Internationale de Pédagogie de l'Enseignement Supérieur, **34**(3), 1-18, 2018, doi.org/10.4000/ripes.1682.

[14] J. Stewart, L. Bleumers, J, Van Looy, I. Mariën, A. All, D. Schurmans, K. Willaert, F. De Grove, A. Jacobs, G.J. Misuraca, L. Stewart, V.-L.J. Bleumers, M. Ilse, A. All, D. Schurman, K. Willaert, "The potential of digital games for empowerment and social inclusion of groups at risk of social and economic exclusion: evidence and opportunity for policy," Institute for Prospective Technological Studies - European Commission, **II**, 172, 2013, doi.org/10.2791/88148.

[15] W.S. Chen, A.Y. Yao, "An empirical evaluation of critical factors influencing learner satisfaction in blended learning: a pilot study," Universal Journal of Educational Research, **4**(7), 1667-1671, 2016, doi:10.13189/ujer.2016.040719.

[16] K.C.M. Bul, I.H.A. Franken, S. Van der Oord, P.M. Kato, M. Danckaerts, L.J. Vreeke, A.Willems, H.J.J. van Oers, R. van den Heuvel, R. van Slagmaat, and A. Maras. "Development and user satisfaction of plan-it commander a serious game for children with ADHD," Games for Health Journal, 4(6), 502-512, 2015, doi.org/10.1089/g4h.2015.0021.

[17] M.A. Khenissi, F. Essalmi, M. Jemni, "Comparison between serious games and learning version of existing games," Procedia - Social and Behavioral Sciences, **191**, 487-494, 2015, doi.org/10.1016/j.sbspro.2015.04.380.

[18] B. Budie, R. Appel-Meulenbroek, A. Kemperman, M. Weijs-Perree, "Employee satisfaction with the physical work environment: the importance of a need based approach," International Journal of Strategic Property Management, **23**(1), 36-49, 2019, doi.org/10.3846/ijspm.2019.6372.

[19] A. Akhmedova, A.Manresa, D. Escobar Rivera, A. Bikfalvi, "Service quality in the sharing economy: a review and research agenda," International Journal of Consumer Studies, **45**(4), 889-910, 2021, doi.org/10.1111/ijcs.12680.

[20] N.M. Martirosyan, D.P. Saxon, R. Wanjohi, "Student satisfaction and academic performance in armenian higher education," American International Journal of Contemporary Research, **4**(2), 1-5, 2014, http://www.aijcrnet.com/journals/Vol_4_No_2_February_2014/1.pdf.

[21] E. Alsadoon, "The impact of social presence on learners' satisfaction in mobile learning," TOJET: The Turkish Online Journal of Educational Technology, **17**(1), 226-233, 2018, https://files.eric.ed.gov/fulltext/EJ1165749.pdf.

[22] R. Vaezi, A. Mills, W.W. Chin, H. Zafar, "User satisfaction research in information systems: historical roots and approaches," Communications of the Association of Information Systems, **38**, 501-532, 2016, doi.org/10.17705/1CAIS.03827

[23] W.H. DeLone, E.R. McLean, "Information systems success: the quest for the dependent variable," Information Systems Research, **3**(1), 60-95, 1992, doi.org/10.1287/isre.3.1.60.

[24] W.J. Doll, G. Torkzadeh, "The measurement of end-user computing satisfaction," MIS Quarterly, **12**(2), 259-274, 1988, doi.org/10.2307/248851.

[25] G. Petri, C. Gresse von Wangenheim, "How to evaluate educational games: a systematic literature review, Journal of Universal Computer Science, **22**(7), 992-1021, 2016, doi.org/10.1016/j.compedu.2017.01.004.

[26] F.D. Davis, "Perceived usefulness, perceived ease of use and user acceptance of information technology," MIS Quarterly, **13**(3), 319-340, 1989, doi.org/10.2307/249008.

[27] Y.H. Al-Mamary, A. Shamsuddin, A.H. Nor Aziati, "Key factors enhancing acceptance of management information systems in yemeni companies," Journal of Business and Management Research, **5**, 108-111, 2014, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1054.2814&rep=rep1&type=pdf.

[28] S. Dhar, V. Ordonez, T.L. Berg, "High level describable attributes for predicting aesthetics and interestingness," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **1**, 1657–1664, 2011, doi:10.1109/CVPR.2011.5995467.

[29] M. Pajusalu, The evaluation of user interface aesthetics, Master Thesis, Rennes/Tallinn: Tallinn University Institute of Informatics, 2012, http://www.cs.tlu.ee/teemad/get_file.php?id=202.

[30] J.A. Abubakar, A.S. Bahrin, M.K. Ahmad, A.N. Zulkifli, "Conceptual model of game aesthetics for perceived learning in narrative games," International Journal on Advanced Science Engineering Information Technology, **7**(3), 993-999, 2016, doi: 0.1109/IUSER.2016.7857944.

[31] Z. Herbert, Sight, Sound, Motion: Applied Media Aesthetics (6th ed.), Wadsworth Publishing Co. Inc., 2010.

[32] C. Crawford, The Art of Computer Design, Berkeley, CA: Osborne/McGraw-Hill, 1984.

[33] J. Pellegrino, A. Scott, "The Transition from Simulation to Game-Based Learning," Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Dec. 6-9, 2004, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.454.5486&rep=rep1&type=pdf

[34] K. McKnight, K. O'Malley, R. Ruzic, M.K. Horsley, J.J. Franey, K. Bassett, "Teaching in a digital age: how educators use technology to improve student learning," Journal of Research on Technology in Education, **48**(3), 194-211, 2016, doi.org/10.1080/15391523.2016.1175856.

[35] D. Djaouti, Serious Game Design: Considération Théoriques et

Techniques sur la Création de Jeux Vidéo à Vocation Utilitaire, PhD Thesis, Toulouse University, 2011.
[36] M. Busch et al., "Personalization in Serious and Persuasive Games and Gamified Interactions," in Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15) ACM, 811-816, 2015, https://uwspace.uwaterloo.ca/bitstream/handle/10012/12804/p811-busch.pdf?sequence=3
[37] N. Whitton, Learning and Teaching with Computer Games in Higher Education, Games-Based Learning Advancements for Multi-sensory Human Computer Interfaces: Techniques and Effective Practices, Chapter II, 18-33, 2009, doi.org/10.4018/978-1-60566-360-9.ch002
[38] D.I. Cordova, M.R. Lepper, "Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization and choice," Journal of Educational Psychology, **88**, 715-730, 1996, doi.org/10.1037/0022-0663.88.4.715.
[39] J.J. Vogel, D.S. Vogel, J. Cannon-Bowers, C.A. Bowers, K. Muse, M. Wright, "Computer gaming and interactive simulations for learning: a meta-analysis," Journal of Educational Computing Research, **34**(3), 229–243, 2006, doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM.
[40] N. Shin, L.M. Sutherland, C.A. Norris, E. Soloway, "Effects of game technology on elementary student learning in mathematics," British Journal of Educational Technology, **43**(4), 540-560, 2011, doi.org/10.1111/j.1467-8535.2011.01197.x.
[41] B. Ingadottir, K. Blondal, D. Thue, S. Zoega, I. Thylen, T. Jaarsma, "Development, usability, and efficacy of a serious game to help patients learn about pain management after surgery: an evaluation study," JMIR Serious Games, **5**(2), 1-15, 2017, doi: 10.2196/games.6894.
[42] S. Tobias, J.D. Fletcher, Computer Games and Instruction, Charlotte: IAP Information Age Publishing, 2011.
[43] D. Drummond, A. Hadchouel, A. Tesnière, "Serious games for health: three steps forwards," Advances in Simulation, **2**(3), 1-8, 2017, doi.org/10.1186/s41077-017-0036-3.
[44] J. Loisier, Étude sur l'Apport des Jeux Sérieux pour la Formation à Distance au Canada Francophone, Réseau d'Enseignement Francophone à Distance (REFAD), 2015, https://www.refad.ca/documents/Etude_Jeux_serieux_en_FAD.pdf.
[45] M.D. Kickmeier-Rust, D. Albert, "Micro-adaptivity: protecting immersion in didactically adaptive digital educational games," Journal of Computer Assisted Learning, **26**(2), 95-105, 2010, doi.org/10.1111/j.1365-2729.2009.00332.x.
[46] N. Iten, D. Petko, "Learning with serious games: is fun playing the game a predictor of learning success?," British Journal of Educational Technology, **47**(1), 151–163, 2016, doi.org/10.1111/bjet.12226.
[47] R. Tamborini, N.D. Bowman, A. Eden, M. Grizzard, A. Organ, "Defining media enjoyment as the satisfaction of intrinsic needs," Journal of Communication, **60**, 758–777, 2010, doi.org/10.1111/j.1460-2466.2010.01513.x.
[48] I. Marfisi-Schottman, S. George, F. Tarpin-Bernard, P. Prevot, "Comment Évaluer la Qualité d'un Learning Game Pendant sa Conception?," in Technologies de l'Information et de la Communication pour l'Enseignement, TICE2012, Lyon: France, 80-90, 2012, https://hal.archives-ouvertes.fr/hal-00963678/document.
[49] G. Calleja, "Revising Immersion: A Conceptual Model for The Analysis of Digital Game Involvement," in Proceedings of the 2007 Digital Games Research Association (DiGRA) International Conference, 82-90, 2007, http://www.digra.org/wp-content/uploads/digital-library/07312.10496.pdf.
[50] A. McMahan, "Immersion, Engagement and Presence: A New Method for Analyzing 3-D Video Games," in M. J. Wolf, & B. Perron (Eds), The Video Game, Theory Reader, New York: Routledge: Taylor & Francis Group, 67-87, 2003.
[51] M. Ebrahimzadeh, S. Alavi, "Digital video games: e-learning enjoyment as a predictor of vocabulary learning," Electronic Journal of Foreign Language Teaching, **14**(2), 145–158, 2017, http://e-flt.nus.edu.sg/v14n22017/ebrahimzadeh.pdf.
[52] J.G. Peery, C. Pasalar, "Designing the learning experiences in serious games: the overt and the subtle-the virtual clinic learning environment," Informatics, **5**(30), 1-14, 2018, doi.org/10.3390/informatics5030030.
[53] M. Ma, A. Oikonomou, L.C. Jain, Serious Games and Edutainment Applications (ed. London Limited). New York: Springer-Verlag, 2011.
[54] Y. Wang, R. Pramod, C. Sankar, P.K. Raju, "Relationships between Goal Clarity, Concentration and Learning Effectiveness When Playing Serious Games," in Proceeding of Twentieth Americas Conference on Information Systems, 2014 https://www.researchgate.net/publication/263426149_Relationships_between_Goal_Clarity_Concentration_and_Learning_Effectiveness_when_Playing_Serious_Games
[55] G.A. Churchill, "A paradigm for developing better measures of marketing constructs," Journal of Marketing Research, **16**, 64-73, 1979, doi.org/10.1177/002224377901600110.
[56] G.C. Moore, I. Benbasat, "Development of an instrument to measure the perceptions of adopting an information technology innovation," Information Systems Research, **2**(3), 192-222, 1991, doi.org/10.1287/isre.2.3.192.
[57] F. Ali, S.M. Rasoolimanesh, M. Sarstedt, C.M. Ringle, K. Ryu, "An Assessment of the use of partial least squares structural equation modeling (pls-sem) in hospitality research," International Journal of Contemporary Hospitality Management, **30**(1), 514-538, 2018, doi.org/10.1108/IJCHM-10-2016-0568.
[58] Chin, W. W., Esposito-Vinzi, V., Henseler, J., & Wang, H. (2009). How to Write Up and Report PLS Analyses. Handbook of Partial Least Squares, pp 655-690
[59] J.C. Nunnally, I.H. Bernstein, Psychometric Theory (ed. 3rd). New York: McGraw-Hill, 1994.
[60] J.F. Hair, C.M. Ringle, M. Sarstedt, "PLS-SEM: indeed a silver bullet," Journal of Marketing Theory and Practice, **19**(2), 139-152, 2011, doi.org/10.2753/MTP1069-6679190202.
[61] J. Henseler, C.M. Ringle, M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," Journal of the Academy of Marketing Science, **43**(1), 115-135, 2015, doi.org/10.1007/s11747-014-0403-8.
[62] C. Schwarz, A. Schwarz, W.C. Black, "Examining the impact of multicollinearity in discovering higher-order factor models," Communications of the Association for Information Systems, **34**(62), 1191-1208, 2014, doi:10.17705/1CAIS.03463.
[63] R. Agler, P. De Boeck, "On the interpretation and use of mediation: multiple perspectives on mediation analysis," Frontiers in Psychology, **15**, 1-11, 2017, doi:10.3389/fpsyg.2017.01984.
[64] V. Venkatesh, "Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model," Information Systems Research, **11**(4), 342–365, 2000, doi.org/10.1287/isre.11.4.342.11872.
[65] L. Michailidis, E. Balaguer-Ballester, X. He, "Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge, Frontiers in Psychology, 05 September 2018, doi: 10.3389/fpsyg.2018.01682.
[66] N. Nazry, M. Nazrina, D.M. Romano, "Mood and learning in navigation-based serious games," Computers in Human Behavior, **73**, 596–604, 2017, doi.org/10.1016/j.chb.2017.03.040.
[67] Y. Zhonggen, "A meta-analysis of use of serious games in education over a decade," International Journal of Computer Games Technology, **1**, 1-8, 2019, doi.org/10.1155/2019/4797032.
[68] D. Rojas, B. Kapralos, S. Cristancho, K. Collins, A. Hogue, C. Conati, A. Dubrowski, "Developing effective serious games: the effect of background sound on visual fidelity perception with varying texture resolution," Studies in Health Technology and Informatics, **173**, 386-392, 2012, doi:10.3233/978-1-61499-022-2-386.

Appendix

Item code	***Aesthetics***	***Loading***
AES1	The simulation is visually appealing	0.827
AES2	The visual design of the simulation is attractive	0.881
AES3	I appreciated the aesthetic side of the simulation	0.841
AES4	The color contrast makes it easier to understand windows	0.758
	Control	

CON1	I was able to complete the simulation at my own pace	0.883
CON2	The simulation gives the user control to better manage his learning	0.870
	Ease of Use	
EOU1	The interface is intuitive	0.815
EOU2	This simulation is easy to use	0.872
EOU4	The simulation controls are logical and consistent	0.791
	Effectiveness	
EFF1	This game has helped me improve the knowledge I acquired in class	0.867
EFF2	The game asks me to put into practice what I've learned in class	0.855
EFF3	The game covers the material and concepts that are related seen in class	0.885
EFF4	I feel like I have acquired new knowledge using this simulation	0.773
EFF5	The simulation allowed me to put into practice all the skills I acquired in class	0.859
	Entertainment	
ENT1	The simulation was fun	0.884
ENT2	I would like to play this game again	0.886
ENT3	The time I spent playing the game was pleasant	0.880
	Fidelity	
FID1	The simulation does reflect real situations	0.935
FID2	There was a very good physical and psychological similarity between the simulation and the environment it sought to represent	0.885
FID3	The simulation encourages learners to use the same cognitive processes they will have to use in the real-world	0.911
FID4	The simulation is faithful to the types of problems encountered in the real-world	0.867
	Goal Clarity	
CLA1	The game's instructions were precise and obvious	0.831
CLA2	Objectives were clearly defined	0.858
CLA3	I understood the learning objectives throughout the simulation	0.733
CLA4	From start to finish, I understood the task to be accomplished in the simulation	0.694
CLA5	The purpose of the game was always clear while I was playing	0.782
	Immersion	
IMM1	While playing this game, I did not feel the time pass	0.835
IMM2	I was very involved in this simulation when I was playing it	0.746
IMM3	I lost track of time while playing the game	0.822
IMM4	When I was playing the game, I became very focused on what I was doing	0.732
IMM5	The simulation had me focused on the task at hand	0.724
	Instructor assistance and support	
INS1	The teacher's mastery with simulation was useful for learners	0.900
INS2	The competence of the teacher with this technology was appreciated by the students	0.891
INS3	The teacher was able to help learners to use technology	0.909
INS4	The teacher's knowledge of the simulation was useful to us	0.898
	Performance / status and feedback	
PERF3	The feedback offered during the simulation was constructive	0.875
PERF4	The information offered by the simulation helped me to make better decisions for the rest of the game	0.768
PERF6	The simulation kept me informed of my status (e.g., score, feedback)	0.837
	Progressive challenge	

CHA4	The increasingly difficult situations presented by the game stimulated my desire to learn	0.699
CHA6	This simulation offers an adequate amount of difficulty to stimulate the interest of the learner	0.795
CHA7	The simulation was able to sharpen my skills allowing me to meet increasingly complex challenges	0.880
CHA8	This game allows different levels of challenges, well suited for different learners	0.797
	Reliability	
REL1	The simulation worked well	0.871
REL2	This technology was reliable	0.820
REL3	I found that the technology worked as expected	0.852
REL4	I had technical problems (reverse item)	0.704
	Satisfaction	
SAT1	In general, I consider that this simulation was well designed	0.809
SAT2	I enjoyed using this game for learning	0.887
SAT3	Overall, I am satisfied with this simulation	0.899

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 44-51 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Remote Patient Monitoring Systems with 5G Networks

Antonio Casquero Jiménez*, Jorge Pérez Martínez

Signal System and Radiocommunication Department, Polytechnic University of Madrid, Madrid, 28040, Spain

ARTICLE INFO

Article history:
Received: 03 May, 2021
Accepted: 15 June, 2021
Online: 10 July, 2021

Keywords:
5G
IoT
Remote patient monitoring
eHealth

ABSTRACT

The new generation of mobile communications and the recent advances in data management are going to enable a fast transformation in the health sector of many countries. 5G networks, with superior technical characteristics, would allow the development of a new set of application and services gathered under the concept of eHealth. In this article we propose a remote monitoring system based on 5G networks that would allow to provide a varied set of medical services from long distance. However, for achieving an optimal performance, the network must guarantee high bandwidths and low latencies, at the time that a massive number of devices and its corresponding generated data are handle efficiently. Consequently, an appropriated system architecture and data model structure are proposed, taking into consideration the high security requirements that any health-related application or service inherently implies.

1. Introduction

Today, health systems in many countries are facing some important sanitary challenges regarding an increasingly aging population, the rise of chronic diseases and the Covid-19 pandemic. In this context, the digital transformation of the health sector with the development of new services supported by the latest technological advances is considered as an indispensable way of saving expenses, optimizing the resources of the sector, and obviously improving the populations welfare. As a result, concepts such eHealth and technologies as 5G or IoT have become essential elements in this digitalization process.

Most of the new and innovative applications and services that are expected to appear in the following years would arise from the combination of 5G and IoT and would be generally supported on Big Data and Artificial Intelligence techniques. Among them, remote patient monitoring systems would be one of the most popular and deployed applications.

This paper is an extension of the paper "5G networks in eHealth services in Spain: remote patient monitoring system", originally presented by the authors in the 2020 IEEE Engineering International Research Conference (EIRCON) [1]. In addition to the results and achievements presented in [1], in this current article we deepen in the theorical concepts behind the solution proposed, emphasising the role of 5G networks. Although the general system architecture is the same as the one presented in [1], this paper extends its study providing a more detailed description of the involved elements and their function within the system. Finally, and in addition to the initial results gathered in [1], the security and data privacy of the remote patient monitoring system proposed is analysed.

1.1. The remote patient monitoring service: definition and communication requirements

Remote patient monitoring (RPM) is an eHealth application that consists in the monitorization of a patient's health state with sensors, wearables or medical devices that measures vital and contextual parameters such as temperature, pulse, sugar or oxygen in the blood, among others [2]. Moreover, the term RPM also includes the applications and platforms that allows the analysis of medical data by means of BigData and artificial intelligence (AI) techniques to provide self-consultation and evaluation by doctors. In general, these applications are especially designed for the monitoring and follow-up in real time of patients with different illnesses, although the uses of RPM systems can be directed to other use cases such as the follow-up of post-operative processes.

According to the predictions done by STL Partners [3] remote patient monitoring applications would be one of the main eHealth services in the next decade, experimenting a tremendous growth, following the trends of a more personalized and continuous home centric health care. In addition, and thanks to the use of 5G networks, the forecasts determine that the adoption of these kind

*Corresponding Author: Antonio Casquero, Polytechnic university of Madrid, Madrid, 28040, Spain, +34649034991, acasquerojimenez@gmail.com

https://dx.doi.org/10.25046/aj060406

of applications could lead to more than 50 Billion USD of global annual cost savings for the health sector by the year 2030 [3].

However, these kinds of applications impose important communications requirements being the availability of a wide coverage area the most important one, as it is mandatory for a remote patient monitoring system to connect all the agents that might be involved such as patients, doctors, caregivers, or even patients´ family members [4]. Nevertheless, it is also crucial to consider the network's capacity to efficiently manage a large number of connected devices, as the network that supports the system must be able to handle an increasing deployment of devices without impacting the service performance. In this line, the security and reliability of the data transmitted must be guaranteed, as well as the continuity of the service so that patients are monitored at all times. Special attention must also be paid to coverage in indoor environments, as most of the sensors and devices deployed will be located in these environments [4]. On the other hand, both energy consumption and battery life of the devices are critical aspects to consider in order to have self-sustaining connected devices, whose batteries are adapted to the duration of the medical processes. Likewise, remote monitoring aplications imposes requirements related to high-speed mobility, especially in the case of emergency situations.

In the following table the main communication requirements can be summarized. However, it is worth mentioning that these would considerably depend on the medical service derived from the system.

Table 1: Technical requirements of RPM systems [5]-[7]

Use case attribute	Value
Throughput	< 1Mbps
Latency	<50ms
Reliability	99,99%
Number of devices	10-10^4/km^2
Battery duration (use case dependant)	10 years
Security	Critical
Mobility	0-500 km/h
Coverage	Important, including indoor
Indoor coverage	No critical 1-10 m horizontal, < 3m vertical

1.2. Organisation of this paper

This paper is structured as follows. In section 44 we have provided the definition of remote patient monitoring as well as an analysis of the requirements imposed by these types of systems. In section 2 we study the main technologies involved in the provision of the connectivity required in RPM systems whereas in section 3 we present a RPM system architecture based on 5G networks. Afterwards, in section 4, a data model for the adequate treatment and management of the data in the system is proposed. In the next section, two crucial aspects such as the security and data privacy of the system are reviewed playing special attention to the advantages of 5G in this context and the possible additional measures that could be taken. Finally, in section 6, the conclusions of the paper are presented.

1.3. Related projects and papers

The development of 5G networks as well as the possible services derived from them are still in an early stage, moving from the commercial verification period to the small-scale deployment of certain solutions. Consequently, to the best of our knowledge, there are no commercial systems or solutions such as those proposed in this paper, although it is true that in the specialized literature, we do find truly interesting proposals.

In [8], the authors present a continuous monitoring system based on 5G networks using wearables and sensors. The devices measure certain vital parameters and via Bluetooth, they send the captured data to the smartphone, which then forwards the data to external servers through 5G networks. In this project, they also propose an intelligent algorithm for decision making and alert generation based on the evaluation of the measured parameters. On the other hand, in [9] they propose a remote monitoring system for smart environments based on IoT and on the existing 4G network infrastructures. In this article, they analyse the communication requirements of the system, especially those referred to the bandwidth, and perform a detailed study of the communication protocols. In this context [10] designs a home monitoring system based on 5G networks and Edge computing with the scope of treating remotely chronic diseases and to promote active aging, in line with the developments and advances reported in [11]. From another perspective, in [12] they also propose the joint use of short-range networks and mobile networks, as an optimal way to develop a remote monitoring system, putting in this case special emphasis on the security of the system and its vulnerabilities. Finally, and with a more general approach, in [13], in addition to a complete review of 5G technology and its applicability to the healthcare sector, they propose a 5G network architecture capable of supporting multiple medical services, including patient monitoring. The system is characterized by the joint use of small cells and macro cells and the deployment of computing and processing servers in the edge. The requirements that eHealth applications impose on telecommunication networks and, in particular, the suitability of 5G networks to meet them are also presented in [13].

2. Connectivity Technologies

The development of eHealth services and applications, like the one we are analysing in this paper, would appear from the combination of 5G and IoT. However, it is worth mentioning the important role that Big Data and Artificial Intelligence techniques and algorithms would play regarding the intelligent analysis of the data generated by these systems. Indeed, the role of data treatment and management platforms is considered indispensable, as they provide a secure and ubiquitous source of information for the consultation of the generated data among all the agents involved in a particular service [2].

However, in this section we focus on the main connectivity technologies involved in the provision of remote patient monitoring systems. These technologies must take into

consideration the requirements defined in section 1.1 and must satisfied them properly.

2.1. 5G

5G is the new paradigm of wireless communications designed to support a wide variety of services. 5G networks improve considerably the performance of the previous generations of mobile communications thanks to the introduction of a new radio interface known as "the New Radio" (NR) and the redefinition of the core of the network resulting in the 5GCN (5G Core Network) [14].

In addition, the 5G system is defined as a service-based architecture (SBA) that through the recent advances in network function virtualisation (NFV) and software-defined networking (SDN) allows a flexible usage and configuration of network functions. As a consequence, different use cases with very diverse requirements can be defined by means of network slices and consequently multiple verticals can develop their services within a single physical infrastructure [15].

The new frequency bands, which include the ultra-high capacity mmWaves, the new waveform used, the utilization of massive MIMO techniques, or the use of heterogeneous access networks, which supports non 3GPP standardised access networks, jointly with the improvements in the network architecture, allows 5G networks to offer the following characteristics [1]:

- Guaranteed user data rate of 100Mbps in DL and 50 Mbps in UL with peak rates of 20Gbps.
- 1/10 X in end-to-end latency, reaching delays of 1 ms.
- Service transmission reliability of >99.999%
- 1000 X in number of IoT devices reaching a density of 1 million terminals/km^2
- 1/10 X in energy consumption

Finally, it is worth mentioning the recent advances in Multi-access edge computing (MEC) technology regarding 5G networks. This concept enables to bring cloud computing processes and storage (typically performed in the core of the network) closer to the final users. Running these processes in close proximity to end users, for example in the 5G base stations, reduces network congestion and improves applications performance by providing faster and more reliable connections with lower latencies [16].

2.1.1. *Why 5G?*

The characteristics and advantages of 5G networks in the provision of medical services can be summarized in the following points:

- Reliability and security. The ultra-reliable connections of 5G networks are the main driver for the utilization of this technology for the development of eHealth applications as patient´s data privacy and service continuity can be guaranteed. Network slices isolation, advanced data encryption techniques and the new and improved authentication mechanisms contribute to this objective.
- High performance. As we have stated in the previous section, 5G networks provide high average data rates and ultra-low latencies, fulfilling the requirement of a wide majority of medical services, including the one we are analysing. Massive MIMO with 3D beamforming and Edge computing and MEC (Multi-access Edge Computing) platforms are some of the innovations included in this field of action.
- Ultra-high Capacity. With up to one million of devices per Km^2,5G networks can efficiently manage the simultaneous connections of multiple medical devices from multiple users. Beside this, these networks provide 90% reduction in energy consumption with 10 years of battery life for low power IoT devices.

2.2. IoT

IoT ("Internet of things") consists in the grouping and interconnection of devices and objects through a network. As we can infer from this definition, RPM systems can be classified as an IoT application, in which the technology chosen for the provision of such connectivity would be 5G [1]. In 5G networks, we can define different categories of IoT according to their requirements [17]: Massive IoT, Broadband IoT, Critical IoT and Industrial Automation IoT. The characteristics of RPM system can be included within the framework of the massive IoT and Broadband IoT.

2.2.1. *Massive IoT*

Massive IoT intends to provide connectivity to a very large number of devices that transmit and/or receive small volumes of data. These devices, that frequently rely on battery power supply, are usually low-cost and can be located in remote places with little coverage [18].

In this context, LTE-M (LTE-MTC) and NB-IoT (Narrowband-IoT), both standardized by the 3GPP [19], fulfil all 5G requirements from both ITU and 3GPP for massive machine type communications. Although they are both LTE technologies, they are designed to coexist within the new 5G NR, thanks to techniques such as Dynamic spectrum sharing (DSS), defined in Release 15, and the numerology used by the sub-carriers, compatible with the NR.

Table 2: Technical comparison of NB-IoT and LTE-M [18].

Characteristics	NB-IoT	LTE-M
Bandwidth	180 KHz	1.4 MHz
Subcarrier bandwidth	15 KHz	15 KHz
Peak data rate	250 Kbps	1 Mbps
Latency	1.5-10 s	50-100 ms
Battery duration (use case dependant)	+10 years	10 years
Operation mode	FDD	FDD/TDD
Voice support	No	Yes
Coverage DL (MCL)	164 dB	164 dB

Coverage UL	Gain: 20dB	Gain: 15 dB
Indoor coverage	Excellent	Good

2.2.2. *Broadband IoT*

Broadband IoT is a category of IoT developed from mobile broadband communications (MBB). The introduction of 5G with the NR and the new frequency bands, allows offering IoT services with higher data rates and lower latencies than those offered by mIoT technologies, reaching peak values of tens of Gbps in transmission and latency values of 5ms [17]. Currently, many broadband IoT applications are being developed due to the already mentioned combination of MBB features (high bandwidth, high transmission speed, low latencies) with the characteristics of IoT services (wide coverage range, energy efficiency).

In this context, it is worth mentioning that the 3GPP as part of Release 17 [20], has introduced a new concept defined as the 5G NR-Light. It consists in a set of modifications of the 5G NR in order to reduce its capabilities to efficiently develop a set of IoT applications that will require devices more complex than those used in mIoT (massive IoT) communications but less complex than those used in 5G NR. These devices are usually going to be wireless industrial sensors (with low latencies and moderate transmission rates), video systems, high-capacity wearables, and patient monitoring systems.

3. RPM system architecture

Remote patient monitoring systems (RPM), as previously discussed, allows the monitoring of an individual's health status using sensors and/or wearables thanks to the control and measure of certain relevant medical indicators and parameters. Currently, we can find some remote monitoring systems models and proposals (section 1.3) in which these devices use personal area networks (PAN) such as Bluetooth and local area networks as Wi-Fi to connect to a hub (e.g. a smartphone) or a gateway. The data are collected and processed in mobile applications for smartphones that allow the patient to monitor and manage relevant medical information.

However, these applications are generally not standardized and therefore the monitoring processes performed does not have the required medical rigor. The systems should be able to send, periodically, or in real time, this medical data to authorized third parties, in order to obtain a more exhaustive analysis involving healthcare professionals and the use of Big Data techniques. This will facilitate diagnosis and follow-up tasks without the patient having to travel to the hospital or primary care center. In addition, it would be possible to detect anomalous patterns in the data and activate alarms in emergency situations.

Taking into consideration these aspects, an once we have described what is a RPM system and the main technologies involved, we present a possible system architecture based on 5G networks [1].

3.1. Functional high-level architecture

The system we propose is an open multiplatform system, based on European standards, which will allow the continuous monitoring of the health state of patients, as well as the possible detection of new pathologies and new therapeutic approaches.

The design of the network, responsible of providing the required connectivity, must be in line with the requirements of the proposed system where we highlight the massive creation of data by the end users (directly or indirectly) from multiple sources. In this context, the system and the network must guarantee the availability and security of the data so that the agents involved in the provision of health services can make use of the multiple applications and services that can be offered from the exploitation of these.

In order to obtain an open and scalable solution, both the data model and structure and the IoT architecture are based on the ETSI reference architecture, particularly on the ITU-T Recommendation Y.2060 [21], which offers the description of a reference model for IoT applications.

In the Figure 1 and taking the multilayer ETSI reference model as a reference, the functional high-level architecture of the system proposed is presented.

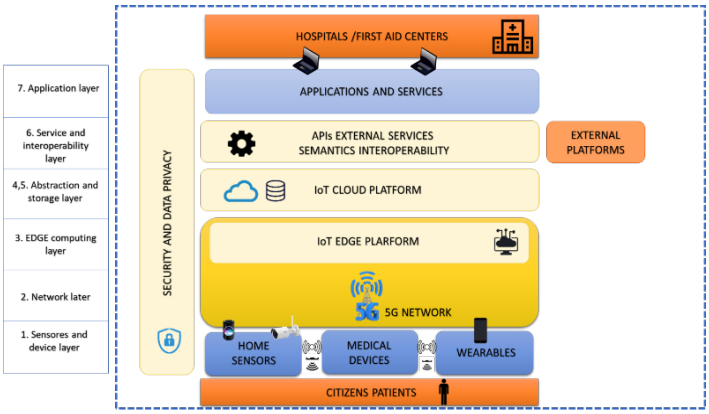


Figure 1. Architecture of the proposed RPM system together with the ETSI reference layer model [1]

3.2. Architecture description

In this subsection we provide a detailed description of the different layers that define the system. The security and data privacy module, which covers all the layers of the system, will be studied in detail in section 5.

3.2.1. *Sensors and device layer*

This is the physical layer of the system that encompass a set of devices and sensors responsible of collecting information about the health status of the patient and of its environment [10]. In this layer we can find contextual sensors, medical devices and wearables that transmit and/or receive information directly over the mobile network without the need of using local networks (LAN OR WLAN) or gateways (routers or smartphones). However, we must also highlight the role that smartphones or virtual assistants can play in the system, due to the considerable number of sensors that they have integrated.

3.2.2. *Network layer*

5G networks will be the telecommunications network in charge of providing the connectivity in the system.

In our solution, we would mainly use low-cost devices and sensors that will perform periodic measurements of certain vital and contextual parameters, using narrow bandwidths. These would deal with very small amounts of data and will not require low latency values. Therefore, the best solution, as this scenario matches the characteristics of mIoT, is to use LTE-M or NB-IoT networks [1].

On the other hand, and in order to provide high performance monitoring applications, we can find another group of devices (advanced medical devices or high capacity wearables) that will perform measurements and monitoring processes of higher capacity. Consequently, they will require higher transmission rates and generally lower latencies. These devices would work with higher data volumes, and their requirements would be similar to those associated with Broadband IoT applications. As a result, we will use the 5G NR or 5G NR-Light [20].

Finally, we must highlight the possibility of using other wireless technologies, such as Wi-Fi 6, for connectivity at indoor environments. The new Wi-Fi standard, officially named as 802.11ax, is capable of providing theorical peak rates of 9.6 Gbps and can quadruplicate the number of simultaneous connected devices [22]. This new Wi-Fi standard, thanks to new security mechanisms and to the definition of network energy efficiency techniques, can effectively provide mIoT communications. The main advantage of this technology is the greater variety of sensors and devices that can be used, since the Wi-Fi 6 access points (AP) allows the integration of other communication modules (by means of external cards) associated to IoT communications such as RFID, ZigBee, and Bluetooth [22]. In this particular case, Wi-Fi 6 could provide indoor connectivity and 5G networks would act as backhaul technology.

This hybrid solution, which offers an optimized performance, is technically feasible thanks to the fact 5G network support multiple access technologies, even if they are not 3GPP standardized.

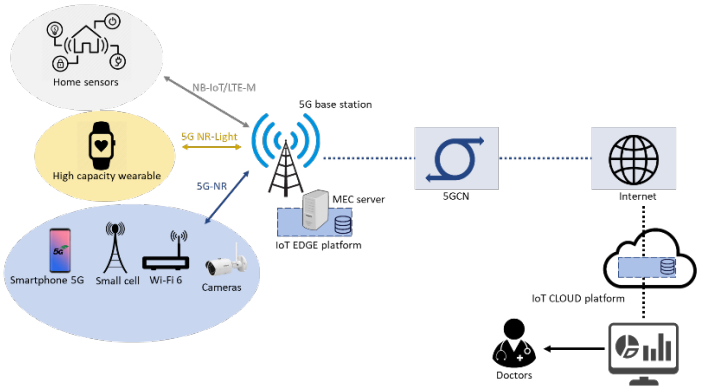


Figure 2: Representation of the global system architecture

3.2.3. *Edge computing layer*

The main function of IoT Edge platform is the integration and aggregation of the data collected in the previous layers and the transmission of the data to the cloud. The platform also carries out filtering and processing tasks in order to reduce the volume of data to be transmitted to the upper layers and consequently the bandwidth required. In this layer, a first analysis of the data can be made as analysing it in the proximity of the users makes it possible to improve the reaction or intervention in real time in cases of risk or emergencies and allows the execution of AI processes based on the collected data with very low latencies.

3.2.4. *Abstraction and storage level*

The abstraction and storage layer is encompassed by the IoT cloud platform, which is in charge of collecting the information processed by the previous levels, gathering the information from all the IoT Edge platforms deployed in a health region (south-bound) and integrating the semantic interoperability layer (north-bound).

Moreover, this platform is responsible of performing cloud computing tasks from the pre-processed data in the previous layers. At this level of the system is where the management of identities is performed as well as storage and administration tasks. The proposed platform is also where we carried out the intelligent analysis of the data [1]. In addition, thanks to the use of data models and information collection standards, the IoT cloud platform can manage information from various different sources such as IoT Edge platforms (user-generated data) or health systems (clinical history databases).

3.2.5. *Semantic interoperability layer*

The concept of semantic interoperability layer was introduced and developed in the ACTIVAGE project [11], a European project which develops IoT solutions for Smart Living Environments. However, the role of this layer is crucial for the proper development of the system and therefore it must also be included in our design.

The semantic interoperability layer allows the described IoT platforms to share data and exchange information with other external platforms thanks to the definition of interfaces and to the conversion of the data to a common model. Data from different sources are harmonized in using HL7 and the SNOMED vocabulary[10]. This layer also includes security and access control functions to the upper layers.

3.2.6. *Application and intelligent services layer*

The application layer enables the provision of a wide range of intelligent services from the captured and processed data. They are categorized as intelligent because in all of them there may be data processes that use artificial intelligence and/or big data techniques. The offered services are intended for patients but also for healthcare personnel or even informal caregivers of the patients. The final services can be adapted to the socioeconomic context of the region in which the system is deployed.

The most relevant medical services that can be derived from the proposed system are now presented. These services can address the main sanitary challenges considered in section 1:

- Chronic care services. Chronic diseases are progressively increasing every year over the population of many countries [23]. The RPM system proposed is expected to allow a personalized management and administration of patients with these conditions thanks to a rigorous and continuous follow up of the medical and contextual parameters that reflect their condition. In addition, the system could favor the capacity for

early detection of diseases thanks to the predictive analysis of data running both in the Edge and in the cloud.

- Elderly care services. In this category we can include any service aimed at maintaining and improving the quality of life of the elderly. In this line monitoring services can facilitate the independence of the elderly allowing them to maintain an autonomous life away from health institutions for a longer period.
- Personal health self-management services. These are services aimed at people whose objective is the prevention of possible chronic diseases based on the detection of one or more conditions but in which there is still no organ damage.

4. System data model

The proposed RPM system, as can be inferred from the functional description of it, can be considered as data centric. The massive amount of data that the system has to handle jointly with the demanding security requirements that any health-related application imposes, demands a reorganization of the traditional data management procedures. The changes required in this respect, side with those introduced in 5G networks.

The aforementioned shift of paradigm implies that beyond the management of the devices involved in the monitoring processes, the system must establish mechanisms and tools for the efficient management of the data itself. In this context, and following the indications reflected in [24], the system should consider the following aspects:

- Data ownership. Patients should have control over the data generated by their monitoring devices. To this end, we must establish an ownership structure for the data and the devices, that allows to establish a relationship and association among them. This structure must allow the patient to manage, locally or remotely, the generated data.
- Data provenance. It is essential to create an immutable record of the source of the data, thus establishing a unique version of it. The system must be able to identify the device that generated the data and at what time it was generated. This information makes possible to guarantee the authenticity and veracity of the information collected.
- Data governance. Patients should be able to manage the access to their data, being able to grant permissions to third parties for external consultation.

Therefore, we can conclude that for data management in 5G networks, and by extension in our system, we must manage the data generated and the users with access to them. One possible solution consists in the utilization of two separate platforms, one in charge of property management and the other of data storage [24].

These concepts are particularly relevant in certain processes and management tasks that the system must perform. In the Figure 3, we show an example which reflects the registration of a device in the system and the access of an authorized third party to the data generated by it. In this scenario we have three entities: the patient, owner of the medical device, the medical device itself, which generates the data, and the healthcare personnel, who wishes to consult the generated medical data.

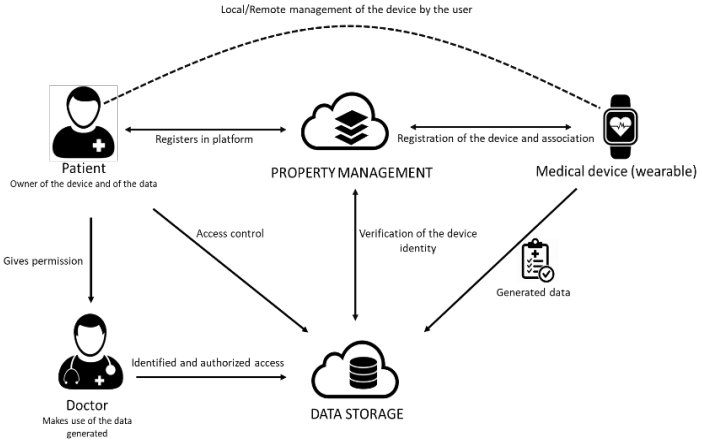


Figure 3: System´s data management model. Based on [24]

As part of this registration process, the patient must register in the system in order to be able to add and associate to his/her identity different medical devices. This process will allow the user to manage the device and control the data generated by the it. Once the registration is completed, and as part of the monitoring process, the data generated will be stored in a data storage platform. This platform will verify the identity of the device and the authenticity of the stored data, which can be consulted by the patient or by authorized third parties.

5. Security and data privacy in the system

The proposed system is responsible of the creation, maintenance and exploitation of the data generated by the patients. In this context, as we are dealing with an application framed within the health sector, there is a series of legal obligations related to data security and privacy that must be considered and therefore, the system must be designed in line with the General Data Protection Regulation (GDPR) established by the European Union. Consequently, the Security and Privacy (S&P) module defined in our system architecture must cover all the layers, from the sensors and devices to the final applications, ensuring three main security principles: confidentiality, availability, and integrity.

In this context, 5G networks introduce a series of improvements in the security domain compared to previous generations of mobile networks. Moreover, we would define a set of specific security measures to enhance the system overall security.

5.1. 5G security improvements

5G networks offers better capabilities than other mobile technologies not only in terms of bandwidth, latency or density of devices that can be supported, but also in terms of security and data privacy. In this context, the main improvements are related to authentication procedures and to the encryption of the data in the network.

Authentication and identity management are fundamental aspects in cellular networks. In 4G networks, authentication tasks are performed establishing a mutual authentication process in which, by means of the authentication and key agreement (AKA)

protocol, both the user's terminals and the network core validate their identities [25]. However, these processes do not include the encryption of the data in the access network in certain signalling processes, for example, when sending the user's identity. To this end, 5G networks encrypt the international mobile subscriber identity (IMSI) and extends the length of the session keys from 128 to 256 bits, thereby improving the protection of communication between the device and the network core [25]. In addition, fifth generation networks introduce a number of specific network functions designed to improve the security of authentication processes, such as the AUSF (Authentication Server Function) or the SEAF (Security Anchor Function) and defines a unified framework with new authentication methods: 5G-AKA, EAP-AKA and EAP-TLS [26]. This framework enables the authentication of a user's equipment independently of the access technology used, whether it is standardized by the 3GPP or not. This fact responds to the heterogenous access capabilities of 5G networks and is a concept extremely important in our system as we propose different access technologies as part of the solution, some of them non 3GPP standardized. Consequently, and in all situations, the user's data privacy would be guaranteed.

On the other hand, and considering the wide variety of medical devices and sensors that can that take part in the monitoring processes, the aforementioned authentication framework includes alternative authentication methods, such as the EAP-TLS, formed by an extensible authentication protocol (EAP) with transport level security (TLS). This method is expected to be considerably used in IoT applications as it does not require the use of (U)SIM cards for identity management and authentication processes. This could facilitate the use of low cost devices with different sizes and shapes [26].

Nevertheless, there are some other important aspects regarding 5G networks security, such as the isolation between the access network and the 5GCN. In this context, in the 5G system there is a clear separation between these two parts, in which the access network performs radio management tasks, and the core manages and controls network resources and security[26]. This separation makes it possible to easily identify and isolate network elements in case they are under attack. As a result, the core is the responsible of ensuring the user authentication and the encryption of the data and signalling traffic, with the advanced encryption standard (AES), whereas the access network sends the encrypted traffic between the user's equipment and the core, thus protecting traffic at the radio interface. In addition, the definition of a service-based architecture in the network core allows protection mechanisms to be applied at higher layers, e.g. transport and application [27].

Finally, we must mention the advantages of defining network slices to provide these kinds of services, not only in terms of network resources allocation, but also in terms of security, as possible failures or attacks in one slice will not affect the services being provided in others. To this end, mechanisms and tools must be developed to allow isolation between slices and to prevent unauthorized users from accessing both the assigned network resources and the information they carry.

5.2. Specific measures

However, and despite of the improvements in terms of security and data privacy that 5G networks include, the highly regulated healthcare sector also imposes the need of the definition of an appropriated data model structure. This model, which has been studied in the previous section, responds to the Edge computing paradigm, so that part of the data, before being send to the cloud, is processed, and managed in a secure space. Therefore, part of the management, processing and storage tasks of the data will be carried out on the IoT Edge platform. This platform will be able to manage, independently and autonomously, the users and the devices associated to them [16].

On the other hand, and as the system will rely on the cloud, the communications between the different platforms and cloud services must be performed securely. In this context, a research group of the Polytechnic University of Madrid called the Life Supporting Technologies group, recently proposed the use of distributed identity systems for this kind of purposes, with the scope of integrating different healthcare services [10]. The proposed distributed system will be the one we will use in our system and will be in charge of managing user identities anonymously. In addition, it will also incorporate smart semantic contracts (SSCs) to control data access, its usage and conditions. The SSCs will be able to unlock encrypted data only when the conditions of the contract are met, so it will be possible to transmit encrypted data between different systems without affecting data privacy.

6. Conclusions

In the following decade, and accelerated by the current situation, the health sector of many countries would experience a deep digitalization process. As part of it, 5G networks would play an important role as they would allow the development of new applications and services. Among them, we highlight remote patient monitoring systems that would permit to offer a more personalized and continuous healthcare from long distance.

In this article, we have presented a possible RPM system architecture based on 5G networks. In this context, the new mobile communication generation offers not only a far more superior performance but also a more secure and reliable environment for the management and treatment of the generated data. In order to improve the general system performance, we also make use of other important technological advances in other fields such as IoT or MEC computing. The development of IoT systems, based on 5G networks, with more accurate and enhanced characteristics and the improvement of MEC computing would allow to considerably upgrade the final performance of the system. As a result, a wide flexible variety of medical devices can be offered, as the system is inherently flexible and adaptable.

Finally, the new enhancement in 5G networks security and data management together with additional security and data privacy measures would allow to give the system the required security. This aspect is crucial as 5G networks can help to achieve the severe regulations and security measures that health systems and applications suffer.

References

[1] A. Casquero, J. Perez, "5G networks in eHealth services in Spain: Remote patient monitoring system," In IEEE Engineering International Research Conference, EIRCON 2020, 1–4, 2020, doi:10.1109/EIRCON51178.2020.9254013.

[2] IDATE, "Ehealth Market trends, players & outlook," DigiWorld Interactive Platform, 1–38, 2018.
[3] D. Singh, "5G'S HEALTHCARE IMPACT: 1 BILLION PATIENTS WITH IMPROVED ACCESS IN 2030," STL Partners, 1–39, 2019.
[4] IDATE, "5G IoT in the healthcare sector," DigiWorld Interactive Platform, 56–67, 2019.
[5] Next Generation Mobile Networks Ltd, Perspectives on Vertical Industries and Implications for 5G, NGMN P1 Requirements & Architecture - Verticals Requirements, 1–41, 2016.
[6] Next Generation Mobile Networks Ltd, Recommendations for NGMN KPIs and Requirements for 5G, NGMN P1 WS#3 Key Performance Indicators to 3GPP TSG RAN, (2016), 1–19, 2016.
[7] 3GPP ETSI 3rd Generation Partnership Project, "5G; Service requirements for next generation new services and markets (Release 15)," 3GPP TS 22.261 Version 15.5.0 Release 15, 1–52, 2018.
[8] R. Singh, R. Garhewal, "Smart M-Health Continuous Monitoring System Using 5G Technology," Research Project University of Ottawa, 1–7, 2018, doi:10.13140/RG.2.2.15946.21440.
[9] Ngo Manh Khoi, S. Saguna, K. Mitra, C. Ahlund, "IReHMo: An efficient IoT-based remote health monitoring system for smart regions," 2015 17th International Conference on E-Health Networking, Application and Services, HealthCom 2015, 563–568, 2015, doi:10.1109/HealthCom.2015.7454565.
[10] MYSPHERA, Life SupportingTechnologies Group, Universidad Politécncia de Madrid, "Public consultation for the remote monitoring of chronic patients," Red.Es, 1–35, 2019.
[11] P. Barralon, B. Charrat, I. Chartier, V. Chirié, G. Fico, S. Guillen, N. Homehr, O. Horbowy, R. Kamenova, F. Lamotte, A. Peine, "IoT for Smart Living Environments Recommendations for healthy ageing solutions," WG5 - Smart Living Environment for Ageing Well – Recommendation Paper Alliance for Internet of Things Innovation, (April), 37–67, 2019.
[12] P.K.D. Pramanik, G. Pareek, A. Nayyar, "Security and privacy in remote healthcare: Issues, solutions, and standards," Telemedicine Technologies: Big Data, Deep Learning, Robotics, Mobile and Remote Applications for Global Healthcare, 201–225, 2019, doi:10.1016/B978-0-12-816948-3.00014-3.
[13] A. Ahad, M. Tahir, K.-L.A. Yau, "5G-Based Smart Healthcare Network: Architecture, Taxonomy, Challenges and Future Research Directions," IEEE Access, **7**, 100747–100762, 2019, doi:10.1109/ACCESS.2019.2930628.
[14] S. Redana, O. Bulakci, "View on 5G ArchitectureView on 5G Architecture - 5G PPP Architecture Working Group," 5G Public Private Partnership (5G PPP), 1–182, 2020, doi:10.5281/zenodo.3265031.
[15] G. Brown, "Cloud RAN & the Next-Generation Mobile Network Architecture," Heavy Reading White Paper Huawei Technologies Co. Ltd., 1–9, 2017.
[16] Y.C. Hu, M. Patel, D. Sabella, N. Sprecher, V. Young, "ETSI White Paper #11 Mobile Edge Computing - A key technology towards 5G," ETSI White Paper No. 11 Mobile, 1–16, 2015.
[17] A. Zaidi, A. Bränneby, A. Nazari, M. Hogan, C. Kuhlins, "Cellular IoT in the 5G era," Ericsson White Paper, 1–17, 2019.
[18] C. Kuhlins, B. Rathony, A. Zaidi, M. Hogan, "Cellular networks for Massive IoT," Ericsson White Paper, 1–16, 2019.
[19] 3GPP Technical Specification Services Group, "Release 15 Description-TR 21.915," 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects, 1–118, 2019.
[20] 3GPP, "New SID on Support of Reduced Capability NR Devices," 3GPP Work Item Description, 1–4, 2019.
[21] UIT-T, "UIT-T Rec. Y.2060 General description of Internet of things," ITU Telecommunication Standardization Sector, 1–20, 2012.
[22] H. Fangming, "Wi-Fi 6 and 5G and Their Application Scenarios," Huawei Campus Network Marketing Support Module, 1–47, 2019.
[23] R. Hanno, R. George, K. Taylor, "Vital Signs: How to deliver better healthcare across Europe," Deloitte Centre for Health Solutions, 1–56, 2016.
[24] G. Horn, "5G End-to-End Data Management," Qualcomm Technologies, 2020.
[25] Huawei Technologies Co LTD, "5G Security: Forward Thinking," Huawei White Paper, 1–12, 2016.
[26] D. Basin, S. Radomirovic, J. Dreier, R. Sasse, L. Hirschi, V. Stettler, "A formal analysis of 5g authentication," Proceedings of the ACM Conference on Computer and Communications Security, 1383–1396, 2018, doi:10.1145/3243734.3243846.
[27] P. Teppo, K. Norrman, "Security in 5G RAN and core deployments," Ericsson White Paper, 6–13, 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 45-53 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Numeric Simulation of Artificial Antigravity upon General Theory of Relativity

Yoshio Matsuki[1,*], Petro Ivanovich Bidyuk[2]

[1]*National University of Ukraine "Kyiv-Mohyla Academy", Kyiv 04070, Ukraine*

[2] *Educational and Scientific Complex "Institute for Applied System Analysis" of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv 03056, Ukraine*

ARTICLE INFO

Article history:
Received: 30 March, 2021
Accepted: 23 April, 2021
Online: 05 May, 2021

Keywords:
Antigravity
Distorted time-space
Curvature tensor

ABSTRACT

This paper is an extended version of the work presented at a conference held in Kyiv, Ukraine, in October 2020, which reported the result of the numeric simulation on the artificial antigravity. This paper further describes the derivation of the idea of the artificial antigravity, and adds the simulation of angular momentum that is needed to describe the antigravity. Also, because the angular momentum is the perpendicular movement to a three-dimensional curved surface in a four-dimensional space-time, this paper challenges the limit of applying the curvature tensor in quantum mechanics; while, current quantum mechanics has been established on the flat surface. The artificial rotation of a hypothetical object is simulated, in which the gravity is so strong that the time-space can be distorted. The spherical polar coordinate system is selected to describe the curvature of the space, and the curvature tensor is formulated. Then the tensor is multiplied by the Euler's rotation matrix to make the inner product for the gravitational energy and the outer cross-product for the angular momentum of the rotation. To simulate the distorted time-space, two cases are selected: the linear distortion and the non-linear distortion upon the distance from the center of the strong gravity; also, the speed of the rotation is set in two options: the slower and the faster. Then the equation of motion is set by the curvature tensor to calculate the coefficient of the gravitational energy on the surface of the sphere in the spherical polar coordinates, and to calculate the coefficient of the angular momentum in the perpendicular direction to the sphere. The result shows that the antigravity can be produced by rotating the object, and the angular momentum can show the opposite directions by the selection of the rotation speed.

1. Introduction

This paper (hereinafter the "extended paper") is an extension of the report presented at the 2020 IEEE 2nd International Conference on System Analysis_Intelligent Computing (SAIC) held in Kyiv, Ukraine, in October 2020 [1] (hereinafter the "conference paper"). The conference paper was a byproduct, on a sideline, of the other series of our research [2–7], which was aimed at finding the origin of the global climate change.

At the beginning of the research [2-5], we assumed that Moon's gravity could be related to the increase of the global temperature, then we calculated the coefficients of several variables such as the distance between Moon and Earth, the global temperature, the emitted carbon dioxide, with the method of econometrics, regarding the distance between Moon and Earth as the surrogate for the energy of the Moon's gravitational field and gravitational waves. Then we reached an assumption that there should be anti-gravitational waves as the antimatter of graviton (gravitational waves) similar to positron as the antimatter of electron; and, we derived the equation of motion of anti-gravitational waves, upon the equation of motion for gravitational waves [8] that was predicted in the flat surface in the rectilinear coordinate system, approximating the special theory of relativity. Then we reported the result of our analysis in a paper [6]. The geometrical relation between the positive and negative flows of gravitational waves (gravitational and anti-gravitational waves) that we made for [6] is shown in Figure 1.

The next question was "How can both positive and negative flows be created?" Then we reviewed the general theory of relativity [8] that explained the gravitational field by the following equations:

*Corresponding Author: Khaoula Naouaoui, naouaoui.khaoula@gmail.com

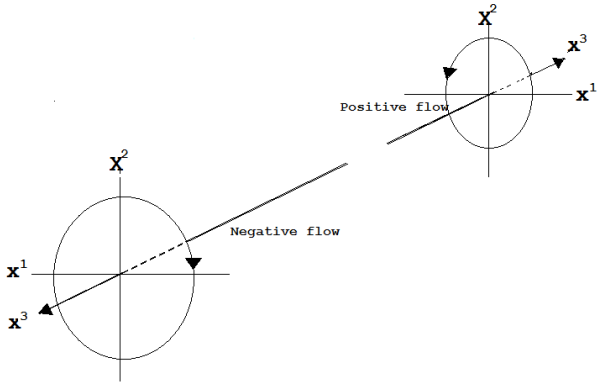


Figure 1: Geometrical relation between positive and negative flows of gravitational waves [6]

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \tag{1}$$

$$\Gamma^{\alpha}_{\mu\nu} = g^{\alpha\alpha}\Gamma_{\alpha\mu\nu} = \frac{1}{2}g^{\alpha\alpha}(g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}) \tag{2}$$

$R_{\mu\nu}$ is called the Ricci tensor, which is a variation of the curvature tensor, and $\Gamma^{\alpha}_{\mu\nu}$ is called a Christoffel symbol. Also, $g^{\mu\nu}$ is called the fundamental tensor that makes the Christofell symbol. The equation (3) shows how the fundamental tensor describes a four-dimensional space-time. Here, the notation for the differential of the tensor is given by $F_{,\mu} = \frac{\partial}{\partial x^{\mu}}F$, where F is any function such as $\Gamma^{\alpha}_{\mu\nu}$ and/or $g_{\mu\nu}$, and x^{μ} is the $\mu - th$ variable (vector) in the given coordinate system.

$$g^{\mu\nu} = \begin{bmatrix} g^{00} & g^{01} & g^{02} & g^{03} \\ g^{10} & g^{11} & g^{12} & g^{13} \\ g^{20} & g^{21} & g^{22} & g^{23} \\ g^{30} & g^{31} & g^{32} & g^{33} \end{bmatrix} = \begin{bmatrix} 1-\frac{2m}{r} & 0 & 0 & 0 \\ 0 & -(1-\frac{2m}{r})^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{bmatrix} \tag{3}$$

The fundamental tensor also makes a geodesic, a path of extremal distance, as shown below.

$$ds^2 = g^{00}dt^2 + g^{11}dr^2 + g^{22}d\theta^2 + g^{33}d\phi^2 \tag{4}$$

In the equation (4), t is time, m is the mass of a planet, r is the distance from the planet, θ is the angle from the axis of r, and ϕ is the angle of the rotation around the axis of r in the spherical polar coordinate system. In (3), there is a singularity at $r = 2m$: therefore, the space is divided into two regions, $r < 2m$ and $r > 2m$. In the region of $r < 2m$, the mass of the planet must be very dense and heavier and it can be a black hole. To connect these two regions, a different coordinate system was invented [8], which makes the distorted time, τ, and the distorted distance, ρ, by the following equations:

$$\tau = t + f(r) \tag{5}$$

$$\rho = t + g(r) \tag{6}$$

The section *2.2.* describes how this idea is taken into account in the numeric simulation.

Then the equation (3) was transformed to the equation (7).

$$g^{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2m}{\mu(\rho-\tau)^{\frac{2}{3}}} & 0 & 0 \\ 0 & 0 & -\mu^2(\rho-\tau)^{\frac{4}{3}} & 0 \\ 0 & 0 & 0 & -\mu^2(\rho-\tau)^{\frac{4}{3}}\sin^2\theta \end{bmatrix} \tag{7}$$

Then the geodesic was also transformed from (4) to (8).

$$ds^2 = g^{00}d\tau^2 + g^{11}d\rho^2 + g^{22}d\theta^2 + g^{33}d\varphi^2 \tag{8}$$

The followings show how to make (1) and (2) by the fundamental tensor:

For example, if $\mu = \nu = 1$,

$$g^{11} = -\frac{2m}{\mu(\rho-\tau)^{\frac{2}{3}}} \tag{9}$$

$$g_{11} = -\frac{\mu(\rho-\tau)^{\frac{2}{3}}}{2m} \tag{10}$$

$$g_{11,\rho} = \frac{\partial}{\partial\rho}(-\frac{\mu}{2m}(\rho-\tau)^{\frac{2}{3}}) \tag{11}$$

According to [8], in the empty space where only the gravitational field of a planet exists, $R_{\mu\nu}$ becomes zero as shown in (12); then (13) is the equation of motion of a particle.

$$R_{\mu\nu} = 0 \tag{12}$$

$$g^{\mu\nu}(g_{\alpha\mu,\nu} - \frac{1}{2}g_{\mu\nu,\alpha}) = 0 \tag{13}$$

Then in order to describe gravitational waves moving in the gravitational field, the equation of motion is differentiated once again as shown below.

$$\frac{d}{dx^{\beta}}g^{\mu\nu}(g_{\alpha\mu,\nu} - \frac{1}{2}g_{\mu\nu,\alpha}) = g^{\mu\nu}_{,\beta}(g_{\alpha\mu,\nu} - \frac{1}{2}g_{\mu\nu,\alpha}) + g^{\mu\nu}(g_{\mu\alpha,\nu\beta} - \frac{1}{2}g_{\mu\nu,\alpha\beta})$$
$$= g^{\mu\nu}(g_{\mu\alpha,\nu\beta} - \frac{1}{2}g_{\mu\nu,\alpha\beta}) = 0 \tag{14}$$

This equation leads to the equation (15), and this is the equation of motion for gravitational waves, according to [8].

$$g^{\mu\nu}g_{\alpha\beta,\mu\nu} = 0 \tag{15}$$

The notation for the secondary differential of $g_{\mu\nu}$, for example by the vectors, x^{α} and x^{β}, is $g_{\mu\nu,\alpha\beta} = \frac{\partial^2}{\partial x^{\alpha}\partial x^{\beta}}g_{\mu\nu}$.

Here, we show this equation only for describing how the gravitational field is related to the creation of gravitational waves; although, this extended paper doesn't include the simulation of gravitational waves.

Then we made the numeric simulation [7] on the curved space, deriving the mathematical forms of the components of the Ricci tensor (1). The following equation is an example of the Ricci tensor in case of $\mu = \nu = 0$, $\alpha = 2$, $\beta = 1$:

$$R_{00} = \Gamma^2_{02,0} - \Gamma^2_{00,2} - \Gamma^2_{00}\Gamma^1_{21} + \Gamma^2_{01}\Gamma^1_{02}$$
$$= g^{22}\Gamma_{202,0} - g^{22}\Gamma_{200,2} - g^{22}\Gamma_{200}g^{11}\Gamma_{121} + g^{22}\Gamma_{201}g^{11}\Gamma_{102}$$
$$= \frac{1}{2}\{-\mu^2(\rho-\tau)^{\frac{4}{3}}\}\{-\frac{4}{3}(-1)(\rho-\tau)^{-\frac{7}{3}}\}_{,\tau} = \{\frac{-2}{3(\rho-\tau)}\}_{,\tau} = \frac{2}{3(\rho-\tau)^2} \quad (16)$$

Then we calculated the coefficient of each component of the tensor, for $\mu = 0, 1, 2, 3$, $\nu = 0, 1, 2, 3$, $\alpha = 0, 1, 2, 3$, and $\beta = 0, 1, 2, 3$, using a personal computer. The algorithm to calculate the coefficients is described in the section *2.3.*

After making the numeric simulation [7] with the curvature tensor, there was still a need to confirm the negative flow of gravitational waves that are to be created by some movement of a strong gravity; however, we could not find any suitable physical object that could be referred to. Therefore, we invented an idea of the hypothetical rotation of an artificial object shown in Figure 2 as a possibility of the movement of a strong gravity. Then we reported the result of the simulation in the conference paper [1] as the "theory and simulation of artificial antigravity".

While the conference paper [1] has reported the result of the simulation on the gravitational energy made by rotating the very strong gravity, this extended paper reports one more feature, the angular momentum, by which we attempt to examine whether there is a consistent explanation of anti-gravitational waves shown in Figure 1, or not.

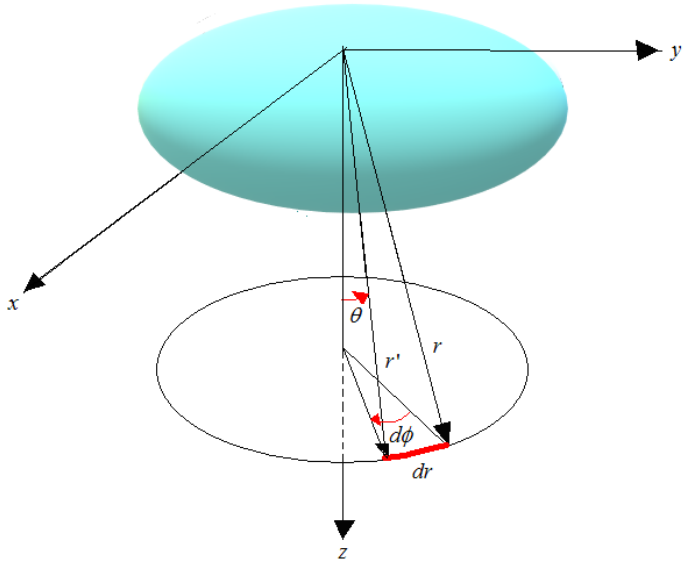


Figure 2: Rotation of an object

It is noted that there is no data taken out from any experiments in laboratories or observations in the cosmos, but this research is made solely on the mathematical model taken out from the general theory of relativity [8] as well as the classical mechanics [9].

For the numeric simulation, we used a personal computer's software, which was developed for the econometrics of geometry [10]. The algorithm of this software was originally developed in the rectilinear coordinate system; but, we used it for calculating the coefficients of the equation of motion in the spherical polar coordinate system as a special case of the curved space. In other words, we used the function of the orthogonal transformation of the matrix algebra as the surrogate for the tensor algebra needed in the general theory of relativity.

2. Method

2.1. Curvature Tensor before the Rotation

For the simulation of this extended paper, we used the same curvature tensor shown in (17) that we used for our previous research for the conference paper [1]. This tensor is for simulating the gravitational field before the rotation. From it, we took the components of R_{11}, R_{22} and R_{33} for simulating the spatial movement of the object, but excluded R_{00} from the simulation because it is for the distorted time coordinate, which is beyond the scope of this research.

$$R_{\mu\nu} = \begin{bmatrix} R_{00} & 0 & 0 & 0 \\ 0 & R_{11} & 0 & 0 \\ 0 & 0 & R_{22} & 0 \\ 0 & 0 & 0 & R_{33} \end{bmatrix} \quad (17)$$

The diagonal components of $R_{\mu\nu}$ are shown in the equations from (18) to (21), which are taken from our previous research [7].

$$R_{00} = \frac{5}{9(\rho-\tau)^2} \quad (18)$$

$$R_{11} = \frac{20}{3(\rho-\tau)^2} + \frac{11\mu}{18m(\rho-\tau)^{\frac{4}{3}}} \quad (19)$$

$$R_{22} = \frac{28}{9\mu^2(\rho-\tau)^{\frac{10}{3}}} + \frac{140m}{9\mu^2(\rho-\tau)^4} + \frac{4}{\sin^2\theta} + \cot^2\theta \quad (20)$$

$$R_{33} = \frac{-28}{9\mu^2(\rho-\tau)^{\frac{10}{3}}\sin^2\theta} + \frac{140m}{9\mu^3(\rho-\tau)^4\sin^2\theta} + \frac{4}{\sin^2\theta} + \frac{11\cot^2\theta}{\sin^2\theta} \quad (21)$$

μ is given by the following equation with the mass of a planet, m:

$$\mu = (\frac{3}{2}\sqrt{2m})^{\frac{2}{3}} \quad (22)$$

2.2. Distortion of time and space in strong gravity

We used the same assumption of our previous research [1] for simulating the distortion of time and space, as shown in Figure 3 and Figure 4. In these figures, r is the distance from the center of the strong gravity, t is the time to travel for the distance, τ is the distorted time (5), which expands and shrinks depending on the distance r and the time t; and, ρ is the distorted distance (6), which expands and shrinks depending on the time t and the distance r.

For the simulation, we created two models, Case-1 (non-linear model) and Case-2 (linear model) as shown in Table 1, which assign the functions of $f(r)$ in the equation (5) and $g(r)$ in the equation (6).

Table 1: Two models for simulating the distorted time-space

	Case-1 (Non-linear model)	Case-2 (Linear model)
$f(r)$	$\log \cdot r$	$1/r$
$g(r)$	e^r	r

Then we made the distributions of r and t, following the description of [8], "Any signal, even a light signal, would take an infinite time to cross the boundary of a black hole". However, we could not set the infinite values for the simulation; therefore, instead we set 24 discrete finite values in order to mock "It takes more time to travel closer to the center of the black hole". The values of r and t are assigned to make input vectors for the numeric simulation with a personal computer.

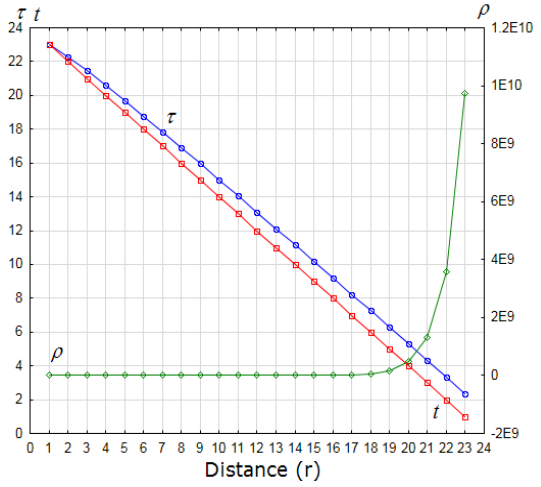


Figure 3: Time and distance from the center of the gravity, Case-1 (non-linear distortion): $f(r) = \log r$ and $g(r) = e^r$

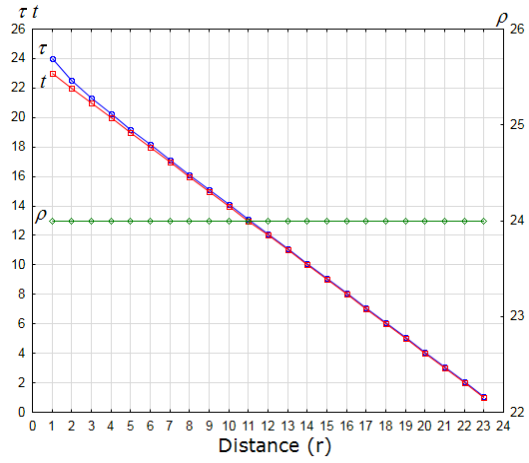


Figure 4: Time and distance from the center of the gravity, Case-2 (linear distortion): $f(r) = \frac{1}{r}$ and $g(r) = r$

2.3. Algorithm for the gravitational energy with no rotation

We used the same algorithm that we used for our previous research [1] to calculate the relative intensity of the gravitational energy with the curvature tensor, which was to be reflected by the stress-energy tensor placed at the end of the distance ϒ in Figure 3 and 4.

The Einstein's equation that rules the motion of particles in the gravitational field is shown below.

$$(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R)_{,\nu} = 0 \tag{23}$$

Then we took the idea of the stress energy tensor from the classical mechanics [9] and set the equation (24), where T is the stress energy tensor and k is a constant.

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = kT \tag{24}$$

In order to calculate the coefficients of the tensor, we made the equation shown below, where c_1, c_2 and c_3 are the coefficients that make a column vector, c.

$$H = kT - R_{\mu\nu} = kT - (c_1 X_1 + c_2 X_2 + c_3 X_3) \tag{25}$$

For calculating c, we formed a 3×3 matrix, X, by three vectors, X_1, X_2 and X_3, as shown in (26) for making the projected image of the gravitational energy on the surface of the sphere in the spherical polar coordinate system.

$$X = [X_1 \quad X_2 \quad X_3] = [R_{11} \quad R_{22} \quad R_{33}] \tag{26}$$

Then we formed H with matrix algebra, as shown below.

$$H = kT - Xc \tag{27}$$

To solve this equation, we set the constraint (28), where X' is the transposed matrix of X.

$$X'H = X'(kT - Xc) = 0 \tag{28}$$

The matrix algebra continues as shown in (29) and (30) to calculate the values of c.

$$X'Xc = X'kT \tag{29}$$

$$c = (X'X)^{-1} X'kT \tag{30}$$

The standard errors of the coefficients are also calculated by (31), where $V(c)$ is the variance of the c.

$$\sqrt{V(c)} = \sqrt{\hat{\sigma}^2 (X'X)^{-1}} \tag{31}$$

$\hat{\sigma}^2$ is calculated by the equations from (32) to (34).

$$\hat{\sigma}^2 = e'e/(n-l) \tag{32}$$

$$e = M \cdot kT \tag{33}$$

$$M = I - X(X'X)^{-1} X' \tag{34}$$

n is the number of rows of each column of X, while in this simulation the value of n is 23 as shown in Figure 3 and 4. l is the number of columns of X, and I is a 23×23 unit matrix that holds 1 (unity) in all diagonal components and 0 in the other components. $(X'X)^{-1}$ is the inverse matrix of $X'X$, and e' is the transposed vector of e.

2.4. Algorithm for the gravitational energy with rotation

If the object rotates as shown in Figure 2, its coordinate system is transformed by the transformation matrix D of the Euler's angles [9] shown below.

$$D=\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix} \tag{35}$$

For the rotation of the object around one axis of ϕ, the tensor of the object's coordinate system, $R_{\mu\nu}$, is multiplied by D; then it is transformed as shown below.

$$D\cdot R_{\mu\nu}=\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}R_{11} & 0 & 0\\ 0 & R_{22} & 0\\ 0 & 0 & R_{33}\end{bmatrix}$$

$$=\begin{bmatrix}\cos\phi\cdot R_{11} & \sin\phi\cdot R_{22} & 0\\ -\sin\phi\cdot R_{11} & \cos\phi\cdot R_{22} & 0\\ 0 & 0 & R_{33}\end{bmatrix} \tag{36}$$

The components, $\sin\phi\cdot R_{22}$ and $-\sin\phi\cdot R_{11}$ are anti-symmetrical, which are perpendicular to the rotation axis, z for ϕ of Figure 2. From the above transformed tensor after the rotation, $D\cdot R_{\mu\nu}$, we took out its diagonal components and formed (37), to calculate the relative intensity of the principal moment of the rotation.

$$\begin{bmatrix}dR_{11} & 0 & 0\\ 0 & dR_{22} & 0\\ 0 & 0 & R_{33}\end{bmatrix}=\begin{bmatrix}\cos\phi\cdot R_{11} & 0 & 0\\ 0 & \cos\phi\cdot R_{22} & 0\\ 0 & 0 & R_{33}\end{bmatrix} \tag{37}$$

Then we formed H as shown below to calculate the coefficients of the diagonal components.

$$H=kT-(c_1\cdot dR_{11}+c_2\cdot dR_{22}+c_3\cdot R_{33}) \tag{38}$$

Henceforth we followed the same procedure explained in the section *2.3.*, but with the matrix shown below.

$$X=\begin{bmatrix}\cos\phi\cdot R_{11} & \cos\phi\cdot R_{22} & R_{33}\end{bmatrix} \tag{39}$$

2.5. Algorithm for the angular momentum of the rotation

We formed a matrix shown in (40), by taking out the anti-symmetrical components of $D\cdot R_{\mu\nu}$ from (36); then formed a column vector shown in (41).

$$\begin{bmatrix}0 & \sin\phi\cdot R_{22} & 0\\ -\sin\phi\cdot R_{11} & 0 & 0\\ 0 & 0 & 0\end{bmatrix}=\begin{bmatrix}0 & R_{22}\cdot d\Omega_3 & 0\\ -R_{11}\cdot d\Omega_3 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{40}$$

$$\begin{bmatrix}dR_{11}\\ dR_{22}\\ dR_{33}\end{bmatrix}=\begin{bmatrix}R_{22}\cdot d\Omega_3\\ -R_{11}\cdot d\Omega_3\\ 0\end{bmatrix} \tag{41}$$

With the above column vector, we formed H shown below, for calculating the coefficients of the vector's components.

$$H=kT-(c_1\cdot R_{22}\cdot d\Omega_3-c_2\cdot R_{11}\cdot d\Omega_3) \tag{42}$$

Then we followed the same procedure as explained above, but with the matrix shown in (43) of the anti-symmetrical components. It is to simulate the angular momentum that is to be projected on the imaginary flat surface, which is perpendicular to the spherical surface.

$$X=\begin{bmatrix}\sin\phi\cdot R_{22} & -\sin\phi\cdot R_{11}\end{bmatrix}=\begin{bmatrix}R_{22}\cdot d\Omega_3 & -R_{11}\cdot d\Omega_3\end{bmatrix} \tag{43}$$

here a little explanation is needed about $d\Omega_3$ of (43). At first, ε of (44) is an infinitesimal rotation operator. But, in general it has a form of (45) according to the Reference [9]. Then a rotated vector as the cross-product of $R_{\mu\nu}$ and $d\Omega$ makes (46).

$$\varepsilon=\begin{bmatrix}0 & d\Omega_3 & 0\\ -d\Omega_3 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{44}$$

$$\varepsilon=\begin{bmatrix}0 & d\Omega_3 & -d\Omega_2\\ -d\Omega_3 & 0 & d\Omega_1\\ d\Omega_2 & -d\Omega_1 & 0\end{bmatrix} \tag{45}$$

$$\begin{bmatrix}dR_{11}\\ dR_{22}\\ dR_{33}\end{bmatrix}=R_{\nu\nu}\times d\Omega=\begin{bmatrix}R_{11}\\ R_{22}\\ R_{33}\end{bmatrix}\times\begin{bmatrix}d\Omega_1\\ d\Omega_2\\ d\Omega_3\end{bmatrix}=\begin{bmatrix}R_{22}d\Omega_3-R_{33}d\Omega_2\\ R_{33}d\Omega_1-R_{11}d\Omega_3\\ R_{11}d\Omega_2-R_{22}d\Omega_1\end{bmatrix} \tag{46}$$

However, in this simulation, we assumed (47) and (48), which make (41) instead of (46).

$$d\Omega_1=d\Omega_2=0 \tag{47}$$

$$d\Omega_3=\sin\phi \tag{48}$$

2.6. Input data

The time t and the distance r are set as shown in Figure 3 for Case-1 and in Figure 4 for Case-2. For simulating the spatial expansion of the gravitational field, we assumed as if θ would become larger in far distance as shown in Figure 5. For simulating the rotation of the object, we set two cases, assuming ϕ_1 (the rotation 1) and ϕ_2 (the rotation 2) also as shown in Figure 5. With these settings, $\sin\theta$, $\cos\theta$, $\cot\theta$, $\sin\phi_1$, $\sin\phi_2$, $\cos\phi_1$ and $\cos\phi_2$ behave as shown in Figure 6.

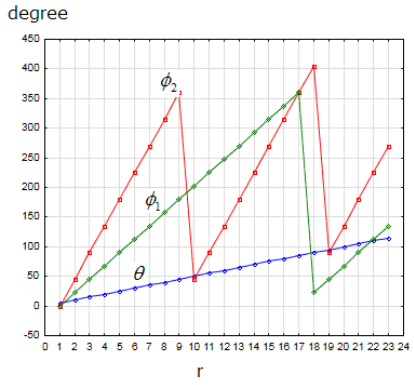


Figure 5: Angles, θ, ϕ_1 and ϕ_2, for the simulation

In addition, we set the stress-energy tensor as 1; because, the purpose of this simulation is to measure the relative intensity of each component of the tensors.

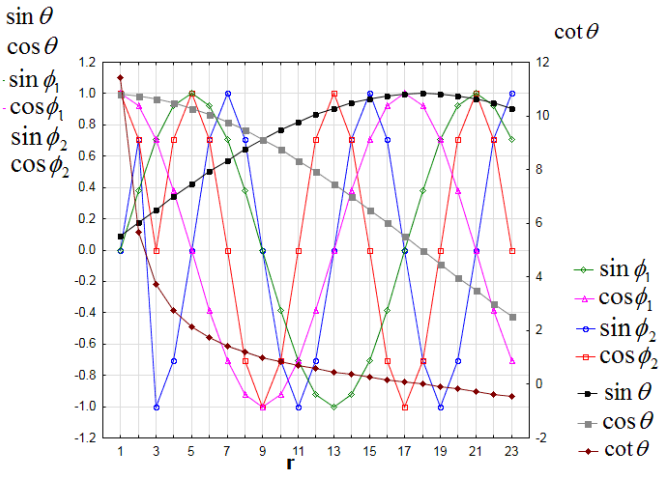


Figure 6: $\sin\theta$, $\cos\theta$, $\cot\theta$, $\sin\phi_1$, $\sin\phi_2$, $\cos\phi_1$ and $\cos\phi_2$

3. Result

3.1. Gravitational energy and angular momentum: overview

Figure 7 (Table 2) shows the relative intensity of the gravitational energy of the object, which is projected on the surface of the sphere of the curved space, and the angular momentum of the perpendicular vector to the surface. In Case-1 (non-linear distortion of the time and space), the gravitational energy (on curved surface) is negative (gravity) before the rotation, but it changes to positive (antigravity) in the rotation 1, and then to negative (gravity) again in the rotation 2. It means that the antigravity appears, depending on the speed of the rotation of the object. The sign of the angular momentum (on perpendicular vector) changes from positive to negative when the rotation becomes faster (from the rotation 1 to the rotation 2). It means that the direction of the angular momentum changes, depending on the speed of the rotation. In Case-2 (linear distortion of time and space), the gravitational energy is positive with no rotation (but smaller than in Case-1 and closer to zero) in Figure 7; while, the gravitational energy (negative) becomes larger when the object rotates faster. The angular momentum of Case-2 changes as it changes in Case-1.

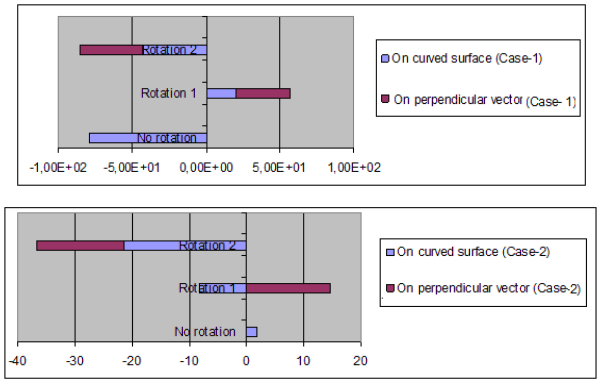


Figure 7: Gravitational energy on curved surface and angular momentum on perpendicular direction to the surface

Note: In Figure 7, "On curved surface" means the gravitational energy, which is the sum of the calculated coefficients in the equation of (25) for the no rotation and the sum of the calculated coefficients in the equation of (38) for each of the rotation 1 and the rotation 2. "On perpendicular vector" means the angular momentum, which is the sum of the calculated coefficients of (42) for each of the rotation 1 and the rotation 2.

Table 2: Intensities of gravitational energy and angular momentum: overview

	Case-1		Case-2	
	On curved surface	On perpendicular vector	On curved surface	On perpendicular vector
No rotation	-78.55	---	1.770	---
Rotation 1	20.00	36.90	-8.178	14.77
Rotation 2	-41.96	-43.00	-21.31	-15.29

3.2. Gravitational energy in three directions

Figure 8 (Table 3) and Figure 9 (Table 4) show the intensities of the gravitational energy, projected on the spherical curved surface in the components of R_{11}, R_{22} and R_{33} with no rotation, and in the components of $\cos\phi\cdot R_{11}$, $\cos\phi\cdot R_{22}$ and R_{33} with the rotation 1 and the rotation 2.

In Figure 8 for Case-1, only the component of R_{11} appears on the surface with no rotation, and only the component of $\cos\phi\cdot R_{11}$ appears with the rotation 1 and the rotation 2. The rotation 1 shows the antigravity (positive).

In Figure 9 for Case-2, the component of $\cos\phi\cdot R_{22}$ also appears in addition to the component of $\cos\phi\cdot R_{11}$ when the object rotates, and they become positive (antigravity) with the rotation 1 and the rotation 2. Here, it is noted that the components of $\cos\phi\cdot R_{22}$ and R_{33} don't appear in Case-1 of Figure 8, but their calculated values are shown in Table 3; and, in Case-2 the component of R_{33} doesn't appear in Figure 9, but its calculated values are shown in Table 4.

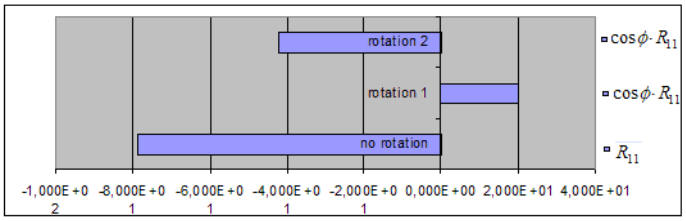


Figure 8: Projection of the gravitational energy in 3 directions on the surface of the curvature, Case-1

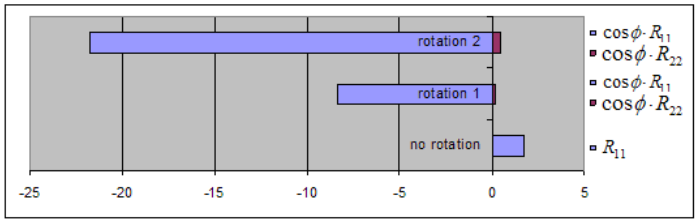


Figure 9: Projection of the gravitational energy in 3 directions on the surface of the curvature, Case-2

Table 3: Intensity of gravitational energy in 3 components, Case-1

Diagonal components of $R_{\mu\nu}$	c and $\sqrt{V(c)}$ of $R_{\mu\nu}$ before the rotation	Diagonal components of rotated $R_{\mu\nu}$	c and $\sqrt{V(c)}$ (Rotation 1)	c and $\sqrt{V(c)}$ (Rotation 2)
R_{11}	-78.68 (26.49)	$\cos\phi \cdot R_{11}$	20.01 (58.22)	-42.05 (52.30)
R_{22}	0.1307 (0.03369)	$\cos\phi \cdot R_{22}$	-5.516×10^{-3} (0.0803)	0.08903 (0.06829)
R_{33}	-6.803×10^{-5} (2.557×10^{-4})	R_{33}	-3.290×10^{-4} (3.869×10^{-4})	-2.990×10^{-4} (5.403×10^{-4})

Note: The value without the bracket is the coefficient c; and, the value in the bracket is the standard error of the coefficient $\sqrt{V(c)}$. By econometrics [10], briefly the calculated coefficient is more significant if its standard error of the coefficient is smaller than the value of the coefficient.

Table 4: Intensity of gravitational energy in 3 components, Case-2

Diagonal components of $R_{\mu\nu}$	c and $\sqrt{V(c)}$ of $R_{\mu\nu}$ before the rotation	Diagonal components of rotated $R_{\mu\nu}$	c and $\sqrt{V(c)}$ (Rotation 1)	c and $\sqrt{V(c)}$ (Rotation 2)
R_{11}	1.767 (7.364)	$\cos\phi \cdot R_{11}$	-8.368 (11.85)	-21.79 (16.36)
R_{22}	2.862×10^{-3} (0.1469)	$\cos\phi \cdot R_{22}$	0.1924 (0.2427)	0.4849 (0.3369)
R_{33}	1.110×10^{-5} (2.224×10^{-3})	R_{33}	-2.854×10^{-3} (3.673×10^{-3})	-7.281×10^{-3} (5.099×10^{-3})

3.3. Angular momentum in two directions

Figure 10 (Table 5) and Figure 11 (Table 6) show the intensities of the rotation's angular momentum in two directions, $\sin\phi \cdot R_{22}$ and $-\sin\phi \cdot R_{11}$, which are perpendicular to the rotation axis, ϕ. In Figure 10 for Case-1, only the vector's component of $-\sin\phi \cdot R_{11}$ appears for the rotation 1. For the rotation 2, the component of $-\sin\phi \cdot R_{11}$ appears, and $\sin\phi \cdot R_{22}$ is very slightly visible in this figure.

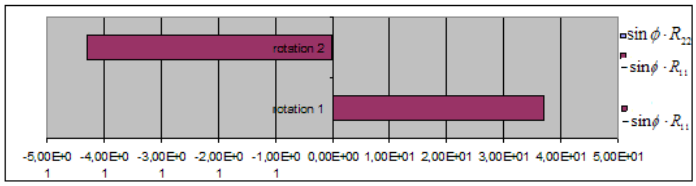


Figure 10: Angular momentum of the rotating object in 2 directions, case-1

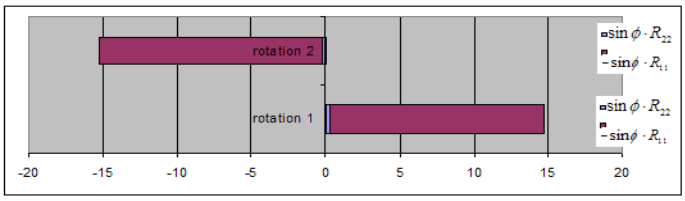


Figure 11: Angular momentum of the rotating object in 2 directions, case-2

In Figure 11 for Case-2, the vector's component of $\sin\phi \cdot R_{22}$ also appears. In both Case-1 and Case-2, when the speed of the rotation increases from the rotation 1 to the rotation 2, the sign of the angular momentum changes from plus to minus. It means that the direction of the angular momentum reverses when the speed of the rotation of the object changes.

This result suggests a consistency with our previous report on the direction of the spin momentum of gravitational waves [6] shown in Figure 1; however, because gravitational waves are beyond the scope of this extended paper, further discussion on the similarity between the direction of the angular momentum of the antigravity and the direction of the spin momentum of anti-gravitational waves should be postponed to the other research.

Table 5: Intensity of the rotation's angular momentum, Case-1

	c and $\sqrt{V(c)}$ (Rotation 1)	c and $\sqrt{V(c)}$ (Rotation 2)
$dx_1 = R_{22} \cdot d\Omega_3 = \sin\phi \cdot R_{22}$	9.077×10^{-2} (5.072×10^{-2})	-4.816×10^{-2} (5.931×10^{-2})
$dx_2 = -R_{11} \cdot d\Omega_3 = -\sin\phi \cdot R_{11}$	36.83 (46.33)	-42.94 (45.44)

Table 6: Intensity of the rotation's angular momentum, Case-2

	c and $\sqrt{V(c)}$ (Rotation 1)	c and $\sqrt{V(c)}$ (Rotation 2)
$dx_1 = R_{22} \cdot d\Omega_3 = \sin\phi \cdot R_{22}$	0.2821 (0.2621)	-0.2285 (0.2597)
$dx_2 = -R_{11} \cdot d\Omega_3 = -\sin\phi \cdot R_{11}$	14.48 (16.65)	-15.06 (16.69)

3.4. Physical meaning of the results

When the idea of quantum mechanics was developed in the early 20th century, there was a discussion [11] to select the coordinate system for quantum mechanics from Einstein's special theory of relativity or his general theory of relativity. He compared two types of the coordinate systems: one was on the flat space-like surface (Figure 12), and another on the curved space-like surface (Figure 13). In each figure, three-dimensional surfaces, S1, S2, S3, S1' in Figure 12 and S in Figure 13, are placed in four-dimensional time-space, where X_0 is for the time, and X_1, X_2, X_3 for the space. The special theory of relativity is explained in Figure 12, while the general theory of relativity is in Figure 13. Figure 13 represents a three-dimensional curved surface in a four-dimensional space-time, which has the property of being everywhere space-like, and the perpendicular vector to the surface has to be in the light-cone of Figure 13, according to [11].

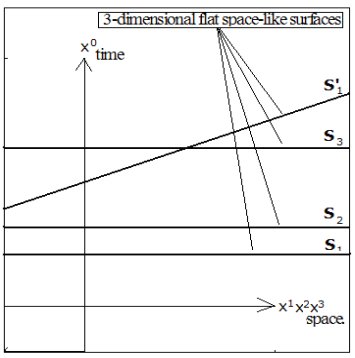


Figure 12: Flat space-like surface (remade from [11]

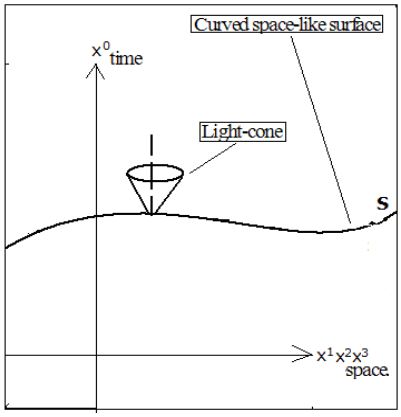


Figure 13: Curved space-like surface (remade from [11]

Then it was predicted by [11] that this perpendicular movement to the curved space must have a physical meaning. However, it was found that the curvature tensor could not satisfy the condition for solving the equation of motion in such a perpendicular direction to the curved surface. Henceforward the theory of quantum mechanics was not developed on the curved surface of Figure 13, but on the flat surface of Figure 12.

In our previous research for the conference paper [1], we made a simulation on the energy of the gravitational field, which was the projection on the spherical surface; but, not the movements of the vectors perpendicular to the spherical surface. However, in this extended paper, we also report the result of the simulation of the angular momentum, which is the perpendicular component of the movement of the curved surface. This part challenges the decision to use the flat surface for quantum mechanics in the early 20th century.

For solving the equation of motion in a three-dimensional curved surface in a four-dimensional space-time, the curvature tensor is needed. But, in our research we used the spherical polar coordinate system as a surrogate of the curved surface so that we could still use the orthogonal transformation of the matrix algebra, which was available originally for the flat space.

The vector components of the spherical surface are the projections of the gravitational energy; and, the movements of these vectors are the movements of the curved surface itself. It looks like Figure 14.

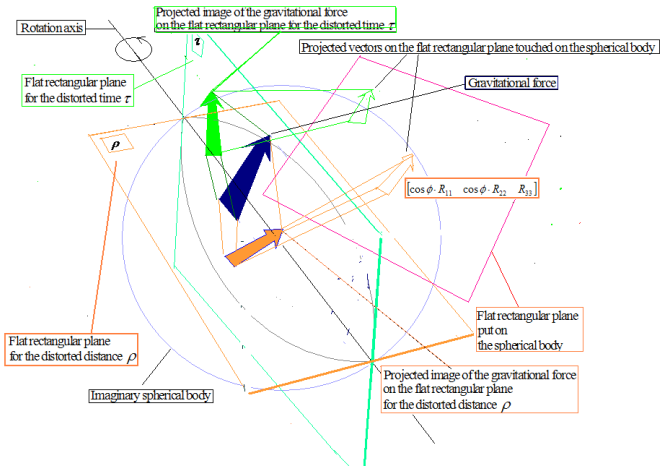


Figure 14: The vector projected on the spherical surface

In Figure 14, a flat rectangular plane is put on the surface of an imaginary spherical body. And, inside of the spherical body, there is one arrow that represents the gravitational force. The rotating object, which is not shown in this figure, is assumed to be located in the center of the spherical body. Also one rotation axis is shown in this figure. On this axis, there are two flat rectangular planes: one is for the distorted time, τ; and, another is for the distorted distance, ρ. First, the arrow of the gravitational force inside of the imaginary spherical body is projected to each of these two flat planes that are with the signs of τ and ρ. Then each of these two arrows is further projected to the flat rectangular plane that touches the surface of the spherical body. Here three vectors of (39) must be on this rectangular plane on the surface, and the calculated coefficients of (38) represent the intensities of the gravitational energy, shown in Figure 8 (Table 3) and Figure 9 (Table 4). If the speed changes in the rotation of the object, the direction of the rotation of the object doesn't change, but the direction of the projected image of the arrow on the rectangular plane of Figure 14 changes. Here it is noted that we didn't include τ in the simulation for this extended paper.

For simulating the vectors of the angular momentum, we used the cross product of anti-symmetrical vectors as the projection of the momentum vector in the perpendicular direction to the curved surface. It looks like Figure 15. The vectors of (43) are projected on the perpendicular plane with the sign of ρ in this figure. The calculated coefficients of (42) are the intensities of the angular momentum, shown in Figure 10 (Table 5) and Figure 11 (Table 6). If the speed changes in the rotation of the object, the direction of the rotation of the object doesn't change; but, the direction of the projected image of the arrow changes on the perpendicular plane for ρ of Figure 15.

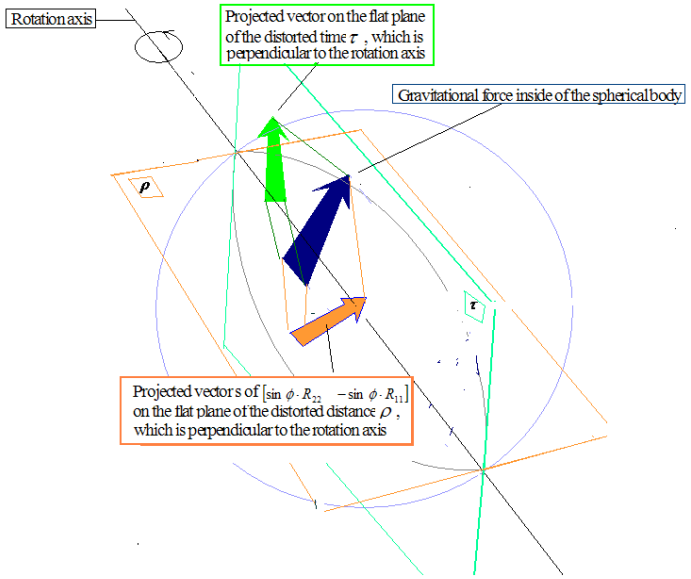


Figure 15: The vectors projected on the perpendicular components of the spherical surface

It is noted that we replaced the generally curved surface of Figure 13 by the spherical polar coordinate system as a surrogate for the simulation. And, we mocked the general movement of the curved surface by the rotation of the curved sphere as shown in Figure 15. In Figure 13, the perpendicular movement of the

generally curved surface is the movement in the light cone, and it is not necessarily the angular momentum. On the other hand, in Figure 15, the perpendicular movement of the spherical polar coordinate system appears as the angular momentum, only when the system rotates.

4. Conclusions and Recommendations

In this research, we investigated how the intensity and the direction change upon the gravitational energy and the angular momentum when the speed of the rotation of the artificial object changes, using Euler's rotation matrix by calculating the coefficients of the equation of motion, which is made of the curvature tensor with the component on the curved surface and the perpendicular component to the surface. The result of the simulation shows that the rotating object, in which time and space are nonlinearly distorted by the strong gravity, can produce the antigravity as the projected image on the curved surface, and change the direction of its angular momentum on the projected perpendicular image to the curves surface.

The change of the direction of the angular momentum upon the emerged antigravity implies our previous prediction [6] on the direction of the spin of anti-gravitational waves, in which anti-gravitational waves have the clockwise spin, while gravitational waves have the anti-clockwise spin. However, the conclusion has not been made on this issue because the analysis we described in [6] was made on the flat space, while we made the simulation in the spherical polar coordinate system for this extended paper. In addition, the discussion about gravitational waves is beyond the scope of this extended paper and it should be deferred to the other research; although, the similarity between the antigravity and the anti-gravitational waves may have been implied by the equation of gravitational waves, which is to be derived as the secondary differential of the equation for the gravitational field.

In addition, our simulation has challenged the limit of the general theory of relativity in its application to quantum mechanics, which is: the perpendicular movement to the generally curved surface could not satisfy the condition to solve the equation of motion. Henceforward the curved space was not used for setting quantum mechanics. However, we challenged this limit, by using the spherical polar coordinate system with the tensor algebra that makes the cross product of anti-symmetrical vectors for simulating the projection of the angular momentum in the perpendicular direction to the spherical surface.

In this research, we used the system of spherical polar coordinates as the surrogate of the generally curved surface; however, in the near future, the developed computer technologies must increase the possibility of simulating the generally curved surface of the Einstein's equations also for solving the equation of motion of quantum particles.

References

[1] Y. Matsuki, P.I. Bidyuk, "Theory and Simulation of Artificial Antigravity," 2020 IEEE 2nd International Conference on System Analysis Intelligent Computing, 2020, doi: 10.1109/SAIC51296.2020.9239195.

[2] Y. Matsuki, P.I. Bidyuk, "Empirical Analysis of Moon's Gravitational Wave and Earth's Global Warming," System Research & Information Technology, N1, 107-118, 2018, doi: 10.20535/SRIT.2308.8893.2018.1.09.

[3] Y. Matsuki, P.I. Bidyuk, "Analysis of Moon's Gravitational-Wave and Earth's Global Temperature: Influence of Time-Trend and Cyclic Change of Distance from Moon," System Research & Information Technology, N3, 19-30, 2018, doi: 10.20535/SRIT.2308.8893.2018.3.02.

[4] Y. Matsuki, P.I. Bidyuk, "Empirical Investigation on Influence of Moon's Gravitational-Field to Earth's Global Temperature," System Research & Information Technology, N2, 18-24, 2019, doi: 10.20535/SRIT.2308.8893.2019.2.02.

[5] Y. Matsuki, P.I. Bidyuk, "Calculating Energy Density and Spin Momentum Density of Moon's Gravitational-Waves in Rectilinear Coordinates," System Research & Information Technology, N3, 7-17, 2019, doi: 10.20535/SRIT.2308.8893.2019.3.01.

[6] Y. Matsuki, P.I. Bidyuk, "Analysis of Negative Flow of Gravitational Waves," System Research & Information Technology, N4, 7-18, 2019, doi: 10.20535/SRIT.2308.8893.2019.4.01.

[7] Y. Matsuki, P.I. Bidyuk, "Numerical Simulation of Gravitational Waves from a Black Hole, using Curvature Tensors," System Research & Information Technology, N1, 54-67, 2020, doi: 10.20535/SRIT.2308.8893.2020.1.05.

[8] P.A.M. Dirac, "General Theory of Relativity", New York: Florida University, A Wiley Inter-science Publication, John Wiley & Sons, 1975.

[9] H. Goldstein, C.P. Poole, J.L. Safko, Classical Mechanics, 3rd Edition published by Pearson Education. Inc., 2002.

[10] A.S. Goldberger, A Course in Econometrics, Harvard University Press, Cambridge, Massachusetts, USA, 1991.

[11] P.A.M. Dirac, Lectures on quantum mechanics, originally published by the Belfer Graduate School of Science, Yeshiva University, New York, 1964.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 52-61 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Advanced Physical Failure Analysis Techniques for Rescuing Damaged Samples with Cracks, Scratches, or Unevenness in Delayering

Yanlin Pan*, Pik Kee Tan, Siong Luong Ting, Chang Qing Chen, Hao Tan, Naiyun Xu, Krishnanunni Menon, Hnin Hnin Win Thoungh Ma, Kyaw Htin

GLOBALFOUNDRIES Singapore Pte. Ltd., QRA-EFA, Singapore, 738406, Singapore

ARTICLE INFO

Article history:
Received: 04 May, 2021
Accepted: 26 June, 2021
Online: 10 July, 2021

Keywords:
Delayering
PFA
Sample damage
Sample rescue

ABSTRACT

This paper is an extended version of work published in IPFA 2020. In the previous paper, advanced physical failure analysis (PFA) techniques for rescuing damaged samples with cracks, scratches, or unevenness in delayering are introduced. In the present work, the techniques will be further exploited and summarized for the potential applications in general devices. The three typical rescue cases will be fully discussed through comprehensive analysis on the failure mechanism and the rescuing process. Compared to the conventional PFA techniques that normally require back-up samples, the novel rescue techniques offer more alternative solutions for coping with sample damage problems in delayering without starting over with a new sample that would waste machine time and human resources. These new PFA techniques involve only basic failure analysis (FA) skills that could be easily manipulated and FA equipment that is commonly available in FA labs, and would extend the scope and capability of the tradition PFA to help the FA engineers deliver FA results with high quality and high success rate in the daily work, especially for handling "one of a kind" devices.

1. Introduction

Delayering by finger polishing is one of the commonly used physical failure analysis (PFA) techniques in the failure analysis (FA) labs. This technique is simple, direct, and flexible, which is basically mechanical polishing using fingers to press the sample against a rotating cloth plate with polishing slurry in between the cloth and the sample [1]. Before the polishing of each layer, reactive ion etch (RIE) is usually used to expose the metal structure by removing the inter-metal dielectric (IMD) for a faster removal rate and a more even surface of the polished sample. For monitoring the polishing progress, FA engineers will then use the scanning electron microscope (SEM) and the optical microscope (OM) to inspect the sample surface. A typical delayering workflow in PFA is shown in Figure 1. Normal process of the delayering completes when the target layer is reached, followed by the defect identification and analysis. However, if the sample comes with the accidental damage or the naturally generated edge rounding during the delayering, the problems have to be fixed to restore the sample to the former condition before the polishing is continued to the target layer.

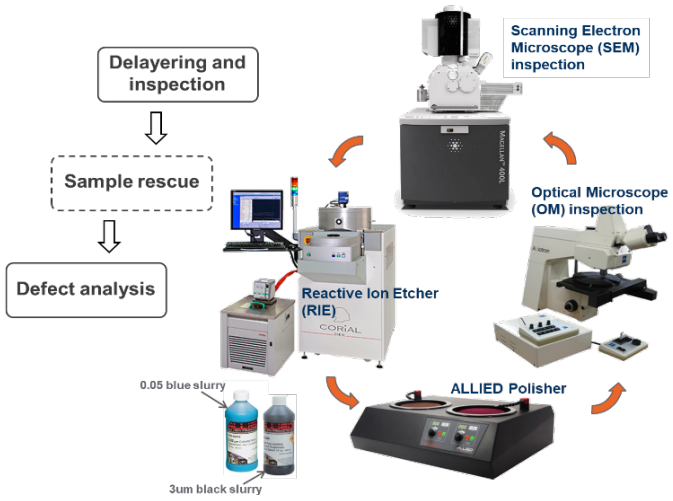


Figure 1: Illustration of a typical delayering workflow in PFA.

The common problems encountered in the delayering include sample cracks, polishing scratches, and surface unevenness [2-6]. The sample cracks refer to the sample die breaking into pieces due to chipping, dropping or crushing, which would pose great challenges to the following polishing if the crack lines extend

*Corresponding Author: Yanlin PAN, 60 Woodlands Industrial Park D Street 2, Singapore 738406. Phone: +65 9114 9926. Email: yanlin.pan@globalfoundries.com

https://dx.doi.org/10.25046/aj060407

across the defect area. The polishing scratches on the sample surface are usually caused by aggregated slurry particles mixed in the slurry solution or diamond lapping films when doing pre-polish sample preparation. The surface unevenness of the samples results from the edge effect which is a normal phenomenon in the delayering by both finger polishing and machine polishing [7-10]. For the small targets such as single memory bits or localized hotspot locations from electrical fault isolation (EFI), the unevenness will not lead to severe problem. But for the big inspection area or the defect location at the edge of the sample, the unevenness must be specifically treated to maintain the feasibility of the FA process. These problems could even become worse with the continuous development of the semiconductor devices, which introduces more layer stacks, smaller transistors, and softer low-k dielectrics. Moreover, accidents due to human error are considered inevitable and hard to control in the real FA work place. Hence, compared to the preventive practice for reducing the sample damage in delayering, finding rescue solutions to the problems is a more pressing need.

In the previous studies, sacrificial dummy as a polishing balance, platinum (Pt) as an etching mask, and lateral delayering with adapted polishing plate were used to counteract the edge effect [11-13]. Other studies focused on the impact of pad properties [14-15], cloth roughness [16], slurry particle size [17-18], and interaction mechanism [19-20] on the material removal rate and polishing uniformity. However, the unevenness problems have not been fully resolved because of the complexity of the issue. Unevenness is related to numerous factors such as force of pressing samples, speed of plate rotation, polishing angle, polishing orientation, slurry type and sample thickness, which are all dependent on the experience of the engineers. These variable parameters and different conditions of different devices would make it nearly impossible to operate in a standard procedure. In addition to surface unevenness, polishing scratches and sample cracks in the delayering are also dealt with in the daily FA work, but most of time there is no effective method of bringing the sample back to the normal condition. In order to fill the skill gaps, new techniques need to be developed.

In this paper, we will introduce three advanced PFA techniques for rescuing damaged samples with cracks, scratches, or unevenness in delayering by finger polishing, through their according typical FA cases. The first case uses the diamond lapping film and the sacrificial dummy samples to fix the cracked sample for continuing the FA process with no effect from the cracks. The second case uses e-beam Pt deposited in SEM machine to repair the scratch pits in the region of interest (ROI). With the polishing progresses to the lower layers, the scratch features gradually diminish until thoroughly removed from the sample surface. In the third case, an innovative technique combining controlled slurry polishing and partial RIE was developed for creating an ultra-large inspection area (> 50000 μm^2) with negligible unevenness. Compared to the conventional PFA techniques that normally require back-up samples, the novel rescue techniques offer more alternative solutions for coping with sample damage problems in delayering without starting over with a new sample that would waste machine time and human resources. These techniques would extend the scope and capability of the tradition PFA, and help the engineers deliver FA results with high success rate, especially for handling "one of a kind" devices.

2. Experiments

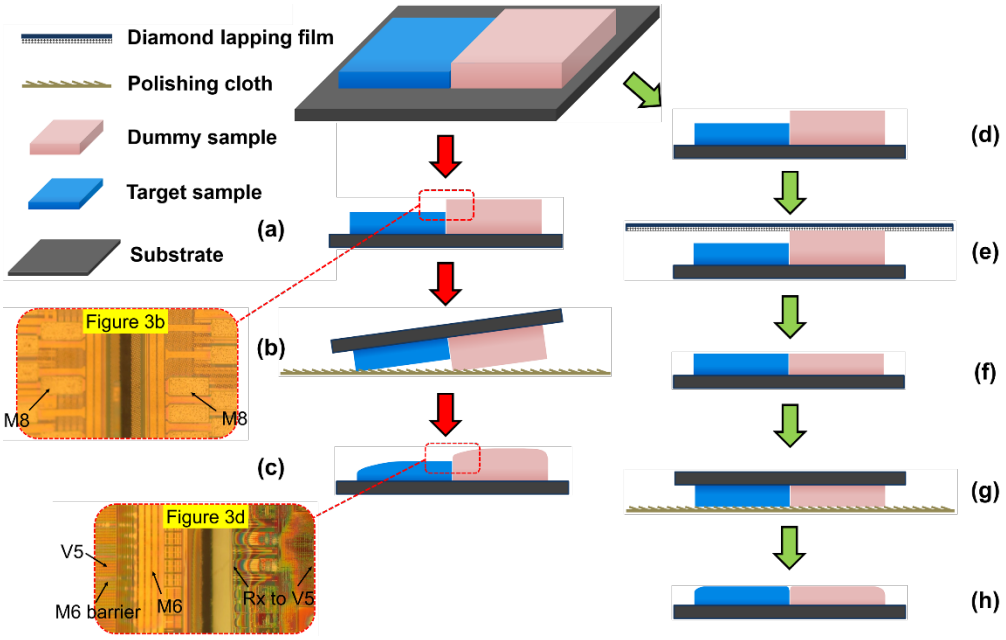
The experiments of the three cases were performed on a 40 nm node logic device, a 40nm node static random access memory (SRAM) device, and a 40 nm node "snake" metal line electrical test (ET) device, respectively. A mechanical polisher (ALLIED HighTech TwinPrep 5) with a rayon flock polishing cloth (ALLIED Spec-Cloth) and 3 μm polishing slurry (ALLIED water based diamond suspension) were used for slurry polishing. A 3 μm and a 15 μm diamond lapping films (ALLIED) were used for the selective polishing on the dummy samples. A SEM system (FEI Magellan™ 400L) equipped with a gas injection system was used for the topography inspection on the sample surface and the e-beam Pt deposition on the scratch pits. An optical microscope (ZEISS Axiotron) was used for the fast inspection during the mechanical polishing and the sample height determination. A RIE system (Corial 200 IL) was used to remove the IMD materials. TEM sample preparation was performed using a FIB-SEM dual-beam system (FEI Helios NanoLab 450S). TEM analysis was performed using a 200 kV Field Emission TEM (JEOL JEM-2100F).

3. Results and Discussion

3.1. Case 1: Using Diamond Lapping Film and Sacrificial Dummy to Save Cracked Sample with Target Defect Area Close to Gaps/Crack Lines

The first case is about a 40 nm node logic device which suffers scan failure with a diagnosed trace path across an area of 283 μm × 71 μm. The sample was delayered halfway (at M8) in the PFA process when it accidentally dropped broken into multiple pieces. Directly continuing the delayering on the piece that contains the defect area would come with the difficulty of polishing and the high risk of missing the defects that could locate anywhere along the whole trace path through each layer. Instead, the broken pieces were glued together with two dummy samples as polishing balance on a substrate using wax (Figure 5a). However, when the normal slurry polishing on the joined sample reached V5, obvious unevenness was observed near the crack line and the wax gap (Figure 5b). The mechanism of the unevenness generating during slurry polishing is illustrated in Figure 2a-2c. The height difference between the dummy sample and the target sample without treatment aggravated with the progress of delayering and caused surface unevenness of the sample.

Figure 3 shows a failed case of delayering where the dummy sample was originally higher than the target sample at M8, according to the focusing conditions in OM. After the sample was delayered from M8 to V5, severe unevenness had generated in the defect area and the dummy side was still higher than the sample side (Figure 3c, 3d), which means that the unevenness would further extend to the lower layers in the following polishing process. In the rescued case, we had to prevent the unevenness from extending to the defect area that would introduce the risk of missing the defect when the defect location or the layers were misjudged due to the unevenness. The unevenness in the defect area also posed great challenges to the layer-by-layer inspection for finding the defects, especially at the critical layers (V5 in this case) which are the interfaces between the low-k IMD layers (M5



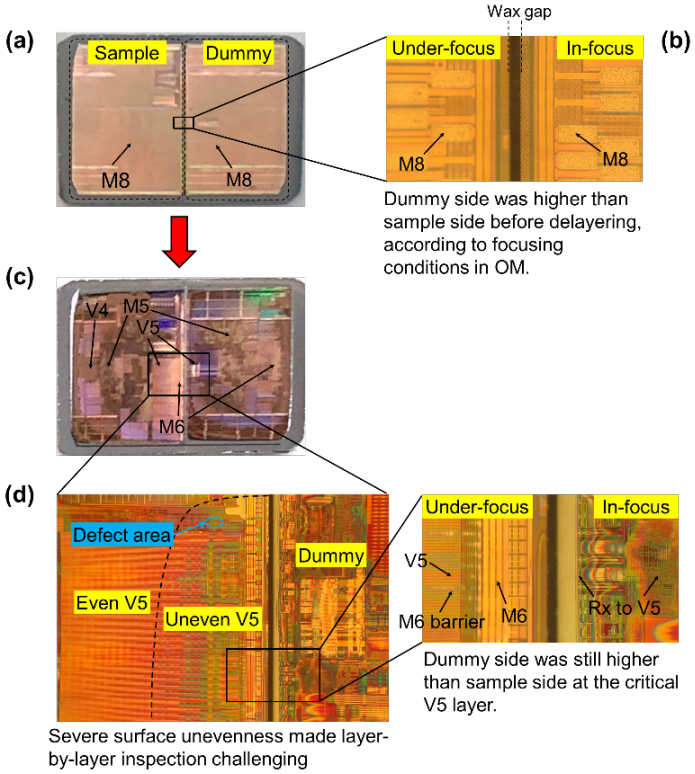


Figure 3: Failed case of delayering where (a) the dummy sample was originally higher than the target sample at M8, according to (b) the sample focusing condition in OM. When the target sample reached V5, (c, d) severe unevenness had generated in the defect area.

uniform surface at the critical layer, the unevenness would significantly deteriorate due to the great difference (up to 10 times) of the removal rate between the low-k materials and the non-low-k materials. Moreover, the sample inspection and the defect finding in SEM would be tediously prolonged because the defect area may cover multiple layers so that the staggered inspections on the different portions of the different layers in the trace path would have to be performed through repeated polish-and-view processes.

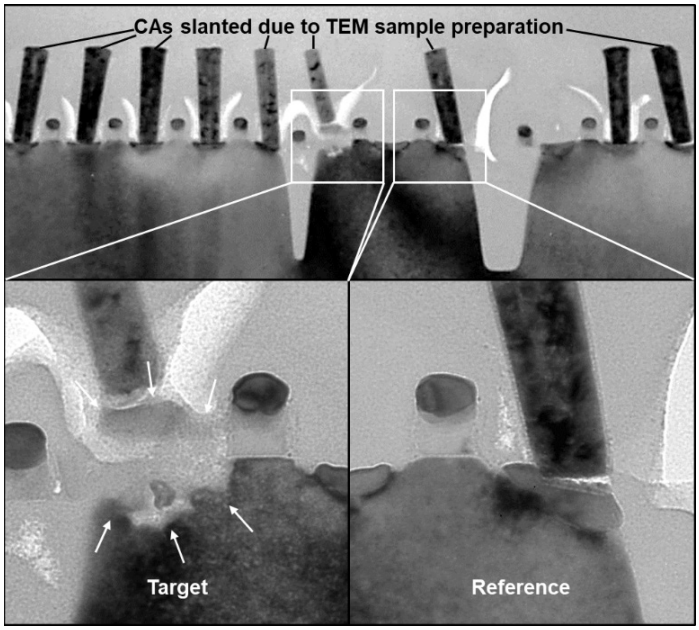


Figure 4: TEM (across PC) images of the target location and the reference location. Damaged Rx and PC-CA short were observed at the target location.

magnification used on the OM was 100× and the numerical aperture (NA) was 0.9. The measurement error regarding to Depth of Field (DOF) was ~ 0.22 μm [21], far smaller than the particle size ~ 3 μm of the polishing slurry. It is worth to mention that the chosen dummy sample should not be thinner than the target sample in the first place.

After removing the height difference between the target sample and the dummy sample, the unevenness was slowly eliminated in the subsequent slurry polishing within the V5 layer (Figure 5d). The sample was restored to its former evenness from V4 downwards, and the defect of damaged Rx and PC-CA short were found at the PC/CA layer (Figure 5e). The mechanism of the rescue process is illustrated in Figure 2d-2h. Figure 4 shows the TEM (across PC) images of the target defect location and the reference location, from which we can see that CA landing on the PC residue caused PC-CA short and the Rx below was damaged. It was suspected that PC residue blocked the PC etching at the defect location. In summary, to rescue the cracked samples with the defect area close to the gap/crack lines, dummy samples as polishing balance need to be used and the surface of the joined samples must be levelled first using diamond lapping films, before the slurry polishing is continued.

3.2. Case 2: Using E-beam Pt Deposition to Repair Scratched Sample with Scratch Pits in the ROI

The second case involved a 40nm node SRAM device. Electrical bench measurement on the failed unit showed source-drain leakage failure in an ET SRAM structure. EFI hotspot analysis was then performed to narrow down the ROI for the root cause finding in PFA. According to the detected defect location, the sample was delayered and viewed from the top layer. To avoid potential unevenness problem subject to the density difference between the target device region and the adjacent scribe lines, the thick top metal layers were removed using 3 μm diamond lapping film. Unfortunately, multiple scratch pits were accidentally created in the ROI during the lapping process (Figure 9a), which would introduce potential damage to the defect. The failure mechanism is illustrated in Figure 6a-6c. The scratch feature as a pit will extend from the scratched layer to the lower layers if the slurry polishing is directly continued without treatment.

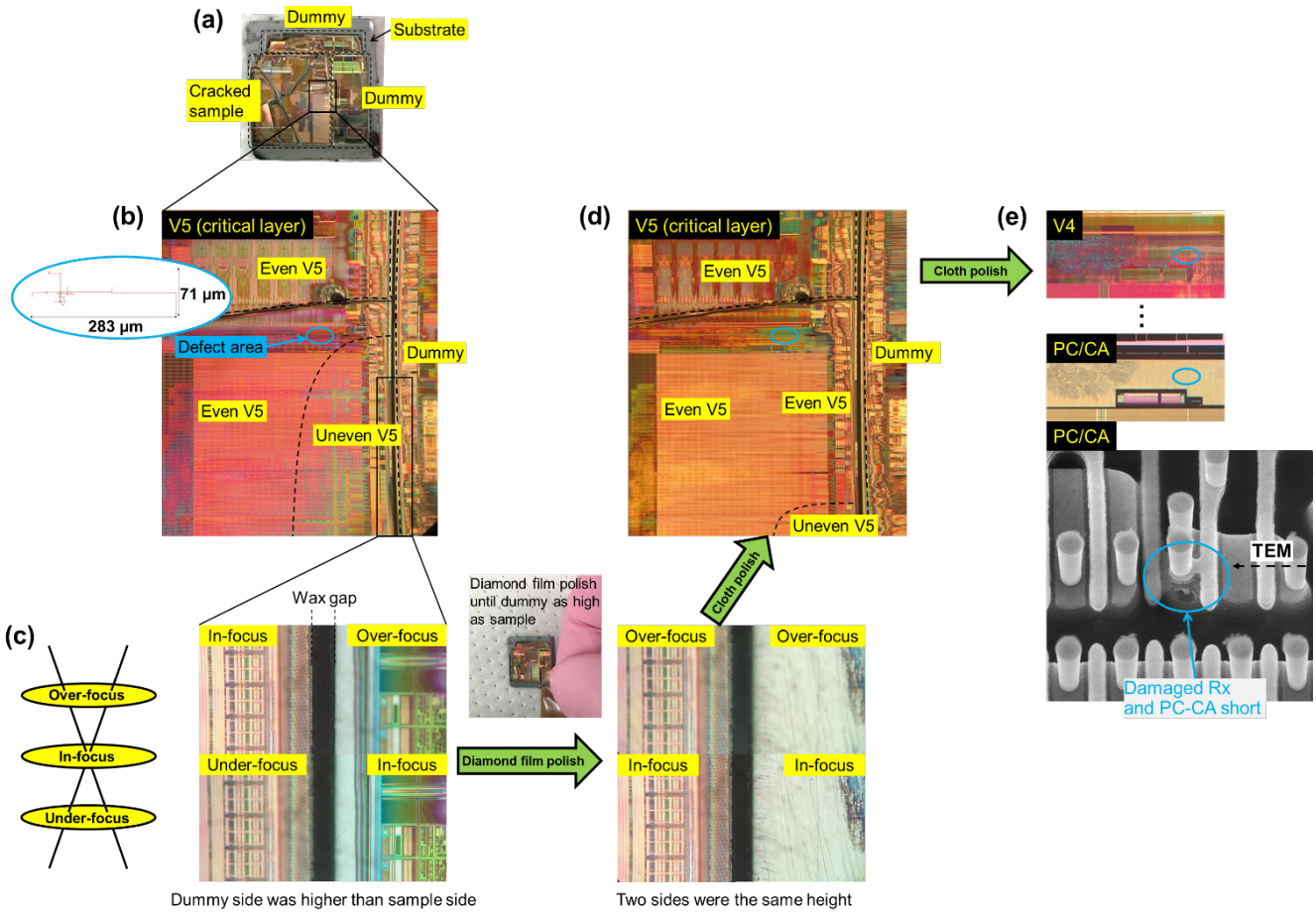


Figure 5: Successful case of sample rescue by using diamond lapping films and sacrificial dummy samples to save the cracked sample with the target defect area close to the crack lines. (a) The cracked samples (halfway at M8) were stuck onto a substrate together with two dummy samples as balance before the slurry polishing is continued. (b) When the joined sample reached V5, obvious surface unevenness was seen near the defect area, generated from the height difference between the dummy sample and the target sample. (c) By selective polishing on the higher dummy side using the diamond lapping films until it was the same height as the target sample, (d) the unevenness was slowly removed in the subsequent slurry polishing within the V5 layer. (e) The sample was restored to its former evenness from V4, and damaged Rx and PC-CA short were found at the PC/CA layer.

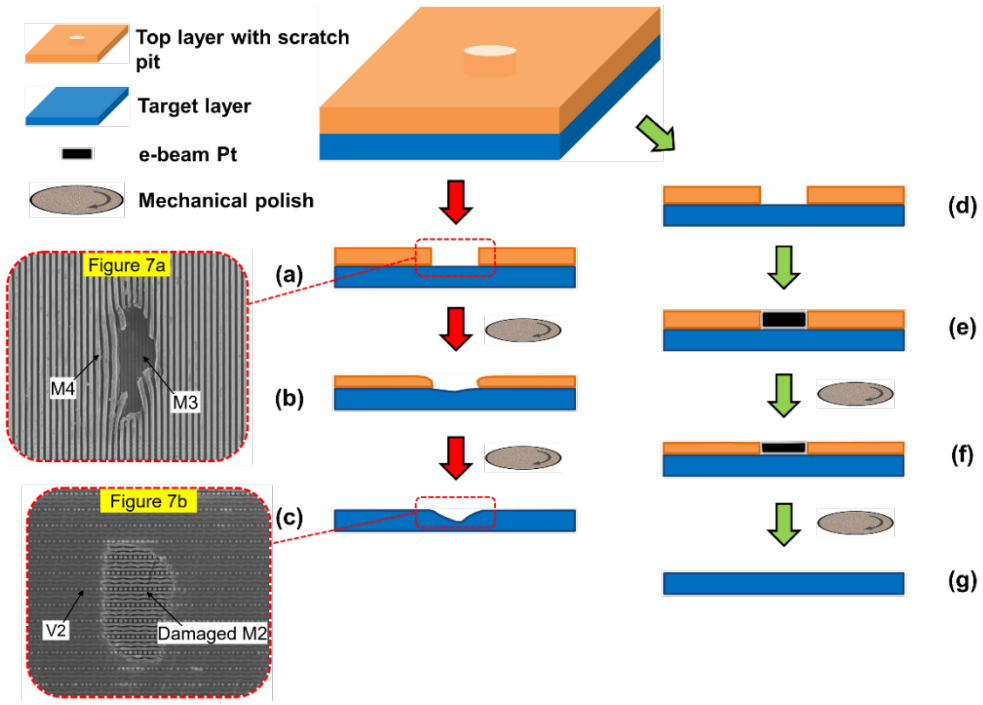


Figure 6: Schematic illustration of the failed case and the successful case of rescuing the scratched samples. (a-c) In the failed case, the scratch feature as a pit will extend to the lower layer if the slurry polishing is continued without treatment, which may damage the defect location. (d-g) In the successful case, by selectively filling the scratch pit with e-beam Pt, the defect location is protected from the damage of over-RIE and over-polish. Insets are the example SEM images from the failed case.

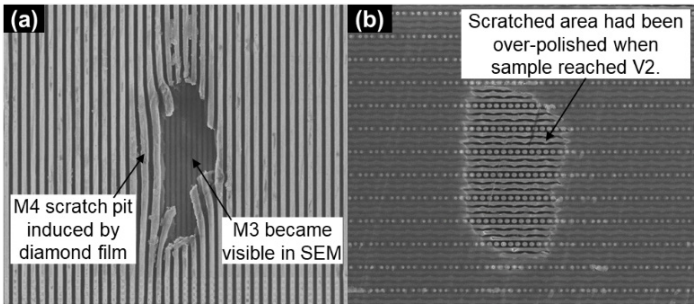
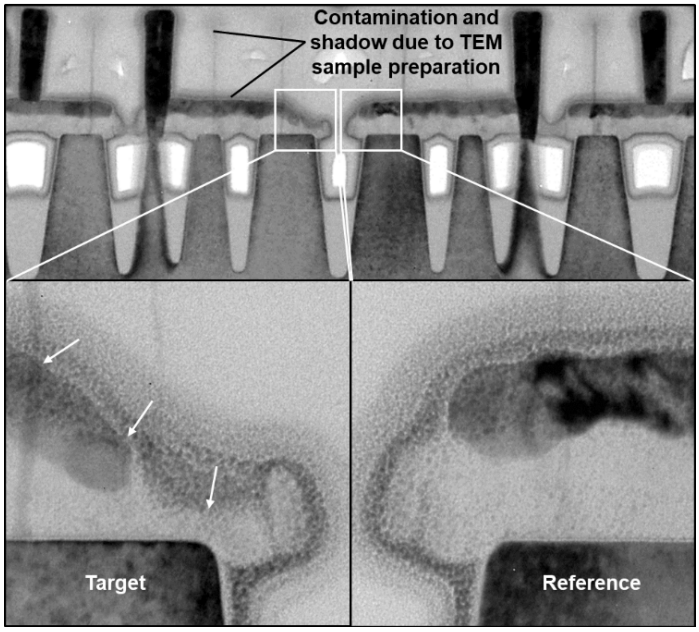


Figure 7: Failed case of delayering where (a) the sample was scratched at M4 by diamond lapping films, and (b) the scratched area had been over-polished when the normal slurry polishing continued and reached V2.

To solve the problem, we used e-beam Pt deposition to fill the scratch pits in SEM. The concept is using Pt metal patch as substitute for Cu to block the RIE and resist the slurry polishing on the next layer. The critical part of this method is the Pt thickness. Over-deposited Pt will leave Pt residue at the ROI resulting in surface unevenness in another way. Therefore, in the beginning we used high kV SEM to deposit thinner Pt film compared to the metal layer of which the thickness was determined by e-beam penetration depth in the SEM imaging. If the Pt were fading faster than the adjacent metal under slurry polishing, we would deposit extra layer of Pt on the ROI. For 40nm node device in this case, the Pt thickness was fixed to 0.1 μm. Voltage of 3 kV and current of 0.8 nA were used for SEM. The deposition rate is 10 μm (length) × 10 μm (width) × 10 μm (thickness) per 10 mins.



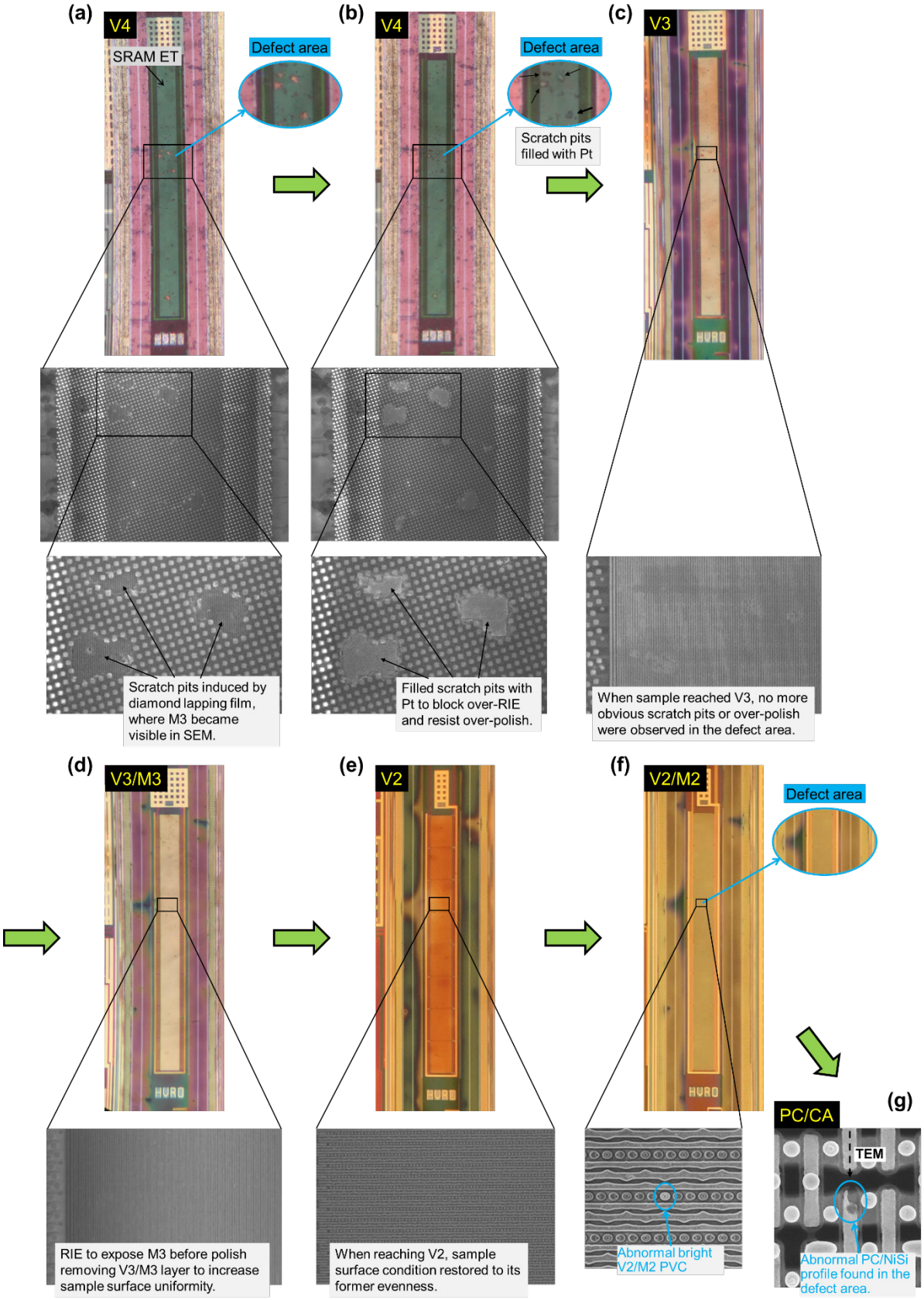


Figure 9. Successful case of sample rescue by using e-beam Pt deposition to repair scratched sample with scratch pits in the ROI. (a) The defect area of the sample got scratch damage when using diamond lapping films to remove top metal layers. (b) To block over-RIE and resist over-polish, e-beam Pt was deposited at the scratch pits to cover the damaged area. (c, d) After the normal RIE and slurry polishing removed V4/M4, the effect of the scratch damage on the V3/M3 layer showed greatly mitigated. (e) At V2 level, the sample had restored to an un-damaged condition, and (f) abnormal bright V2/M2 PVC was observed in the ROI. (g) The defect of abnormal PC/NiSi profile was finally found at the EFI spot location.

With protective Pt filling the scratch pits (Figure 9b), the ROI slowly restored to its former evenness and no more damage feature was seen when the sample reached V2 (Figure 9c-9e), as shown in the successful rescue case. After the M2 was exposed by RIE, optimal surface condition for SEM inspection was obtained and abnormal bright V2/M2 passive voltage contrast (PVC) was observed at the spot location (Figure 9f). The rescue mechanism is illustrated in Figure 6d-6g. The delayering on the sample was then continued from M2 downwards by normal delayering process and finally found the abnormal PC/NiSi at PC/CA level (Figure 9g). TEM images (along PC) are shown in Figure 8. Compared to the reference location, the defect location has abnormal PC/NiSi profile and missing portion of PC which are correlated with the source-drain leakage. PC formation issue in the wafer process is suspected as the root cause. In summary, this case demonstrated simple method of e-beam Pt deposition for tackling the common sample scratching problems induced by diamond film lapping or other incidents in delayering.

3.3. Case 3: Using Controlled Slurry Polishing Combined With Partial RIE to Remove Sample Unevenness and Create Ultra-Large (> 50000 μm^2) Inspection Area

The third case is about an ultra-large "snake" metal line ET structure in a 40 nm node device that failed M4 high resistance. To perform PFA on the failed ET, we tried to polish the sample to V4 and identify the physical defect that resulted in the high resistance. The area to be inspected is as big as > 50000 μm^2 where a single metal line "snake" runs through the whole ET structure in the wafer scribe line. PVC inspection in SEM is needed to locate the defect position which requires intact connection in the metal line without any damage such as metal bridging or metal broken from the sample preparation. However, when the normal slurry polishing on the sample reached V5, obvious evenness at the ET corners was observed, which would potentially lead to metal smear or break if we continued to polish until V4.

A failed case is shown in Figure 11. The "snake" metal line ET sample had uneven IMD at V5 due to the edge effect, while slurry polish directly proceeded upon full time RIE removing IMD. After M5 was removed, severe M4 unevenness and damage were generated. At the corner areas, the sample was even over-polished to M3. The failure mechanism is shown in Figure 10a-10c, where the IMD unevenness is consider as the root cause. Without intact metal line connection, PVC inspection was not possible. Even if the defect was not damaged, tedious inspection in SEM would be needed to check every piece of metal through the whole large area for the defect.

Figure 10d-10h illustrates how we coped with the issue in the successful case. By smearing the metal lines to form an etch-stop area after the first partial RIE (Figure 10f), the IMD was evened

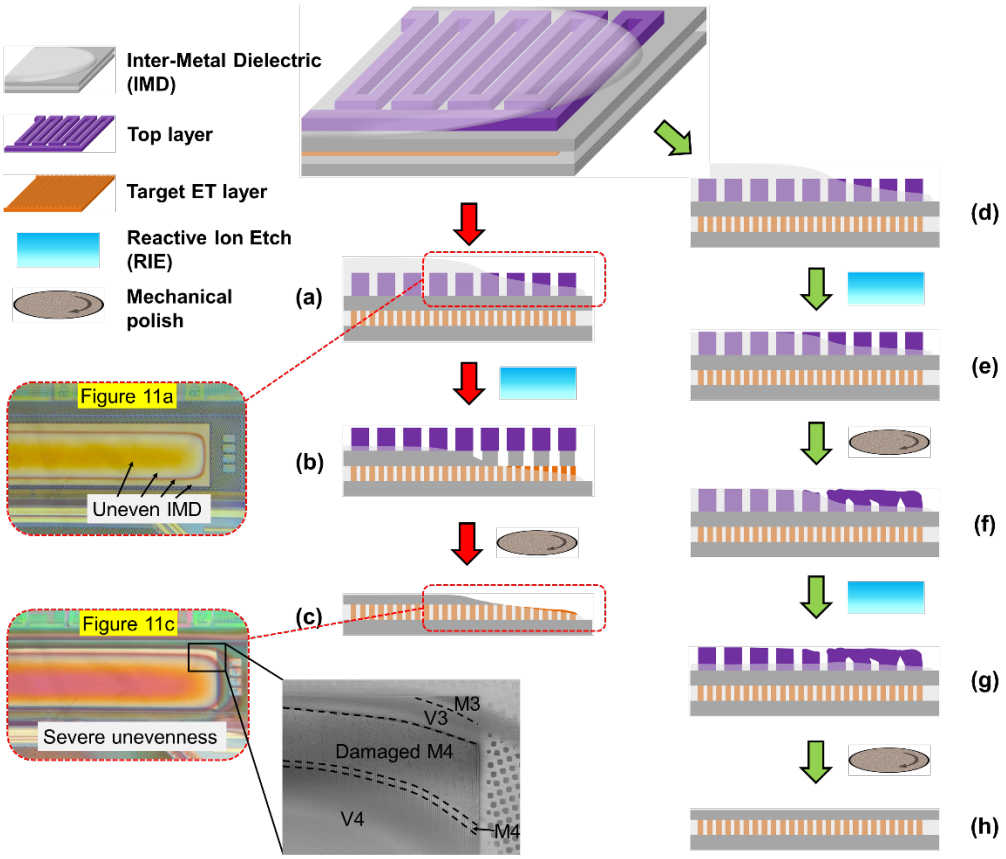


Figure 10: Schematic illustration of the failed case and the successful case of rescuing the uneven samples. (a-c) In the failed case, the edge effect induced unevenness will aggravate along with the slurry polishing, especially for the structure in the scribe lines. (d-g) In the successful case, by smearing the metal lines to form an etch-stop area after the first partial RIE, the unevenness of the IMD will be removed, which will result in an even surface in the next layer. Insets are the OM and SEM images from the failed case.

out by another partial RIE (Figure 10g) to obtain an even surface in the next layer (Figure 10h). The SEM and OM results of the successful rescue case are shown in Figure 13. When the uneven IMD was spotted as circular halo at the ET edge, partial RIE using half of the full RIE time was performed to expose the outer M5 with thinner IMD (Figure 13a). The different division of the RIE time is dependent on different sample conditions including via length and degree of unevenness. More divisions deliver better results in flattening the IMD. In this case, the full RIE time was equally divided into two halves (e.g., if the full time is 20 s, each half will be 10 s). The first partial RIE shrank the IMD halo by around 50% (Figure 13b), which means half of the M5 was exposed. Then we slightly slurry polished the surface until the exposed M5 metal lines collapsed and smeared together. The halo area further shrank smaller because the V5 was polished thinner or nicely fully removed. As a result, three different regions (smeared M5, exposed M5, and un-exposed M5) were generated (Figure 13c). Since the RIE for oxide etching doesn't chemically react with metal structure, the smeared M5 worked as a mask to block the RIE impact on the IMD4. After the second partial RIE was applied onto the sample, the IMD5 in the areas of exposed M5 and un-exposed M5 was etched down until the same level as that in the smeared M5 area (Figure 13d). The cross-sectional view is illustrated in Figure 10g.

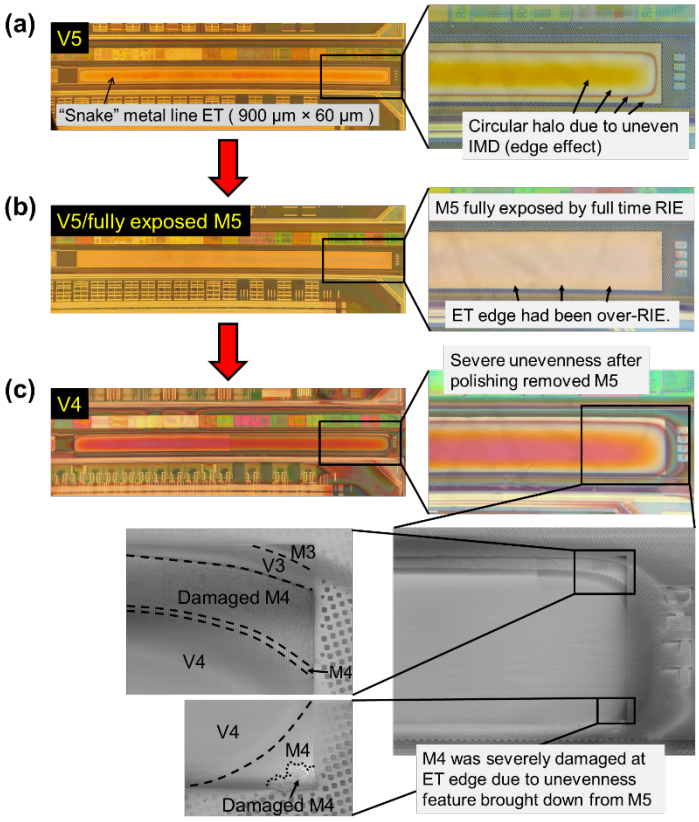


Figure 11: Failed case of delayering ultra-large "snake" metal line ET where (a) the uneven V5/M5 was performed with (b) full time RIE which induced over-exposure of M5 and in turn (c) over-polish on M4.

Finally, after removing M5 by normal slurry polishing, the sample achieved great evenness at V4 across the whole ET structure without any damage to M4. PVC inspection in SEM was then successfully performed and the bright signal was traced along the single ET metal line to quickly locate the defect site where the PVC stopped (Figure 13e). TEM (across the metal lines) analysis on the defect location showed abnormal M4 pattern which was responsible for the high resistance failure, and lithography issue in the wafer process was suspected (Figure 12). In summary, slurry polishing combined with partial RIE is able to generally remove the unevenness induced by edge effect within large areas.

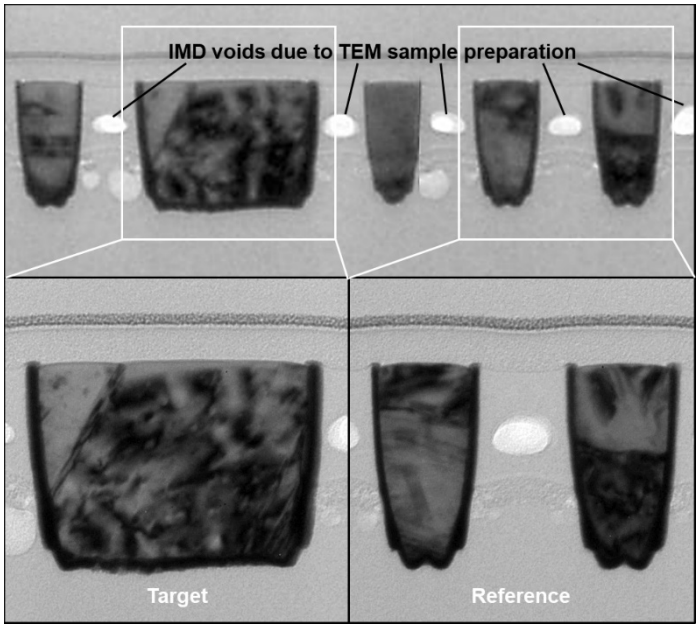


Figure 12: TEM (across metal line) images of the target location and the reference location. Abnormal M4 pattern was observed at the target location.

Table 1: Summary of PFA techniques for rescuing damaged samples with cracks, scratches, or unevenness in delayering.

Types	Problems	Solutions
Cracks	Polishing difficulty, Edge effect	Balance polishing using dummy; Level the joined samples using diamond lapping films
Scratches	Over-RIE, Over-polishing	Deposit Pt on the scratched area to block RIE and resist polishing
Unevenness	Edge effect, Limited inspection area	Smear the partial RIE exposed metal, and then even out the IMD by another partial RIE

In this paper, advanced PFA techniques for rescuing damaged samples with cracks, scratches, or unevenness in delayering by finger polishing have been discussed, through three typical FA cases. The first case is on rescuing cracked samples where diamond lapping film and sacrificial dummy were used to cope with the polishing difficulty and the edge effect with the target defect area close to the gaps/crack lines. The second case is about using e-beam Pt deposition to repair the damage in ROI to the scratched samples and restore the sample to its even condition. The third case studied using controlled slurry polishing combined with partial RIE to remove sample unevenness and create ultra-large (> 50000 μm^2) inspection areas. These techniques are very useful in helping the FA engineers to tackle accidents, solve problems and deliver high-quality FA results, especially for handling "one of a kind" devices.

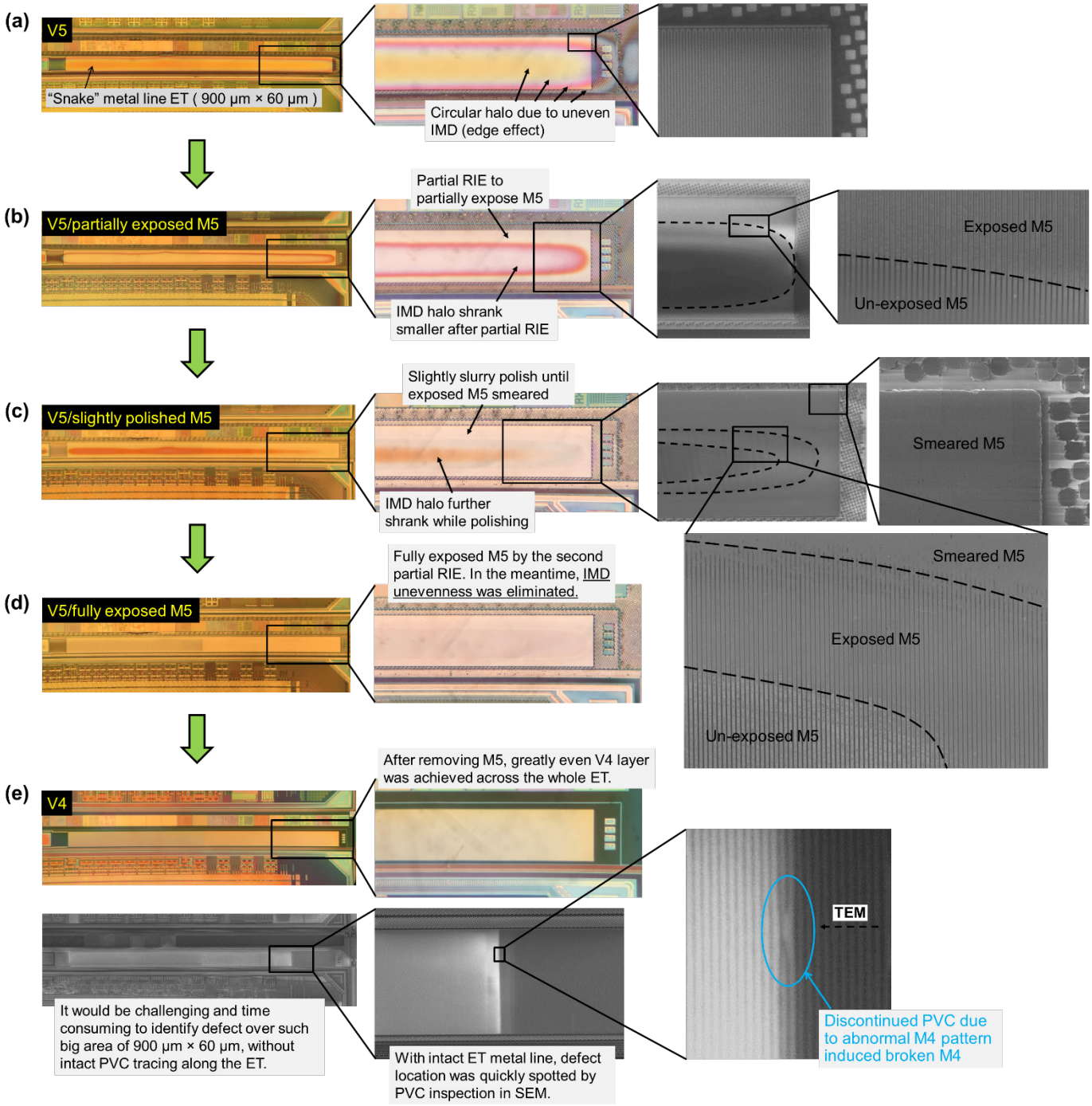


Figure 13: Successful case of sample rescue by using controlled slurry polishing combined with partial RIE to remove sample unevenness and create ultra-large (> 50000 µm^2) inspection area in the "snake" metal line ET. (a) A circular halo due to the uneven IMD at the ET edge was observed at V5. (b) M5 was partially exposed by the first partial RIE, and the sample was (b) slightly slurry polished until (c) the exposed M5 was smeared. (d) The second partial RIE was then performed to get rid of the IMD unevenness. (e) After M5 was removed by slurry polishing, the sample achieved great evenness globally through the whole ET structure. With an intact ET metal line, PVC inspection easily located the defect of broken M4.

Conflict of Interest

The authors declare no conflict of interest.

Abbreviations

DOF	Depth of Field
EFI	Electrical Fault Isolation
ET	Electrical Test
FA	Failure Analysis
FIB	Focused Ion Beam
IMD	Inter-Metal Dielectric
Mx	Metalx, x is layer number
NA	Numerical Aperture
OM	Optical Microscope
PFA	Physical Failure Analysis
RIE	Reactive Ion Etch

ROI	Region of Interest
SEM	Scanning Electron Microscope
SRAM	Static Random Access Memory
TEM	Transmission Electron Microscope
Vx	Viax, x is layer number

References

[1] J. Leo, H. Tan, Y. Z. Ma, S. M. Parab, Y. M. Huang, D. D. Wang, L. Zhu, J. Lam, Z. H. Mai, "Key Issues for Implementing Smart Polishing in Semiconductor Failure Analysis," Journal of Applied Mathematics and Physics, **5**, 1668-1677, 2017, doi:10.4236/jamp.2017.59139.

[2] M. E. Weldy, L. Serrano, "Increasing Planarity for Failure Analysis Using Blocked Reactive Ion Etching Combined With Planar Polish," in 2005 International Symposium for Testing and Failure Analysis (ISTFA), 206-208, 2005, doi:10.31399/asm.cp.istfa2005p0206.

[3] H. Feng, P. K. Tan, H. H. Yap, G. R. Low, R. He, Y. Z. Zhao, B. Liu, M. K. Dawood, J. Zhu, Y. M. Huang, D. D. Wang, H. Tan, J. Lam, Z. H. Mai, "A Sample Preparation Methodology to Reduce Sample Edge Unevenness and Improve Efficiency in Delayering the 20-nm Node IC Chip," in 2015 IEEE International Symposium on Physical and Failure Analysis of Integrated Circuits (IPFA), 2015, doi:10.1109/IPFA.2015.7224432.

[4] K. S. Wills, S. Perungulam, "Delayering Techniques: Dry Processes, Wet Chemical Processing, and Parallel Lapping," Microelectronics Failure Analysis Desk Reference Fifth Edition, 444-463, 2004.

[5] Y. M. Sub, B. T. H. Yap, F. I. Lee, A. B. Minhar, K. W. Tan, H. J. Looi, T. M. Foo, "A Case of Charging Induced Damage into the Common Metal Interconnect During Chemical Mechanical Polishing," in 2017 IEEE International Conference on Opto-Electronic Information Processing (ICOIP), 2017, doi:10.1109/OPTIP.2017.8030704.

[6] K. S. Wills, "Planar Deprocessing of Advanced VLSI Devices," in 2006 International Symposium for Testing and Failure Analysis (ISTFA), 393-397, 2006, doi:10.31399/asm.cp.istfa2006p0393.

[7] S. Bott, R. Rzehak, B. Vasilev, P. Kucher, J. W. Bartha, "A CMP Model Including Global Distribution of Pressure," in 2011 IEEE Transaction on Semiconductor Manufacturing, 304-314, 2011, doi:10.1109/TSM.2011.2107532.

[8] Y. Y. Meng, L. Zhang, Y. B. Li, W. Zhang, H. F. Zhou, J. X. Fang, "Impact of Bevel Condition on STI CMP Scratch," in 2020 China Semiconductor Technology International Conference (SCTIC), 2020, doi:10.1109/CSTIC49141.2020.9282450.

[9] X. Wu, Z. Huang, Y. J. Wan, H. T. Liu, X. Chen, "A Novel Force-Controlled Spherical Polishing Tool Combined With Self-Rotation and Co-Rotation Motion," in 2020 IEEE Access, **8**, 108191-108200, 2020, doi:10.1109/access.2020.2997968.

[10] A. Wieters, P. Thieme, "Wafer Edge / Bevel Treatment of Device Wafers by Means of CMP," in IEEE 2007 International Conference on Planarization / CMP Technology (ICPT), 2007.

[11] T. Moor, E. Malyanker, E. R. Moyal, "Single Die 'Hand-Free' Layer-by-Layer Mechanical Deprocessing for Failure Analysis or Reverse Engineering," in 2008 International Symposium for Testing and Failure Analysis (ISTFA), 2008, doi:10.31399/asm.cp.istfa2008p0363.

[12] H. H. Yap, P. K. Tan, G. R. Low, M. K. Dawood, H. Feng, Y. Z. Zhao, R. He, H. Tan, J. Zhu, B. H. Liu, Y. M. Huang, D. D. Wang, J. Lam, Z. H. Mai, "Top-down Delayering to Expose Large Inspection Area on Die Side-edge with Platinum (Pt) Deposition Technique," Microelectronics Reliability, 55, 1611-1616, 2006, doi:10.1016/j.microrel.2015.06.037.

[13] H. G. Ong, C. L. Gan, "Alternative Lapping Method to Reduce Edge Rounding Effect," in 2012 International Symposium for Testing and Failure Analysis (ISTFA), 462-464, 2012.

[14] K. P. Park, H. S. Kim, O. Chang, H. S. Jeong, "Effects of Pad Properties on Material Removal in Chemical Mechanical Polishing," Journal of Materials Processing Technology, **187-188**, 73-76, 2007, doi:10.1016/j.jmatprotec.2006.11.216.

[15] L. X. Wu, C. F. Yan, "Effects of Polishing Parameters on Evolution of Different Wafer Patterns During Cu CMP," in 2015 IEEE Transaction on Semiconductor Manufacturing, 106-116, 2015, doi:10.1109/TSM.2014.2387211.

[16] D. Lim, H. Kim, B. Jang, H. Cho, J. Kim, H. Hwang, "A Novel Pad Conditioner and Pad Roughness Effects on Tungsten CMP," in IEEE 2014 International Conference on Planarization / CMP Technology (ICPT), 352-355, 2014, doi:10.1109/ICPT.2014.7017318.

[17] C. Wang, P. Sherman, A. Chandra, D. Dornfeld, "Pad Surface Roughness and Slurry Particle Size Distribution Effects on Material Removal Rate in Chemical Mechanical Planarization," CIRP Annals-Manufacturing Technology, **54**, 309-312, 2005, doi:10.1016/S0007-8506(07)60110-3.

[18] T. F. Zeng, T. Sun, "Size Effect of Nanoparticles in Chemical Mechanical Polishing - A Transient Model," in 2005 IEEE Transaction on Semiconductor Manufacturing, 655-663, 2005, doi:10.1109/TSM.2005.858508.

[19] C. J. Evans, E. Paul, D. Dornfeld, D. A. Lucca, G. Byrne, M. Tricard, F. Klocke, O. Dambon, B. A. Mullany, "Material Removal Mechanisms in Lapping and Polishing," CIRP Annals, **52**, 611-633, 2003, doi:10.1016/S0007-8506(07)60207-8.

[20] L. Shan, C. H. Zhou, S. Danyluk, "Mechanical Interactions and Their Effects on Chemical Mechanical Polishing," in 2001 IEEE Transaction on Semiconductor Manufacturing, 207-213, 2001, doi:10.1109/66.939815.

[21] "Microscopy Basics: Depth of Field and Depth of Focus", Nikon, https://www.microscopyu.com/microscopy-basics/depth-of-field-and-depth-of-focus.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 5, 53-58 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Model Reduction H_∞ Finite Frequency of Takagi-Sugeno Fuzzy Systems

Rim Mrani Alaoui[1], Abderrahim El-Amrani[2,*], Ismail Boumhidi[2]

[1]*Higher Institute of Engineering and Business, ISGA, Fez, 30000, Morroco*

[2]*LISAC Labs, Sidi Mohamed Ben Abdellah University, Fez, 30000, Morocco*

ARTICLE INFO

Article history:
Received: 14 June, 2021
Accepted: 01 September, 2021
Online: 16 September, 2021

Keywords:
nonlinear models
MR
FF
LMIs
fuzzy systems

ABSTRACT

The daily treats model reduction finite frequency (FFMR) design for Takagi Sugeno (T S) systems. This work is to FFMR design in such a way whether augmented model is steady get a reduced H_∞ index in FF areas with noise is established as a prerequisite. To highlight the importance of suggested process, a practical application has been made.

1 Introduction

Fuzzy models T-S [1] call for wide observation of various practical industrial applications, mainly as the recognized T-S samples actually approximate nonlinear shapes. The essential characteristic on the sample TS is its general estimation of a nonlinear function. There are large number of results of literature that treat the globally difficulties utilizing the TS fuzzy samples, see [2]–[7].

The our existing sources on model reduction problem and disturbances are based on the whole full frequency (EF) area, which will give several types of model reduction design [8]–[13]. However, most practical industrial applications work in a FF domain. So far, a few applications have been made [14]–[19]. Thus, for this we will present new approaches to solve these problems.

The primary goal of our work is to define a fuzzy model reduction of discrete Model over FF ranges such a way that the augmented model is steady get a reduced H_∞ index in FF areas with disturbance is established as a prerequisite. we have also presented an example of simulation in order to exemplify the efficiency of the suggested method.

Notations :

- "T" : Matrix transposition
- " $*$ " : Matrix symmetry
- $M > 0$: matrix M is positive
- $\mathbf{sym}[\mathcal{D}] : \mathcal{D} + \mathcal{D}^*$
- $\mathbf{He}[\mathcal{D}] : \frac{\mathcal{D}+\mathcal{D}^*}{2}$

2 Problem statement

2.1 System formulation

Envisage the nonlinear model presented by :

Rule *x*: IF $\zeta_1(\mu)$ is T_1^i,... $\zeta_n(\mu)$ is $\mathbf{T}_n^i$ THEN

$$\begin{aligned} \mathbf{z}_{\mu+1} &= \mathbf{M}_x\mathbf{z}(\mu) + \mathbf{N}_x\mathbf{d}(\mu) \\ \mathbf{w}(\mu) &= \mathbf{J}_x\mathbf{z}(\mu) \end{aligned} \quad (1)$$

$$(2)$$

with $\mathbf{z}(\mu)$ is the input; $\mathbf{w}(\mu)$ is deliberate output; $\mathbf{M}_x$, $\mathbf{N}_x$ and $\mathbf{J}_x$ are system parametres; $\zeta_1(\mu)$,..., $\zeta_n(\mu)$ the premise variables. $\mathbf{d}(\mu)$ is a known disturbance signal located in a following FF areas

$$\varrho = \left\{ \tilde{\varrho} \in \mathbb{R} \;\; \tilde{\varrho}_1 \leq \tilde{\varrho} \leq \tilde{\varrho}_2; \quad \tilde{\varrho}_1, \tilde{\varrho}_2 \in [-\pi, +\pi], \right. \quad (3)$$

We describe the nonlinear system (1) employ singleton fuzzifer, center-average and inference product by the following relation :

$$\begin{aligned} \mathbf{z}_{\mu+1} &= \mathbf{M}(\mathbf{l})x(\mu) + \mathbf{N}(\mathbf{l})\mathbf{d}(\mu) \\ \mathbf{w}(\mu) &= \mathbf{J}(\mathbf{l})\mathbf{z}(\mu) \end{aligned} \quad (4)$$

*A. EL-AMRANI & Email : abderrahim.elamrani@usmba.ac.ma

where

$$\begin{aligned} \mathbf{M}(\mathbf{l}) &= \sum_{x=1}^{q} \mathbf{l}_x \mathbf{M}_x; \quad \mathbf{N}(\mathbf{l}) = \sum_{x=1}^{q} \mathbf{l}_x \mathbf{N}_x; \\ \mathbf{J}(\mathbf{l}) &= \sum_{x=1}^{q} \mathbf{l}_x \mathbf{J}_x \end{aligned} \tag{5}$$

In this work, a fuzzy MR function is represented by :

Rule x: if $\zeta_1(\mu)$ is $\mathbf{T}_1^x$... $\zeta_n(\mu)$ is $\mathbf{T}_n^x$ then

$$\begin{aligned} \hat{\mathbf{z}}(\mu+1) &= \hat{\mathbf{M}}_x \hat{\mathbf{z}}(\mu) + \hat{\mathbf{N}}_x \mathbf{d}(\mu) \\ \hat{\mathbf{w}}(\mu) &= \hat{\mathbf{J}}_x \hat{\mathbf{z}}(\mu) \end{aligned} \tag{6}$$

with $\hat{\mathbf{z}}(\mu)$ is state MR vector; $\hat{\mathbf{w}}(\mu)$ is the output MR function, $\hat{\mathbf{M}}_x$, $\hat{\mathbf{N}}_x$ and $\hat{\mathbf{J}}_x$ are parameters should exist defined.

We get defuzzified for system (6) as following :

$$\begin{aligned} \hat{\mathbf{z}}(\mu+1) &= \hat{\mathbf{M}}(\hat{\mathbf{l}}) \hat{\mathbf{z}}(\mu) + \hat{B}(\hat{\mathbf{l}}) \mathbf{d}(\mu) \\ \hat{\mathbf{w}}(\mu) &= \hat{\mathbf{J}}(\hat{\mathbf{l}}) \hat{\mathbf{z}}(k) \end{aligned} \tag{7}$$

with

$$\begin{aligned} \hat{\mathbf{M}}(\hat{\mathbf{l}}) &= \sum_{x=1}^{q} \hat{\mathbf{l}}_x \hat{\mathbf{M}}_x, \quad \hat{B}(\hat{\mathbf{l}}) = \sum_{x=1}^{q} \hat{\mathbf{l}}_x \hat{\mathbf{N}}_x, \\ \hat{\mathbf{J}}(\hat{\mathbf{l}}) &= \sum_{x=1}^{q} \hat{\mathbf{l}}_x \hat{\mathbf{J}}_x \end{aligned} \tag{8}$$

Consider $\nabla(\mu) = \begin{bmatrix} \mathbf{z}(\mu) \\ \hat{\mathbf{z}}(\mu) \end{bmatrix}$, $\mathbf{g}(\mu) = \mathbf{w}(\mu) - \hat{\mathbf{w}}(\mu)$. Then, the error system is as follows:

$$\begin{aligned} \nabla(\mu+1) &= \bar{\mathbf{M}}(\mathbf{l}, \hat{\mathbf{l}}) \nabla(\mu) + \bar{\mathbf{N}}(\mathbf{l}, \hat{\mathbf{l}}) \mathbf{d}(\mu) \\ \mathbf{g}(\mu) &= \bar{\mathbf{J}}(\mathbf{l}, \hat{\mathbf{l}}) \nabla(\mu) \end{aligned} \tag{9}$$

where

$$\begin{aligned} \bar{\mathbf{M}}(\mathbf{l}, \hat{\mathbf{l}}) &= \begin{bmatrix} \mathbf{M}(\mathbf{l}) & 0 \\ 0 & \hat{\mathbf{M}}(\hat{\mathbf{l}}) \end{bmatrix}; \quad \bar{\mathbf{N}}(\mathbf{l}, \hat{\mathbf{l}}) = \begin{bmatrix} \mathbf{N}(\mathbf{l}) \\ \hat{\mathbf{N}}(\hat{\mathbf{l}}) \end{bmatrix}; \\ \bar{\mathbf{J}}(\mathbf{l}, \hat{\mathbf{l}}) &= \begin{bmatrix} \mathbf{J}(\mathbf{l}) & -\hat{\mathbf{J}}(\hat{\mathbf{l}}) \end{bmatrix} \end{aligned} \tag{10}$$

Definition 2.1 *[20] Consider $\beta > 0$, if the following inequality is verified:*

$$\sum_{\mu=0}^{\infty} \boldsymbol{g}^T(\mu) \boldsymbol{g}(\mu) \leq \beta^2 \sum_{\mu=0}^{\infty} \boldsymbol{d}^T(\mu) \boldsymbol{d}(\mu) \tag{11}$$

(12)

From Parseval function [21], we get :

$$\int_{-\pi}^{\pi} \boldsymbol{g}^T(\varrho) \boldsymbol{g}(\varrho) d\varrho \leq \beta^2 \int_{-\pi}^{\pi} \boldsymbol{d}^T(\varrho) \boldsymbol{d}(\varrho) d\varrho \tag{13}$$

(14)

We express the question of this work by : we design an appropriate fuzzy MR system (7) such that a error model is well-posed, stable satises the FF index:

$$\int_{\varrho \in \varrho} \mathbf{g}^T(\varrho) \mathbf{g}(\varrho) d\varrho \leq \beta^2 \int_{\varrho \in \varrho} \mathbf{d}^T(\varrho) \mathbf{d}(\varrho) d\varrho \tag{15}$$

2.2 Preliminaries

Lemma 2.2 *[22] Let $\Phi \in \mathbb{R}^m$, $\mathcal{L} \in \mathbb{R}^{m \times m}$ and $\mathcal{S} \in \mathbb{R}^{p \times m}$. Then, The following equations are the same:*

- $\Phi^* \mathcal{L} \Phi < 0, \forall \Phi \neq 0 : \mathcal{S}\Phi = 0$
- $\mathcal{S}^{\perp T} \mathcal{L} \mathcal{S}^{\perp} < 0$
- $\exists \theta \in \mathbb{R} : \mathcal{L} - \theta \mathcal{S}^T \mathcal{S} < 0$
- $\exists \mathcal{D} \in \mathbb{R}^{m \times p} : \mathcal{L} + \mathcal{D}\mathcal{S} + \mathcal{S}^T \mathcal{D}^T < 0$

Lemma 2.3 *Error system (10) is stable and FF in (15) is fulfilled, on condition that there are $\mathcal{B}$, $0 < C$, satisfying*

$$\begin{aligned} &\begin{bmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) & \bar{\boldsymbol{N}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) \\ I & 0 \end{bmatrix}^T \begin{bmatrix} \mathcal{B} & e^{j\tilde{\varrho}_3} C \\ e^{-j\tilde{\varrho}_3} C & -\mathcal{B} - 2cos(\tilde{\varrho}_4) C \end{bmatrix} \\ &\begin{bmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) & \bar{\boldsymbol{N}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) \\ I & 0 \end{bmatrix} + \begin{bmatrix} \bar{\boldsymbol{J}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \bar{\boldsymbol{J}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) & 0 \\ 0 & -\beta^2 I \end{bmatrix} < 0 \end{aligned} \tag{16}$$

with $\tilde{\varrho}_3 = \frac{\tilde{\varrho}_2 + \tilde{\varrho}_1}{2}$, $\tilde{\varrho}_4 = \frac{\tilde{\varrho}_2 - \tilde{\varrho}_1}{2}$.

Proof 2.4 *The substantiation of Lemma 2.3 is defined in appendix.*

3 FF performance analysis

Theorem 3.1 *Error model (10) is stable, FF index (15) is fulfilled, on condition that there are $\mathcal{X}$, $\mathcal{B}$, C, $\mathcal{E}$, $\mathcal{Z}$, $\mathcal{T}$ satisfying $C > 0$, $\mathcal{X} > 0$ and*

$$\begin{bmatrix} \urcorner_1 & \urcorner_2 & \mathcal{Z}\bar{\boldsymbol{N}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) - \mathcal{T}^T & 0 \\ * & \urcorner_3 & \urcorner_4 & \bar{\boldsymbol{J}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \\ * & * & \urcorner_5 & 0 \\ * & * & * & -I \end{bmatrix} < 0 \tag{17}$$

$$\begin{bmatrix} \mathcal{X} - \mathcal{Z} - \mathcal{Z}^T & \mathcal{Z}\boldsymbol{M}(\boldsymbol{l}, \hat{\boldsymbol{l}}) - \mathcal{E}^T \\ * & \Psi_1 \end{bmatrix} < 0 \tag{18}$$

with

$$\begin{aligned} \urcorner_1 &= \mathcal{B} - \mathcal{Z} - \mathcal{Z}^T; \\ \urcorner_2 &= e^{j\tilde{\varrho}_3} C - \mathcal{E}^T + \mathcal{Z}\bar{A}(\boldsymbol{l}, \hat{\boldsymbol{l}}); \\ \urcorner_3 &= -\mathcal{B} - 2cos(\tilde{\varrho}_4) C + \mathcal{E}\bar{\boldsymbol{M}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) + \bar{\boldsymbol{M}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \mathcal{E}^T; \\ \urcorner_4 &= \mathcal{E}\bar{\boldsymbol{N}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) + \bar{\boldsymbol{M}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \boldsymbol{l}^T(\boldsymbol{l}) \\ \urcorner_5 &= -\beta^2 I + \mathcal{T}\bar{\boldsymbol{N}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) + \bar{\boldsymbol{N}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \boldsymbol{l}^T; \\ \triangle_1 &= -\mathcal{X} + \mathcal{E}\bar{\boldsymbol{M}}(\boldsymbol{l}, \hat{\boldsymbol{l}}) + \bar{\boldsymbol{M}}^T(\boldsymbol{l}, \hat{\boldsymbol{l}}) \mathcal{E}^T. \end{aligned}$$

Proof 3.2 *First, Condition (17) maybe written as*

$$\begin{pmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & \bar{\boldsymbol{N}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \\ I & 0 \\ 0 & I \end{pmatrix}^T [\begin{pmatrix} 0 & \bar{\boldsymbol{J}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & 0 \\ 0 & 0 & I \end{pmatrix}^T$$
$$\begin{pmatrix} I & 0 \\ 0 & -\beta^2 I \end{pmatrix}\begin{pmatrix} 0 & \bar{\boldsymbol{J}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & 0 \\ 0 & 0 & I \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix}$$
$$\begin{pmatrix} \mathcal{B} & e^{j\varrho_c}C \\ e^{-j\tilde{\varrho}_3}C & -\mathcal{B}-2cos(\tilde{\varrho}_4)C \end{pmatrix}\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}]$$
$$\begin{pmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & \bar{\boldsymbol{N}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \\ I & 0 \\ 0 & I \end{pmatrix} < 0 \tag{19}$$

Denote

$$\begin{aligned} \mathcal{D} &= \begin{bmatrix} \mathcal{E} \\ \mathcal{Z} \\ \mathcal{T} \end{bmatrix}; \quad \theta = \begin{bmatrix} \nabla(\mu+1) \\ \nabla(\mu) \\ \boldsymbol{d}(\mu) \end{bmatrix}; \\ \mathcal{S} &= \begin{bmatrix} -I & \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & \bar{\boldsymbol{N}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \end{bmatrix} \end{aligned} \tag{20}$$

With the help of Lemma 2.2, we have

$$\mathcal{D}\mathcal{S} + \mathcal{S}^T\mathcal{D}^T + \begin{bmatrix} 0 & \bar{\boldsymbol{J}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & 0 \\ 0 & 0 & I \end{bmatrix}^T$$
$$\begin{bmatrix} I & 0 \\ 0 & -\beta^2 I \end{bmatrix}\begin{bmatrix} 0 & \bar{\boldsymbol{J}}(\boldsymbol{l},\hat{\boldsymbol{l}}) & 0 \\ 0 & 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} \mathcal{B} & e^{j\tilde{\varrho}_3}C \\ e^{-j\tilde{\varrho}_3}C & -\mathcal{B}-2cos(\tilde{\varrho}_4)C \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} < 0 \tag{21}$$

With the help of Lemma 2.2, given (16).

Consider the Lyapunov equation :

$$\begin{bmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \\ I \end{bmatrix}^T \begin{bmatrix} \mathcal{X} & 0 \\ 0 & -\mathcal{X} \end{bmatrix}\begin{bmatrix} \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \\ I \end{bmatrix} < 0 \tag{22}$$

Let

$$\begin{aligned} \theta &= \begin{bmatrix} \nabla(\mu+1) \\ \nabla(\mu) \end{bmatrix}; \quad \mathcal{L} = \begin{bmatrix} \mathcal{X} & 0 \\ 0 & -\mathcal{X} \end{bmatrix}; \\ \mathcal{S} &= \begin{bmatrix} -I & \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \end{bmatrix}; \quad \mathcal{D} = \begin{bmatrix} \mathcal{E}^T & \mathcal{Z}^T \end{bmatrix}^T \end{aligned} \tag{23}$$

Using the conditions in 2.2, then

$$\begin{bmatrix} \mathcal{X} & 0 \\ 0 & -\mathcal{X} \end{bmatrix} + \begin{bmatrix} \mathcal{E} \\ \mathcal{Z} \end{bmatrix}\begin{bmatrix} -I & \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \end{bmatrix} +$$
$$\begin{bmatrix} -I & \bar{\boldsymbol{M}}(\boldsymbol{l},\hat{\boldsymbol{l}}) \end{bmatrix}^T \begin{bmatrix} \mathcal{E} \\ \mathcal{Z} \end{bmatrix}^T < 0 \tag{24}$$

which is nothing but (18).

4 FF performance Design

Theorem 4.1 *Error model (10) is stable, FF index (15) is fulfilled, on condition that there are* $\mathcal{B} = \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ * & \mathcal{B}_3 \end{bmatrix}$, $C = \begin{bmatrix} C_1 & C_2 \\ * & C_3 \end{bmatrix} >$ 0, $\mathcal{X} = \begin{bmatrix} \mathcal{X}_1 & \mathcal{X}_2 \\ * & \mathcal{X}_3 \end{bmatrix} > 0$, $\check{\boldsymbol{M}}(\hat{\boldsymbol{l}})$, $\check{\boldsymbol{N}}(\hat{\boldsymbol{l}})$, $\check{\boldsymbol{J}}(\hat{\boldsymbol{l}})$, $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{Z}_1$, $\mathcal{Z}_2$, $\mathcal{T}_1$, $\mathcal{V}$, *satisfying*

$$\begin{bmatrix} \bar{\Delta}_{11} & \bar{\Delta}_{12} & \bar{\Delta}_{13} & -\mathcal{Z}_2^T + E\hat{\boldsymbol{M}}(\hat{\boldsymbol{l}}) \\ * & \bar{\Delta}_{22} & \bar{\Delta}_{23} & \hat{\boldsymbol{M}}(\hat{\boldsymbol{l}}) \\ * & * & \bar{\Delta}_{33} & -\mathcal{X}_2 + \boldsymbol{M}^T(\boldsymbol{l})\mathcal{Z}_2^T \\ * & * & * & -\mathcal{X}_3 \end{bmatrix} < 0 \tag{25}$$

$$\begin{bmatrix} \bar{\neg}_{11} & \bar{\neg}_{12} & \bar{\neg}_{13} & \bar{\neg}_{14} & \bar{\neg}_{15} & 0 \\ * & \bar{\neg}_{22} & \bar{\neg}_{23} & \bar{\neg}_{24} & \bar{\neg}_{25} & 0 \\ * & * & \bar{\neg}_{33} & \bar{\neg}_{34} & \bar{\neg}_{35} & \boldsymbol{J}^T(\boldsymbol{l}) \\ * & * & * & \bar{\neg}_{44} & \bar{\neg}_{45} & -\hat{\boldsymbol{J}}^T(\hat{\boldsymbol{l}}) \\ * & * & * & * & \bar{\neg}_{55} & 0 \\ * & * & * & * & * & -I \end{bmatrix} < 0 \tag{26}$$

with

$$\begin{aligned} \bar{\Delta}_{11} &= \mathcal{B}_1 - \mathcal{E}_1 - \mathcal{E}_1^T;\ \bar{\Delta}_{12} = \mathcal{X}_2 - E\mathcal{V} - \mathcal{E}_2^T; \\ \bar{\Delta}_{13} &= \mathcal{E}_1\boldsymbol{A}(h) - \mathcal{Z}_1^T;\ \bar{\Delta}_{22} = \mathcal{X}_3 - \mathcal{V} - \mathcal{V}^T; \\ \bar{\Delta}_{23} &= \mathcal{E}_2\boldsymbol{M}(\boldsymbol{l});\ \bar{\Delta}_{33} = -\mathcal{X}_1 + \mathcal{Z}_1\boldsymbol{M}(\boldsymbol{l}) + \boldsymbol{M}^T(\boldsymbol{l})\mathcal{Z}_1^T; \end{aligned}$$

$$\begin{aligned} \bar{\neg}_{11} &= \mathcal{B}_1 - \mathcal{E}_1 - \mathcal{E}_1^T;\ \bar{\neg}_{12} = \mathcal{B}_2 - E\mathcal{V} - \mathcal{E}_2^T; \\ \bar{\neg}_{13} &= e^{j\tilde{\varrho}_3}C_1 - \mathcal{Z}_1^T + \mathcal{E}_1 A(\boldsymbol{l});\ \bar{\neg}_{22} = \mathcal{B}_3 - \mathcal{V} - \mathcal{V}^T; \\ \bar{\neg}_{14} &= e^{j\tilde{\varrho}_3}C_2 - \mathcal{Z}_2^T + E\tilde{\mathcal{M}}(\hat{\boldsymbol{l}}); \\ \bar{\neg}_{24} &= e^{j\tilde{\varrho}_3}C_3 + \tilde{\mathcal{M}}(\hat{\boldsymbol{l}}); \\ \bar{\neg}_{15} &= \mathcal{E}_1\boldsymbol{N}(\boldsymbol{l}) + E\tilde{\boldsymbol{N}}(\hat{\boldsymbol{l}}) - \mathcal{T}_1^T; \\ \bar{\neg}_{25} &= \mathcal{Z}_2\boldsymbol{N}(\boldsymbol{l}) + \tilde{\boldsymbol{M}}(\hat{\boldsymbol{l}}); \\ \bar{\neg}_{33} &= -\mathcal{B}_1 - 2cos(\tilde{\varrho}_4)C_1 + \mathcal{Z}_1\boldsymbol{M}(\boldsymbol{l}) + \boldsymbol{M}^T(\boldsymbol{l})\mathcal{Z}_1^T; \\ \bar{\neg}_{34} &= -\mathcal{B}_2 - 2cos(\tilde{\varrho}_4)C_2 + \boldsymbol{M}^T(\boldsymbol{l})\mathcal{Z}_2^T; \\ \bar{\neg}_{35} &= \mathcal{Z}_1\boldsymbol{N}(\boldsymbol{l}) + \boldsymbol{M}^T(\boldsymbol{l})\mathcal{T}_1^T; \\ \bar{\neg}_{23} &= e^{-j\tilde{\varrho}_3}C_2 + \mathcal{E}_2\boldsymbol{M}(\boldsymbol{l}); \\ \bar{\neg}_{44} &= -\mathcal{B}_3 - 2cos(\tilde{\varrho}_4)C_3;\ \bar{\neg}_{45} = \mathcal{Z}_2\boldsymbol{N}(\boldsymbol{l}); \\ \bar{\neg}_{55} &= -\beta^2 I + sym[\mathcal{T}_1\boldsymbol{N}(\boldsymbol{l})] \end{aligned}$$

The following parameters as

$$\begin{aligned} \hat{\boldsymbol{M}}(\hat{\boldsymbol{l}}) &= \mathcal{V}^{-1}\check{\boldsymbol{M}}(\hat{h});; \\ \hat{\boldsymbol{N}}(\hat{\boldsymbol{l}}) &= \mathcal{V}^{-1}\check{\boldsymbol{N}}(\hat{h}); \\ \hat{\boldsymbol{J}}(\hat{\boldsymbol{l}}) &= \check{\boldsymbol{J}}(\hat{\boldsymbol{l}}) \end{aligned} \tag{27}$$

Proof 4.2 *Parameterise slack matrices* $\mathcal{E}$, $\mathcal{Z}$, $\mathcal{T}$ *in Theorem 3.1 as*

$$\begin{aligned} \mathcal{E} &= \begin{bmatrix} \mathcal{E}_1 & E\mathcal{V} \\ \mathcal{E}_2 & \mathcal{V} \end{bmatrix}; \quad \mathcal{Z} = \begin{bmatrix} \mathcal{Z}_1 & 0 \\ \mathcal{Z}_2 & 0 \end{bmatrix}; \\ \mathcal{T} &= \begin{bmatrix} \mathcal{T}_1 & 0 \end{bmatrix} \end{aligned} \tag{28}$$

where $E = \begin{pmatrix} I \\ 0 \end{pmatrix}$.

Moreover, the Theorem 4.3 is to solve the FFMR problem in the FF index (15).

Theorem 4.3 *Error model (10) is stable, FF index (15) is fulfilled, on condition that there are* $\mathcal{B} = \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ * & \mathcal{B}_3 \end{bmatrix}$, $C = \begin{bmatrix} C_1 & C_2 \\ * & C_3 \end{bmatrix} > 0$, $\mathcal{X} = \begin{bmatrix} \mathcal{X}_1 & \mathcal{X}_2 \\ * & \mathcal{X}_3 \end{bmatrix} > 0$, $\breve{\boldsymbol{M}}_x$, $\breve{\boldsymbol{N}}_x$, $\breve{\boldsymbol{J}}_x$, $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{Z}_1$, $\mathcal{Z}_2$, $\mathcal{T}_1$, $\mathcal{V}$, *satisfying, with* $x, t \in \{1, 2, ..., m\}$

$$\begin{bmatrix} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} & \tilde{\Delta}_{13} & -\mathcal{Z}_2^T + E\breve{\mathcal{A}}_x \\ * & \tilde{\Delta}_{22} & \tilde{\Delta}_{23} & \breve{\mathcal{A}}_x \\ * & * & \tilde{\Delta}_{33} & -\mathcal{X}_2 + A_t^T\mathcal{Z}_2^T \\ * & * & * & -\mathcal{X}_3 \end{bmatrix} < 0 \tag{29}$$

$$\begin{bmatrix} \tilde{\daleth}_{11} & \tilde{\daleth}_{12} & \tilde{\daleth}_{13} & \tilde{\daleth}_{14} & \tilde{\daleth}_{15} & 0 \\ * & \tilde{\daleth}_{22} & \tilde{\daleth}_{23} & \tilde{\daleth}_{24} & \tilde{\daleth}_{25} & 0 \\ * & * & \tilde{\daleth}_{33} & \tilde{\daleth}_{34} & \tilde{\daleth}_{35} & \boldsymbol{J}_t^T \\ * & * & * & \tilde{\daleth}_{44} & \tilde{\daleth}_{45} & -\breve{\boldsymbol{J}}_x^T \\ * & * & * & * & \tilde{\daleth}_{55} & 0 \\ * & * & * & * & * & -I \end{bmatrix} < 0 \tag{30}$$

with

$$\begin{aligned}
\tilde{\Delta}_{11} &= \mathcal{B}_1 - \mathcal{E}_1 - \mathcal{E}_1^T;\ \tilde{\Delta}_{12} = \mathcal{X}_2 - E\mathcal{V} - \mathcal{E}_2^T; \\
\tilde{\Delta}_{13} &= \mathcal{E}_1\boldsymbol{M}_t - \mathcal{Z}_1^T; \\
\tilde{\Delta}_{22} &= \mathcal{X}_3 - \mathcal{V} - \mathcal{V}^T; \\
\tilde{\Delta}_{23} &= \mathcal{E}_2\boldsymbol{M}_t;\ \tilde{\Delta}_{33} = -\mathcal{X}_1 + \mathcal{Z}_1\boldsymbol{M}_t + \boldsymbol{M}_t^T\mathcal{Z}_1^T; \\
\tilde{\daleth}_{11} &= \mathcal{B}_1 - \mathcal{E}_1 - \mathcal{E}_1^T;\ \tilde{\daleth}_{12} = \mathcal{B}_2 - E\mathcal{V} - \mathcal{E}_2^T; \\
\tilde{\daleth}_{13} &= e^{j\tilde{\varrho}_3}C_1 - \mathcal{Z}_1^T + \mathcal{E}_1\boldsymbol{M}_t;\ \tilde{\daleth}_{22} = \mathcal{B}_3 - \mathcal{V} - \mathcal{V}^T; \\
\tilde{\daleth}_{14} &= e^{j\tilde{\varrho}_3}C_2 - \mathcal{Z}_2^T + E\breve{\boldsymbol{M}}_x; \\
\tilde{\daleth}_{24} &= e^{j\tilde{\varrho}_3}C_3 + \breve{\boldsymbol{M}}_x; \\
\tilde{\daleth}_{15} &= \mathcal{E}_1\boldsymbol{N}_t + E\breve{\boldsymbol{N}}_x - \mathcal{T}_1^T;\ \tilde{\daleth}_{25} = \mathcal{Z}_2\boldsymbol{N}_t + \breve{\boldsymbol{M}}_x; \\
\tilde{\daleth}_{33} &= -\mathcal{B}_1 - 2cos(\tilde{\varrho}_4)C_1 + \mathcal{Z}_1\boldsymbol{M}_t + \boldsymbol{M}_t^T\mathcal{Z}_1^T; \\
\tilde{\daleth}_{34} &= -\mathcal{B}_2 - 2cos(\tilde{\varrho}_4)C_2 + \boldsymbol{M}_t^T\mathcal{Z}_2^T; \\
\tilde{\daleth}_{35} &= \mathcal{Z}_1\boldsymbol{N}_t + \boldsymbol{M}_t^T\mathcal{T}_1^T;\ \tilde{\daleth}_{23} = e^{-j\tilde{\varrho}_3}C_2 + \mathcal{E}_2\boldsymbol{M}_t; \\
\tilde{\daleth}_{44} &= -\mathcal{B}_3 - 2cos(\tilde{\varrho}_4)C_3;\ \tilde{\daleth}_{45} = \mathcal{Z}_2\boldsymbol{N}_t; \\
\tilde{\daleth}_{55} &= -\beta^2 I + sym[\mathcal{T}_1\boldsymbol{N}_t].
\end{aligned}$$

The following reduced-order parameters as :

$$\begin{aligned}
\hat{\mathbf{M}}_t &= \mathcal{V}^{-1}\breve{\mathbf{M}}_x \\
\hat{\mathbf{N}}_t &= \mathcal{V}^{-1}\breve{\mathbf{N}}_x; \\
\hat{\mathbf{J}}_t &= \breve{\mathbf{J}}_x
\end{aligned} \tag{31}$$

Proof 4.4 *We propose the following equations :*

$$\begin{aligned}
&\sum_{x=1}^{q}\sum_{t=1}^{m}\hat{\boldsymbol{l}}_x\boldsymbol{l}_t\tilde{\Delta}_{xt} \\
&\sum_{x=1}^{q}\sum_{t=1}^{m}\hat{\boldsymbol{l}}_x\boldsymbol{l}_t\tilde{\daleth}_{xt};
\end{aligned} \tag{32}$$

Therefore, us find Theorem 4.1.

Remark 4.5 *If us pick* $C = 0$, *we mastery employ theorem 4.3 to settle the* H_∞ *MR nonlinear systems over EF range.*

5 Numerical Example

Consider a fuzzy system in discrete time, represents in [12] :

$$\begin{aligned}
\mathbf{J}_1 &= \begin{bmatrix} 1.4419 & 0.6720 & 0.1387 & -0.8595 \end{bmatrix} \\
\mathbf{J}_2 &= \begin{bmatrix} 1.3329 & 0.6720 & 0.1387 & -0.8478 \end{bmatrix}; \\
\mathbf{M}_1 &= \begin{bmatrix} 0.1612 & 0.0574 & -0.0144 & 0.1846 \\ 0.0434 & -0.3638 & 0.5258 & -0.0357 \\ -0.0747 & -0.3146 & -0.0487 & -0.1043 \\ -0.1664 & 0.4031 & 0.0347 & 0.2864 \end{bmatrix}; \\
\mathbf{M}_2 &= \begin{bmatrix} 0.1312 & 0.0474 & -0.0044 & 0.1546 \\ 0.0234 & -0.3018 & 0.4258 & -0.0357 \\ -0.0554 & -0.2421 & -0.0367 & -0.0843 \\ -0.1551 & 0.3031 & 0.0247 & 0.1864 \end{bmatrix}; \\
\mathbf{N}_1 &= \begin{bmatrix} 0.2023 \\ -0.2313 \\ -0.1137 \\ 0.1279 \end{bmatrix};\ \mathbf{N}_2 = \begin{bmatrix} 0.0123 \\ -0.1313 \\ -0.1138 \\ 0.1179 \end{bmatrix}; \\
E &= \begin{bmatrix} I & 0 \end{bmatrix}.
\end{aligned} \tag{33}$$

The normalized membership function :

$$\begin{aligned}
\mathbf{T}_1(\mathbf{z}_1(\mu)) &= \frac{1 - \sin(\mathbf{z}_1(\mu))}{2} \\
\mathbf{T}_2(\mathbf{z}_1(\mu)) &= 1 - \mathbf{T}_1(\mathbf{z}_1(\mu))
\end{aligned} \tag{34}$$

Via using Theorem 4, the obtained matrix parameters of FF H_∞ reduced order systems are the following :

- EF area : $(-\pi \le \tilde{\varrho} \le \pi)$

$$\begin{aligned}
\left[\begin{array}{c|c} \hat{\mathbf{M}}_1 & \hat{\mathbf{N}}_1 \\ \hline \hat{\mathbf{J}}_1 & - \end{array}\right] &= \left[\begin{array}{cc|c} -0.7415 & -0.7001 & -0.1925 \\ 0.7145 & 0.1925 & 0.4512 \\ \hline -0.8874 & -1.2145 & - \end{array}\right]; \\
\left[\begin{array}{c|c} \hat{\mathbf{M}}_2 & \hat{\mathbf{N}}_2 \\ \hline \hat{\mathbf{J}}_2 & - \end{array}\right] &= \left[\begin{array}{cc|c} -0.8840 & 0.7415 & 0.1325 \\ -0.6245 & 0.2415 & -0.4125 \\ \hline -1.2021 & -1.1325 & - \end{array}\right]
\end{aligned} \tag{35}$$

- FF area : $(\frac{\pi}{7} \le \tilde{\varrho} \le \frac{\pi}{4})$

$$\left[\begin{array}{c|c} \hat{\mathbf{M}}_1 & \hat{\mathbf{N}}_1 \\ \hline \hat{\mathbf{J}}_1 & - \end{array}\right] = \left[\begin{array}{cc|c} -0.4415 & 0.2145 & -0.1125 \\ -0.1745 & 0.1325 & 0.1954 \\ \hline -0.2314 & 0.7458 & - \end{array}\right];$$

$$\left[\begin{array}{c|c} \hat{\mathbf{M}}_2 & \hat{\mathbf{N}}_2 \\ \hline \hat{\mathbf{J}}_2 & - \end{array}\right] = \left[\begin{array}{cc|c} -0.6614 & 0.2845 & -0.1354 \\ -0.3325 & 0.1325 & 0.2157 \\ \hline -0.3458 & 0.5842 & - \end{array}\right] \quad (36)$$

The comparison result with the technique proposed in Theorem 4.3 illustrate in Table 1, that indicate the little conservation of the method suggest in the paper.

The initial cases are proposed null. Figure 1 illustrate the errors $e(k)$ from various approaches. It illustrate the FF method have the best performance in comparison with the EF approach.

Table 1: H_∞ performance apply from various approaches

Frequency	*Methods*	β_{min}	*Max error*
$-\pi \leq \varrho \leq \pi$	[12]	0.9715	0.0915
$-\pi \leq \varrho \leq \pi$	**Th 4.3** (Q=0)	0.5514	0.0423
$\frac{\pi}{7} \leq \varrho \leq \frac{\pi}{3}$	**Th 4.3**	**0.1305**	**0.0088**

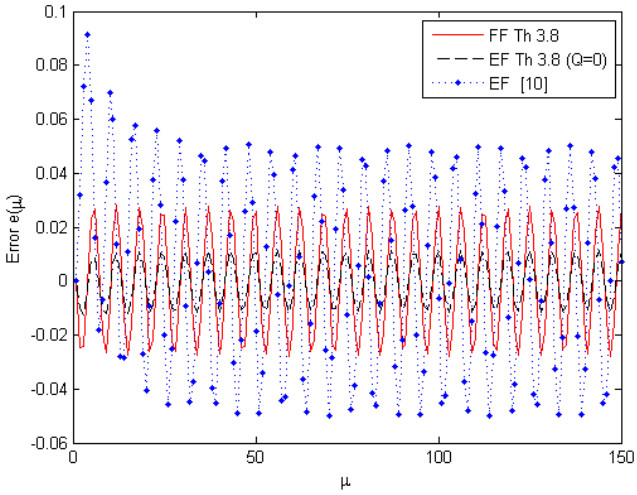


Figure 1: Error response of $\mathbf{g}_\mu$.

The ratio $\mathbf{W}(\mu)$ is presented as :

$$\mathbf{W}(\mu) = \sqrt{\sum_{\mu=0}^{\infty} \mathbf{g}^T(\mu)\mathbf{g}(\mu) / \sum_{\mu=0}^{\infty} \mathbf{d}^T(\mu)\mathbf{d}(\mu)} \quad (37)$$

Figures 2 indicate the the ratio in (37), we could notice whether the error model is stable, knowing that the initial Condition are null, that indicate the little conservation of the method suggest in the paper.

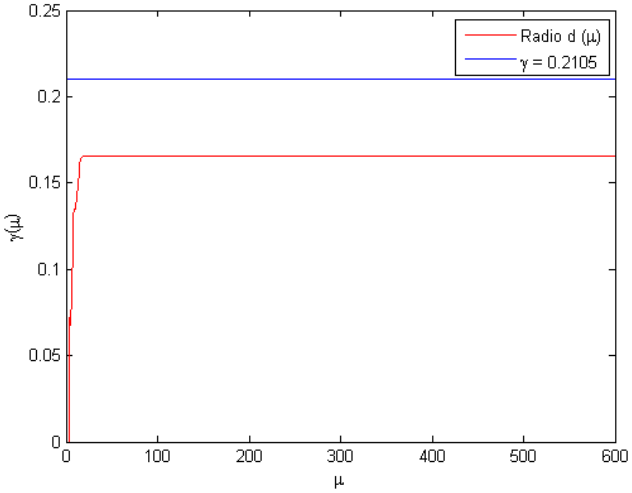


Figure 2: Value of $\mathbf{W}(\mu)$.

6 Conclusion

This work, we dealt with problematic for the FF model reduction design of nonlinear systems over FF ranges. We have suggested a model reduction process in order to minimize the conservatism design using the frequency information of the disturbances and we have assumed that the disturbances are known in a recognized FF domain. Also, systematic techniques have been suggested for the generation of a model reduction which ensures asymptotic stability and FF H_∞ index, at the basis of a more general linearization procedure.

References

[1] T. Takagi, M. Sugeno, " Fuzzy identification of systems and its applications to modeling and control," IEEE transactions on systems, man, and cybernetics, SMC-**15**(1), 116-132, 1985, doi: 10.1109/TSMC.1985.6313399

[2] H. Zhang, Y. Shi, A. S. Mehr, "On H_∞ filtering for discrete time Takagi Sugeno fuzzy Systems," IEEE Transactions on Fuzzy Systems, **20**(2), 396-401, 2011, doi: 10.1109/TFUZZ.2011.2175933

[3] A. Benzaouia, A. Hmamed, A. EL Hajjaji, " Stabilization of controlled positive discrete-time T-S fuzzy systems by state feedback control," International Journal of Adaptive Control and Signal Processing, **24**(12), 1091-1106, 2010, doi: 10.1002/acs.1185

[4] H. Gao, Y. Zhao, J. Lam, K. Chen, "H_∞ fuzzy filtering of nonlinear systems with intermittent measurements," IEEE Transactions on Fuzzy Systems, **17**(2), 289-299, 2009.

[5] L. Zhang, H. Yang, Q. Mao, "T S fuzzy system identification based on bayesian inference," ICIC Express Letter, **13**(1), 51-58, 2019.

[6] M. Nentwig, P. Mercorelli. "hrottle valve control using an inverse local linear model tree based on a fuzzy neural network," 7th International Conference on Cybernetic Intelligent Systems, IEEE, 1-6, 2008, doi: 10.1109/UKRICIS.2008.4798943

[7] M. Schimmack, F. Pöschke, H., Schulte, P. Mercorelli, "Observer Strategies for Virtual Sensing of Embroidered Metal-Polymer Heater Structure," 23rd International Conference on Methods & Models in Automation & Robotics (MMAR) IEEE, 887-892, 2018, doi: 10.1109/MMAR.2018.8486033

[8] A. C. Antoulas, "An overview of approximation methods for large-scale dynamical systems," Annual reviews in Control, **29**(2), 181-190, 2005, doi: 10.1016/j.arcontrol.2005.08.002

[9] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their $L,^{\infty}$-error bounds". International journal of control," **39**(6), 1115-1193, 1984, doi: 10.1080/00207178408933239

[10] G. Lastman, "Reduced-order aggregated models for bilinear time-invariant dynamical systems," IEEE transactions on automatic control, **29**(4), 359-361, 1984, doi: 10.1109/TAC.1984.1103534

[11] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," IEEE transactions on automatic control, **26**(1), 17-32, 1981, doi: 10.1109/TAC.1981.1102568

[12] L. Wu, X. Su, J. Qiu, "Model approximation for discrete-time state-delay systems in the T–S fuzzy framework," IEEE Transactions on Fuzzy Systems, **19**(2), 366-378, 2011, doi: 10.1109/TFUZZ.2011.2104363

[13] D. Wang, A. Zilouchian. "Model reduction of discrete linear systems via frequency-domain balanced structure," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, **47**(6), 830-837, 2000, doi: 10.1109/81.852936

[14] X. Du, G. H. Yang, "H_{∞} model reduction of linear continuous-time sys-tems over finite-frequency interval," IET control theory & applications, **4**(3), 499-508, 2020, doi : 10.1049/iet-cta.2008.0537

[15] X. Li, S. Yin, H. Gao, "Passivity-preserving model reduction with finite frequency H_{∞} approximation performance," Automatica, **50**(9), 2294-2303, 2014, doi: 10.1016/j.automatica.2014.07.001

[16] D. W. Ding, X. J. Li, X. Du, X. Xie," Finite-frequency model reduction of Takagi-Sugeno fuzzy systems," IEEE Transactions on Fuzzy Systems, **24**(6), 1464-1474, 2016, doi: 10.1109/TFUZZ.2016.2540060

[17] A. El-Amrani, B. Boukili, A. El-Hajjaji, A. Hmamed, "H_{∞} model reduc-tion for T-S fuzzy systems over finite frequency ranges," optimal control applications and methods, **39**(4), 1479-1496, 2018, doi : 10.1002/oca.2422

[18] D. W. Ding, X. Xie, X. Du, X. J. Li, "Finite-frequency model reduction of discrete-time T–S fuzzy state-delay systems," Neurocomputing, **203**, 121-128, 2016, doi: https://doi.org/10.1016/j.neucom.2016.03.053

[19] A. El-Amrani, A. El Hajjaji, I. Boumhidi, A. Hmamed, A. Aitouche, "Finite Frequency Approach for H_{∞} model reduction of 2D continuous systems," 27th Mediterranean Conference on Control and Automation (MED) IEEE, 177-182, 2019, doi: 10.1109/MED.2019.8798569

[20] X. Su, L. Wu, P. Shi, Y. D. Song, "H_{∞} model reduction of Takagi-Sugeno fuzzy stochastic systems", IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics: a publication of the IEEE Systems, Man, and Cybernetics Society, **42**(6), 1574-1585, 2012, doi: 10.1109/tsmcb.2012.2195723

[21] A. V. Oppenheim, A. S. Willsky, S. H. Nawab. "Signals and Systems", 2nd Edition,. Prentice Hall, Boston University, 1-987, 1996.

[22] M. J., Lacerda, R. C. F., Oliveira, P. L. D., Peres, "Robust H_2 and H_{∞} filter design for uncertain linear systems via LMIs and polynomial matrices," Signal Processing, **91**(5), 1115-1122, 2011, doi: 10.1016/j.sigpro.2010.10.013

7 Appendix

Multiply equation (16) of $\begin{bmatrix} \nabla(\mu) \\ \mathbf{d}(\mu) \end{bmatrix}$ for right and by its transposition to the left, we get

$$\begin{aligned} & \beta^2 \mathbf{d}(\mu)^T \mathbf{d}(\mu) + tr(C(\mathbf{g}^{j\tilde{\varrho}_3} \nabla(\mu)\nabla(\mu+1)^T \\ - \quad & e^{-j\tilde{\varrho}_3} \nabla(\mu+1)\nabla(\mu)^T - 2cos(\tilde{\varrho}_4)\nabla(\mu)\nabla(\mu)^T)) \\ - \quad & \nabla^T(\mu+1)\mathcal{B}\nabla(\mu+1) - \nabla^T(\mu)\mathcal{B}\nabla(\mu) + \mathbf{g}^T(\mu)\mathbf{g}(\mu) \leq 0 \end{aligned} \tag{38}$$

By taking the sum of $\mu = 0$ in π,seeing that $\nabla(0) = 0$ and $\lim_{\mu \to pi} \nabla(\mu) = 0$ we get

$$\begin{aligned} & \sum_{\mu=0}^{\pi} \mathbf{g}^T(\mu)\mathbf{g}(\mu) - \beta^2 \sum_{\mu=0}^{\pi} \mathbf{d}(\mu)^T \mathbf{d}(\mu) \\ + \quad & Tr[C \sum_{\mu=0}^{\pi} (e^{j(\tilde{\varrho}_3} \nabla(\mu)\nabla(\mu+1)^T + e^{-j\tilde{\varrho}_3} \nabla(\mu+1)\nabla(\mu)^T \\ - \quad & 2\cos(\tilde{\varrho}_4)\nabla(\mu)\nabla(\mu)^T)] \leq 0 \end{aligned} \tag{39}$$

From Parseval function [21], we get :

$$\begin{aligned} & \beta^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{d}^T(\tilde{\varrho})\mathbf{d}(\tilde{\varrho}) d\tilde{\varrho} \\ - \quad & Tr[\frac{\cos(\tilde{\varrho} - \tilde{\varrho}_c) - \cos\tilde{\varrho}_w}{\pi} \int_{-\pi}^{\pi} \mathcal{M}^T(\tilde{\varrho})C\mathcal{M}(\tilde{\varrho}) d\tilde{\varrho}] \\ - \quad & \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{g}^T(\tilde{\varrho})\mathbf{g}(\tilde{\varrho}) d\tilde{\varrho} \geq 0 \end{aligned} \tag{40}$$

Then :

$$\cos(\tilde{\varrho} - \tilde{\varrho}_3) - \cos(\tilde{\varrho}_4) \geq 0 \tag{41}$$

have toward all gifted $\tilde{\varrho}_1 \leq \tilde{\varrho} \leq \tilde{\varrho}_2$. From $C > 0$ and (41), it attends that

$$Tr[\frac{\cos(\tilde{\varrho} - \tilde{\varrho}_c) - \cos(\tilde{\varrho}_4)}{\pi} \int_{-\pi}^{\pi} \mathcal{M}^T(\tilde{\varrho})C\mathcal{M}(\tilde{\varrho}) d\tilde{\varrho}] \geq 0 \tag{42}$$

seeing that, (40) and the FF of signal input $\tilde{\varrho}_1 \leq \tilde{\varrho} \leq \tilde{\varrho}_2$, thus (15) is fulfilled.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 6, 54-59 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Hiragana and Katakana minutiae based recognition system

Piotr Szymkowski[*,1], Khalid Saeed[1], Nobuyuki Nishiuchi[2]

[1]*Bialystok University of Technology, Department of Computer Science, Bialystok, 15-351, Poland*

[2]*Tokyo Metropolitan University, Department of Computer Science, Tokyo, 191-0065, Japan*

ARTICLE INFO

Article history:
Received: 12 July, 2021
Accepted: 12 October, 2021
Online: 20 November, 2021

Keywords:
Handwriting recognition
Biometric recognition
Minutiae

ABSTRACT

The Japanese writing system is unique due to the number of characters employed and the methods used to write words. It consists of three different 'alphabets', which may result in the methods used to process Latin script not being sufficient to obtain satisfactory results when attempting to apply them to a recognition of the Japanese script. The authors present an algorithm based on minutiae, i.e., feature points, to recognise the hiragana and katakana characters. A method using image processing algorithms is compared with a method using a neural network for the purpose of automating this process. Based on the distribution and type of minutiae, vectors of features have been created to recognise 96 different characters. The authors conducted a study showing the effect of the chosen segmentation method on the accuracy of the character recognition. The proposed solution has achieved a maximum accuracy at the level of 65.2%.

1 Introduction

This paper is an extension of work originally presented in The International Conference on Graphics and Signal Processing [1]. The Japanese writing system consists of three types of characters: kanji, hiragana and katakana. The hiragana and katakana characters are found only in Japan. They are syllabaries, which means that a single written character stands for a single spoken syllable. Hiragana is mainly used to write prefixes and suffixes of the Japanese words, particles and words of Japanese origin which for some reason are not written in the kanji characters. Hiragana is also used to write furigana, a reading aid for the difficult or unusual kanji characters. It is often found in books for children and teenagers.

Katakana is used to write first names, surnames and words of foreign origin. It is also applied for onomatopoeia (mimetic words), which frequently occur in Japanese. The two previously discussed syllabaries each contain 46 characters each.

The last group of characters are the kanji characters. These characters have been borrowed from the Chinese language. The number of the kanji characters in Japan is estimated to be approximately 50,000, but the average Japanese needs circa 3,000 characters for the use on an everyday basis. The kanji characters are used to write the subject of words. Unlike the hiragana and katakana, the kanji characters have different manners of reading. The way they are read depends on whether they stand alone or are accompanied by other characters. One kanji character can have even more than 10 different ways of reading.

The quantity and complexity of the writing system in relation to the 26 characters of the Latin alphabet results in the fact that recognition methods and algorithms for the purpose of writing should be subject to a change.

2 State of the Art

For the purpose of the recognition of the Japanese writing system, publications on the on-line recognition are easier to find than the off-line ones. The key publications presenting a cross-section of ongoing studies on this very subject matter will be introduced below.

2.1 Deep Convolutional Recurrent Network

The first solution presented uses deep convolutional recurrent neural networks in order to recognise single lines of text written in Japanese [2]. The network used consists of three layers: a convolutional feature extractor, recurrent layers and transcription layers [3]. Prior to the recognition itself, a preprocessing of the images is required. This consists of normalising, i.e., homogenising the samples, and binarizing them with the use of the Otsu's method. The authors use the TUAT Kondate database. Another aspect presented in this paper constitute the artificially generated samples, developed on the basis

[*]Corresponding Author: Piotr Szymkowski, prszymkowski@gmail.com

of the previously mentioned database. The model presented here shows a character recognition rate at the level of 96.35% when using only the data from the database, and at the level of 98.05% when using also the artificially generated data. The generation of these data is based on the use of the local and global elastic distortions.

2.2 Extended Peripheral Direction Contributivity

The recognision method of the Japanese writing system presented by the authors [4] adopts the Extended Peripheral Direction Contributivity method. It is one of the more commonly used directional features. The feature vector is developed in the following manner. First of all, the image is scanned from their peripheral to inner part in the eight directions. Then, on each black pixel of the contour of the first, second and third line encountered in each direction, the run-length l_i in each of the eight directions is calculated and the direction contributivity d_i (i = 1, 2, ..., 8) is calculated for the pixel using (1) [5].

$$d_i = \frac{l_i}{\sqrt{\sum_{j-1}^{8} l_j^2}} \tag{1}$$

Thirdly, the image is divided into several zones in each direction of the peripheral scan. Finally, a feature vector is obtained by recording the directional components for each zone, in each direction and in each order encountered. The information that can be obtained by the aforementioned method are as follows [6]:

- A quantity of lines and their orientation,
- A position of the lines,
- The relations between the lines.

2.3 Address Recognition

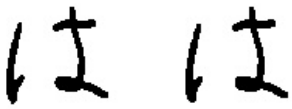
The algorithm presented, represents the use of a system that reads information from the letters and postal shipments concerning the information on a destination address [7]. For the purpose of reducing the information to be extracted, a priori knowledge about the structure of the Japanese address has been used. The introduced system is based on five processes: segmentation, segment combination, character recognition, character combination and path search. The system is supported by a glossary of the possible addresses.

3 Image Processing Algorithms

The algorithm presented by the authors uses many image processing methods. Some of the currently used solutions use a priori knowledge, which also reduces a probability of misclassification of the characters. The principles of operation of the most important of them will be specified below.

3.1 Morphological Closing

One of the essential operations performed on the images are the morphological ones. As their name suggests, their application results in a change in the appearance of the result image. In case of the algorithm presented by the authors, the image on which these operations are employed is black and white. The two most basic morphological operations are the image erosion and image dilation. The result of applying the filter is shown in Figure 1. Their application is equivalent to the minimum and maximum of the statistical filters [8].



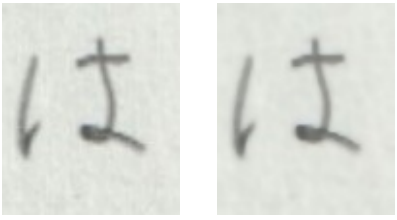
(a) (b)

Figure 1: Original image (**a**) and image after morphological closing applied (**b**).

The closing filter consists in the performance of the dilation and erosion operations on the image using the same structural element. As a result of this operation, objects lying close to each other are merged, their edges are filled and smoothed.

3.2 Bilateral Filter

The bilateral filter is one of the blurring filters used to remove the initial noise present in the images. In contrast to the Gauss filter, it is not linear, which means that it blurs the image while retaining transparency of the edges [9].



(a) (b)

Figure 2: Original image (**a**) and image after bilateral filter application (**b**).

This is achieved by varying the blur rate, which depends on the size of the structural object and the value of the local gradient. It results in stronger blurring of homogenous areas and lesser blurring in the regions of sharp local changes, i.e. edges, what can be seen at Figure 2. The bilateral filter is presented in (2).

$$I_{filter} = \frac{1}{W_p} \sum_{x_i \subseteq \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|)) \tag{2}$$

The normalising formula is presented in (3).

$$W_p = \sum_{x_i \subseteq \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|)) \tag{3}$$

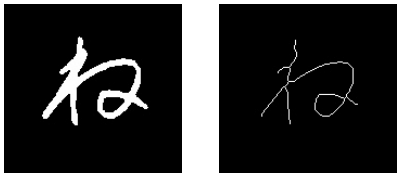
where:

- I filter – result image,
- I – input image,

- x - coordinates of pixel under analysis,
- Ω - structural object,
- f_r - smoothing factor (intensity),
- g_s - smoothing factor (distance).

3.3 Modified K3M Algorithm

Thinning is one of the operations used to simplify the description of the shapes of objects in the image. The operation is performed on a binarized image. As shown in Figure 3, all lines in the image are narrowed until they reach a width of 1 pixel. The modified K3M algorithm used in this publication has been created by one of the authors [10].



(a) (b)

Figure 3: Original image (**a**) and sign thinned using Modified K3M algorithm (**b**).

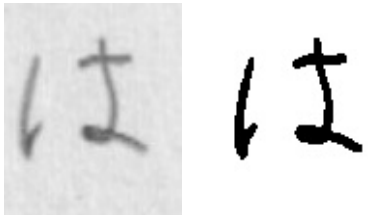
3.4 Otsu's Method

The Otsu's method is one of the binarization techniques, which uses the thresholding what means that the algorithm uses a predetermined threshold to segment the entire input image. An example of how the method works is shown in Figure 4. It is a technique based on the discriminant analysis [11]. The optimal threshold is chosen based on one of the presented in (4) equivalent criterion functions.

$$\eta = \frac{\delta_B^2}{\delta_T^2}, \lambda = \frac{\delta_B^2}{\delta_W^2}, \kappa = \frac{\delta_T^2}{\delta_W^2} \tag{4}$$

Where:

- $\delta_B^2 = P_0(\mu_0 - \mu_T)^2 + P_1(\mu_1 - \mu_T)^2 = P_0P_1(\mu_0 - \mu_1)^2$ - between class variation,
- $\delta_W^2 = P_0\delta_0^2 + P_1\delta_1^2$, - within class variation,
- $\delta_T^2 = \sum_{i=0}^{L-1}(i - \mu_T)^2 P_i$ - global variation,
- $\delta_T^2 = \delta_W^2 + \delta_B^2$.



(a) (b)

Figure 4: Original image (**a**) and image after Otsu method application (**b**).

4 Algorithm Description

To achieve the results, a Python application was prepared by the authors using keras library to apply artificial neural networks and sklearn to apply the SVM algorithm. The presented proprietary solution uses two artificial intelligence algorithms for the purpose of automating the recognition process of Japanese handwriting, therefore two types of feature vectors are created. It is similar to another used in the biometric algorithms, more specifically while recognising a user by their fingerprints. In order to create feature vectors for the character recognition, characteristic points - minutiae - are used. Figure 5 shows a diagram demonstrating how the presented solution works.

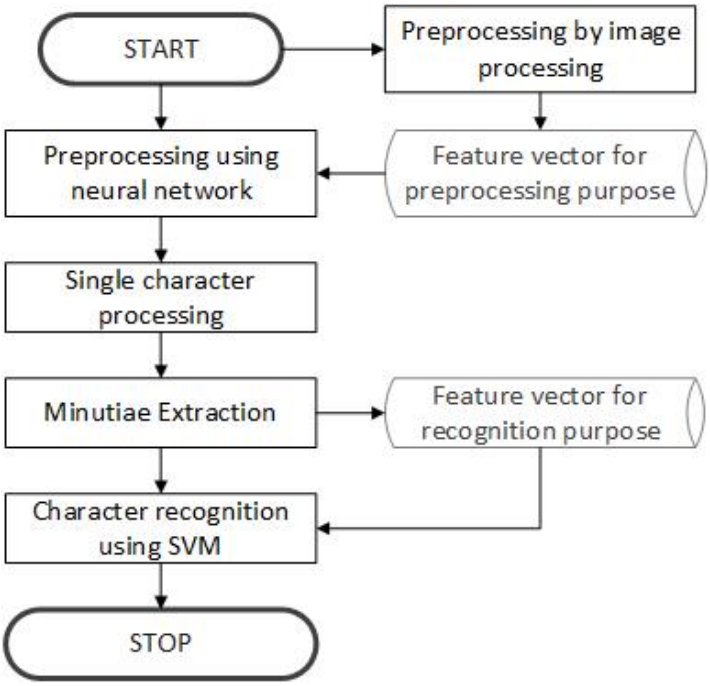


Figure 5: Algorithm diagram

4.1 Preprocessing Using Image Processing

This step is followed by the creation of feature vectors further used during the learning of the neural network in order to perform an automatic segmentation. The steps of this process are shown at Figure 6. The first stage consists in the use a bilateral filter [12]. As a consequence of the aforementioned, some of the noise that appeared in the image is eliminated. However, this has failed to affect the edge condition.

The next step has been to apply the background subtraction algorithm. It is usually applied when segmenting a series of photographs or films. After subtracting two consecutive photographs or films frames, the result image shows changes in the image, movement of animals, objects. In case of the presented solution, an image on which an averaging filter has been applied using a structural object with a diameter of 5 pixels is subtracted from the original image upon removing noise. The result image shows the differences in the image frequency.

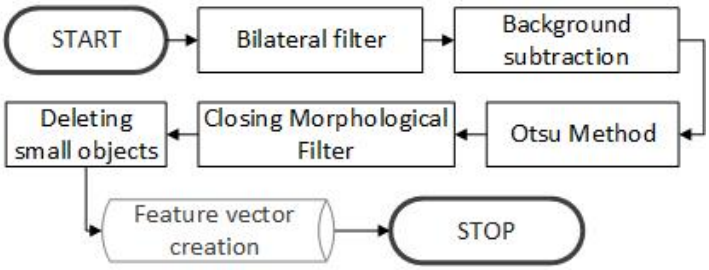


Figure 6: Image processing approach to preprocessing.

Binarization is the third preprocessing step. For this purpose, the algorithm according to the Otsu's method has been used. As in the event of the biometric algorithms, the green channel has been used for this purpose. Upon completing this stage, the image shows a partially or fully completed character. Then a closing morphological filter is applied. After this operation, the edges in the image are smoothed and some of the noise is deleted. Finally, the noise is removed by applying the last algorithm.

An algorithm has been developed that calculates the area of each object appearing in the image. If it is less than a threshold one, the object is deleted (Figure 12).

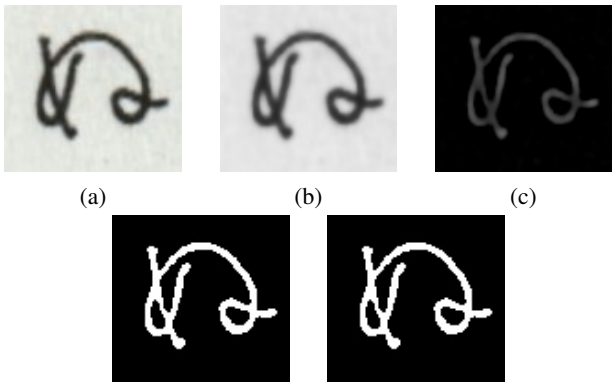


Figure 7: Character after each operation of preprocessing: (**a**) original image, (**b**) image after bilateral filter application, (**c**) image after background subtraction,(**d**) binarized image and(**e**) image after small object deletion and closing filter application.

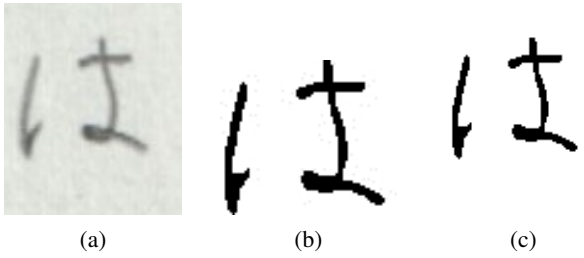


Figure 8: Sample (**a**), sample preprocessed by hand(**b**) and sample preprocessed using image processing approach.

The final step is to create a feature vector based on the output image. A square structural object with a length and a width of 11 pixels is used. Each vector possesses information about the central pixel, its 120 neighbouring ones, as well as information about its classification upon preprocessing (1 - background, 0 - object). The result of these operations can be seen in Figure 8.

4.2 Segmentation Using a Neural Network

A neural network has been used to automate the preprocessing process. The large illustrations (2900 x 1200 pixels) representing Japanese writing system have been used to develop the sufficient feature vectors. The principle of operation of this part of the algorithm is presented in Figure 9.

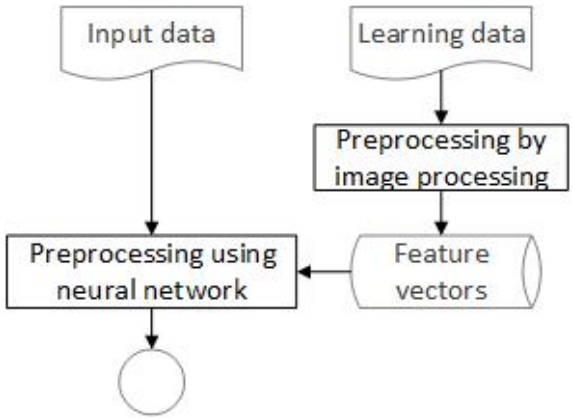


Figure 9: Neural network approach to preprocessing.

A sequential neural network model with the ReLU activation function has been employed for the research purposes. The neural network consists of four layers: an input layer, two hidden layers, and an output layer. The input layer has 363 nodes, the first hidden layer possesses 120 nodes, the second hidden layer has 30 nodes, and the output layer possesses 1 node. As the output, information is received as to whether the pixel under study (located in the centre of the structural object) should be classified as a background or as an object [13]. A comparison of manually processed samples and those processed using neural networks is presented in Figure 10.

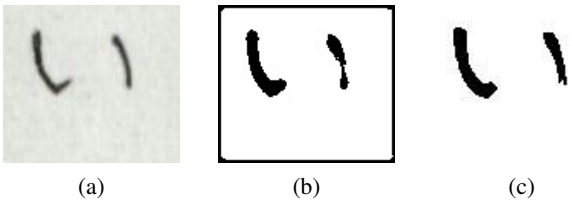


Figure 10: Original sample (**a**),sample after image processing based preprocessing (**b**) and sample preprocessed using neural network(**c**).

4.3 Processing of Single Characters

Upon preprocessing the image containing the object of interest, you obtain the image that requires the further adjustments to extract the features of interest. The next steps of single character processing are shown in Figure 11. The first step is to remove the part of the image that does not carry any information. This is done with the help of the vertical and horizontal histograms. Commencing from the ends of the image, you search for the maximum value of a distance from the edge, where the number of pixels of the object is 0.

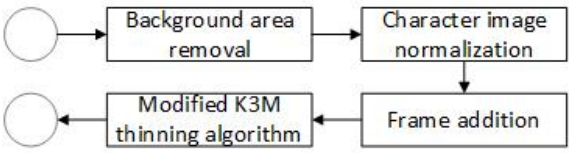


Figure 11: Single character processing

4.4 *Development of Feature Vectors and Character Recognition*

The next step is the normalisation, which consists in standardising the size of the images [14]. The operations are performed on the images to adjust their size to 62 by 62 pixels. A one-pixel frame is then added to enable to perform further operations on the images (Figure 12). The last operation performed at this stage of the image processing is skeletonization using a modified K3M algorithm. The choice of the skeletonization algorithm has been made using an experimental method.

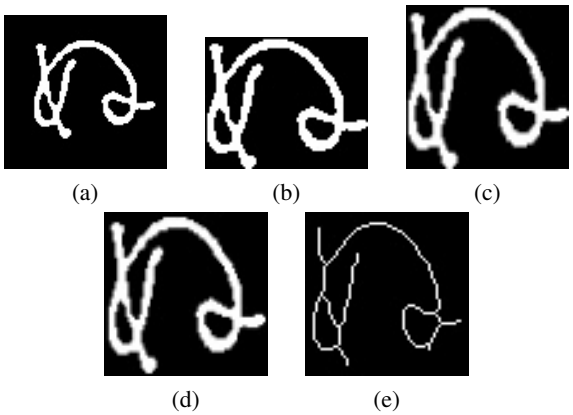


Figure 12: Character after each operation of single character processing step: (**a**) original image, (**b**) image after background area removal, (**c**) image resized,(**d**) one pixel-wide frame added and(**e**) image after thinning algorithm application.

While determining the feature vectors, the authors have used the methods applied in the biometric recognition using fingerprints. They use minutiae - characteristic points present in the fingerprints (Figure 13). Among all available types, three types have been selected:

- Ends and beginnings of lines,
- Bifurcations,
- Intersections or Trifurcations.

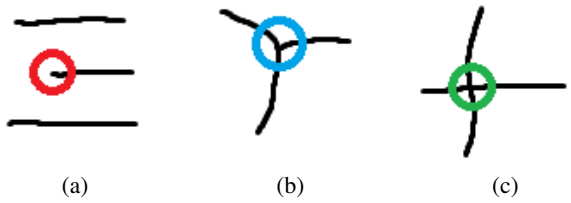


Figure 13: Minutiae types: (**a**) beginning or ending, (**b**) biforcation and (**c**) triforcation.

For the purpose of developing a feature vector, the image has been divided into 16 equal areas (16 by 16 pixels) and the sums of each type of minutiae are stored in the appropriate places. Each feature vector contains the feature values and the decision class.

On the basis of an in-house database containing approximately 50 samples for each of the 96 hiragana and katakana characters, experiments have been conducted to test the ability of the SVM (Support Vector Machine) to recognise the characters [15].

5 Results

During preprocessing, a prediction match of 86.20% with the intended result has been obtained. It is a relatively high value, but even the smallest deviation from an exemplar one can result in, and has resulted in, the creation of the additional minutiae or a loss of some of them in the image. In the course of the study, the division of samples into a learning set and a test set has been carried out in the ratio 3:2.

Two experiments have been conducted during the character recognition based on the feature vectors of minutiae occurrences. The first experiment has constituted in testing the effect of the choice of preprocessing method on the results of the recognition of the Japanese characters. Three trials have been carried out: data upon a manual segmentation has been used within the first trial, data preprocessed with the use of an algorithm based on the image processing algorithms has been applied in the second trial, and data preprocessed using artificial intelligence has been used in the third one. Table 1 shows the results of the first experiment. Depending on the technique used, the results vary. It can be noted that the more complicated the method processing of a single character is complicated, the lower recognition results were obtained. The highest precision was achieved when using the authors' algorithm is 46.6% when applying image processing approach to preprocessing.

Table 1: First experiment results.

Circumstances	Precision of recognition
Hand preprocessed samples	50,3%
Samples preprocessed using image processing	46,6%
Samples preprocessed using neaural network	42,6%

Table 2: Second experiment results.

Circumstances	Precision
Text preprocessed using neural network with added artificial samples	63,3%
Text preprocessed using image processing approach with added artificial samples	65,2%

In the second experiment, a 50 percent chance of development of a descendant from it has occurred. The newly created sample has been subject to twisting and stretching by random parameters. It has been further adapted to the requirements of the algorithm and a new feature vector has been developed on the basis of it. In both

cases the results have been significantly improved by more than 15%, reaching 63.3% when using a neural network and 65.2% when using image processing for preprocessing, as presented in the Table 2.

Table 3: Solution compparison.

Authors	Approach	Database	Declared precision (Best)
Author's Solution	Minutia based	Own database	65,2%
N.T. Ly, C.T. Nguyen, M.Nakagawa [2]	Deep Convolutional Recurent Network	TUAT Kondate	98,05%
M.Mori, T.Wakahara, K.Ogura [4]	Extended Peripheral Direction Contrubutivity	ETL9B	95%
X. Xia, X. Yu, W. Liu, et al. [7]	Over segmentation strategy	Own database	83.4%
S. Tsuruoka, M. Hattori, M. F. b. A. Kadir, et al. [16]	Modified quadratic discriminate function and users dictionaries	ETL9B	82.5%

6 Conclusions

The presented results enable to conclude that the introduced method is not sufficient to allow the unaided character recognition at a high enough level. One of the reasons may be the use of a non-uniform database. The obtained result (62.5%) is the lowest among those compared 3. However, those results allow to suppose that combining this method with other will allow to obtain better solution. The future papers of the authors will focus on the development of the presented solution in order to obtain a system allowing for the off-line text recognition with the use of the hybrid methods.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This work was supported by grant PROM from the Polish National Agency for Academic Exchange NAWA, supported by the Bialystok University of Technology and partially funded with resources for research by the Ministry of Science and Higher Education in Poland, under Grant WZ/WI-IIT/4/2020.

References

[1] P. Szymkowski, K. Saeed, N. Nishiuchi, "SVM Based Hiragana and Katakana Recognition Algorithm with Neural Network Based Segmentation," in Proceedings of the 2020 The 4th International Conference on Graphics and Signal Processing, ICGSP 2020, 51–55, Association for Computing Machinery, New York, NY, USA, 2020, doi:10.1145/3406971.3406978.

[2] N. T. Ly, C. T. Nguyen, M. Nakagawa, "Training an End-to-End Model for Offline Handwritten Japanese Text Recognition by Generated Synthetic Patterns," in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), 74–79, 2018, doi:10.1109/ICFHR-2018.2018.00022.

[3] S. D. Budiwati, J. Haryatno, E. M. Dharma, "Japanese character (Kana) pattern recognition application using neural network," in Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, 1–6, 2011, doi:10.1109/ICEEI.2011.6021648.

[4] T. Wakahara, Y. Kimura, M. Sano, "Handwritten Japanese character recognition using adaptive normalization by global affine transformation," in Proceedings of Sixth International Conference on Document Analysis and Recognition, 424–428, 2001, doi:10.1109/ICDAR.2001.953825.

[5] M. Mori, T. Wakahara, K. Ogura, "Measures for structural and global shape description in handwritten Kanji character recognition," 1998, doi: 10.1117/12.304621.

[6] B. Lyu, R. Akama, H. Tomiyama, L. Meng, "The Early Japanese Books Text Line Segmentation base on Image Processing and Deep Learning," in 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), 299–304, 2019, doi:10.1109/ICAMechS.2019.8861597.

[7] X. Xia, X. Yu, W. Liu, C. Zhang, J. Sun, S. Naoi, "An Efficient off-Line Handwritten Japanese Address Recognition System," in 2019 International Conference on Document Analysis and Recognition (ICDAR), 714–719, 2019, doi:10.1109/ICDAR.2019.00119.

[8] M. Nixon, A. S. Aguado, Feature Extraction and Image Processing for Computer Vision, Third Edition, Academic Press, Inc., USA, 3rd edition, 2012.

[9] M. Buczkowski, P. Szymkowski, K. Saeed, "Segmentation of Microscope Erythrocyte Images by CNN-Enhanced Algorithms," Sensors, **21**(5), 2021, doi:10.3390/s21051720.

[10] M. Tabedzki, K. Saeed, A. Szczepański, "A modified K3M thinning algorithm," International Journal of Applied Mathematics and Computer Science, **26**(2), 439–450, 2016, doi:doi:10.1515/amcs-2016-0031.

[11] C. Solomon, T. Breckon, Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab, Wiley Publishing, 1st edition, 2011.

[12] Y. Seki, "Collection of Online and Offline Handwritten Japanese Charcters and Handwriting Classification Using the Data," in 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), 264–269, 2020, doi:10.1109/ICFHR2020.2020.00056.

[13] N.-T. Ly, C.-T. Nguyen, K.-C. Nguyen, M. Nakagawa, "Deep Convolutional Recurrent Network for Segmentation-Free Offline Handwritten Japanese Text Recognition," in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 07, 5–9, 2017, doi: 10.1109/ICDAR.2017.357.

[14] I. Uddin, D. A. Ramli, A. Khan, J. I. Bangash, N. Fayyaz, A. Khan, M. Kundi, "Benchmark Pashto Handwritten Character Dataset and Pashto Object Character Recognition (OCR) Using Deep Neural Network with Rule Activation Function," Complexity, **2021**, 6669672, 2021, doi:10.1155/2021/6669672.

[15] K. Nguyen, C. Nguyen, M. Nakagawa, "A Segmentation Method of Single- and Multiple-Touching Characters in Offline Handwritten Japanese Text Recognition," IEICE Transactions on Information and Systems, **E100.D**, 2962–2972, 2017, doi:10.1587/transinf.2017EDP7225.

[16] S. Tsuruoka, M. Hattori, M. F. b. A. Kadir, T. Takano, H. Kawanaka, H. Takase, Y. Miyake, "Personal Dictionaries for Handwritten Character Recognition Using Characters Written by a Similar Writer," in 2010 12th International Conference on Frontiers in Handwriting Recognition, 599–604, 2010, doi: 10.1109/ICFHR.2010.98.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 6, 60-65 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Modelling and Testing Services with Continuous Time SRML

Ning Yu, Martin Wirsing*

Institute for Informatics, LMU Munich, 80538 Munich, Germany

ARTICLE INFO

Article history:
Received: 31 July, 2021
Accepted: 26 October, 2021
Online: 20 November, 2021

Keywords:
continuous time SRML
Service-Oriented Hybrid Doubly Labelled Transition System
differential equations
traffic control system

ABSTRACT

The SENSORIA Reference Modeling Language (SRML) aims at modelling composite services at a high level. Continuous time SRML extends SRML so that it can model services whose components can perform both discrete processes and continuous processes. In order to show how continuous time SRML is applied, in this paper, we systematically introduce our study on continuous time SRML in the following approach: First we introduce the theoretical foundation of continuous time SRML, the Service-Oriented Hybrid Doubly Labeled Transition Systems. This is the semantic domain over which continuous time SRML is defined and interpreted. Then we design a case study of a traffic control system. In the case study, we illustrate the scenario of the system, explain the continuous time SRML model of the system, and show how to transform the model to a kind of Deterministic Finite Automata that can be used for testing and verification. Finally, we show our idea of testing our model with the IBM WebSphere Process Server. With this approach, we come to a conclusion that continuous time SRML can be used to model certain systems in the real-world, and can be tested with proper tools.

1 Introduction

Service-Oriented Computing (SOC) [1] is a paradigm for distributed computing, in which computation units are abstracted as services. The SENSORIA Reference Modelling Language (SRML) [2, 3] has been developed in the IST-FET integrated project SENSORIA [4], and it is a prototype domain-specific language for abstracting service-oriented systems at a high level abstraction. Hybrid systems arise in embedded control when discrete components are coupled with continuous components. Continuous time SRML [5] is an extension of SRML, and it aims at abstracting service-oriented systems in which hybrid systems are embedded. That is: the discrete components and continuous components are abstracted as components of services. Being different from SRML, continuous time SRML can be used to formally model services, whose components can perform both discrete processes and continuous processes that can be describe by differential equations.

In [6], we introduced the application of continuous time SRML through the case study of a Traffic Control System. In this paper, we give a novel systematic presentation of our approach from the semantic domain of continuous time SRML, to the case study, and finally to the test environment. Continuous time SRML is interpreted over Service-Oriented Hybrid Doubly Labeled Transition Systems [5] (SO-HL^{2}TS). This kind of transition systems extends Service-Oriented Doubly Labeled Transition Systems [7] (SO-L^{2}TS), so that each state of a SO-L^{2}TS is lifted to a function mapping a time interval to an infinite trace of state, over which the differential equations in continuous time SRML can be interpreted. Our approach includes two parts: specifying the system with continuous time SRML modules, and transforming the modules to a hybrid automaton-like Deterministic Finite Automaton (DFA). In order to test the system module, we combine IBM Websphere Integration Designer [8] (WID) and Matlab to form the test environment. WID is used to implement the continuous time SRML module, and Matlab [9] is used to compute the variables controlled by differential equations. Thus, the main contribution of this paper is to provide a way to test or verify the complex systems (like the Traffic Control System) that is specified with continuous time SRML.

This paper is arranged as follows: In Section 2, we introduce the related work such as hybrid automata, model checking tools and the early research of this topic. In Section 3, we introduce HL^{2}TS, the semantic domain of continuous time SRML. In Section 4, we show the case study of the traffic control system. In Section 5, we introduce the testing environment and show the idea of implementing the traffic control system. In Section 6, we make a conclusion of this paper and introduce our future work.

*Corresponding Author: Martin Wirsing, Oettingenstraße 67,80538 Munich (Germany), +49 89 2180 9154 & wirsing@pst.ifi.lmu.de

2 Related work

SRML is brought forward in [7], and it is used to model important properties of services and the business conversations within and among services in SOC. A service specified with SRML can only perform discrete activities. Continuous time SRML extends the semantic domain and syntax of SRML, such that it could express differential equations which represents the continuous process performed by services. This essential property of continuous time SRML enables it be able to model the combination of service-oriented systems and hybrid systems. We will introduce this in Section 3.

Hybrid automata [10] are formal models for hybrid systems. They abstract hybrid systems with graphs and workable representations. A hybrid automaton consists of a set of variables, a control graph, some essential conditions and a set of events. These compositions match the compositions of a service component that is specified with continuous time SRML. Thus, a service specified by continuous time SRML can be transformed to a hybrid automata-like DFA, this will be introduced in Section 4. Hybrid automata are interpreted over timed transition systems, while continuous time SRML is interpreted over SO-HL^{2}TS. Timed transition systems abstract the duration of a continuous process as a parameter of a transition. In this way, it is simple and clear for abstracting continuous processes. SO-HL^{2}TS specify continuous processes as traces of infinite states with definite initial states and finial states. This enables us to know the values of all the variable of an arbitrary state.

In [11], the author provides another way for modelling hybrid systems. In that paper, continuous processes are also interpreted over traces of states. This is the original idea from which the traces of states in SO-HL^{2}TS come. But in [11], discrete processes are not modelled between traces of states. In [12], the author also provides rule schemata for verifying the hybrid systems. Moreover, based on this work, a model checking tools for hybrid systems, KeYmaera X, is brought forward in [13]. In [14], provides an approach for simulating embedded systems with Real-Time Maude. Compared with our approach for testing, this can be another way to test the traffic control model, but the interactions of a service would not be implemented as easy and explicate as that in WID.

3 Continuous time SRML and its semantic domain

Continuous time SRML models composite services, whose elementary components may involve in continuous time executions. The services in continuous time SRML are represented with service modules. A service module is composed of service components, external service interfaces and internal wires linking between them. In Figure 1, we see the compositions of a service module. In the service module, SC is a service component, EX-P is an external service interface (provides-interface), every EX-R is an external service interface (requires-interface), and every IW is an internal wire.

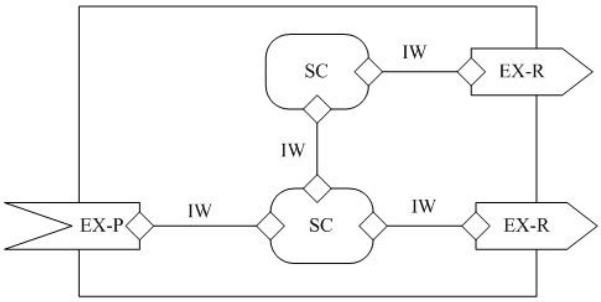


Figure 1: A service module of continuous time SRML

In a service module of continuous time SRML, a service component is specified with a business role, an external service interface is specified with a business protocol, and an internal wire is specified with a connector. The full syntax of continuous time SRML can be found in [5], and they are interpreted over SO-HL^{2}TSs.

SO-HL^{2}TSs are extensions of HL^{2}TSs [5, 15] such that they enrich HL^{2}TSs with service-oriented features. HL^{2}TSs extends the L^{2}TS [16] in that it defines a set of functions Σ. These functions map from the real number domain to the state domain of the HL^{2}TS and can be used to interpret the continuous processes described by differential equations. The following definition of HL^{2}TS is from [5, 15].

Definition 1 (Hybrid Doubly Labeled Transition System) *A Hybrid Doubly Labeled Transition System* (HL^{2}TS) is a tuple

$$\langle S, s_0, \Sigma, Act, R, AP, L\rangle$$

where:

- S is a set of states;
- $s_0 \in S$ is the initial state;
- Σ is a set of functions and for every function $\sigma \in \Sigma$ there is $r_\sigma \in \mathbb{R}$ and $r_\sigma \geq 0$, such that $\sigma : [0, r_\sigma] \to S$;
- Act is a finite set of observable actions;
- $R \subseteq \{\sigma(r_\sigma) : \sigma \in \Sigma\} \times 2^{Act} \times \{\sigma(0) : \sigma \in \Sigma\}$ is the transition relation. A transition $\sigma(r_\sigma) \xrightarrow{\alpha} \sigma'(0)$ with $\alpha \subset Act$, is denoted by $(\sigma(r_\sigma), \alpha, \sigma'(0)) \in R$;
- AP is a set of atomic propositions;
- $L : S \to 2^{AP}$ is a labeling function such that $L(s)$ is the subset of all atomic propositions that are true in state s.

In Definition 1, the set S of state can be finite or infinite; for every $\sigma \in \Sigma$, the real number r_σ is unique, and function σ on the interval $[0, r_\sigma]$ represents the prolongation of states in the continuous process with the duration r_σ; function L is a state labeling function that defines which propositions are true in each state, and these propositions are from the fixed set of atomic propositions AP.

In order to add service-oriented features to HL^{2}TSs, we refine HL^{2}TSs with Service-Oriented Transition Systems (SO-TSs)

[2]. SO-TSs are used that models the communications in service-oriented systems during the computing processes. In a SO-TS, states and transitions are labelled with certain sub-sets of events that appear in the computing processes. With such refinements, we can obtain SO-HL^{2}TSs. The following definition of SO-HL^{2}TSs is from [5, 15].

Definition 2 (Service-Oriented Hybrid L^2TS) A *Service-Oriented Hybrid L^2TS (SO-HL^{2}TS)* that refines a SO-TS $\langle S, \rightarrow, s_0, G\rangle$ is a tuple

$$\langle S, s_0, \Sigma, Act, R, AP, L, TIME\rangle$$

where:

- $Act = \{e! : e \in E\} \cup \{e¡ : e \in E\} \cup \{e? : e \in E\} \cup \{e¿ : e \in E\}$;
- $R \subseteq \{\sigma(r_\sigma) : \sigma \in \Sigma\} \times 2^{Act} \times \{\sigma(0) : \sigma \in \Sigma\}$ is such that:,

  $\sigma(r_\sigma) \rightarrow \sigma'(0)$ *iff* $(\sigma(r_\sigma), \alpha, \sigma'(0)) \in R$ for some $\alpha \in 2^{Act}$
- $AP = Act \cup PLG$;
- *For every* $\sigma \in \Sigma$ *and every* $\zeta \in [0, r_\sigma]$, $L : S \rightarrow 2^{AP}$ is such that:

  $$L(\sigma(\zeta)) = \{e!, e¡, e?, e¿ : e \in HST^{\sigma(\zeta)}\} \cup PLG^{\sigma(\zeta)}$$

In Definition 2, the set Act is refined with the sets of events that are published (denoted by e!), delivered (denoted by e¡), executed (denoted by e?)and discarded (denoted by e¿) respectively, E is the set of all the events. The set AP of atomic propositions is labelled with the set of actions and the set of pledges that associate with every interaction (the propositions that hold during the time when the interaction is valid). Every state is labelled with the actions whose associating events are from the history of the events of that state, and the pledges that hold in the state. We also specify the time in each state with the function $TIME$. The full definition can be seen in [5]

When applying a function $\sigma \in \Sigma$ to its interval $[0, r_\sigma]$, we obtain a infinite *trace of states* $\sigma_0, \ldots, \sigma_{r_\sigma}$. By connecting traces of states with possible actions, *paths of states* are formed. Because a trace of states represent the evolution of a continuous process, actions which represent the discrete processes can only take place at the initial and finial states of each trace. Thus, a transition in a path only exists between the last state of one trace and the first state of another trace. This can be seen in the definition of transition relation R in Definition 2. The following definition of traces and paths of states is from [5, 15].

Definition 3 (Paths of SO-HL^{2}TSs) *Given a SO-HL^{2}TS* $m = \langle S, s_0, \Sigma, Act, R, AP, L, TIME, \Pi\rangle$, *paths of m are defined as follows:*

- *For every* $\sigma \in \Sigma$, $\sigma(0)\ldots\sigma(r_\sigma)$ *denotes the trace of states of* σ *where* $\sigma(0)$ *is the first state of the trace and* $\sigma(r_\sigma)$ *is the last state of the trace.* $\sigma(0)\ldots\sigma(r_\sigma)$ *includes all the states in the set* $\{\sigma(\xi)|0 \leq \xi \leq r_\sigma\}$, *which can be finite or infinite.*
- $\rho = (\sigma_1(0)\ldots\sigma_1(r_{\sigma_1}), \sigma_2(0)\ldots\sigma_2(r_{\sigma_2}), \ldots)$ *is a path of m if there exists an* $\alpha \in 2^{Act}$ *such that for every* $\sigma_i(r_{\sigma_i})$ *and* $\sigma_{i+1}(0)$ *with* $i \in \mathbb{N}$, *there exists an* $\alpha \in 2^{Act}$ *such that* $(\sigma_i(r_{\sigma_i}), \alpha, \sigma_{i+1}(0)) \in R$;
- *The states in a path are ordered lexicographically such that, for every* $i, j = 1, 2, \ldots$ *and* $\zeta \in [0, r_{\sigma_i}], \xi \in [0, r_{\sigma_j}]$, *there is* $\sigma_i(\zeta) \prec \sigma_j(\xi)$ *iff either* $i < j$, *or* $i = j$ *and* $\zeta < \xi$;
- *For the states in a path, there is:*
  - *for every* $i, j = 1, 2, \ldots$ *and* $\zeta \in [0, r_{\sigma_i}], \xi \in [0, r_{\sigma_j}]$, *there is* $TIME^{\sigma_i(\zeta)} \leq TIME^{\sigma_j(\xi)}$ *if* $i < j$, *and* $TIME^{\sigma_i(\zeta)} < TIME^{\sigma_j(\xi)}$ *if* $i = j$ *and* $\zeta < \xi$;

  In particular, if $TIME^{\sigma_i(r_{\sigma_i})} < TIME^{\sigma_{i+1}(0)}$ *then the transition* $(\sigma_i(r_{\sigma_i}), \alpha, \sigma_{i+1}(0))$ *takes time, otherwise if* $TIME^{\sigma_i(r_{\sigma_i})} = TIME^{\sigma_{i+1}(0)}$ *the transition is executed in zero time;*
- *A path* ρ *terminates if it is a finite sequence* $\sigma_1(0)\ldots\sigma_1(r_{\sigma_1}), \ldots, \sigma_n(0)\ldots\sigma_n(r_{\sigma_n})$. *In such case the first state of the trace* $\sigma_1(0)$ *is denoted by* $first\rho$ *and the last state* $\sigma_n(r_{\sigma_n})$ *is denoted by* $last\rho$;
- *The concatenation of traces* $\rho_1 = (\sigma_1(0)\ldots\sigma_1(r_{\sigma_1}), \sigma_2(0)\ldots\sigma_2(r_{\sigma_2}), \ldots)$ *and* $\rho_2 = (\varsigma_1(0)\ldots\varsigma_1(r_{\varsigma_1}), \varsigma_2(0)\ldots\varsigma_2(r_{\varsigma_2}), \ldots)$, *denoted by* $\rho_1 \circ \rho_2$, *is defined as follows:*
  - $\rho_1 \circ \rho_2 = (\sigma_1(0)\ldots\sigma_1(r_{\sigma_1}), \ldots, \sigma_n(0)\ldots\sigma_n(r_{\sigma_n}), \varsigma_1(0)\ldots\varsigma_1(r_{\varsigma_1})\ldots)$ *iff* ρ_1 *terminates at* $\sigma_n(r_{\sigma_n})$ *and* $(\sigma_n(r_{\sigma_n}), \alpha, \varsigma_0(0)) \in R$;
  - $\rho_1 \circ \rho_2 = \rho_1$ *iff* ρ_1 *does not terminate;*
  - $\rho_1 \circ \rho_2$ *is not defined in other cases;*
- λ *is an empty hybrid trace such that for any arbitrary hybrid trace* ρ, $\rho \circ \lambda = \lambda \circ \rho = \rho$.

The full semantics of continuous time SRML is defined based on Definition 2 and Definition 3, and can be found in [5].

4 Traffic control system and the transformation process

In order to show the application of continuous time SRML, we design a case study of a traffic control system to control the self-driving trains. A traffic control system includes a local traffic control center, a built-in intelligent driving system on a self-driving train, and a monitoring center. Suppose depending on the data obtained from the sensors on the train and the built-in intelligent driving system, the self-driving train can move on their own in most of the driving time without communicating to the outside. But when approaching train stations, they need to get additional information from the local traffic control center, e.g. to know if the station is free, in order to guarantee safe driving. In addition, the monitoring center needs to know the status of the train in some condition, so that they can react in time if the train is not driving safely. In our case study, we show in a traffic control system, how the intelligent driving system on a self-driving train communicates with the local traffic control center and the monitoring center when the train is approaching a train station. The control scheme of the intelligent driving system is adapted from the model of a train system in [17], which extends ETCS [18] level 3 for rail-road crossings.

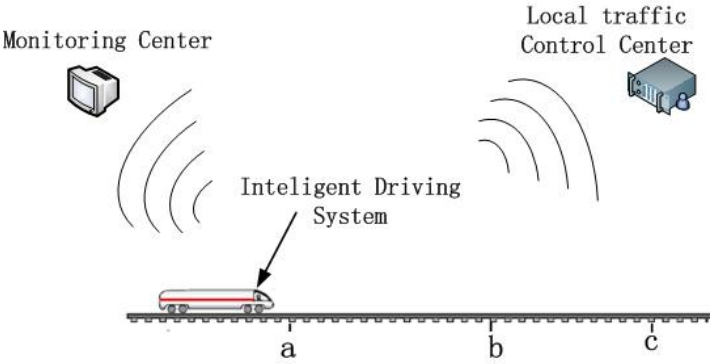


Figure 2: A brief scenario of the traffic control system

```
BUSINESS ROLE IDS is
    INTERACTIONS
        s&r VControl
            _receive newSpeed:speed
        rcv VMoveOn
            _receive newSpeed:speed
        s&r Brake
            _receive dec:acc
        r&s BrakeInfo
            _send currPos:position
                 currTime:time
    ORCHESTRATION
    var C:position, B:acc, d²C:acc, dC:speed, v₀:speed, v₁:speed, t:time
    init C=0∧v₀=100∧v₁=0∧B=0∧t=0 ∧dC=v₀∧d²C=B
    transition Nego
            trigger VControl_receive
            guard dC=v₀
            effect dC=v₁
    transition Corr
            trigger Brake_receive
            guard dC=v₁
            effect d²C=B
    transition Cont
            trigger VMoveOn_receive
            guard dC=v₁
            effect dC=v₀
    transition Stop
            trigger  BrakeInfo_receive
            guard d²C=B
            effect d²C=B
```

Figure 3: Business Role *IDS*

Figure 2 shows a brief view of the scenario. In this figure, we assume that the train station is at point c, and on reaching point a, the train moves with a constant speed v_0. When the train reaches point a, as it needs to move at a lower speed, the intelligent driving system sends a request to the local traffic control center, to ask with which speed the train can move on. Based on the condition of the current traffic condition, the local traffic control center replies with the value of the proper speed. On receiving this message, the intelligent driving system sets a constant speed v_1 of this value to the train. If the local traffic control center finds the traffic is lighter and the train can move with a higher speed, it sends a message to the intelligent driving system with the value of the higher speed. On receiving this message, if the train hasn't passed point b, the intelligent driving system sends another request to the local traffic control center, to ask with which lower speed it can move on in the future. Before the train reaches point b, this process can repeat for several times. When the train reaches point b, the intelligent driving system sends a request to the local traffic control center, to ask with which deceleration the train can brake. Based on the condition of the current traffic condition, the local traffic control center replies with the value of the proper deceleration. On receiving this message, the intelligent driving system sets the deceleration of the train to the required deceleration. After the train brakes, if the intelligent driving system receives the request from the monitoring center, it sends the current condition (position, speed, deceleration, and so on) of the train to the monitoring center.

We abstract the whole system with continuous time SRML as a traffic control service module, and abstract the intelligent Driving System with business role *IDS*, the Local Traffic Control Center with business protocol *TCC*, and the Monitoring Center with business protocol *MOC*. In this paper, we concentrate on the computational reality of a service module, and ignore the interactions caused by communications with the outside parties at the service interfaces, thus, we only make study to the business role *IDS*.

A business role includes a set of interactions and an orchestration in which these interactions are organized. An interaction consists of a type (e.g. *s&r*) and a name (e.g. *VControl*) and attributes of event types (e.g. *_receive*) and parameters (e.g. *newSpeed:speed*). With this structure, we can specify events that associate with an interaction (e.g. *VControl_receive*), and parameters that associate with an event (e.g. *VControl_receive.newSpeed*). An orchestration of a business role includes a set of variables that is local to the service component abstracted by the business role, a set of initial values of the variables that identifies the initial state of the service component, and a set of transitions that model the activities performed by the service component. A transition consists of a *trigger*, a *guard* and an *effect*. A *trigger* is an event or a state condition that specifies the condition for the transition to occur. A *guard* is a condition that identifies the states in which the transition can occur. An *effect* is the proposition that is true in the state that follows the state specified by the guard condition, and represents the effect of the transition. In the declaration of a transition, we don't model the actions caused by the transition (e.g. the assignments of local variables) because they are not considered in the verification procedure.

Figure 3 shows the specification of business role *IDS*. *IDS* contains four interactions and four transition rules. Interaction *VControl* denotes the communication between *IDS* and *TCC*: when the train reaches point a, the *IDS* sends a request to *TCC* to ask for a lower speed, and it can receive the reply from *TCC* with parameter *newSpeed* when the train is between point a and point b, the value of *newSpeed* is then assigned to local variable v_1; interaction *VMoveOn* denotes the communication between *IDS* and *TCC*: when the train is between point a and point b and the traffic is lighter, *IDS* can receive a message from *TCC* with parameter *newSpeed*, indicating that the train can move with a higher speed, the value of *newSpeed* is then assigned to local variable v_0; interaction *Brake* denotes the communication between *IDS* and *TCC*: when the train reaches point b, the *IDS* sends a request to *TCC* to ask for a braking deceleration, and it can receive the reply from *TCC* with parameter *dec* when the train is after point b, the value of *dec* is then assigned to local variable B; interaction *BrakeInfo* denotes the communication between *IDS* and *MOC*: when the train is between point b and point c, and it brakes, *IDS* can receive a request from *MOC* asking the current position of the train and the current time, and it replies to

MOC with parameter *currPos* and *currTime*, which take the values of local variable C and t respectively. Transition *Nego* describes that the train changes its speed from v_0 to v_1 when it receives the slow down message from the local traffic control center; transition *Cont* describes that the train changes its speed from v_1 to v_0 when it receives the speed up message from the local traffic control center; transition *Corr* describes that the train brakes with the deceleration B when it receives the braking message from the local traffic control center; transition *Stop* describes that the train receives the request message from the monitoring center, and replies with its current position and the current time. In business roles, the continuous time executions are specified in the form of differential equations. (e.g. $dC = v_0$ in Figure 3). In general, in continuous time SRML, the n *th* derivative of variable a to time variable t, $\frac{\partial^n a}{\partial t^n}$, is denoted by $d^n a$. The formal semantics of business roles can be found in [5].

Next we take transition *Nego* as an example to illustrate how a transition in *IDS* can be interpreted over a SO-HL^{2}TSs: Suppose a SO-HL^{2}TSs $h = \langle S, s_0, \Sigma, Act, R, AP, L, TIME, \Pi\rangle$ satisfies *IDS* (the satisfaction of a business role over a SO-HL^{2}TS can be found in [5]), $\sigma(r_\sigma) \xrightarrow{\alpha} \sigma'(0)$ is a transition in h ($\sigma, \sigma' \in \Sigma$ and $\alpha \in Act$), if $VControl_receive \in \alpha$ and $\sigma(r_\sigma) \models dC = v_0$, then $\sigma'(0)\cdots\sigma'(r_{\sigma'}) \models dC = v_1$. Particularly, formula $dC = v_1$ represents the differential equation $\frac{\partial C}{\partial t} = v_1$, where t denotes the continuous time variable. Such a formula expresses that variable C changes according to the differential equation $\frac{\partial C}{\partial t} = v_1$ along the trace $\sigma'(0)\cdots\sigma'(r_{\sigma'})$ of h.

In order to be able to test or verify the Traffic-Control module that is explained in this section, we show how to transform the business role *IDS* to a hybrid automata-like DFA. A Deterministic Finite Automata (DFA) [19] consumes a string of input symbols over a set of states, and it operates as follows: when inputting a symbol, the DFA transits from one state to another state, and it does the transitions until all input symbols have been consumed. In a business role, if we take the trigger and guard of every transition as an input symbol, and label the effect of every transition to a state, then we can construct a DFA based on the business role. According to the scenario of the traffic control system and the specification of business role IDS, we can construct the DFA shown in Figure 4.

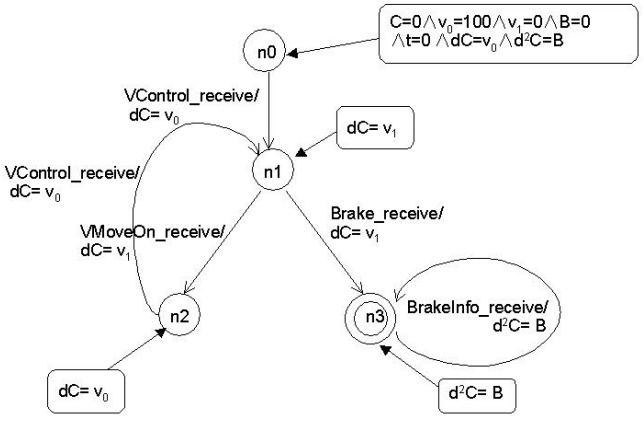


Figure 4: The DFA for business role *IDS*

5 The testing environment

IBM WebSphere Process Server (WPS) is an integration platform for Service Oriented Architecture (SOA), and supports the Service Component Architecture (SCA) programming model. As a process engine, WPS provides a hosting environment for business processes and several Web-based applications. These features of WPS match the structure and the requirements of continuous time SRML modules, so we choose WPS as the basic testing environment.

As the implementing tools, we combine WID with Matlab. In WID, an assembly editor is the central panel where service modules are built by creating and linking services components and service interfaces. A service component is specified with an implementation, some interfaces defining the inputs, outputs and faults of the service component, and zero or several references which denote the interface of other service components connecting to this service component. Service components can be implemented in various types, such as Java objects, business processes, and business state machines. We choose business state machines to implement our Traffic-Control module, because from Figure 5, we find that the hybrid automaton based DFA can be mapped to WID business state machines almost directly. Figure 5 shows the mapping relations between a continuous time SRML module and a WPS service module.

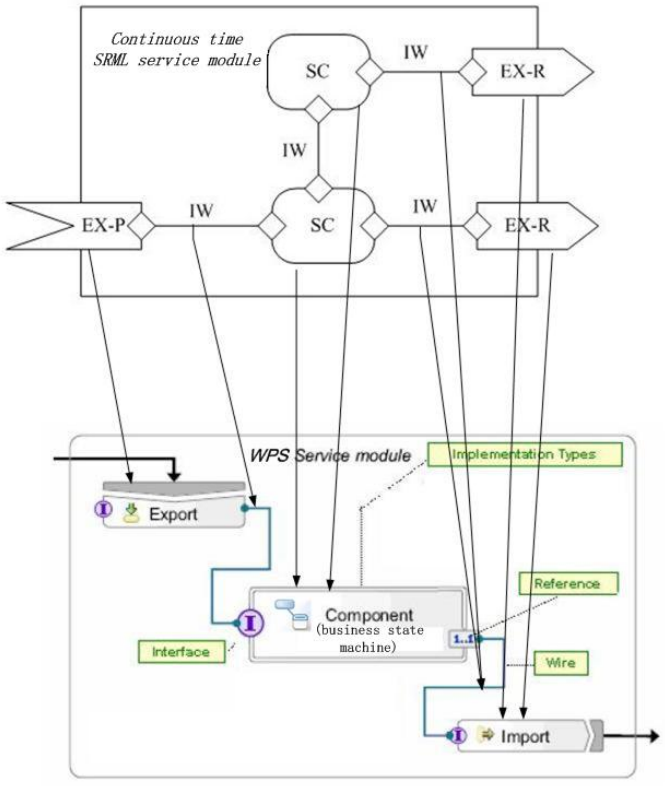


Figure 5: Mappings from continuous time SRML module to WPS service module

In a service component, the continuous processes described with differential equations can be directly implemented with Java code when there are exact solutions of the differential equations. Otherwise, we need to solve the differential equations numerically.

In such case, we combine WID with Matlab to achieve our goal. Matlab [9] is a multi-paradigm numerical computing environment, and it allows interfacing with programs written in languages such as C, C++ and Java. In the latest versions of Matlab, there is a function to wrap programs to Java packages. This function enables us to write Matlab programs to solve the differential equations in the service components in WID, and pass the solutions to the service components by importing the programs packed into Java packages to WID. We'll continue to do the implementation and testing in our future work.

6 Conclusion and future work

In this paper, we give a complete introduction to continuous time SRML, including the semantic domain of continuous time SRML, the application of continuous time SRML, and the testing environment. Continuous time SRML abstracts hybrid system-embedded service-oriented systems at a high level. The semantic foundation of continuous time SRML is SO-HL^{2}TS, over which this modelling language is interpreted. The transitions of SO-HL^{2}TSs enable continuous time SRML to model the communications between services or within a service; and the traces of states of SO-HL^{2}TSs enable continuous time SRML to model the continuous processes performed by service components. Systems modelled with continuous time SRML can be transformed to hybrid automata-like DFA for testing and verification. This can be demonstrated by the case study of the Traffic Control System. By combining IBM WID and Matlab, we are able to implement the hybrid automata-like DFA.

In our future work, we will show the implementation of the hybrid automata-like DFA for the Traffic Control model, and provide the result of testing. We will also try to verify the Traffic Control system with KeYmaera X, which is a newly developed model checking tool for hybrid systems. For the transformation presented in Section 4, we are also working on a systematic approach so that this can be done automatically by programs. Further, we will try to enrich continuous time SRML with the expressions of actions, so that models specified with continuous time SRML match the business state machines of IBM WID better.

References

[1] D. Georgakopoulos, M. Papazoglou, Service-Oriented Computing, The MIT Press, Cambridge, Massachusetts, 2009.

[2] J. Fiadeiro, A. Lopes, J. Abreu, "A formal model for service-oriented interactions," Science of Computer Programming, **77**, 577–608, 2012, doi:10.1016/j.scico.2011.12.003.

[3] J. Fiadeiro, A. Lopes, L. Bocchi, J. Abreu, "The Sensoria Reference Modelling Language," Lecture Notes in Computer Science, **6582**, 61–114, 2011, doi:10.1007/978-3-642-20401-2_5.

[4] M. Wirsing, M. Hölzl, Rigorous Software Engineering for Service-Oriented Systems, Springer, 2011.

[5] N. Yu, Injecting Continuous Time Execution into Service-Oriented Computing, Munich University , Ph.D. thesis, 2016.

[6] N. Yu, M. Wirsing, "The Application of Continuous Time SRML," in Proceedings of 2020 ICICSE&ICACTE, 99–103, 2020, doi:10.1145/3424311.3424316.

[7] J. Abreu, Modelling Business Conversations in Service Component Architectures,University of Leicester, Ph.D. thesis, 2009.

[8] IBM, IBM Integration Designer V8.5.5 documentation, https://www.ibm.com/support/knowledgecenter/SSTLXK_8.5.5/com.ibm.wbpm.wid.main.doc/kc-welcome.html, 2014.

[9] MathWorks, https://www.mathworks.com/,2014-2021.

[10] T. A. Henzinger, "The theory of hybrid automata," in Proceedings of 11th Annual IEEE Symposium on Logic in Computer Science, 1996, doi:10.1109/LICS.1996.561342.

[11] K. Cordwell, A. Platzer, "Towards Physical Hybrid Systems," in Proceedings of International Conference on Automated Deduction, CADE'19, 216–232, 2019, doi:10.1007/978-3-030-29436-6_13.

[12] A. Platzer, Logical Foundation of Cyber-Physical Systems, Springer,Cham, 2018.

[13] A. Platzer, KeYmaeraX: An aXiomatic Tactical Theorem Prover for Hybrid Systems, http://www.ls.cs.cmu.edu/KeYmaeraX/index.html, 2021.

[14] M. Fadlisyah, E. Ábrahám, P. C. Lepri, D. Ölveczky, "A Rewriting-Logic-Based technique for modelling Thermal Systems," in Proceedings of First International Workshop on Rewriting Techniques for Real-Time Systems, PTRTS 2010, 82–100, 2010, doi:10.4204/EPTCS.36.5.

[15] N. Yu, M. Wirsing, "A SOC-Based Formal Specification and Verification of Hybrid Systems," in Proceedings of the 22nd International Workshop on Algebraic Development Techniques (WADT2014), 151–169, 2015, doi:10.1007/978-3-319-28114-8_9.

[16] R. De Nicola, F. Vaandrager, "Three Logics for Branching Bisimulation," Journal of the ACM, **42(2)**, 458–487, 1995, doi:10.1145/201019.201032.

[17] A. Platzer, J. Ousel, "European Train Control System: A Case Study in Formal Verification," in Proceedings of 11th International Conference on Formal Engineering Methods (ICFEM 2009), 246–265, 2009, doi:10.1007/978-3-642-10373-5_13.

[18] The Worldwide Railway Organization (UIC), The European Train Control System, https://uic.org/etcs#Standards-and-Specifications, 2015.

[19] J. E. Hopcroft, R. Motwani, J. D. Ullman, Introduction to Automata Theory, Languages, and Computation (3rd Edition)., Addison-Wesley Longman publishing Co., Inc., USA., 2006.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 66-75 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Intermediation in Technology Transfer Processes in Agro-Industrial Innovation Systems: State of Art

Leidy Dayhana Guarin Manrique[1,*], Hugo Ernesto Martínez Ardila[2], Luis Eduardo Becerra Ardila[2]

[1]*School of Electrical, Electronic and Telecommunications Engineering, Industrial University of Santander, Bucaramanga, 680002, Colombia*

[2]*INNOTEC Research Group, School of Industrial and Business Studies, Industrial University of Santander, Bucaramanga, 680002, Colombia*

ARTICLE INFO

Article history:
Received: 16 July, 2021
Accepted: 21 October, 2021
Online: 23 November, 2021

Keywords:
Intermediation
Technology transfer
Innovation
Innovation systems
Agro-industrial innovation systems

ABSTRACT

Considering the importance of addressing innovation issues that impact the development of economic sectors, this document presents the research work aimed at establishing a state of the art related to technology transfer and intermediation issues, which can be adopted in the innovation systems. In this case, the agro-industrial innovation system is taken as a reference from the perspective of a country, Colombia. Likewise, it is proposed to consider the synergy between actors such as the university and the industry, from a holistic vision of the systems. In this sense, and making use of the Scopus, Web of Science and Google Scholar databases, through the implementation of a methodology in which three main phases of search, selection and reading of scientific publications were generated, a set of documents was obtained, and through these it was possible to identify: concepts on innovation systems, aspects that intervene in transfer processes of technology especially related to the articulation of the actors that are part of the innovation systems, as well as the way in which these issues can be adopted, taking the agro-industrial sector as a reference. Thus, through this research, the existence of structural gaps in the networks of actors is highlighted as a key factor, which, when trying to be moderated through the intervention of government actors, also require the participation of intermediary innovation agents, that facilitate the articulation and flow of data, information and communication between those who develop technology and those who require these developments to mitigate a need in the productive sectors.

1. Introduction

This paper is an extension of the work "Innovation systems supported in technology transfer processes" that is published in 2020 *International Congress of Innovation and Trends in Engineering* (CONIITI) [1].

In Colombia, agribusiness is identified as a key sector for the sustainable development of the country [2], which is why the Law 1876 of 2017 has been approved, to support the creation of the National Agricultural Innovation System of Colombia where articulation, market orientation, sustainable development, associativity, intellectual property, development and innovation, and producers as research agents are highlighted as principles [3]. Additionally, this study seeks to contribute from the approach of technology transfer processes associated to the promotion of innovation, initiatives that seek the consolidation of actions related to the fulfillment of the Sustainable Development Goals [4], and the adoption of challenges that try to contribute to the food assurance of the population by 2050, which is contemplated to be close to 9,100 million people [5].

In this sense, it is proposed that it is from the academic and industrial actors, as well as the identification of intermediary agents, mechanisms, or tools, that actions aimed at the articulation of actors in innovation systems are adopted; and the generation of research results, as well as their dissemination and appropriation, contribute both, to the sustainable development of the territory and the strengthening of innovation activities in the agro-industrial sector. Dissemination of research results is understood as the activity that makes known to the public or interested third parties, the results of the research carried out by a knowledge-generating actors [6].

*Corresponding Author: Leidy Dayhana Guarin Manrique, leidy2127825@correo.uis.edu.co

https://dx.doi.org/10.25046/aj060610

Therefore, and based on the importance that technology transfer processes have in the development of innovation systems of different geographical or spatial location, and sectoral scope, this work presents a set of concepts highlighting as a key aspect identified in the research carried out, the need to take into account the incorporation of innovation intermediary agents who through their contacts, networks, capacities, and experience are seen as those facilitators that can mitigate the gap of social distancing between the actors that can be part of the technology transfer processes in different sectors. Likewise, the novelty of this document is also associated with the analysis of the case of an agroindustrial innovation system that is visualized from government initiatives as a potential system to generate opportunities for articulation around innovation initiatives, which may finally arrive to have an impact in the markets.

2. Analysis Scenario

In the understanding of the complex nature of innovation systems in the agribusiness sector, within the framework of technology transfer processes, the intermediary agents are identified as actors of special interest for the analysis of synergies, because they assume a key role in the articulation of the various actors that make up these systems [7]. In this regard, universities should disseminate their research results through their technology transfer offices, which can contribute or be aligned with the needs and requirements of the industry. However, there are barriers related to the connections or articulation between actors of an innovation system [8], which hinder the development of dissemination activities.

Likewise, and based on the interaction between the university-industry-goverment and civil society actors that are part of the innovation systems [9], there are other barriers identified in the literature, which at a technological level hinder the implementation of activities of innovation-oriented to generate products and services with added value in the markets. One of the barriers is the asymmetry of information, which refers to the fact that both, the receiver and the sender in a technology transfer process do not have complete information on the technology developed and its exploitation possibilities to be able to reach a transfer agreement that satisfies the parties involved [10].

Based on the above, it is possible to identify problems associated with the knowledge and information necessary to establish a common language among the actors that participate in the technology transfer processes [9], which can lead, for example to lack of financing for technological development by potential investors, and the low opportunity for commercialization of technologies from universities.

Consequently, it is proposed that the lack of articulation of actors such as the university and industry in an innovation system, can negatively influence the generation of information flows between these two actors and, consequently, raise the levels of information asymmetry in a technology transfer process.

3. Methodology

The results obtained in this work have been framed in the execution of the phases presented in Figure 1.

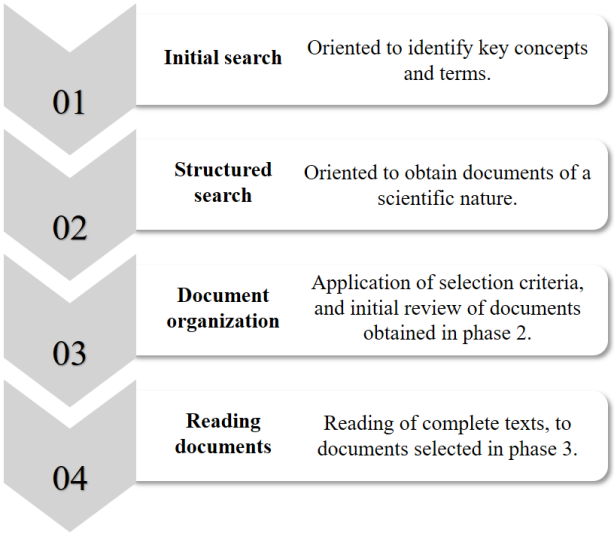


Figure 1: Research methodology..

In the first phase, the search for information was carried out with the purpose of identify the first set of documents related to technology transfer and intermediation in innovation systems. Consequently, from this information, key terms were identified for structuring the search equation and relevant authors on the aforementioned topics. In this search, the Google Scholar[1] tool was used, understanding that it is a global search engine that is constantly being updated, and in which recognized authors and leading institutions in the publication of various topics can be identified.

In the second phase, the selection of the specialized databases Scopus and Web of Science was carried out[2]. These databases are considered robust in the global scope of publications they handle, as well as the multidisciplinary of the areas they cover, which allows addressing the issue of intermediation in technology transfer processes in a broad way. Likewise, it is highlighted that both Scopus and Web of Science handle quality documents, that is, the documents that are published are reviewed and validated by experts.

The search was executed on October 26, in 2019, and the used general equation was: TITLE-ABS-KEY (("technolog* transfer" OR "know* transfer" OR "technolog* exchang*" OR "know* exchang*") AND ("structur* mediat*" OR intermediar* OR gatekeep* OR broker* OR liaison* OR betwenness OR bridg* OR mediat* OR coordinat*) AND (universit* OR academic)). The selection criteria applied to the results of this review were: 1) Spanish and English languages; 2) Include indexed articles in peer-reviewed scientific journals; 3) Articles belong to the citation databases: Social Sciences Citation Index (SSCI); 4) Select documents that are classified as articles or reviews, foreseeing that

[1] For more information: https://guiasbuh.uhu.es/Google_Scholar_Citations/Ventajas_y_desventajas_de_GSC

[2] For more information: https://www.recursoscientificos.fecyt.es/licencias/productos-contratados/wos

they must necessarily be reviewed by peer reviewers, representing an aspect of quality.

In the third phase, the results were grouped into two sets, which were: (i) Documents in the period from 2017 to 2019, which were read title and abstract, to select those associated with the subject under study, giving a total of 15 documents to be read completely; and, (ii) Documents obtained by adding (rev * OR meta *) to the equation, to which the title and summary were read to select those associated with the subject under study, giving a total of 52 documents that would be completely read.

In the last phase of the methodology, the complete documents were read, and information related to the topics of innovation systems, technology transfer, agro-industrial innovation systems, and intermediation was obtained. Finally, government-type sources associated with science, technology, and innovation issues and the agro-industrial sector in Colombia were reviewed.

In short, the protocol of the review to identify works associated with technology transfer and intermediation processes, included the following steps:

- Preliminary search in the Google Scholar database in order to identify relevant studies and keywords.
- Definition of keywords to search, establish and refine the search equation.
- Search for information in Web of Science and Scopus.
- Categorization and tabulation of results, including the number of publications according to countries, years, areas, in addition to the analysis of the most relevant journals, authors and keywords, in order to identify the guidelines followed by research on the subject of interest.
- Identification of articles to be analyzed applying the selection criteria, by reading the abstracts, titles and keywords.
- Reading and analysis of the documents.

4. State of the Art

Based on the results obtained from the methodology, two approaches have been found; the first, associated with the behavior of scientific knowledge as well as those publication sources that are relevant when reviewing works associated to the subject, and secondly, the most relevant issues associated with the action of the intermediary agents in technology transfer processes.

4.1. Scientific publications

On the one hand, regarding scientific publications, as reflected in Figure 2, technology transfer issues and intermediation actions related to the aforementioned processes have gained momentum during the last decade, highlighting the years 2017 and 2018 (until the moment of the search) as those spaces of time where the researchers found the topic more relevant. An example is found in the university context, where through intermediary agents such as their technology transfer offices, they seek to bring the inventions developed in the academy to the markets [9].

On the other hand, with respect to the most prominent journal sources obtained from the Scopus and Web of Science search, are shown in Table 1.

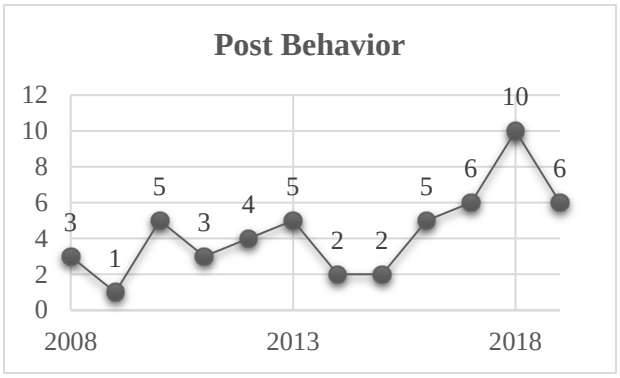


Figure 2: Publication in the time. Own elaboration based in Scopus and Web of Science.

Table 1: Most Outstanding Publication Sources. Own elaboration based in Scopus and Web of Science.

Source title	Number of Publications
Journal of Technology Transfer	5
Journal of Knowledge Management	4
Administrative Sciences	2
Technology Analysis & Strategic Management	2
Technological Forecasting and Social Change	2

Regarding the Journal of Technology Transfer[3], it is identified that its focus is given by issues related to the study of technology transfer, from multiple dimensions of analysis, which can range from the identification of cases to issues of regulations or policies associated to technology transfer processes.

Another journal that has been highlighted is the Journal of Knowledge Management [4] which relates mainly to knowledge management and innovation derived from different analysis contexts.

the following section is presented based on the results obtained from reading the documents found in the search for information, including those mentioned in the previous journals and during the period from 2008 to 2021.

4.2. Key information identified

For this second part, the following information compendium has been proposed, related to the topics of technology transfer, intermediation and agro-industrial innovation systems.

4.2.1. Technology Transfer

Technology transfer is understood as the process in which knowledge and technologies are disseminated between a holder and an interested party to be appropriated, used, or exploited [11],

[3] For more information: https://www.springer.com/journal/10961

[4] For more information: https://www.emeraldgrouppublishing.com/journal/jkm

[12], [13], generally seeking to be brought to the market [14], in exchange for an incentive or monetary value [15].

Within the framework of these technology transfer processes, the following factors are taken into account: the generation of innovation in organizations [16], social capital [17] involved in the social networks of the actors who participate in the technology transfer [18], their interactions [19] and their previous relationships [20]; the university-industry relationships [21], the formality of the process (which can be formal or informal) according to knowledge management and the interactions between actors [17], [22], the limits of organizations and incentives [23], economic benefits e.g. profits on patents [24], technological cooperation between organizations [24], patents seen as a transfer mechanism [18], [25], the business culture that universities can adopt [21], the option of allowing students to participate in industry [17], the culture of the parties [17], legal frameworks [24], the uncertainties of innovations involved in technology transfer processes [26], and the importance of considering the intellectual property in the processes [13], [27]. This intellectual property can be seen materialized in the adoption of technology transfer mechanisms such as the creation of spin-off, patent licensing, collaboration actions between academia and industry [28], and the training of professionals with knowledge that contribute to the solution of industry needs [29].

Besides, when seeking to identify whether the technology transfer has been successful or not, it is found that depends of the accumulation or assimilation of knowledge in organizations that adopt the technologies that have been transferred [17]. In turn, at the macroeconomic level, technology transfer can be measured through business creation and job creation [30]. Additionally, other ways of identifying the success of technology transfer have been through the licensing of inventions [20], [24], incentives for personnel who participate in technology transfer processes [23], the indicators of patent families [24], and knowledge intensive entrepreneurship at universities [30].

However, in addition to factors in the technology transfer process, barriers are also identified that should be mitigated, some of these are: the lack of incentives from universities [9] and industry [16], cognitive and organizational distances between the university and industry [20], the time required to implement technological development activities, absorption capacity, the characteristics of knowledge, context and actors, and the culture of the parties participating in the process [17], the common language between the actors [9], the level of trust [17], [31], and the asymmetries of information between the parties [32], [33].

Finally, based on both the factors and the barriers that must be managed, it is necessary to understand that among the actors and medium involved in the process, it is possible to distinguish the sender that seeks to transfer the technology to a receptor, through a medium or mechanism, known as a channel [12]. In some scenarios, the channel may not exist, generating in the aforementioned process, a structural void that indicates the need for an agent to mitigate the gap in the flow of information, knowledge, and technologies between those who participate in the technology transfer processes. This agent is known as an intermediary [34], [35].

4.2.2. Intermediaries in the technology transfer

Intermediaries are agents that connect the components that structure an innovation system [36], also known as facilitators [37], who within their objectives seek to optimize interaction [38] to generate trust among the system's actors [39], strengthen interrelation processes, and at the same time mitigate the structural gaps that are identified in the social network of its actors [34].

At a general level, intermediaries are recognized as actors of special interest in innovation systems, for their intervention in actions aimed at mitigating barriers in technology transfer processes. They have key elements that help to the technology transfer process such as contacts, previous relationships with inventors, and experiences in technology commercialization [20]. In the case of the academic context that seeks to transfer technology to the industrial sector, the commercialization processes of the research results of a university, are more probable to achieve if there are links with industry [40]. In some cases, these links can be facilitated through the intervention of intermediary agents.

Thus, with regard to intermediation, the following variables have been analyzed in studies associated to the articulation between the university and industry actors: the experience of the technology transfer offices, which influences research and development agreements in the university [41], the previous relationships between intermediary agents and actors that participate in technology transfer processes [20], the performance of knowledge transfer offices measured through a number of supported patent applications, monetary income obtained from university-industry collaborations, and the creation of spin-offs [42], the efficiency of the technology transfer offices of the universities based on their age [43], [44], and the performance of technology transfer offices measured through support in licensing, patenting and Spin-off creation activities [45].

An example of intermediation is the case of technology and science parks. They support the creation of networks, especially those where university plays an important role due to both its link to the park and the legitimacy they provide to companies through their image of reliability and reputation [46]. Another example are the technology transfer offices, which within their functions provide advice on legal protection issues [45] of the technologies that are waiting to be brought to the market, in addition to providing support in the creation of technology-based companies, technological marketing [45], and assessment and due diligence in commercialization tasks of technological developments.

However, some studies have focus on the idea that an intermediary agent could only have a particular interest in obtaining economic benefits from technological results [37], beyond considering its action as positive in aspects of knowledge transfer, articulation [7] and interaction [47]. Consequently, the purpose of the intermediary agent could be an important barrier that has been associated to their nature, whether they really seek to generate linking actions, or simply act with individual commercial or political interests [48].

Therefore, it is from the nature of the intermediary agents and the objective they pursue, that they act in the configuration of the different innovation systems, especially when taking into account

the different barriers that limit the successful result in terms of technology transfer [20].

4.2.3. Agro-industrial innovation systems

The concept of agro-industrial innovation systems considers the definition of a system that refers to the set of components that are interrelated with a defined objective in a given context [49], in accordance with the synergy that it experiences [50], [51]. Furthermore, when systems are complemented with the term innovation, they open the space to identify the possible achievement of competitive advantages [52]. When it comes to competitive development issues in the countries, innovation should be aimed at promoting the transfer of knowledge to reality considering the results of research, observation, discoveries, or intuitions that are generated in the nation [53].

Therefore, the conceptualization of innovation systems is open and unfinished [54]. However, some characteristica are critical such as it incorporates actors of different nature with the purpose of exchange of technologies [55], have different scope at the geographical or sectoral level [51], [56], [57] it is associated with social networks [58], and other aspects such as synergy management [58], knowledge management processes [59], [60], and the procurement of economic advantages [61].

In the case of sectoral innovation systems, which are those of interest for this research associated with the agro-industrial sector, the following stand out:

- Actors are related based on the economic sector to which their activities are associated [62].
- A differential element of these systems is the knowledge associated with innovation processes [63] in the sector.
- The key factor for the development of these innovation systems is the group of companies involved in accordance to the conditions of the sector [64].
- A set of elements of interest that sustains the development of innovation systems at the sectoral level are related to social networks, capacities of industry actors, the policies that direct their development [65], assimilation and knowledge dissemination [66], and finally, the supply chain [67].

According to the analysis level different study dimensions can be proposed as possible approaches to understand the innovation processes and its possible obstacles [53].

A representation of the set of dimensions and elements that can be considered for study purposes of innovation systems in the agro-industrial sector, and that have been previously adopted in the research activities [68], [69] in Colombia, can be seen in Figure 2.

Figure 2 shows that the dynamics of the innovation system is supported by the integration of different elements, such as a specific context, dimensions that make it possible to study the dynamics of the system from different perspectives, agents that externally intervene in the system, system behavior, and a set of interrelated actors represented under the concept of the quadruple helix [70] within the framework of a knowledge economy [71], [72]. These interrelationships are shown as a key attribute of the system, since from them, it is possible to identify realities of the different sectors [53] and generate networks or strengthen those that already exist, thus enabling information flows, knowledge and technology transfers, the realization of agreements, the establishment of collaborations, and articulation actions organized around a purpose. However, these interrelationships may require the intervention of facilitating agents, which generate links between the actors of the system [39], [73], [74]. These agents identified as intermediaries, facilitate the creation of networks of actors promoting collaboration and the flow of knowledge and information [20], [36], [37], [39], [42], [75]-[82].

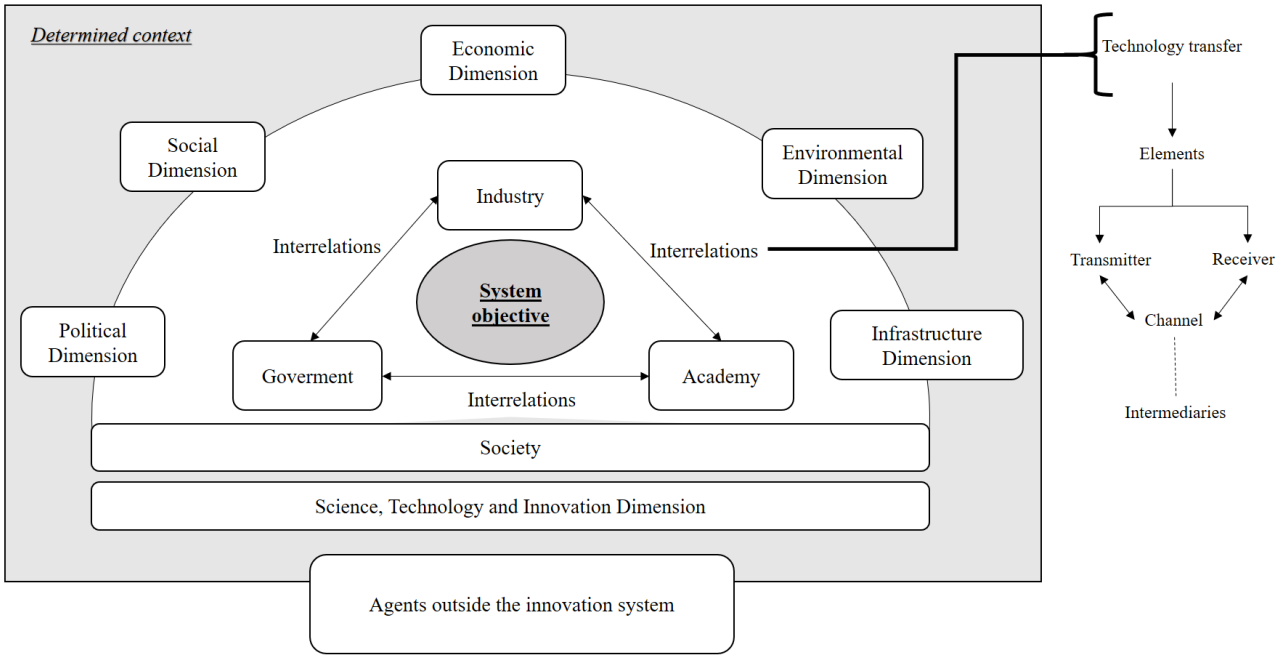


Figure 3: Dimensions and elements in innovation systems. Based on [53], [69].

Based on what is illustrated in Figure 2, it is found that there are similarities with other models of innovation systems, (e.g. the Tropical Agriculture Platform) [83]. The following are highlighted as common key elements:

- The proposal of a specific context, made up of institutional actors and innovation policies, which interact to achieve innovation objectives applied to the agricultural sector. In this regard, public and private research and technological development organizations stand out as key actors [22].
- The existence of institutions that act as intermediaries between the actors that carry out research, and the industry that commercialize the research output. An example is the technology transfer offices, whose mission is to promote technology-based entrepreneurship [74], [77], [84], [85], facilitate the commercialization of knowledge, manage intellectual property [9], [86], provide advice on inventions, manage academic human capital, promote contract research and promote articulation between the different actors of the university, industry, and government [86].
- Interaction with actors from other sectors, which according to their own purposes come to influence the synergy of innovation activities that take place in the sector system.
- The proposal of a political dimension, which guides both the planning and the execution of the guidelines that regulate the synergy of the system.

From the articulation of actors in the innovation systems, the sectoral innovation system like the agro-industrial, can achieve results in processes of diffusion, adoption, and appropriation of technologies among the university, industry, government, and society in general.

In addition, the value of including intermediary agents in technology transfer processes is also recognized, anticipating the possible existence of gaps in terms of social networks that emerge from the dynamics of these systems, understanding that within their role as articulatory agents, they allow generating a higher level of cohesion in the relationship of the various actors in the system.

5. Discussion

Faced with the development of innovation processes and technological management, in 2019 Colombia ranks 67th out of 129 countries in the global innovation index, from which aspects such as infrastructure, market, business sophistication, institutions, knowledge and technology production, human capital, researching, and creative production, were analyzed [87], [88]. In addition, as a result of this evaluation, in terms of collaboration activities between the university and the industry to develop research, the country ranks third in the Latin American region [88]. In the same orientation towards technology management processes, at the national level, it finds that most of the innovations in Colombia are developed from the research actions of the universities [89].

However, despite identifying that innovation represents an opportunity for economic development, the country requires an optimal infrastructure and necessary conditions to carry out technology transfer processes. In addition, it was identified in the science, technology, and innovation policies, that there is a low level of intermediary agents such as technology transfer offices, incubators, and science and technology parks, which by their nature support the linkage, interaction, and flow of information and knowledge between the actors of the quadruple helix [90]. The existence of a low capacity of the industrial sector to articulate and transfer knowledge and technologies among the actors of the innovation system in Colombia is pointed out in this policies. Thus, in the national development plan of Colombia, it is identified that only 3.8% of companies work collaboratively with universities, a figure that has influenced the proposal to reach a goal of 59 technology transfer agreements between the university and the industry in the country, in the period 2018-2021 [91].

From the approach of sectoral innovation systems associated with agribusiness in Colombia, the Strategic Plan for Agricultural Science, Technology and Innovation has been established [92], which takes into account the need to generate strategies aimed at promoting programs of transfer and adaptation of technologies associated with the different productive sectors, in which the proposal of mechanisms oriented to generate articulation between the different actors of the agricultural innovation system should be included. In this sense, from the Colombian Agricultural Research Corporation (Agrosavia), the Ministry of Agriculture and Rural Development of Colombia, and the Ministry of Science, Technology, and Innovation of Colombia, in the formulation of the Strategic Plan for Agricultural Science, Technology and Innovation [92], formulate two actions focusing on science, technology and innovation in the area of technology transfer, technical assistance and innovation [92], which are: (i) “the development, adjustment, and validation of extension tools and specialized comprehensive technical assistance ”, and (ii) “the generation of collaborative workspaces through territorial innovation systems”.

Based on the previous approaches, and understanding the value chain made up of the primary sector associated with agricultural activities, and the secondary sector related to agribusiness activities, the Santander agro-industrial innovation system is proposed as an example, in which numerous technological developments and research results carried out by universities have been identified, which tend to be unknown by the sector and have not had the possibility of being commercialized, particularly due to the lack of articulation between the actors that are part of the innovation system [93]. In addition, there are shortcomings in the information flows required from both the receiver and the sender side, to develop technology transfer processes from the university to the agro-industrial sector.

Another case of the agricultural sector in which similar problems have been studied in technology transfer processes, and in which the intermediary actions of science parks, technology transfer offices, consultants, and in general, intermediary innovation agents have favored its mitigation in the productive and market contexts, is reflected in the Dutch case [37] where intermediaries act as facilitators or forgers of networks of actors that interact under the synergy of transfer processes.

Based on the above, it is necessary to understand that when technology transfer processes are developed, elements associated with information and knowledge are involved. Therefore, it is

necessary to take into account the existence of cognitive differences between the actors who seek to interrelate [80] around these processes. These cognitive differences can be related to the information asymmetry barrier, since it represents the gaps in terms of the information that a technology owner requires to determine the interest of a second party in its appropriation, as well as the information that the recipient needs to adopt it [10] and therefore use, exploit or commercialize it. In this sense, this barrier can lead to inefficiencies in technology commercialization activities between the university and industry [94].

Overcoming the asymmetry of information can facilitate the investment activities of industry players in the technological developments of the university, which can be considered as university ventures [95], thus opening the possibility of promoting commercial management actions among inventors or holders of technological developments and representatives of the industry, where intellectual property rights play a key role [20].

Additionally, it is important to bear in mind how the information asymmetry barrier is mitigated in the technology transfer process between the university and the industry, and how it will depend on the inventors' ability to disseminate [74], the absorption capacity of industry representatives [96], and the participation of intermediary agents that moderate information flows.

Therefore, it is proposed that the experience in commercialization and the social networks of intermediaries, as well as their previous relationships with inventors, can help reduce the information asymmetry between the senders and recipients of a technology transfer process [20]. In the case of universities among the factors that can be taken into account in the process of reducing the level of asymmetry of information is its prestige, since it generates more trust among the actors of the different economic sectors [97].

6. Conclusions

Based on the information presented above, it is necessary to highlight the importance of encouraging the different actors of the agribusiness innovation systems, to develop articulation that in addition of being focused on technology transfer processes, also generate alliances between the university and the industry, both in research processes and in the dissemination and exploitation of their results. Likewise, and anticipating the possible existence of structural gaps in the social networks of these systems, the option of identifying and interacting with intermediary actors is proposed as an alternative, understanding that their intervention in the technology transfer processes between the university and the industry can be assumed as a strategic factor that would contribute to the economic development of a region.

In front of this scenario, it is also necessary for government sector actors to contribute to the strengthening of the synergy of innovation systems, through regulatory mechanisms focused on expanding articulation actions and alliances between the actors of the aforementioned systems, to promote territorial development from the different economic sectors of the regions.

When it comes to contributing to economic sectors, including the agro-industrial sector, innovation is evidenced as a key factor in strengthening the markets. In addition to generate competitive advantages in the market, it can be associated to the processes of technology transfer between the university and the industry. A whole development takes place that starts from the scientific research and matures until reaching a viable product or service to be offered to the industry. Therefore, in this scenario, it is proposed that the university should have capacities to disseminate its technologies, and the industry must have capacities to adopt the technologies that are transferred to it, the foregoing to achieve, as a result, the use, exploitation, or commercialization of inventions and technological developments generated by universities.

In addition, the level of connectivity that the actors representing the university and industry can achieve also affects technology transfer processes, which is influenced by the asymmetry of information generated by the parties involved in the process of technology transfer, which is why an agent that acts as an intermediary between the university and the industry is identified as a third actor, that manages to moderate the asymmetry of information and at the same time make the transfer process effective by facilitating the flow of information between the parties involved in the process.

Based on the above, and on what was stated during the document, it is necessary to take into account that technology transfer processes can use different mechanisms that range from the creation of spin-offs to the generation of agreements between universities and companies, where the human capital of the universities brings their knowledge to the industry and reinforces their capacities through the presumption of real market scenarios. However, it is also relevant to understand that technology transfer processes include a set of factors and barriers that must be viewed from a multidisciplinary approach, for which hybrid capabilities between the university and industry require an effort to be developed by the actors that participate in these processes, thus allowing the intervention of intermediary agents who, due to their nature, can support the management of actions against the factors and barriers mentioned.

7. Future work

It is recommended to generate research actions around topics such as:

The evaluation of the effect generated by the different types of intermediary actors in the technology transfer processes in the agro-industrial sector, and also in other sectors of interest to the various actors that make up the innovation systems at the national, regional, or sectoral level.

Generate alliance options around various research and technological development projects, executed between the university and the industry, in such a way that their effectiveness can be evaluated in terms of technology transfer processes.

The dissemination of the results that the different types of intermediary agents have achieved in technology transfer processes, in such a way that in the second phase of this activity, an initial set of indicators can be proposed that allow measuring the impact of the intervention of these actors in the development of the systems of innovation.

Regarding future lines of research, it is important to focus towards the study of the nature of the actors that are part of the technology transfer processes, in such a way that it can be evidenced the types of academic, industrial and hybrid capacities that allow facilitating the transfer of inventions that are developed in the university to the market.

Another alternative study is identified in the need to generate university-industry articulation methodologies, in which the objective is to identify market challenges that can be developed by the academy.

References

[1] L. Guarin, H. Martínez, "Innovation systems supported in technology transfer processes," in 2020 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI), 1–9, 2020, doi:10.1109/CONIITI51147.2020.9240319.

[2] DANE, "Boletín técnico - Producto Interno Bruto PIB -Cuarto trimestre de 2017," 28, 2018.

[3] MinAgricultura, Ley 1876 por medio de la cual se crea el Sistema Nacional de Innovación Agropecuaria y se dictan otras disposiciones. Congreso de Colombia, 29 de diciembre de 2017, 1–28, 2017.

[4] PNUD, Proyecto desarrollo económico incluyente, 2011.

[5] FAO, Cómo alimentar al mundo en 2050, 2009.

[6] J.J.S. Gil, "La difusión del conocimiento como un elemento clave de la innovación y la competitividad en el entorno universitario bajo el sistema Open Journal System (OJS)," International Journal of Information Systems and Software Engineering for Big Companies (IJISEBC), **4**(1), 63–71, 2017.

[7] M. Ramirez, I. Clarke, L. Klerkx, "No Analysing intermediary organisations and their influence on upgrading in emerging agricultural clusters," Environment and Planning A: Economy and Space, **50**(6), 1314–1335, 2018, doi:10.1177/0308518X17741316.

[8] AGROSAVIA, Misión para la transformación del campo. Diagnóstico en CTeI, 2015.

[9] E.S. Olaya, J. Berbegal-Mirabent, O.G. Duarte, "Performance of university transfer offices as intermediary for knowledge market development [Desempeño de las oficinas de transferencia universitarias como intermediarias para la potencialización del mercado de conocimiento]," Intang. Cap, 2014, doi:10.3926/ic.497.

[10] OMPI, Comité permanente sobre el derecho de patentes. Decimocuarta sesión., 2010.

[11] B. Bozeman, H. Rimes, J. Youtie, "The evolving state-of-the-art in technology transfer research: Revisiting the contingent effectiveness model," Research Policy, **44**(1), 34–49, 2015, doi:10.1016/j.respol.2014.06.008.

[12] C. Battistella, A.F. De Toni, R. Pillon, "Inter-organisational technology/knowledge transfer: a framework from critical literature review," The Journal of Technology Transfer, **41**(5), 1195–1234, 2016, doi:10.1007/s10961-015-9418-7.

[13] C.S. Hayter, E. Rasmussen, J.H. Rooksby, "Beyond formal university technology transfer: innovative pathways for knowledge exchange," The Journal of Technology Transfer, **45**(1), 1–8, 2020, doi:10.1007/s10961-018-9677-1.

[14] F. Donneys González, B. Blanco Campins, "La transferencia de tecnología en universidades colombianas," Economía y Desarrollo, **157**(2), 182–198, 2016.

[15] F.C. Sercovich, "Dependencia tecnológica en la industria Argentina," Desarrollo Económico, 33–67, 1974, doi:10.2307/3466047.

[16] T.D.H. Gusberti, A.C. Bretas, "Diagnosis of the market for ideas and the role of industrial associations as intermediaries in the Brazilian context," Industry and Higher Education, **32**(1), 9–22, 2018, doi:10.1177/0950422217733087.

[17] C. Van Horne, D. Vincent, "Challenges in technology transfer: An actor perspective in a quadruple helix environment," The Journal of Technology Transfer, **42**(2), 285–301, 2017, doi:10.1007/s10961-016-9503-6.

[18] Y. Ye, K. De Moortel, T. Crispeels, "Network dynamics of Chinese university knowledge transfer," The Journal of Technology Transfer, **45**(4), 1228–1254, 2020, doi:10.1007/s10961-019-09748-7.

[19] P. D'Este, O. Llopis, F. Rentocchini, A. Yegros, "The relationship between interdisciplinarity and distinct modes of university-industry interaction," Research Policy, **48**(9), 1–57, 2019, doi:10.1016/j.respol.2019.05.008.

[20] Y.-C. Kim, M. Rhee, R. Kotha, "Many hands: The effect of the prior inventor-intermediaries relationship on academic licensing," Research Policy, **48**(3), 813–829, 2019, doi:10.1016/j.respol.2018.11.007.

[21] P. Centobelli, R. Cerchione, E. Esposito, "Exploration and exploitation in the development of more entrepreneurial universities: A twisting learning path model of ambidexterity," Technological Forecasting and Social Change, **141**, 172–194, 2019, doi:10.1016/j.techfore.2018.10.014.

[22] S. Zaichenko, "The human resource dimension of science-based technology transfer: lessons from Russian RTOs and innovative enterprises," The Journal of Technology Transfer, **43**(2), 368–388, 2018, doi:10.1007/s10961-017-9567-y.

[23] S. Horner, D. Jayawarna, B. Giordano, O. Jones, "Strategic choice in universities: Managerial agency and effective technology transfer," Research Policy, **48**(5), 1297–1309, 2019, doi:10.1016/j.respol.2019.01.015.

[24] C. Martinez, L. Bares, "The link between technology transfer and international extension of university patents: Evidence from Spain," Science and Public Policy, **45**(6), 827–842, 2018, doi:10.1093/scipol/scy008.

[25] D. Chatterjee, L. Leisyte, S. Dasappa, B. Sankaran, "University research commercialization in emerging economies: A glimpse into the 'black box,'" Science and Public Policy, **45**(3), 361–372, 2018, doi:10.1093/scipol/scx066.

[26] N. Baines, H.L. Smith, "Key driving factors for product and service innovations in UK university spin-offs," Industry and Higher Education, **33**(3), 161–171, 2019, doi:10.1177/0950422219825579.

[27] G.A. Van Norman, R. Eisenkot, "Technology transfer: From the research bench to commercialization. Part 1: Intellectual property rights —basics of patents and copyrights," JACC: Basic to Translational Science, **2**(1), 85–97, 2017, doi:10.1016/j.jacbts.2017.01.003.

[28] J. González, Manual de transferencia de tecnologia y conocimiento, The Transfer Institute, 2011.

[29] A. Anduaga, "The engineer as a 'linking agent' in international technology transfer: The case of Basque engineers trained in Liège," Engineering Studies, **3**(1), 45–70, 2011, doi:10.1080/19378629.2010.550922.

[30] F. Iqbal, P.C. Hung, F. Wahid, S.M.Q.A. Mohammed, "A glance at research-driven university's technology transfer office in the UAE," International Journal of Technology Management, **78**(1–2), 70–87, 2018, doi:10.1504/IJTM.2018.093939.

[31] E. De Wit-de Vries, W.A. Dolfsma, H.J. Van der Windt, M.P. Gerkema, "Knowledge transfer in university–industry research partnerships: a review," The Journal of Technology Transfer, **44**(4), 1236–1255, 2019, doi:10.1007/s10961-018-9660-x.

[32] N.T. Gallini, B.D. Wright, "Technology transfer under asymmetric information," The RAND Journal of Economics, 147–160, 1990, doi:10.2307/2555500.

[33] J.-W. Min, N.S. Vonortas, Y. Kim, "Commercialization of transferred public technologies," Technological Forecasting and Social Change, **138**, 10–20, 2019, doi:10.1016/j.techfore.2018.10.003.

[34] M. Fritsch, M. Kauffeld-Monz, "The impact of network structure on knowledge transfer: An application of social network analysis in the context of regional innovation networks," The Annals of Regional Science, **44**(1), 21, 2010, doi:10.1007/s00168-008-0245-8.

[35] S. Parjanen, H. Melkas, T. Uotila, "Distances, knowledge brokerage and absorptive capacity in enhancing regional innovativeness: A qualitative case study of Lahti region, Finland," European Planning Studies, **19**(6), 921–948, 2011, doi:10.1080/09654313.2011.568804.

[36] J. Howells, "Intermediation and the role of intermediaries in innovation," Research Policy, **35**(5), 715–728, 2006, doi:10.1016/j.respol.2006.03.005.

[37] L. Klerkx, C. Leeuwis, "Establishment and embedding of innovation brokers at different innovation system levels: Insights from the Dutch agricultural sector," Technological Forecasting and Social Change, **76**(6), 849–860, 2009, doi:10.1016/j.techfore.2008.10.001.

[38] L. Klerkx, C. Leeuwis, "Matching demand and supply in the agricultural knowledge infrastructure: Experiences with innovation intermediaries," Food Policy, **33**(3), 260–276, 2008, doi:10.1016/j.foodpol.2007.10.001.

[39] W.L. Ruíz Castañeda, S. Quintero Ramírez, J. Robledo Velásquez, "Impacto de los intermediarios en los sistemas de innovación," Journal of Technology Management & Innovation, **11**(2), 130–138, 2016, doi:10.4067/S0718-27242016000200013.

[40] R. Welsh, L. Glenna, W. Lacy, D. Biscotti, "Close enough but not too far: Assessing the effects of university-industry research relationships and the rise of academic capitalism," Research Policy, **37**(10), 1854–1864, 2008, doi:10.1016/j.respol.2008.07.010.

[41] A. Caldera, O. Debande, "Performance of Spanish universities in technology transfer: An empirical analysis," Research Policy, **39**(9), 1160–1173, 2010, doi:10.1016/j.respol.2010.05.016.

[42] J. Berbegal-Mirabent, F. Sabaté, A. Cañabate, "Brokering knowledge from universities to the marketplace: The role of knowledge transfer offices," Management Decision, **50**(7), 1285–1307, 2012, doi:10.1108/00251741211247012.

[43] W. Chapple, A. Lockett, D. Siegel, M. Wright, "Assessing the relative performance of U.K. university technology transfer offices: Parametric and non-parametric evidence," Research Policy, **34**(3), 369–384, 2005, doi:10.1016/j.respol.2005.01.007.
[44] D.S. Siegel, D.A. Waldman, A.N. Link, "Assessing the impact of organizational practices on the productivity of university technology transfer offices: An exploratory study," 1999, doi:10.1016/S0048-7333(01)00196-2.
[45] D.S. Siegel, R. Veugelers, M. Wright, "Technology transfer offices and commercialization of university intellectual property: performance and policy implications," Oxford Review of Economic Policy, **23**(4), 640–660, 2007, doi:10.1093/oxrep/grm036.
[46] L. Lecluyse, M. Knockaert, A. Spithoven, "The contribution of science parks: a literature review and future research agenda," The Journal of Technology Transfer, **44**(2), 559–595, 2019, doi:10.1007/s10961-018-09712-x.
[47] J.M. Ekboir, A.O. Vera-Cruz, "Intermediary organisations to foster the agricultural system of innovation: The Mexican Produce Foundation," International Journal of Technological Learning, Innovation and Development, **5**(1–2), 111–125, 2012, doi:10.1504/IJTLID.2012.044879.
[48] H. Yang, L. Klerkx, C. Leeuwis, "Functions and limitations of farmer cooperatives as innovation intermediaries: Findings from China," Agricultural Systems, **127**, 115–125, 2014, doi:10.1016/j.agsy.2014.02.005.
[49] B. Wilson, Sistemas: Conceptos, Metodología y Aplicaciones, Editorial Limusa S.A. De C.V., 1993.
[50] Y.A. Cáceres, L.C. Flórez, "Una reflexión sistémica sobre los fundamentos conceptuales para sistemas de información," Revista Colombiana de Computación, **8**(1), 1–19, 2007.
[51] J. Heijs, Sistemas nacionales y regionales de innovación y política tecnológica: Una aproximación teórica, 2001.
[52] L. Mathison, J. Gándara, C. Primera, L. García, "Innovación : Factor clave para lograr ventajas competitivas," Negotium, **3**(7), 65–83, 2007.
[53] J. Sábato, N. Botana, La ciencia y la tecnología en el desarrollo futuro de América Latina, Lima, 1970.
[54] C. Freeman, Technology, policy, and economic performance: lessons from Japan, Frances Printer Publishers, 1987.
[55] E.M. Balmaseda, I.Z. Elguezabal, G.I. Clemente, "Evolución de los modelos sobre el proceso de innovación: desde el modelo lineal hasta los sistemas de innovación," in Decisiones basadas en el conocimiento y en el papel social de la empresa: XX Congreso anual de AEDEM, Asociación Española de Dirección y Economía de la Empresa, 2007.
[56] Minciencias, Actores del Sistema Nacional de Ciencia, Tecnología e Innovación, 2016.
[57] C. Freeman, "The 'national system of innovation' in historical perspective," Cambridge Journal of Economics, **19**, 15–24, 1995.
[58] L.J. Quintero-Campos, "Aportes teóricos para el estudio de un sistema de innovación," INNOVAR. Revista de Ciencias Administrativas y Sociales, **20**(38), 57–76, 2010.
[59] D.J. Spielman, Innovation systems perspectives on developing-country agriculture: A critical review, 2005, doi:10.22004/ag.econ.59692.
[60] R. Rajalahti, W. Janssen, E. Pehu, Agricultural innovation systems: From diagnostics toward operational practices, 2008.
[61] B.-Å. Lundvall, The learning economy and the economics of hope, Anthem Press, 2016.
[62] P. Adams, F. Malerba, Sectoral Systems of Innovation, Oxford University Press, 2014.
[63] J. Jasso Villazul, Los sistemas de la innovación como espacios regionales, sectoriales y empresariales: Características y taxonomía, 2000.
[64] M.N. Arancegui, "Los sistemas regionales de innovación: una revisión crítica," EKONOMIAZ. Revista Vasca de Economía., **70**(1), 25–59, 2009.
[65] F. Malerba, R. Nelson, "Sistemas sectoriales, alcance y desarrollo económico," Economía: Teoría y Práctica, **SPE1**, 41–62, 2009.
[66] M. Amaro Rosales, M. Morales Sánchez, "La Biotecnología en México, una aproximación desde los sistemas sectoriales de innovación," Ide@ s CONCYTEG, **5**(64), 1224–1246, 2010.
[67] T. Mets, "Developing the Sectorial Innovation System of Estonian Biotechnology," Engineering Economics., **5**(50), 73–79, 2006.
[68] UIS, UNAB, UPB, INAL, Diseño de un modelo de trabajo colaborativo entre actores del sector agropecuario para el desarrollo de la Agrópolis de Santander Magdalena Medio, 2016.
[69] UIS, Findeter, BID, Agrópolis de Santander Magdalena Medio - Estudios de profundización escala diamante, 2016.
[70] O. Afonso, S. Monteiro, M. Thompson, "A growth model for the quadruple helix," Journal of Business Economics and Management, **13**(5), 849–865, 2012, doi:10.3846/16111699.2011.626438.
[71] H. Etzkowitz, La triple hélice: universidad , industria y gobierno. Implicaciones para las políticas y la evaluación, 2002.
[72] H.G.C. Castillo, "El modelo de la triple hélice como un medio para la vinculación entre la universidad y empresa," Revista Nacional de Administración, **1**(1), 85–94, 2010, doi:10.22458/rna.v1i1.286.
[73] W.L. Ruíz Castañeda, Análisis del impacto de los intermediarios en los sistemas de innovación: Una propuesta desde el modelado basado en agentes, Universidad Nacional de Colombia, 2016.
[74] W.L. Castañeda Ruíz, J. Robledo Velásquez, "Evaluación del Impacto de los Intermediarios en los Sistemas de Innovación: Marco de Análisis," in ALTEC, 1–17, 2013.
[75] H. Van Lente, M. Hekkert, R. Smits, B.A.S. Van Waveren, "Roles of systemic intermediaries in transition processes," International Journal of Innovation Management, **7**(03), 247–279, 2003, doi:10.1142/S1363919603000817.
[76] J. Vereijssen, T.M. Williams, N.M. Park, M. Nielsen, N.M. Agnew, "Evaluating co-innovation principles in a fundamental bioprotection research programme addressing challenges to potato production," New Zealand Plant Protection, **70**, 16–24, 2017, doi:10.30843/nzpp.2017.70.20.
[77] E. Villani, E. Rasmussen, R. Grimaldi, "How intermediary organizations facilitate university–industry technology transfer: A proximity approach," Technological Forecasting and Social Change, **114**, 86–102, 2017, doi:10.1016/j.techfore.2016.06.004.
[78] N. Aguilar-Gallegos, E.G. Martínez-González, J. Aguilar-Ávila, H. Santoyo-Cortés, M. Muñoz-Rodríguez, E.I. García-Sánchez, "Análisis de redes sociales para catalizar la innovación agrícola: de los vínculos directos a la integración y radialidad," Estudios Gerenciales, **32**(140), 197–207, 2016, doi:10.1016/j.estger.2016.06.006.
[79] M. De Silva, J. Howells, M. Meyer, "Innovation intermediaries and collaboration: Knowledge–based practices and internal value creation," Research Policy, **47**(1), 70–87, 2018, doi:10.1016/j.respol.2017.09.011.
[80] C.W. Kilelu, L. Klerkx, C. Leeuwis, A. Hall, "Beyond knowledge brokering: an exploratory study on innovation intermediaries in an evolving smallholder agricultural system in Kenya," Knowledge Management for Development Journal, **7**(1), 84–108, 2011, doi:10.1080/19474199.2011.593859.
[81] P. Kivimaa, W. Boon, S. Hyysalo, L. Klerkx, "Towards a typology of intermediaries in sustainability transitions: A systematic review and a research agenda," Research Policy, **48**(4), 1062–1075, 2019, doi:10.1016/j.respol.2018.10.006.
[82] L. Klerkx, M. Schut, C. Leeuwis, C. Kilelu, "Advances in Knowledge Brokering in the Agricultural Sector: Towards Innovation System Facilitation," IDS Bulletin, **43**(5), 53–60, 2012, doi:10.1111/j.1759-5436.2012.00363.x.
[83] P. de A.T. TAP, Marco Común sobre el Desarrollo de Capacidades para los Sistemas de Innovación Agrícola: Documento de Síntesis, CAB International, Wallingford, Reino Unido., 2016.
[84] J.M.B. Garmendia, A.R. Castellanos, "Estructuras de intermediación para la transferencia de conocimiento universitario: las oficinas de transferencia tecnológica," Propiedad Intelectual, **9**(13), 152–176, 2010.
[85] J. Maira, J. Etxabe, P.A. Serena, "Strategies on Technology Transfer and Patents Commercialization for Nanotechnology at the Spanish National Research Council," Recent Patents on Nanotechnology, **12**(1), 45–58, 2018, doi:10.2174/1872210511666170601115516.
[86] F. Alvarado-Moreno, "El Papel de las Oficinas de Transferencia Tecnológica (OTT) en las Universidades: Una Perspectiva de la Última Década," Journal of Technology Management & Innovation, **13**(3), 104–112, 2018, doi:10.4067/S0718-27242018000300104.
[87] Portafolio, "Colombia cae en clasificación global de los más innovadores," Portafolio, 2019.
[88] J. García, C. Gómez, C. Jiménez, J.C. Castellanos, La eficiencia de la Innovación en Colombia frente al mundo: Un análisis desde el Global Innovation Index, 2016 - 2019, 2019.
[89] Minciencias, Transferencia de conocimiento, una estrategia para potencializar la investigación, Minciencias, 2016.
[90] CONPES, Política nacional de ciencia, tecnología e innovación 2015 - 2025, 2015.
[91] DNP, Plan Nacional de Desarrollo 2018 - 2022, Https://Www.Dnp.Gov.Co/DNPN/Paginas/Plan-Nacional-de-Desarrollo.Aspx, 2018.
[92] Agrosavia, MADR, Minciencias, PECTIA: Plan Estratégico de Ciencia, Tecnología e Innovación del sector Agropecuario colombiano (2017-2027), 2016.
[93] L. Guarin, E. Romero, L.E. Becerra, "Agrópolis de Santander Magdalena - Medio en Colombia: Alternativa para el desarrollo rural en las regiones," in Congreso Internacional de Investigación e Innovación - Multidisciplianrio, 7382–7393, 2018.
[94] G. Abramo, C.A. D'Angelo, F. Di Costa, M. Solazzi, "The role of

information asymmetry in the market for university–industry research collaboration," The Journal of Technology Transfer, **36**(1), 2011, doi:10.1007/s10961-009-9131-5.

[95] P. Gubitta, A. Tognazzo, F. Destro, "Signaling in academic ventures: the role of technology transfer offices and university funds," The Journal of Technology Transfer, **41**(2), 2016, doi:10.1007/s10961-015-9398-7.

[96] A.C. Schulze-Krogh, "Firms' Absorptive Capacity for Research-Based Collaboration—an Analysis of a Norwegian R&D Brokering Policy Program," Science and Public Policy, **45**(4), 2018, doi:10.1093/scipol/scx081.

[97] M. Cattaneo, M. Meoli, S. Vismara, "Cross-border M&As of biotech firms affiliated with internationalized universities," The Journal of Technology Transfer, **40**(3), 2015, doi:10.1007/s10961-014-9349-8.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 83-89 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

The Effect of Myocardial Fat's Thickness and Myocardial Impedance on Bipolar Radiofrequency Catheter Ablation Using Computer Simulation

Yao Sun[1], Keijiro Nakamura[2], Xin Zhu[*,1]

[1]*Biomedical Information Engineering Lab, The University of Aizu, Aizu-Wakamatsu, Fukushima, 960-0000, Japan*

[2]*Division of Cardiovascular Medicine, Toho University Ohashi Medical Center, Meguro, Tokyo, 100-003, Japan*

ARTICLE INFO

Article history:
Received: 03 July, 2021
Accepted: 05 September, 2021
Online: 16 September, 2021

Keywords:
Ventricular tachycardia
Bipolar ablation
Unipolar ablation
Computer simulation
Fat thickness
Myocardial impedance
Radio frequency ablation

ABSTRACT

Radiofrequency catheter ablation is routinely used for the therapy of cardiac arrhythmias. Compared with the traditional unipolar ablation, bipolar ablation may improve the controllability of treatment, and prevent side effects and complications caused by catheter ablation. In addition, the variations of myocardial fat's thickness and myocardial impedance may have significant influence on the performance of bipolar ablation. In this study, computer simulation was performed to study the effects of myocardium fat's thickness and myocardial impedance on unipolar and bipolar ablation. The simulation demonstrates similar results with experimental ones using a swine heart. We observed that when the myocardial fat's thickness increases, bipolar ablation's heating effect and controllability may decrease. However, the final heating effect of bipolar ablation is invariably better than that of unipolar ablation. The ablation effects of unipolar and bipolar ablation are both reduced when myocardial impedance increases, while the heating effects of bipolar ablation are more sensitive to the variation of myocardial impedance and fat layers' thickness compared with unipolar ablation. The unipolar ablation is more stable in terms of fat, impedance and ablation time.

1. Introduction

Ventricular tachycardia (VT) is generally related to severe structural heart diseases (SHD) and accounts for an elevated risk of sudden cardiac death in patients with the heart failure [1]. It is also reported that for patients hospitalized with AMI, 5% to 10% of them have ventricular fibrillation (VF) or sustained VT prior to hospital presentation, and another 5% have VF or sustained VT after hospital arrival, most within 48 hours of admission [2]. A report about the incidence and case-fatality of ventricular events in hospital mentioned that the hospital case-fatality rates (CFRs) of VT declined from 27.7% to 6.9% and the incidence rates (IRs) declined from 14.3% to 10.5% from 1986 to 2009. This report concluded that the CFRs and IRs of VT are decreased, but it's still necessary to timely identify and treat patients with these serious ventricular arrhythmias [3]. Implantable cardioverter defibrillators (ICDs) remain the main method for the primary and secondary prevention of sudden cardiac death associated with these arrhythmias [4]. However, ICDs can only be used for treatment, not for the prediction or prevention of ventricular arrhythmia [4]. ICDs are effective but their shocks are painful and can lead to post-traumatic stress disorders [5]. As a first line drug for arrhythmias, amiodarone can reduce arrhythmias for some patients but have limited efficacy for long-term management [6].

Radiofrequency (RF) catheter ablation has been proved as an effective treatment method for ventricular arrhythmias [7]. Particularly, the use of RF to treat atrial fibrillation (AF) has a history of more than 20 years [8]. Catheter ablation causes partial necrosis of lesion and therefore restore normal heart rhythms. Unipolar ablation (UA) is the standard configuration of catheter ablation, i.e., a singular electrode and a sheet ground on body surfaces [9]. UA is routinely used in clinical practice, but the effect and the treatment process of UA are expected to be improved [10][11]. Bipolar ablation (BA) can deliver concentrated energy between two electrodes. BA's focused energy delivery shortens ablation durations and minimizes lesion width; therefore, this reduces the potential of adjacent tissue injury. Although BA is still under study, it has been shown to be more effective than UA [12].

[*]Corresponding Author: Xin Zhu, Email: zhuxin@u-aizu.ac.jp

Computer simulation is an additional method for the study on the mechanisms of cardiac arrhythmias and the procedure of catheter ablation [13]. In [14], the author reported a systematic study on the effectiveness of different ablation line patterns based on a biophysical model. Analysis with a two-dimensional finite element method can also predict temperature distributions in ablation or hyperthermia [15], [16].

Due to the difference in density and other properties, the thermal and electrical conductivities of myocardial fat are significantly different in myocardium, blood and pericardial fluid [10]. In view of the low electrical conductivity of fat, the therapeutic effects of UA and BA may vary in different myocardial fat's thickness and myocardial impedance. Theoretically, when the thickness of cardiac fat increases, the voltage of myocardium would decrease because the impedance of fat is larger, and the electric current decreases as the impedance increases when the ablation power keeps the same value, and *vice versa*. As a result, the performance of ablation is reduced correspondingly. At present, the main research on radiofrequency ablation is the occurrence of complications and improvement of catheter ablation. However, there are few studies about the effect of myocardial fat's thickness and myocardial impedance on catheter ablation.

In this study, we conducted computer simulation to investigate the heating effects of UA and BA at different values of myocardial impedance and fat layer's thickness. To validate the results of computer simulation, we performed an *in vitro* experiment by using a swine's heart for comparison.

2. Methods and Materials

2.1. Configurations of Myocardial Fat Thickness and Impedance

We created a two-dimensional (2D) myocardial model based on the cross-section of the heart with a fat layer between the pericardial fluid and the myocardium as illustrated in Figure 1. To reflect most clinical conditions, fat thickness of 0-2mm was set [17]. Considering the pressure of the electrode on myocardium, the fat layer under the electrode invades has a distortion as the semidiameter of the electrode [18]. In this model, we set up a thin layer of pericardial fluid and blood with a thickness of 3 mm, and for myocardium, the thickness is 10 mm [10], [19]. The ablation

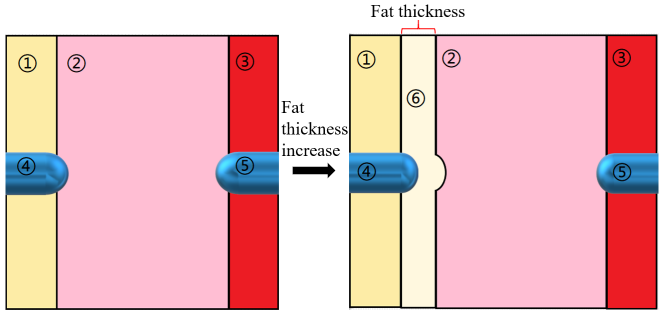


Figure 1: 2D model. ① Pericardial fluid, ②Myocardium, ③Blood, ④ Positive electrode, ⑤ Negative electrode, ⑥ Fat.

electrodes have a radius of 1.155 mm (7 Fr). The influence from blood and water flow was not considered to simplify this study.

In general, as the heating temperature of radiofrequency ablation does not exceed 100°C, the conductivity of body tissue in the model increases with temperature up to 100° C as follows,

$$z = z_0[1 + max \times (T - 37)/(100 - 37)], \quad (1)$$

where z is impedance, max the increasing percentage of z at 100°C, l the length and s the area. As the conductivity is proportional to impedance [18], we adopted the function of electricity in radiofrequency ablation, and impedance was considered increasing by 5%, 10%, 15% and 20% during the heating period, *i.e.*, max = 5%, 10%, 15%, and 20%. By

$$p = U^2/Re(z) \quad (2)$$

$$\frac{U^2}{Re(z)} = U^2 \frac{1}{Re(z_0)}\left(1 - \frac{\Delta Re(z)}{Re(z_0)}\right) \quad (3)$$

$$\frac{\Delta Re(z)}{Re(z_0)} = max \times (T - 37)/(100 - 37) \quad (4)$$

where p is power, and U voltage [20]. We calculated p from z_0 and U, and we also hypothesized that the rate of increase of impedance approximately equals to the decreasing rate of power at a constant voltage. In this study, $Re(z_0) = 100\Omega$ was set to calculate power. Figure 2 showed the variation of power with impedance.

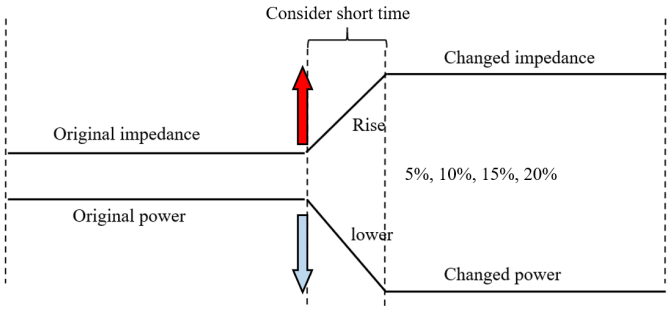


Figure 2: The power changed by increased myocardial impedance.

2.2. Numerical Analysis

In this study, we used the Finite Element Toolbox in MATLAB™ R2020b (Mathworks Inc., USA) to construct computer models and perform computer simulation. The simulation is converted to a procedure resolving partial differential equations, i.e., parabolic and elliptic functions in partial differential equation (PDE). Boundary conditions have been set to create electric and thermal fields [21]. The electrical frequency in radio frequency ablation is about 500 kHz. The biological medium can be considered almost totally resistive [10][15] because the human's body scales is about 0.5 m, which is much less than the wavelength of radio frequency wave. This means we can use a quasi-static approach to solve the electrical problem. Laplacian equation serves as the governing equation for calculating the electric potential within a 2D heterogeneous model [10].

$$\nabla(-\varepsilon\nabla V) = 0, \quad (5)$$

where V is the potential, and ε the permittivity varying spatially due to the heterogeneity of tissues. From the solution of (5), the

electric field E is got, and the specific absorption rate (SAR) Q can be expressed as follows,

$$E = -\nabla V \tag{6}$$

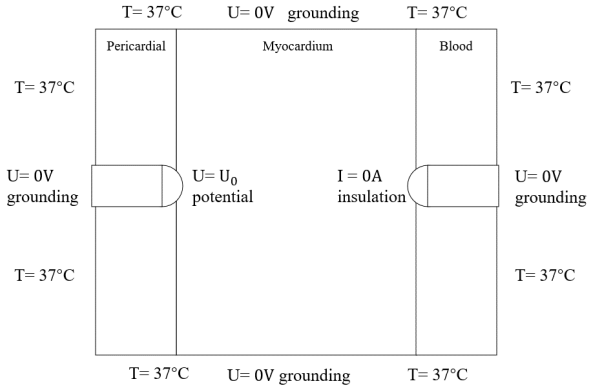
$$Q = \frac{1}{2}\sigma|E|^2, \tag{7}$$

where σ is the electric conductivity. Based on SAR, i.e., Q obtained above, and Penne's bioheat transfer equation as follows[15][22], we can obtain temperature distributions.

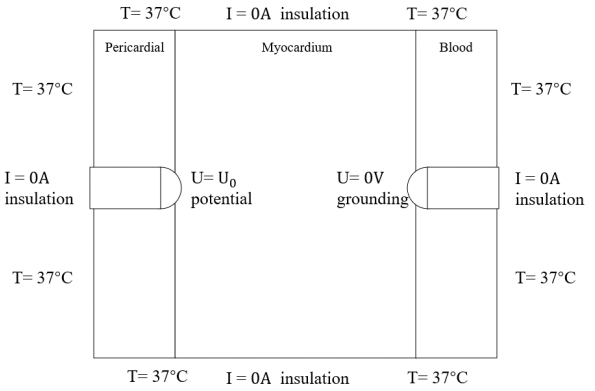
$$\rho c_p \frac{\partial T}{\partial t} - \nabla(\kappa \nabla T) = Q - \omega \tag{8}$$

$$\omega = (\mathrm{F}\rho)_t \left(\rho c_p\right)_b (T - T_b) \tag{9}$$

In equation (8), ρ is the density, c_p the constant pressure specific heat capacity, κ thermal conductivity, T temperature, and Q SAR calculated by equation (7). In equation (9), ω is the heat taken away by the blood flow rate F, and T_b blood temperature, with an initial temperature of myocardium (37°C). 2-D FEM can be used to solve equations (5)-(9) numerically. Spatially and temporally variable physical parameters like current, potential and heat are mapped to first-order triangular elements [23]. Then, equations (5)-(8) are discretized, converted into a matrix notation and solved with finite element method (FEM). Equation (8) is discretized spatially and temporally; therefore, a transient state is hypothesized to resolve this problem. Finally we got the temperature distribution [15].



a



b

Figure 3: The boundary conditions of electric and thermal fields for UA (a) and BA (b).

We employed two boundary conditions including Dirichlet and Neumann boundary conditions. For the configurations of electric field, Dirichlet boundary condition is used to constrain the potential of electrode, and Neumann boundary condition is used to calculate electric currents, mainly setting the outer boundaries to be of air insulation. Figure 3 shows the boundary conditions in computer simulation, including the initial temperature and electrodes boundary conditions.

The initial temperature in BA and UA is 37°C. It is hypothesized that the temperature in the perpendicular direction of the model is the same, so that the temperature conduction is only affected by the electrode and the initial temperature in the 2D model. For the bipolar model in Figure 3(b), the left and right electrodes are positive and grounded (U=0V), respectively; for the unipolar model shown in Figure 3(a), the positive electrode is the same as that of the bipolar model, but the negative electrode is insulated (I=0A). Table 1 listed the parameters used in the computer simulation [10], [18].

Table 1: Parameters used in the computer simulation

Parameters	$\boldsymbol{\sigma}$ (S/m)	$\boldsymbol{\varepsilon}$	$\boldsymbol{\kappa}$ (W/m/K)	$\boldsymbol{\rho}$ (kg/m^3)	$\boldsymbol{c}$ (J/kg/K)
Pericardial effusion	1.35	118.00	0.63	1.02×10^3	4.18×10^3
Myocardial	0.61	113.00	0.60	1.02×10^3	8.30×10^{-6}
Fat	0.05	20.00	0.22	0.90×10^3	5.00×10^{-7}
Blood	1.10	118.00	0.54	1.06×10^3	4.18×10^3

$\boldsymbol{\sigma}$, Conductivity; $\boldsymbol{\varepsilon}$, electric constant; $\boldsymbol{\kappa}$, Thermal Conductivity; $\boldsymbol{\rho}$,Density; $\boldsymbol{c}$, Specific heat capacity

2.3. Experiment

We conducted an *in vitro* experiment to perform catheter ablation using a swine's heart, thus we can validate the results of computer simulation. A glass tank filled with water at normal temperature(22℃) was employed to simulate the environment of human tissue, and a swine's myocardium was used to perform UA and BA. The power of 30W lasts for 30s and 50W for 10s. Figure 4 shows the experiment results of swine's myocardium after ablation.

Table 2 lists the experiment results for the models with 2 mm's fat layer and without fat, including the lesion depth and width in swine's myocardium under different parameters of ablation durations and powers. The power and duration of heating commonly used in clinical practice were used in the experiment. Based on the experimental results, even though the cooling effect of water may reduce the efficiency of ablation, the transmural damage was obtained in the myocardium with a thickness of 8 mm. The heating effects of BA are better than those of UA in any ablation configurations. The existence of myocardial fat has a significantly stronger effect on the results of BA, and we usually

don't use 50W power in BA heating for 30s because the power is too strong, and may put patients at risk if the heating time is too long.

Table 2: Experiment Result Compared with Simulation Result

Experiment Result (environmental temp.: 22℃)				
Power/ time	**Lesion thickness (without fat)**			
	Uni. Diameter(mm)	**Uni. Depth(mm)**	**Bi. Diameter(mm)**	**Bi. Depth(mm)**
30W/ 30s	6.0	3.0	6.0	8.0
50W/ 10s	7.0	5.0	—	—
Power/ time	**Lesion thickness (with 2mm fat)**			
	Uni. diameter(mm)	**Uni. depth(mm)**	**Bi. diameter(mm)**	**Bi. depth(mm)**
30W/ 30s	6.5	3.0	5.5	5.0
50W/ 10s	7.5	4.5	—	—
Simulation Result (environmental temp.: 22℃)				
Power/ time	**Lesion thickness (without fat)**			
	Uni. diameter(mm)	**Uni. depth(mm)**	**Bi. diameter(mm)**	**Bi. depth(mm)**
30W/ 30s	6.5	3.5	6.5	10.0
50W/ 10s	7.0	5.5	—	—
Power/ time	**Lesion thickness (with 2mm fat)**			
	Uni. diameter(mm)	**Uni. depth(mm)**	**Bi. diameter(mm)**	**Bi. depth(mm)**
30W/ 30s	6.5	3.5	6.0	5.5
50W/ 10s	7.5	5.0	—	—



Figure 4: The results of a 30s 30w unipolar ablation experiment performed by a physician using an in vitro swine's myocardium. The lesion has a diameter of 6.00mm, and a depth of 3.00mm.

3. Results

Table 2 lists the simulation results obtained under the same configurations. Through comparison, the lesion thicknesses in the simulation are nearly the same but better than those in experiments. This may attribute to that the ablation parameters were more ideal in simulation. As the lesion thickness is the main therapeutic effect index in catheter ablation, we may believe the results are basically the same.

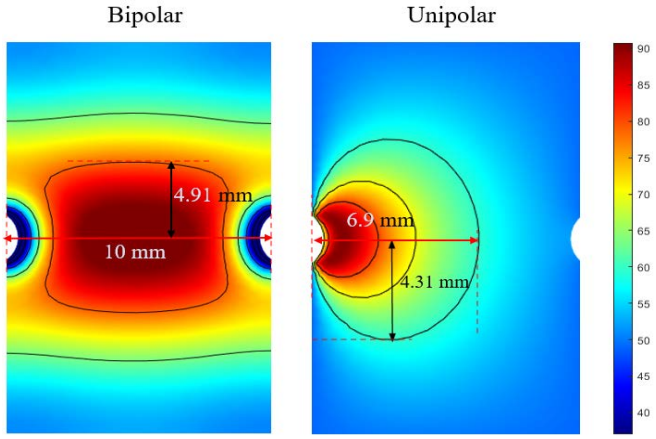


Figure 5: Simulation under 16W ablation for 120 s.

Simultaneously, in order to further ensure the accuracy of the simulation results, we also set the same ablation parameters based on a research report published in 2014 to validate the simulation [10]. We set an ablation configuration with a power of 16W ablation for 120 s. The results of the study are shown in Figure 5. Our simulation results are similar to those in [10], so we can ensure that our simulation results are basically consistent with the experiment and related research.

In this study, we selected central regions from the experimental results for calculating transmural lesion ratios (TLR), a 10 mm × 5 mm red rectangle shown in Figure 6. Thus we have:

$$TLR = \frac{area\ of\ temperature\ over\ 60\ ℃\ in\ 10\ mm \times 5\ mm}{the\ area\ of\ 10\ mm \times 5\ mm}. \quad (10)$$

We also visualized the temperature distribution in order to better observe heating areas as shown in Figure 6. The colormap was changed from colorful to gray levels from 0 to 255, and we found the appropriate threshold value to differentiate the temperature after testing, regions at 37℃ to 59℃ corresponded to the gray levels from 0 to 93 are marked in black, 60 ℃ to 100℃ corresponded to the gray levels from 94 to 254 marked in gray, and more than 100℃ corresponded to the gray level 255 marked in white.

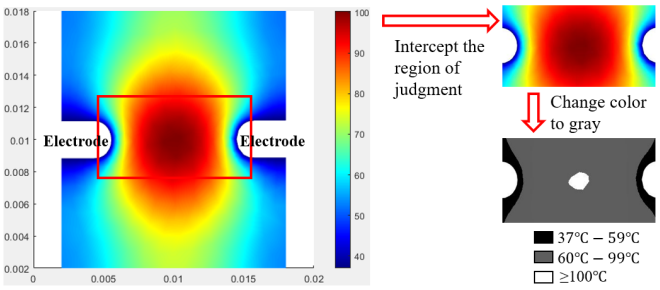


Figure 6: The visualization of temperature distribution.

At regions above 100℃, the blood vaporizes and clots to form a thrombus. This is a situation that needs to be avoided during

therapy. In the study, the simulation temperature may be too high due to the absence of cooling liquid, and we treat this as a tissue necrosis.

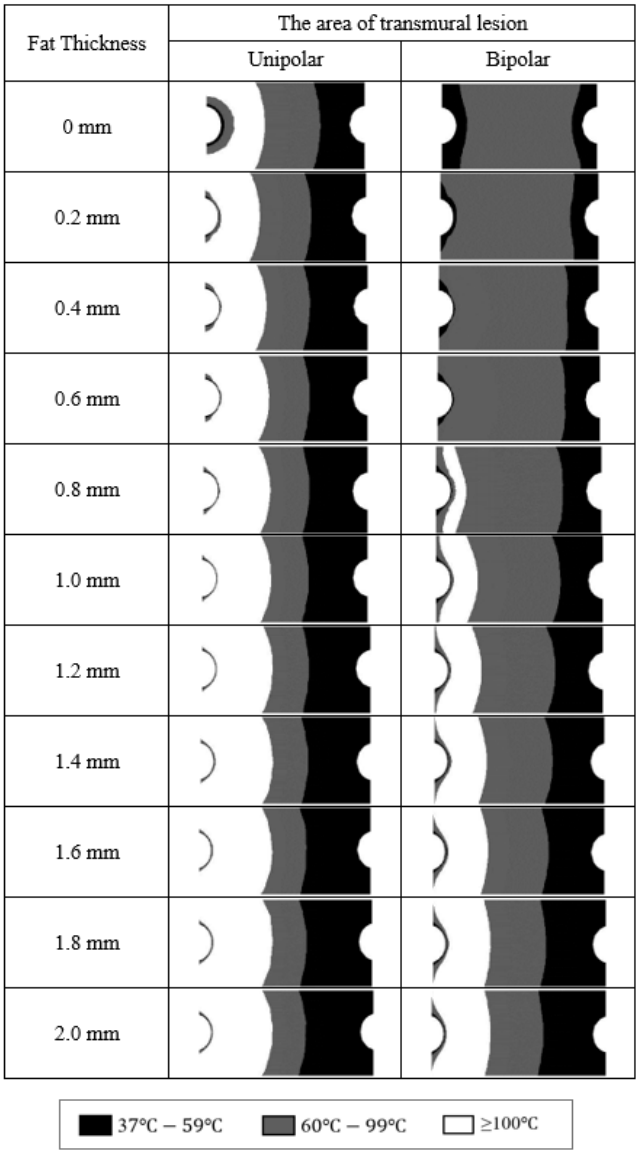


Figure 7: The area of transmural lesion of UA and BA under different fat thickness.

Figure 7 shows simulated heating results under the heating power of 30 W in UA and BA with ablation time of 60s. We observed that when the fat thickness is less than 1 mm, BA has obvious advantages over UA. BA can effectively control the high temperature range and maintain a wall penetration effect. When the fat thickness is greater than 1 mm, the controllability of BA gradually decreases. This heating results of UA are almost unaffected by fat thickness, but the local temperature is too high and transmural effect is not ideal in the UA.

Figure 8 shows the TLR of simulation under UA and BA with the variation of myocardial fat's thickness. Both UA and BA's heating effects show an overall reduced trend. The TLR of BA are better than that of UA, but with the increasing of the fat's thickness, the TLR of BA decline slightly faster. When the fat's thickness is 0.2 mm, as heat begins to transfer to the fat, the temperature around the fat rises, leading to a sudden improvement in transmural heating effects. The TLR of UA also tends to be stable when the fat's thickness is over 0.2 mm.

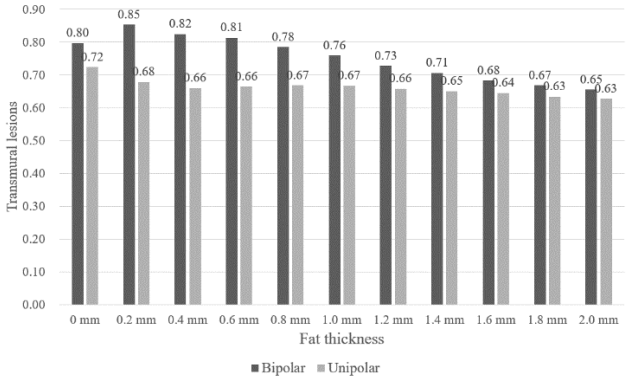


Figure 8: TLR of UA and BA under different fat thickness.

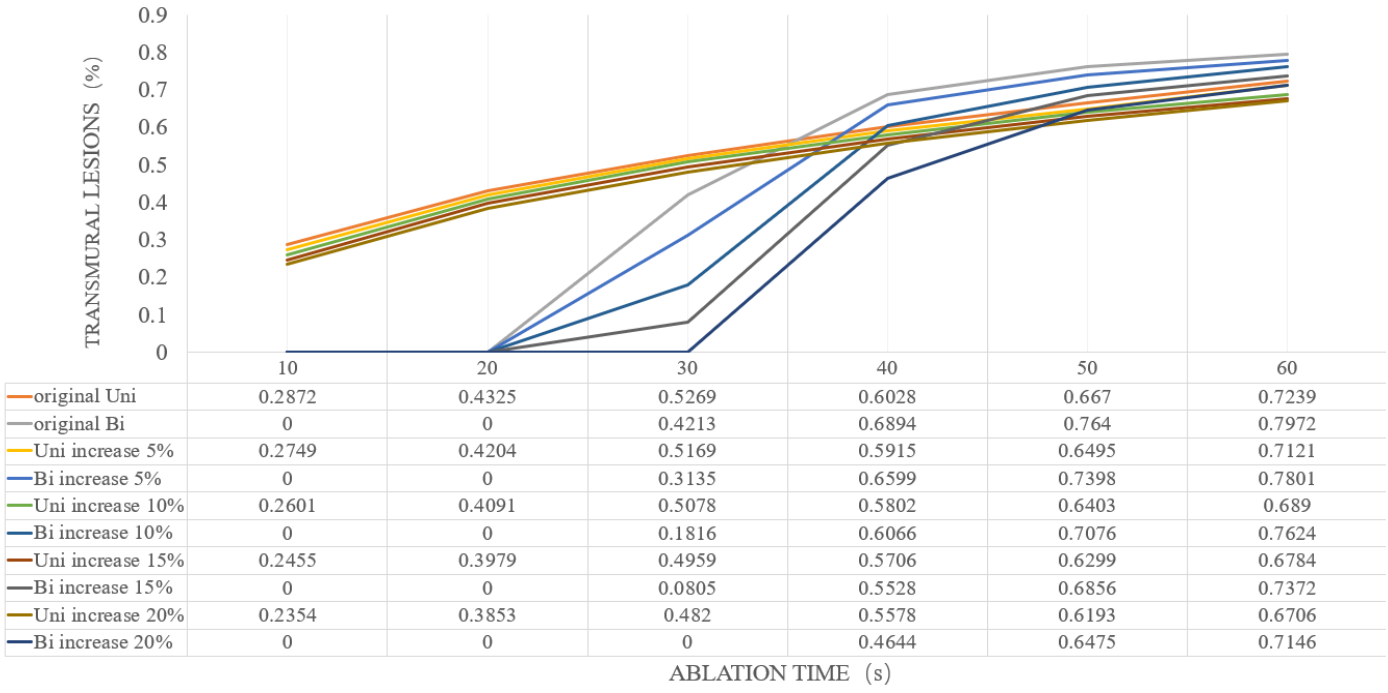
In this study, we considered the myocardial impedance increased 5%, 10%, 15% and 20% from the original impedance when the ablation power is 30W and sustains for 60 s. Figure 8 shows the TLR of simulated UA and BA under different heating time with increased impedance.

Figure 9 shows that the transmural lesions of both UA and BA generally decrease as the impedance increases. For BA, heating effect is poor at the beginning because the heat is still in the accumulation stage. When the heating duration is 30s to 40s, the transmural lesion increases rapidly and stabilizes after more than 40s. The final transmural effect of BA is stronger than that of UA. For UA, transmural lesion increases slowly and steadily. This indicates UA did not change drastically when impedance changed. As a result, the ablation effect of BA is more sensitive to myocardial impedance.

4. Discussion

This study compared the impact of myocardial fat's thickness and impedance changes on the heating effect of catheter ablation. From the simulation, we roughly identified a range of fat thickness, i.e., a fat thickness less than 1 mm, that BA has obvious advantages, including controllable temperature and a well transmural effect over UA. Because the UA has a radial range of action, fat does not change the direction of current conduction and UA is basically stable from the variation of fat's thickness and myocardial impedance. Therefore, UA is nearly unaffected by external conditions in case of excessive myocardial fat or impedance uncertainty.

In terms of ablation effect, we found BA demonstrated better results. The heat diffuses from the middle to ensure higher transmural ability, and the dispersed heat also reduces the burning temperature and ensures safety. Thus, for patients with lower BMI, BA might be a better choice. Less myocardial fat can reduce heat accumulation and lead to lower myocardial impedance. On the other hand, since the conductivity of fat is lower than that of



	10	20	30	40	50	60
original Uni	0.2872	0.4325	0.5269	0.6028	0.667	0.7239
original Bi	0	0	0.4213	0.6894	0.764	0.7972
Uni increase 5%	0.2749	0.4204	0.5169	0.5915	0.6495	0.7121
Bi increase 5%	0	0	0.3135	0.6599	0.7398	0.7801
Uni increase 10%	0.2601	0.4091	0.5078	0.5802	0.6403	0.689
Bi increase 10%	0	0	0.1816	0.6066	0.7076	0.7624
Uni increase 15%	0.2455	0.3979	0.4959	0.5706	0.6299	0.6784
Bi increase 15%	0	0	0.0805	0.5528	0.6856	0.7372
Uni increase 20%	0.2354	0.3853	0.482	0.5578	0.6193	0.6706
Bi increase 20%	0	0	0	0.4644	0.6475	0.7146

Figure 9: TLR of UA and BA under different impedance during ablation time.

myocardium, pericardial effusion and blood, fat thickness may be positively correlated with myocardial impedance, but their relationship remains to be determined by further studies.

In addition to the influence from myocardial fat's thickness and myocardial impedance, we also studied the ablation time to further understand the heating duration and characteristics of UA and BA. We found that BA is more sensitive to time variations. As far as the simulation conclusion is concerned, UA has a better ablation than effect 30s ago, and the BA effect is significantly improved, which is better than UA after 30 s. Time may be an important variable that affects BA. High-power and short-term ablation is the goal pursued by catheter ablation, but cooling water is necessary under UA in case of increasing power. Otherwise, blood steam evaporation may occur. Thus BA may avoid blood steam evaporation.

Using a 2D model for biological research has some advantages over 3D model because the calculation time of 2D model is greatly shortened. In contrast, the calculation time of a 3D model is longer, but the simulation results are closer to actual situations. Furthermore, flow of blood and cooling water can be easily introduced in a 3D model [24], [25]. These are hard to be set in a 2D model [26]; therefore, they are not considered in this study.

5. Conclusions

We perform a study about the effect of myocardial fat's thickness and myocardial impedance on catheter ablation using computer simulation. The results indicate that the thickness of fat layer and myocardial impedance change may have influence on the heating performance of BA to a certain degree. Therefore, for patients with lower BMI, BA might be a better choice. The effect of UA is more stable for patients with higher BMI. We also proved that the BA has better ablation effects compared with those of UA when there are small variations of myocardial fat's thickness and myocardial impedance.

6. Future Work

In the future, we will extend the 2D cardiac model to a 3D cardiac model and consider the influence of blood flow and electrode types on the performance of UA and BA.

Acknowledgements

This research is partially supported by Japan Society For The Promotion of Science Kakenhi Basic Research Fund 18K11532& 21K10287, and 2021 Competitive Research Fund of The University of Aizu P-5.

References

[1] W.G. Stevenson, K. Soejima, "Catheter Ablation for Ventricular Tachycardia," Circulation, **115**(21), 2750–2760, 2007, doi:10.1161/CIRCULATIONAHA.106.655720.

[2] S.M. Al-Khatib, W.G. Stevenson, M.J. Ackerman, W.J. Bryant, D.J. Callans, A.B. Curtis, B.J. Deal, T. Dickfeld, M.E. Field, G.C. Fonarow, A.M. Gillis, C.B. Granger, S.C. Hammill, M.A. Hlatky, J.A. Joglar, G.N. Kay, D.D. Matlock, R.J. Myerburg, R.L. Page, "2017 AHA/ACC/HRS Guideline for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society," Journal of the American College of Cardiology, **72**(14), e91–e220, 2018, doi:10.1016/J.JACC.2017.10.054.

[3] H. V. Tran, A.S. Ash, J.M. Gore, C.E. Darling, C.I. Kiefe, R.J. Goldberg, "Twenty-five year trends (1986-2011) in hospital incidence and case-fatality rates of ventricular tachycardia and ventricular fibrillation complicating acute myocardial infarction," American Heart Journal, 208, 1–10, 2019, doi:10.1016/J.AHJ.2018.10.007.

[4] K. Shivkumar, "Catheter Ablation of Ventricular Arrhythmias," Https://Doi.Org/10.1056/NEJMra1615244, **380**(16), 1555–1564, 2019, doi:10.1056/NEJMRA1615244.

[5] M.O. Sweeney, "The Contradiction of Appropriate Shocks in Primary Prevention ICDs," Circulation, **122**(25), 2638–2641, 2010,

doi:10.1161/CIRCULATIONAHA.110.000208.
[6] G.H. Bardy, K.L. Lee, D.B. Mark, J.E. Poole, D.L. Packer, R. Boineau, M. Domanski, C. Troutman, J. Anderson, G. Johnson, S.E. McNulty, N. Clapp-Channing, L.D. Davidson-Ray, E.S. Fraulo, D.P. Fishbein, R.M. Luceri, J.H. Ip, "Amiodarone or an Implantable Cardioverter–Defibrillator for Congestive Heart Failure," Http://Dx.Doi.Org/10.1056/NEJMoa043399, **352**(3), 225–237, 2009, doi:10.1056/NEJMOA043399.
[7] L.S. Klein, H.T. Shih, F.K. Hackett, D.P. Zipes, W.M. Miles, "Radiofrequency catheter ablation of ventricular tachycardia in patients without structural heart disease.," Circulation, **85**(5), 1666–1674, 1992, doi:10.1161/01.CIR.85.5.1666.
[8] M. Haïssaguerre, P. Jaïs, D.C. Shah, A. Takahashi, M. Hocini, G. Quiniou, S. Garrigue, A. Le Mouroux, P. Le Métayer, J. Clémenty, "Spontaneous Initiation of Atrial Fibrillation by Ectopic Beats Originating in the Pulmonary Veins," Http://Dx.Doi.Org/10.1056/NEJM199809033391003, **339**(10), 659–666, 2009, doi:10.1056/NEJM199809033391003.
[9] X. Zhu, D. Yang, W. Lu, W. Chen, D. Wei, K. Fukuda, H. Shimokawa, "Computer simulation of cathode ablation for atrial fibrillation," Proceedings - 2014 IEEE International Conference on Computer and Information Technology, CIT 2014, 721–725, 2014, doi:10.1109/CIT.2014.37.
[10] A. González-Suárez, M. Trujillo, J. Koruth, A. d'Avila, E. Berjano, "Radiofrequency cardiac ablation with catheters placed on opposing sides of the ventricular wall: Computer modelling comparing bipolar and unipolar modes," Http://Dx.Doi.Org/10.3109/02656736.2014.949878, **30**(6), 372–384, 2014, doi:10.3109/02656736.2014.949878.
[11] R.R. Tilz, A. Rillig, A.M. Thum, A. Arya, P. Wohlmuth, A. Metzner, S. Mathew, Y. Yoshiga, E. Wissner, K.H. Kuck, F. Ouyang, "Catheter ablation of long-standing persistent atrial fibrillation: 5-year outcomes of the hamburg sequential ablation strategy," Journal of the American College of Cardiology, **60**(19), 1921–1929, 2012, doi:10.1016/J.JACC.2012.04.060.
[12] G. SIVAGANGABALAN, M.A. BARRY, K. HUANG, J. LU, J. POULIOPOULOS, S.P. THOMAS, D.L. ROSS, A. THIAGALINGAM, P. KOVOOR, "Bipolar Ablation of the Interventricular Septum is More Efficient at Creating a Transmural Line than Sequential Unipolar Ablation," Pacing and Clinical Electrophysiology, **33**(1), 16–26, 2010, doi:10.1111/J.1540-8159.2009.02602.X.
[13] B. Avitall, "Computer simulation of ventricular tachyarrhythmias during coronary artery ligation and release," Journal of Electrocardiology, **12**(1), 17–22, 1979, doi:10.1016/S0022-0736(79)80040-0.
[14] A.F. Members, A.J. Camm, G.Y.H. Lip, R. De Caterina, I. Savelieva, D. Atar, S.H. Hohnloser, G. Hindricks, P. Kirchhof, E.C. for P.G. (CPG), J.J. Bax, H. Baumgartner, C. Ceconi, V. Dean, C. Deaton, R. Fagard, C. Funck-Brentano, D. Hasdai, A. Hoes, P. Kirchhof, J. Knuuti, P. Kolh, T. McDonagh, C. Moulin, B.A. Popescu, Ž. Reiner, U. Sechtem, P.A. Sirnes, M. Tendera, et al., "2012 focused update of the ESC Guidelines for the management of atrial fibrillationAn update of the 2010 ESC Guidelines for the management of atrial fibrillationDeveloped with the special contribution of the European Heart Rhythm Association," European Heart Journal, **33**(21), 2719–2747, 2012, doi:10.1093/EURHEARTJ/EHS253.
[15] N. Tsuda, K. Kuroda, Y. Suzuki, "An inverse method to optimize heating conditions in RF-capacitive hyperthermia," IEEE Transactions on Biomedical Engineering, **43**(10), 1029–1037, 1996, doi:10.1109/10.536904.
[16] P.M. Van Den Berg, A.T. De Hoop, A. Segal, N. Praagman, "A Computational Model of the Electromagnetic Heating of Biological Tissue with Application to Hyperthermic Cancer Therapy," IEEE Transactions on Biomedical Engineering, BME-**30**(12), 797–805, 1983, doi:10.1109/TBME.1983.325081.
[17] J.-W. Jeong, M.H. Jeong, K.H. Yun, S.K. Oh, E.M. Park, Y.K. Kim, S.J. Rhee, E.M. Lee, J. Lee, N.J. Yoo, N.-H. Kim, J.C. Park, "Echocardiographic Epicardial Fat Thickness and Coronary Artery Disease," Circulation Journal, **71**(4), 536–539, 2007, doi:10.1253/CIRCJ.71.536.
[18] M. Trujillo, E.B.-I.J. of Hyperthermia, undefined 2013, "Review of the mathematical functions used to model the temperature dependence of electrical and thermal conductivities of biological tissue in radiofrequency," Taylor & Francis, **29**(6), 590–597, 2013, doi:10.3109/02656736.2013.807438.
[19] C.E. Hamner, D.D. Potter, K.R. Cho, A. Lutterman, D. Francischelli, T.M. Sundt, H. V. Schaff, "Irrigated radiofrequency ablation with transmurality feedback reliably produces cox maze lesions in vivo," Annals of Thoracic Surgery, **80**(6), 2263–2270, 2005, doi:10.1016/J.ATHORACSUR.2005.06.017.
[20] Callegaro, Luca. Electrical impedance: principles, measurement, and applications. CRC Press, 2012.
[21] Pinder, George F., and William G. Gray. Finite element simulation in surface and subsurface hydrology. Elsevier, 2013.
[22] M. Ezzat, N. AlSowayan, … Z.A.-M.-H. and M., undefined 2014, "Fractional modelling of Pennes' bioheat transfer equation," Springer, **50**(7), 907–914, 2014, doi:10.1007/s00231-014-1300-x.
[23] T. Nakata, N. Takahashi, "Direct finite element analysis of flux and current distributions under specified conditions," IEEE Transactions on Magnetics, **18**(2), 325–330, 1982, doi:10.1109/TMAG.1982.1061870.
[24] A. González-Suárez, E. Berjano, J.M. Guerra, L. Gerardo-Giorda, "Computational Modeling of Open-Irrigated Electrodes for Radiofrequency Cardiac Ablation Including Blood Motion-Saline Flow Interaction," PLOS ONE, **11**(3), e0150356, 2016, doi:10.1371/JOURNAL.PONE.0150356.
[25] J.M. GUERRA, E. JORGE, S. RAGA, C. GÁLVEZ-MONTÓN, C. ALONSO-MARTÍN, E. RODRÍGUEZ-FONT, J. CINCA, X. VIÑOLAS, "Effects of Open-Irrigated Radiofrequency Ablation Catheter Design on Lesion Formation and Complications: In Vitro Comparison of 6 Different Devices," Journal of Cardiovascular Electrophysiology, **24**(10), 1157–1162, 2013, doi:10.1111/JCE.12175.
[26] L. Zou, L. Jing, V. Cvetkovic, "Shear-enhanced nonlinear flow in rough-walled rock fractures," International Journal of Rock Mechanics and Mining Sciences, **97**, 33–45, 2017, doi:10.1016/J.IJRMMS.2017.06.001.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 88-97 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Innovations in Recruitment—Social Media

Lucie Böhmová*, Antonín Pavlíček

Department of Systems Analysis, Prague University of Economics and Business, Prague, 130 67, Czech Republic

ARTICLE INFO

Article history:
Received: 16 July, 2021
Accepted: 02 October, 2021
Online: 30 November, 2021

Keywords:
Human Resources Management
Innovations in Recruitment
Recruitment
Social Networks
Data Analysis
Social Network Analysis
Model

ABSTRACT

The main objective and contribution of the paper is to describe the creation of a model to support recruitment using social media information and its deployment in practice. The model includes the design of an automated solution for downloading social media data and a proposal for the subsequent analysis and creation of a predictive model based on the evaluation of user behavior on a social network. A final assessment of the effectiveness of the proposed model was made through formal validation and a case study. For the case study validation, the proposed model was implemented through use of the recruitment application called Prace Na Miru (PM; Tailored Work) for Facebook data extraction. A Myers Briggs Type Indicator (MBTI) personality test was used to determine the predictors of user social network behavior. Using cluster analysis and machine learning (or decision trees), a stochastic predictive model was developed to determine the personality type of candidates—the accuracy of MBTI personality category prediction ranges between 68% and 84% for individual cases, with a confirmation rate ranging between 43% and 81%. The case study confirmed the model's usefulness for supporting recruitment in a real-life deployment of the PM model.

1. Introduction

At the beginning of the twenty-first century, communication platforms that allow users to interact with each other have proliferated globally. These platforms gradually began to replicate the social structures of the real world. Within barely five years (2001–2005), social networks as we understand them today took on an established form.

These networks are no longer mere communication channels but are permeating more and more areas and blurring the boundaries between private and professional life. Social networks have also shown their potential in human resources management (HRM), especially in recruitment. Research [1] has found that 43% of organizations use social media sites to research candidates, and approximately 50% of employers report rejecting job applicants because of social media content. [2]

The current state of the labor market is not very favorable for recruitment from the point of view of organizations struggling to find suitable employees. Low unemployment and high demand for employees are forcing organizations to look for new and more effective recruitment solutions. In this respect, the use of social networks brings an advantage in two ways: they are economically advantageous, and they are an innovative solution [3]-[5].

Social media such as Facebook, Twitter, and LinkedIn are used by recruiters to attract talent by sharing the company culture and information related to career opportunities. For example, Microsoft has developed a specific website for recruitment, using job blogs and life at work videos to provide relevant information to both active and passive job seekers. Similarly, Deloitte places employee testimonials on social media (e.g., YouTube videos) to allow potential job applicants to learn more directly from the insiders. A shortcoming is the question of whether these job messages can reach the target audiences [6].

The results of research [7] confirmed that recruiters intended to use social media as it provides pre-hire benefits. Similarly, results showed social media recruitment is more compatible and less complex, and it ensures better trialability and observability. The study proved that social media recruitment was cost-effective while attracting and retaining better talent [8].

The combination of HRM and social networks in the context of recruitment issues expects an interdisciplinary approach based on applied informatics (as it is primarily based on the use of ICT),

*Corresponding Author: Lucie Böhmová, lucie.bohmova@vse.cz

but its implications extend to the fields of management and human resources (HR) [4].

In the field of applied informatics, our research primarily uses data analysis of social networks through our own designed application called **P**race Na **M**iru (abbreviation **PM**; literal translation into English—Tailored Work), which allows for automated data acquisition. It is further related to applied informatics by the selected research design, called "design science research" [9], and the resulting artifact, which is a model verified not only by meeting the theoretical assumptions but, largely, by its application in practice [10].

1.1. Research problem

The generally defined research problem is the support of recruitment in modern HRM through the use of social networks.

The authors propose a solution in the form of a recruitment support model—the PM model.

1.2. Objectives and research questions

The main objective was to create and verify the effectiveness of a model for supporting recruitment through social networks which would be generally valid and applicable to any type of organization and job position.

In order to achieve the main objective, sub-objectives (SOs) and sub-research questions (SRQs) were identified as follows:

- **SO 1**: Create and deploy the PM model to support recruitment using a specific social network within a selected problem context and validate its effectiveness.
- **SRQ 1**: What is the design of the PM model and the process of its deployment in practice?
  - **SO 1.1**: Create a custom application to extract data from a specific social network.
  - **SRQ 1.1**: How can we create an application to mine data from social networks?
  - **SO 1.2**: Identify appropriate parameters to evaluate user behavior on social networks.
  - **SRQ 1.2**: What are the appropriate parameters to evaluate user behavior on social networks?
  - **SO 1.3**: Identify predictors of user behavior on social networks using appropriate analytical methods.
  - **SRQ 1.3**: What are the predictors of user behavior on social networks? And what are the appropriate analytical methods to determine them?

1.3. Methods used

In this paper, the authors used methods according to [10] and [11] in the publication *Design Science Research*, which consists of five phases:

- **Phase 1—problem awareness**: This is based on an analysis of the current state of knowledge on the issue, especially best practices and literature research. The output is the basis for the design of the artefact.
- **Phase 2—design/theme**: the result is a general tentative design of the artefact in Figure 1.
- **Phase 3—development**: the development of the indicative design of the artefact from phase 2 into the final form of the artefact, the result of this phase.

  In terms of conceptual development, the development and design/idea phases are intertwined—the case here.
- **Phase 4—evaluation**: the artifact is evaluated in order to obtain information indicative of the artifact's behavior and quality and useful to the artifact's future development. Scientific methods are used in the evaluation of the artefact — formal verification—by creating an instance in the form of the PM model. These scientific methods are of three types:
  - data analysis;
  - cluster analysis;
  - predictive modelling using decision trees (or machine learning).

  Validation is done through quantitative research in the form of a questionnaire survey among experts in practice.

  Evaluation is done through qualitative research in the form of a case study and subsequent practical deployment.
- **Phase 5—Conclusion**: the last phase summarizes the research's conclusions.

This paper is primarily concerned with description of the fourth phase.

Phase 1—Problem Awareness (Background)

The main starting point for the development of the model is a recruitment application for extracting user data from a social network (or custom work).

Another important starting point for model construction is the parameters for evaluating user behavior on the social network. In the context of recruitment, the most appropriate parameters are those based on a personality test.

Therefore, it was necessary to specify requirements for selecting an appropriate test according to the purpose of the PM model, which are [2, 12] as follows:

- assessment of personality characteristics;
- assessment of interpersonal characteristics;
- evaluation of work characteristics;
- relevance for recruitment;
- speed;
- transparency;
- the ability to complete the online test from anywhere;
- immediate evaluation without the need for multiple specialist (e.g., psychologist) costs.

There are several psychological models and personality tests for assessing jobseekers that can serve as parameters to evaluate

behavior [13]-[16]. Classic tests include the Eysenck's Personality Inventory (EPI), which is primarily aimed at identifying temperament based on measures of introversion (or extroversion) and stability (or lability); the Achievement Motivation Inventory (AMI), used to measure various dimensions of performance motivation in a professional context and useful in the selection of employees, team creation, or sport psychology; the Personal Competence Profile, which takes into account findings on emotional intelligence and social competence; and the Myers-Briggs Type Indicator (MBTI) model of personality types, which is widespread among organizations and is mainly used by HRM in recruitment, by health professions, and by educational programs to address the diversity of personalities that exist [2], [17].

Based on the analysis of the abovementioned characteristics, the MBTI personality test was chosen as the most suitable and was therefore used in the PM model [17]-[19]. The MBTI personality test meets all the above requirements—this is explored in more detail in section 1.6 "Predictors for user behavior evaluation on social media."

This implies the need to identify predictors using appropriate analytical methods to assess user behavior on the social network. The most suitable predictors in this work were found to be areas of interest, such as favorite book, TV series, film, or sports personality.

The PM model development procedure is based on a model abstraction technique which allows the projection of alternatives through a model experiment. This, in effect, helps to track changes in individual parameters and their relationships at the desired level of simplification. The model is then used to explore, in particular, relationships, causes, and analogies. The methodological framework for the development of the recruitment support model is based on the CRISP-DM methodology [20].

The fundamental constraints in the creation of the model are mainly based on legislation. Organizations must comply with the General Data Protection Legislation (GDPR) and other employment legislation when identifying and verifying candidate information on social media. At the same time, it must not engage in discriminatory behavior. Therefore, such sensitive information about candidates is omitted from the model, even though it can be obtained from social networks relatively easily.

1.4. Validity and reliability of MBTI

We know from discussions of psychological and personality tests that critics have pointed out these tests may not be universally valid, and the results may not be reliable. For example, the popular-science journal *Scientific American* has published articles in recent years (e.g., "How Accurate Are Personality Tests?" [21] and "Personality Tests with Deep-Sounding Questions Provide Shallow Answers about the True You" [22]) that question the validity of personality tests.

However, peer-reviewed scientific journals, most recently the *Journal of Best Practices in Health Professions Diversity: Education, Research and Policy*, have published systematic (meta-analysis) reviews of the literature concerning its validity and reliability. The article "Validity and Reliability of the Myers-Briggs Personality Type Indicator: A Systematic Review and Meta-analysis," on the contrary, confirms the validity of MBTI tests on a large sample of 221 studies [23], [24].

Authors agree that the MBTI instrument has reasonable construct validity. Their results indicate that the extravert-introvert, sensing-intuition, and judging-perceiving subscales have satisfactory reliabilities of .75 or higher and that the thinking-feeling subscale has a reliability of 61 [23].

The only mentioned problem is that the majority of studies are usually conducted on college-age students; thus, the evidence to support the tool's utility applies more to this group, and careful thought should be given when applying it to other individuals.

This limitation is not a problem in our research since the respondents in our case were also primarily students and young alumni.

Phases 2 and 3: Design and Development

The authors' application Prace Na Miru is the essential component in developing the PM model. Running on the website prace-na-miru.eu/en/, the application harvests user data.

The collected data is saved in JSON format within the MongoDB database. Heroku is used to host both the page and the database. JavaScript and Node.js frameworks were used to write the code. Candidates use the Facebook login button to access the website and to approve the download/export of his data into our database.

The sample selection is derived through self-selection among students and graduates of the Prague University of Economics and Business. Information about the application Prace Na Miru was distributed through an email newsletter to the target student and alumni groups. The application was also published on social media in suitable groups, during a university job fair called *Šance*, and on the web page of the college career center. There were 960 unique candidates who registered with and logged in to the application from October 2016 to January 2017. 624 of them responded to the subsequent email invitation to complete the MBTI test and 484 completed the questionnaire in full. This data has been used for further analysis.

1.5. Data analysis

The data collected by the Prace Na Miru application was processed in the SW called Knime [25]. The spectrum of publicly available data was organized into categories. Gender, log-in device, list of friends, date and year of birth, and other details were provided by users. Some of this data, such as favorites, are important for recruitment (or, more accurately, for the construction of the model), for example, favorite TV series, favorite athletes, and so forth. Some basic demographic data about users, such as gender, age, and education, was also available.

Table 1: Correlation of Publicly Accessible User Information to Categories

Categories of publicly available data	Friend lists	TV series	Athletes	Books	Events	Games	Music	Likes	Films
Friend lists	1	0.09	0.11	0.09	0.31	-0.06	0.16	0.29	0.04
TV series	0.09	1	0.17	0.32	0.11	0.23	0.57	0.09	0.57
Athletes	0.11	0.17	1	0.15	0.07	0.25	0.16	0.03	0.12

Books	0.09	0.32	0.15	1	-0.09	0.32	0.27	0.07	0.52
Events	0.31	0.11	0.07	-0.09	1	-0.01	0.14	0.47	0.03
Games	-0.06	0.23	0.25	0.32	-0.01	1	0.11	0.04	0.31
Music	0.16	0.57	0.16	0.27	0.14	0.11	1	0.22	0.49
Likes	0.29	0.09	0.03	0.07	0.47	0.04	0.22	1	0.12
Films	0.04	0.57	0.12	0.52	0.03	0.31	0.49	0.12	1

Spearman's correlation coefficient was used to detect correlations within the collected data (Table 1). The existence of low bond tightness for the categories "favorite book" and "series" (0.32) and "favorite game" and "book" (0.32) was proven. Medium-bond tightness was found in the categories "likes" and "events" (0.47) and "favorite movie" and "book" (0.52). Significant-bond tightness was in the categories "favorite music" and "series" (0.57) and "favorite music" and "movie" (0.57). It turns out that users sharing information about their favorite series also share information about their favorite music and movies.

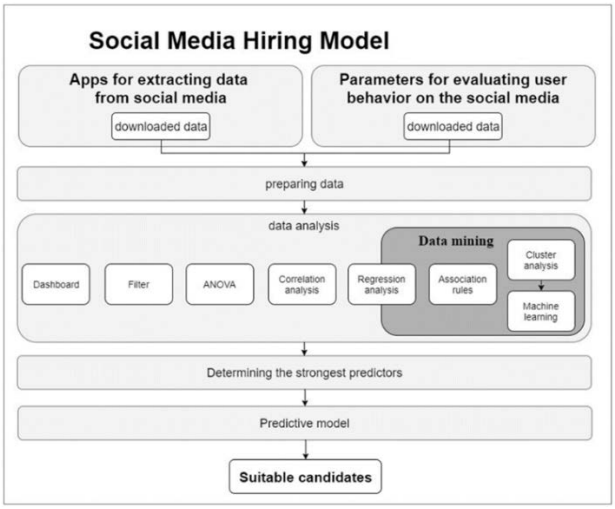


Figure 1: Social media hiring model

1.6. Predictors for user behavior evaluation on social media

An important part of the specification for the construction of a functional PM model is identifying predictors to evaluate user behavior within the social media environment. The MBTI personality test [12], [26], [27] was used in the PM model. The MBTI personality test combines four criteria:

- approach to the surrounding environment—introversion (I) or extroversion (E);
- way of obtaining information—intuition (N) or sense (S);
- way of evaluating information—feeling (F) or thinking (T);
- lifestyle—perception (P) or judging (J).

Merging the four criteria mentioned above results in 16 personality types. Figure 2 shows the number of persons represented by MBTI type. The most represented type is ENTJ, for whom employment in leadership positions such as leader or manager is suitable. This is followed by INTJ (scientist, builder), ENFJ (teacher), and INFJ (writer, creator).

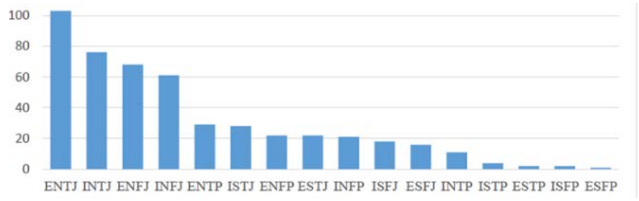


Figure 2: Number of persons represented by MBTI type

1.7. MBTI test and its evaluation

All users logged into Prace Na Miru were tested by the MBTI test. The response rate at this point was 50.4%. This fact helped to determine personality types and characteristic features of the questioned users and compared them with their online approach to social media. The sample size was 484 unique respondents.

1.8. PM model creation

For the creation of the PM model, the authors used training data (N=960) upon which the model learned how to decide the personality types of the candidates. The result of the modeling phase—Model PM—is shown in Figure 3.

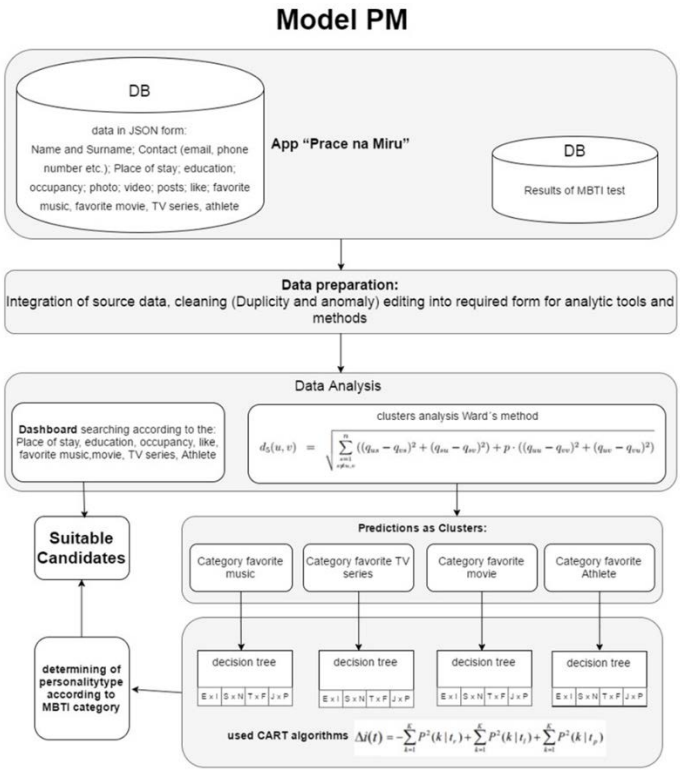


Figure 3: PM model

Phase 4: Model PM Evaluation

The model validation is realized by its implementation in a specific problem context: the recruitment of suitable candidates for positions at the University Development and Counselling Centre and internship programs at the University xPORT Business Accelerator. The target user group upon which the model is validated are students and recent university graduates.

The evaluation is performed by obtaining valuable and reliable information proving the quality of the PM model.

1.9. Formal verification

The formal verification is based on the construction of the PM model. Scientific methods included are as follows:

- data analysis and cluster analysis;
- predictive modelling using decision trees (or machine learning).

In developing the PM model, the authors used the MBTI personality test to determine predictors of user behavior on the social network, which helps to diagnose candidate characteristics such as perception of the environment or the way they acquire and process information. In addition, the PM model includes instructions for an automated solution for downloading user data from Facebook using a customized recruitment application called Jobs. Also included is a proposal for data analytics. Specifically, the PM model describes cluster analysis, decision trees, and predictive model(s) for determining personality type (which is based on the MBTI test).

After preprocessing the data, the authors segmented users into clusters, which are used in the PM model as predictors of behavior according to MBTI categories. Due to the large amount of data, it was necessary to select only some clusters that have predictive power. Out of 28 possible attributes, the author identified four with the highest predictive power as predictors:

- favorite music;
- favorite TV series;
- favorite movie; and
- favorite athlete.

For each attribute, the author created decision trees using the BigML [28] software tool, which allowed for their categorization according to the MBTI personality categories.

For each decision tree, the author includes a graphic illustration of the main tree branch that is the strongest and a ray graph that shows the overall decision tree with the relative representation of number of persons (according to the circle plot) and the confidences for the given personality category. The individual circles represent the path of the tree through the number of occurrences in the clusters. The color represents the amount of confidence of the decision tree. The darker, more saturated the color in the ray graph, the more confident is the decision tree. If confidence is 0%, the result is equivalent to random selection [29], [30].

PM model for the Attribute "Favorite Music"

For the creation of decision trees, clusters were created for individual attributes using the software tool Pajek [31]. As an example, the author presents in Table 2 the first five most frequently occurring items in each cluster.

Table 2: Cluster Overview for the Attribute "Favorite Music"

Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Cluster F	Cluster G	Cluster H	Cluster I
AC/DC	Bruno Mars	Snoop Dogg	Ed Sheeran	Electro Swing	Walk Off The Earth	Linkin Park	The Beatles	Ben Cristovao
Calvin Harris	Wiz Khalifa	Britney Spears	Fedde Le Grand	OneRepublic	Universal Music CZ	Tomáš Klus	Radiohead	Taylor Swift
Yellow card	Monika Bagárová	Fergie	INNA	Philip Glass	George Ezra	Nightwish	Kings Of Leon	Rihanna
The Prodigy	Adam Levine	T-Pain	Johny Machette	Coldplay	Lindsey Stirling	Apocalyptica	Queens of the Stone Age	Beyoncé
Christina Perri	Tyrese Gibson	Sean Paul	Ellie Goulding	River Hollywood Undead	Thirty Seconds to Mars	Avril Lavigne	The Black Keys	Kryštof

1.10. Decision tree for category "extroverted (E) or introverted (I) perception of the environment"

Below is the text notation of the predictive model for the category "extroverted (E) or introverted (I) perception of the environment":

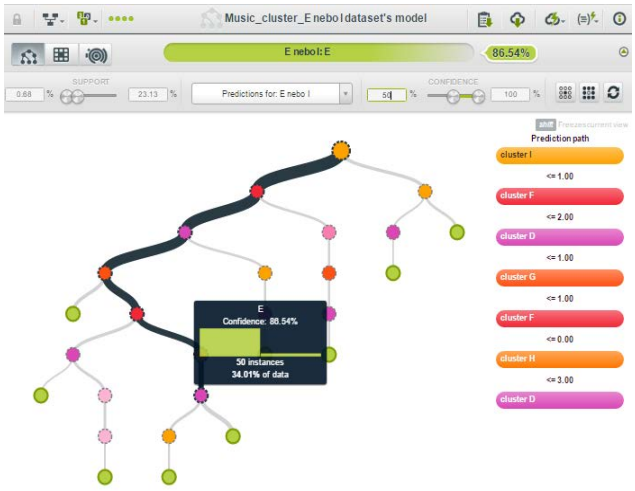


Figure 4: PM model for the attribute "favorite music"—decision tree for category "extroverted (E) or introverted (I) perception of the environment"

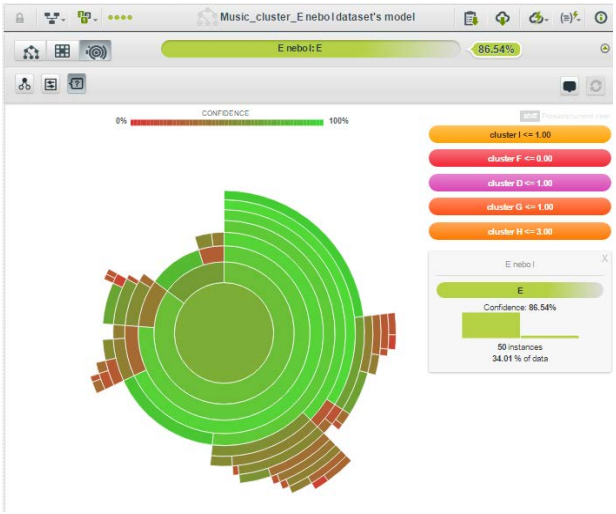


Figure 5: PM model for the attribute "favorite music"—decision tree bar graph for category "Perception of the environment extroverted (E) or introverted (I)"

- **Data distribution**

  E: 70.75% (104 instances)

  I: 29.25% (43 instances)

- **Predicted distribution**

  E: 73.47% (108 instances)

  I: 26.53% (39 instances)

- **Field importance**

1. cluster I: 28.47%
2. cluster F: 23.07%
3. cluster D: 19.41%
4. cluster G: 11.36%
5. cluster C: 5.59%
6. cluster E: 5.16%
7. cluster H: 4.48%
8. cluster B: 2.45%

1.11. Decision tree for category "sensory (S) or intuition (N) information acquisition"

Below is the text notation of the predictive model for the category "sensory (S) or intuition (N) information acquisition":

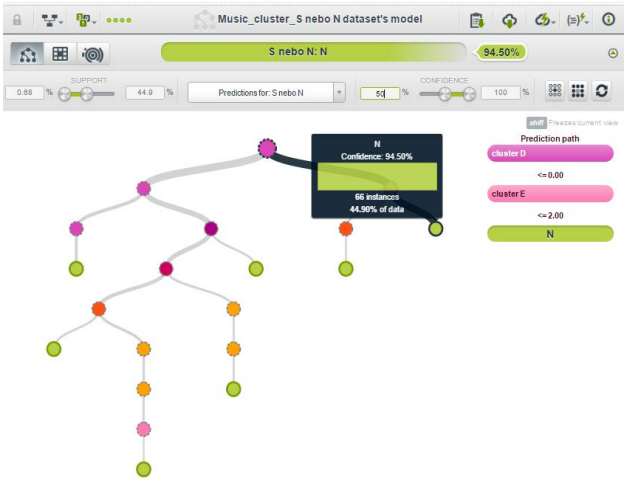


Figure 6: PM model for the attribute "favorite music"—decision tree for category "sensory (S) or intuition (N) information acquisition"

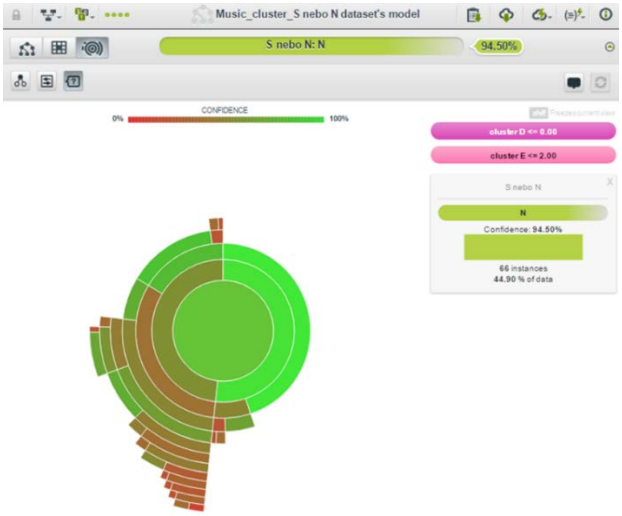


Figure 7: PM model for the attribute "favorite music"—decision tree bar chart for category "Sensory (S) or intuition (N) information acquisition"

- **Data distribution**

  N: 82.31% (121 instances)

  S: 17.69% (26 instances)

- **Predicted distribution**

  N: 82.99% (122 instances)

  S: 17.01% (25 instances)

- **Field importance**

  1. cluster D: 47.21%
  2. cluster A: 12.80%
  3. cluster I: 11.08%
  4. cluster B: 9.43%
  5. cluster E: 7.05%
  6. cluster G: 6.97%
  7. cluster C: 2.82%
  8. cluster F: 1.88%
  9. cluster H: 0.75%

1.12. Decision tree for category "thinking (T) or feeling (F) information processing"

Below is the text notation of the predictive model for the category "thinking (T) or feeling (F) information processing":

- **Data distribution**

  F: 49.66% (73 instances)

  T: 50.34% (74 instances)

- **Predicted distribution**

  F: 48.30% (71 instances)

  T: 51.70% (76 instances)

- **Field importance**

  1. cluster F: 18.69%
  2. cluster E: 16.51%
  3. cluster C: 15.07%
  4. cluster G: 14.95%
  5. cluster H: 8.88%
  6. cluster D: 8.50%
  7. cluster I: 6.85%
  8. cluster B: 6.21%
  9. cluster A: 4.32%

1.13. Decision tree for category "inference (J) or perception (P) lifestyle"

What follows is the textual notation of the predictive model for the category "inference (J) or perception (P) lifestyle"

- **Data distribution**

  J: 76.87% (113 instances)

  P: 23.13% (34 instances)

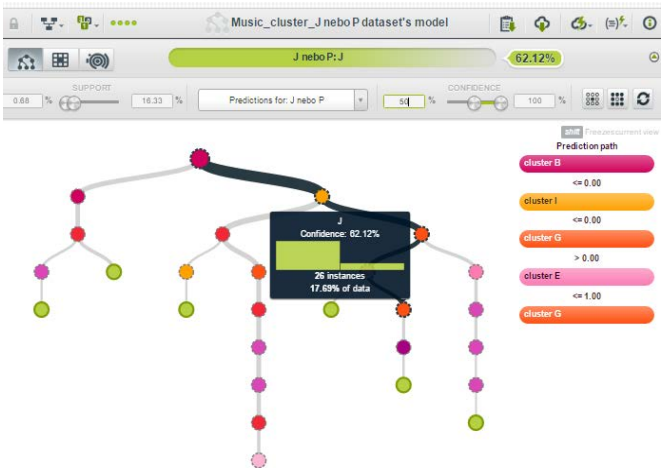


Figure 10: PM model for the attribute "favorite music"—decision tree for category "inference (J) or perception (P) lifestyle"



Figure 11: PM model for the attribute "favorite music"—decision tree bar chart for category "inference (J) or perception (P) lifestyle"

- **Predicted distribution**

  J: 76.87% (113 instances)

  P: 23.13% (34 instances)

- **Field importance**

  1. Cluster D: 17.31%
  2. Cluster F: 16.94%
  3. Cluster I: 16.50%
  4. Cluster B: 14.91%
  5. Cluster G: 13.03%
  6. Cluster E: 9.45%
  7. Cluster C: 7.37%
  8. Cluster H: 3.43%
  9. Cluster A: 1.06%

2. PM model Validation

The validation of the PM model was carried out on 198 people from the target group.

The results of the individual predictive models are specified below in terms of the percentage of cases confirming the accuracy of the PM model and the average confidence of each predictive model.

2.1. Predictive PM model for attribute "favorite music":

- Category "extroverted (E) or introverted (I) perception of the environment"

  The model confirmed the correct result 83% of the time. The average confidence of the model is 66%.

- Category "sense (S) or intuition (N) information acquisition"

  The model confirmed the correct result in 82% of cases. The average confidence of the model is 81%.

- Category "thinking (T) or feeling (F) information processing"

  The model confirmed the correct result in 77% of cases. The average confidence of the model is 63%.

- Category "inference (J) or perception (P) lifestyle"

  The model confirmed the correct result in 80% of cases. The average confidence of the model is 51%.

2.2. Predictive model PM for attribute "favorite TV series":

- Category "extroverted (E) or introverted (I) perception of the environment"
- The model confirmed the correct result in 80% of cases. The average confidence of the model is 56%.
- Category "sense (S) or intuition (N) information acquisition"
- The model confirmed the correct result in 70% of cases. The average confidence of the model is 57%.
- Category "thinking (T) or feeling (F) information processing"
- The model confirmed the correct result 74% of the time. The average model confidence is 43%.
- Category "inference (J) or perception (P) lifestyle"
- The model confirmed the correct result in 81% of cases. The average confidence of the model is 49%.

2.3. Predictive model PM for attribute "favorite movie":

- Category "extroverted (E) or introverted (I) perception of the environment"

The model confirmed the correct result in 79% of cases. The average confidence of the model is 56%.

- Category "sense (S) or intuition (N) information acquisition"

  The model confirmed the correct result in 84% of cases. The average model confidence is 49%.

- Category "thinking (T) or feeling (F) information processing"

  The model confirmed the correct result in 78% of cases. The average confidence of the model is 51%.

- Category "inference (J) or perception (P) lifestyle"

  The model confirmed the correct result in 75% of cases. The average confidence of the model is 59%.

2.4. Predictive model PM for attribute "favorite athlete":

- Category "extroverted (E) or introverted (I) perception of the environment"

  The model confirmed the correct result in 71% of cases. The average confidence of the model is 53%.

- Category "sense (S) or intuition (N) information acquisition"

  The model confirmed the correct result in 80% of cases. The average confidence of the model is 44%.

- Category "thinking (T) or feeling (F) information processing"

  The model confirmed the correct result in 79% of cases. The average confidence of the model is 53%.

- Category "inference (J) or perception (p) lifestyle"

  The model confirmed the correct result in 68% of cases. The average confidence of the model is 44%.

Using cluster analysis and machine learning (or decision trees), a stochastic predictive model (see Figure 12) was developed to determine the personality type of the candidates—the accuracy of the MBTI personality category prediction ranges from 68% to 84% for individual cases, with a confirmation rate ranging from 43% to 81%. The case study confirmed the model's usefulness for supporting recruitment, and it has been put into practice.



Figure 12: Decision tree bar chart for category "inference (J) or perception (P) lifestyle"

2.5. Validation and evaluation

Validation and evaluation are an important part of the PM model deployment for actual use at the University Counselling Centre (CC UEP) and the Universtity xPORT Business Accelerator (xPORT) as described in the case study below. In so doing, we used the processing method described in [32] to create the case study. The proposal and case study preparation process were separated into four steps:

1. At the March 2017 employment fair *Šance*, we reached out to employers on behalf of CC UEP with an offer to use the PM model to assist in the recruitment of interns.
2. Companies sent out job descriptions and completed surveys for each vacancy. There was a link in the survey to a dashboard that searched the app (Prace Na Miru) data for certain qualities that are available on Facebook.
3. The PM model was tested and verified on an agreed-upon employment position from *Šance* and the internship program. According to the survey's specifications, the authors contacted chosen users from the application Prace Na Miru and offered them a job. Users interested in the opportunity submitted their CV to the perspective employers.
4. Feedback from employers about the quality of the proposed candidates was collected via an online survey containing 11 questions. The feedback response rate was 54%.

Implementation of steps one through four occurred between March and May 2017. In total, 95 companies were involved.

The survey provided feedback from professional HR recruiters on the PM model. 88% of respondents were positive/liked the PM model.

For recruiters the primarily relevant demographic information from social media was, for example, birth date and place, education, and social media network behavior, such as wall posts, photos, and videos.

Other information (favorite series, movie, music, athlete) which we use in the PM model as predictors were perceived by respondents as less important. Therefore, the PM model has an additional value in the user's behavior evaluation.

3. Discussion

The article has shown how the PM model can be successfully deployed in practice. The case study has proven that the PM model is generally valid and feasible for various companies and job positions.

It has been shown that the use of PM model requires extensive technical and analytical knowledge, especially in the data analysis phase. Therefore recruiters (specifically HR departments) should receive a functional version of the PM model, matching the needs of their particular segment, job position, and candidate target group.

Theoretically, the recruiting app Prace na Miru (Tailored Work) can gather data from any social media network with an open API. Different personality tests or other typological evaluation tests can be chosen as well.

Practically any reasonable test can be used to set parameters to evaluate predictors, providing the authors select the correct analytical method as well.

This article has shown the same results as [33].

Benefits of the Model

The PM model predicts the inclusion of a candidate within a personality type according to the MBTI categorization based on the candidate's behavior on Facebook. The PM model is assembled upon the sophisticated creation of clusters that are based on data categorization automatically obtained via a custom-created Facebook app, Prace Na Miru; its data analysis; and MBTI test evaluation.

The establishment of predictors is a major benefit of this research. Employers will be able to choose and target more accurate applicants on social media. Another advantage of the research is the model's usage in practice. The authors have verified this within the case study and also deployed it in the recruiting solution for CC UEP and the xPORT internship program.

Finally, this proposed solution has the benefit of being adaptable to multiple target groups depending on the needs of the organization.

Research limitations:

- Because social media is a rapidly changing and continually evolving medium, the major problem is the application's long-term viability.
- Only applicants with social media profiles can be evaluated by the PM model.
- The focus is limited to the Czech/Slovak labor market and its legal framework [34].
- Despite testing the model on real data, the authors are aware of possible model distortions [33].

Possible further research and activity:

- Involve students and alumni from additional Czech universities, as well as from other countries.
- Include other social media such as Instagram or Twitter.

4. Conclusion

The main output of the thesis is an artefact in the form of a model for recruitment support, its validation, and a practical application (the PM model).

The article primarily describes the fourth phase of verification, which consists of two stages: formal verification and case study. In the formal verification, a concrete instance of the artifact (in the form of a recruitment support model) is created based on the artifact—the PM model.

The stochastic predictive model confirmed the accuracy of the MBTI personality category prediction as ranging between 68% and 84% for individual cases, with confounding ranging between 43% and 81%. The case study confirmed the effectiveness of the model at supporting recruitment and the real-life deployment of the model.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper was processed with a contribution from the Prague University of Economics and Business, IG Agency, OP VVV IGA/A, CZ.02.2.69/0.0/0.0/19_073/0016936, grant number 05/2021 and grant IGS 13/2021.

References

[1] J. B. Becton, H. J. Walker, J. B. Gilstrap, P. H. Schwager, Social media snooping on job applicants: The effects of unprofessional social media information on recruiter perceptions. Personnel Review, **48**(5), 1261-1280, 2019, doi:http://dx.doi.org/10.1108/PR-09-2017-0278.

[2] N. Ahmed, S. Khan, K. Latif. Job description ontology, Paper presented at the Proceedings - 14th International Conference on Frontiers of Information Technology, FIT 2016, 217-222.

[3] M. Armstrong. Řízení lidských zdrojů: nejnovější trendy a postupy. Praha: Grada Publishing, 2015, ISBN 9788024752587.

[4] L. Gelinas, R. Pierce, S. Winkler, I. G. Cohen, H. F. Lynch, B. E. Bierer, Using social media as a research recruitment tool: Ethical issues and recommendations, American Journal of Bioethics, **17**(3), 3-14, 2017.

[5] LinkedIn, TALENT TRENDS 2014. LinkedIn Talent Solutions, 2014, available: https://business.linkedin.com/content/dam/business/talent-solutions/global/en_US/c/pdfs/linkedin-talent-trends-2014-en-us.pdf.

[6] D. Smith, Research shows role of pre-hire and post-hire outcomes in effectiveness of social media recruitment, Human Resource Management International Digest, **29**(5), 17-19, 2021, doi:http://dx.doi.org/10.1108/HRMID-02-2021-0033.

[7] A. Muduli, J. J. Trivedi, Social media recruitment: The role of credibility and satisfaction. Evidence - Based HRM, **8**(2), 237-251, 2020, doi:http://dx.doi.org/10.1108/EBHRM-08-2019-0069.

[8] J. L. Bender, A. B. Cyr, L. Arbuckle, L. E. Ferris, Ethics and privacy implications of using the internet and social media to recruit participants for health research: A privacy-by-design framework for online recruitment. Journal of Medical Internet Research, **19**(4), 2017.

[9] A. Dresch, D. P. Lacerda, Jr. Antunes, Design Science Research: A Method for Science and Technology Advancement, New York: Springer, 2015, ISBN 978-3-319-07373-6.

[10] V. K. Vaishnavi, W. Kuechler, Design Science Research Methods and Patterns, CRC Press, Boca Raton, FL, 2015.

[11] Z. Molnár, S. Mildeová, H. Řezanková, R, Brixí, J. Kalina, Pokročilé metody vědecké práce, Zeleneč: Profess Consulting, Věda pro praxi (Profess Consulting), 2012, ISBN 978-80-7259-064-3.

[12] Y. Melanthiou, F. Pavlou, E. Constantinou, The use of social network sites as an E-recruitment tool, Journal of Transnational Management, **20**(1), 31-49, 2015.

[13] C. Khatri, J. S. Chapman, J. Glasbey, M. Kelly, D. Nepogodiev, A. Bhangu, J. E. Fitzgerald, Social media and internet driven study recruitment: Evaluating a new model for promoting collaborator engagement and participation. PLoS ONE, **10**(3), 2015.

[14] D. Muntinga, M. Moorman, E. Smit, 'Introducing COBRAs', International Journal of Advertising, **30**(1), 13-46, Business Source Ultimate, EBSCOhost, 2011, ISSN: 0265-0487.

[15] Oracle, Social Recruiting Guide: How to Effectively Use Social Networks, 2012, available: http://www.oracle.com/us/media1/effectively-use-social-networks-1720586.pdf.

[16] H. Yin, B. Cui, L. Chen, Z. Hu, X. Zhou, Dynamic user modeling in social media systems, ACM Transactions on Information Systems, **33**(3), 2015.

[17] M. Miller, Face-to-face with the digital folk: The ethics of fieldwork on facebook, Folk culture in the digital age: The emergent dynamics of human interaction (pp. 212-232), 2012.

[18] A. Ladkin, D. Buhalis, Online and social media recruitment: Hospitality employer and prospective employee considerations. International Journal of Contemporary Hospitality Management, **28**(2), 327-345, 2016.

[19] M. Mattare, Revisiting Understanding Entrepreneurs Using the Myers-Briggs Type Indicator®, Journal of Marketing Development and Competitiveness, **9**(2), 114-119, 2015, ISSN: 21552843.

[20] P. Chapman, J. C. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C, Shearer, R. Wirth, CRISP-DM 1.0, Step-by-step data mining guide, 2020, Available: ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf.

[21] A. Chen, How Accurate Are Personality Tests?, Scientific American, Oct. 10, 2018, accessed: Sep. 14, 2021, available: https://www.scientificamerican.com/article/how-accurate-are-personality-tests/.

[22] R. Stein, Personality Tests with Deep-Sounding Questions Provide Shallow Answers about the True You, Scientific American, 2018, accessed: Sep. 14, 2021, available: https://www.scientificamerican.com/article/personality-tests-with-deep-sounding-questions-provide-shallow-answers-about-the-true-you/.

[23] K. Randall , M. Isaacson, C. Ciro, Validity and Reliability of the Myers-Briggs Personality Type Indicator: A Systematic Review and Meta-analysis, Journal of Best Practices in Health Professions Diversity: Education, Research & Policy, vol. 10, no. 1, pp. 1–27, Spring 2017.
[24] Jobvite, Social Recruiting Survey, 2014, available: https://www.jobvite.com/wp-content/uploads/2014/10/Jobvite_SocialRecruiting_Survey2014.pdf.
[25] Knime, 2017, available: https://www.knime.com/.
[26] I. B. Myers, M. H. Mccaulley, R. Most, Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator, Palo Alto, Ca.: Consulting Psychologists Press, 1985, ISBN 08-910-6027-8.
[27] C.E. Fretwell, C. C. Lewis, M. Hannay, Myers-Briggs Type Indicator, A/B Personality Types, and Locus of Control: Where Do They Intersect?, American Journal of Management, **13**(3), 57-66, 2013, ISSN: 21657998.
[28] BigML, 2017, available: https://bigml.com/.
[29] M. N. Manu, K. R. Anandakumar, A current trends in big data landscape. Paper presented at the 2015 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2015.
[30] Q. Tran, Using ANOVA to analyze modified gini index decision tree classification, Paper presented at the Proceedings of the 2008 International Conference on Data Mining, DMIN 2008, 164-170, 2008.
[31] Pajek, Analysis and visualization of large networks, 2017, available: http://mrvar.fdv.uni-lj.si/pajek/.
[32] R.K. Yin, Case study research: design and methods, 5. Edice, Los Angeles, Calif. [u.a.]: Sage Publ, 2014, ISBN 14-522-4256-9.
[33] L. Böhmová, A. Pavlíček, 'How the Modern Human Resources Management Can Take Advantage of Information From Social Media while Recruiting', in 2020 International Conference on Engineering Management of Communication and Technology (EMCTECH), Oct. 2020, 1–6. doi: 10.1109/EMCTECH49634.2020.9261522.
[34] Official Journal of the European Union, Ec.europa, 2016, available: http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 90-100 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Emotion Mining from Speech in Collaborative Learning

Nasrin Dehbozorgi[*,1], Mary Lou Maher[2], Mohsen Dorodchi[3]

[1]*Department of Software Engineering, Kennesaw State University, GA, 30004, USA*

[2]*Department of Software and Information Systems, UNC Charlotte, NC, 27012, USA*

[3]*Department of Computer Science, UNC Charlotte, NC, 27012, USA*

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 27 August, 2021
Online: 16 September, 2021

Keywords:
Active learning
Low-stake teams
Collaborative learning
Emotion mining
Aspect-based sentiment analysis
NLP

ABSTRACT

Affective states, a dimension of attitude, have a critical role in the learning process. In the educational setting, affective states are commonly captured by self-report tools or based on sentiment analysis on asynchronous textual chats, discussions, or students' journals. Drawbacks of such tools include: distracting the learning process, demanding time and commitment from students to provide answers, and lack of emotional self-awareness which reduces the reliability. Research suggests speech is one of the most reliable modalities to capture emotion and affective states in real-time since it captures sentiments directly. This research, which is an extension of the work originally presented in FIE conference'20, analyses students' emotions during teamwork and explores the correlation of emotional states with students' overall performance. The novelty of this research is using speech as the source of emotion mining in a learning context. We record students' conversations as they work in low-stake teams in an introductory programming course (CS1) taught in active learning format and apply natural language processing algorithms on the speech transcription to extract different emotions from conversations. The result of our data analysis shows a strong positive correlation between students' positive emotions as they work in teams and their overall performance in the course. We conduct aspect-based sentiment analysis to explore the themes of the positive emotions and conclude that the student's positive feelings were mostly centered around course-related topics. The result of this analysis contributes to future development of predictive models to identify low-performing students based on the emotions they express in teams at earlier stages of the semester in order to provide timely feedback or pedagogical interventions to improve their learning experience.

1. Introduction

Social constructs have an essential role in the learning process which highlights the importance of teamwork and improving students' social skills in educational settings [1], [2]. Evaluating individual student's performance in teams is a complex process, especially in low-stake teams in which the product of teamwork does not have a large contribution to the students' grade [3]. In such teams, the main goal of teamwork is peer learning and developing interpersonal skills [1], [4]. Practicing low-stake teams is common in collaborative active learning classes [3].

In collaborative active learning, class time is utilized to create an engaging learning experience for students and improve their social skills [5]. There are multiple types of pedagogy that could be classified as active learning, such as team-based learning (TBL) [6], cooperative learning [7], [8], collaborative learning [5], problem-based learning [5], or studio-based learning [9]. Most forms of active learning emphasize collaboration and the social construction of knowledge. During the past decade, active learning has been applied more often in higher education, and it highlights the importance of having structured protocols for team formation, team size, and assessment [10], [11]. The inherently complex nature of teamwork in active learning calls for methods to assess students' performance at both individual and team levels [12], [13]. This is more critical in low-stake teams which are used in introductory-level courses. To address this evaluation issue theories on team performance, converge in identifying attitude components such as emotional states that influence team performance [13]-[18].

[*]Corresponding Author: Nasrin Dehbozorgi, Dnasrin@kennesaw.edu

Here the question is how to operationalize attitude constructs as factors to be measured. Research shows that a survey is commonly used for this purpose [13]. Attitude has been traditionally measured by having students fill out self-report forms, typically using a Likert scale to express their feelings towards a specific subject [13]. However, the self-report tools have some drawbacks such as: lack of student commitment to fill them out, not providing precise answers, and lack of awareness about their emotion at the moment. More advanced tools exist to allow researchers to retrieve emotional information by observing facial expressions, gestures, posture, and periods of silence [19]-[22]. Although these tools may provide more precise results on capturing emotion compared to self-reports, still they are not error-free, and deploying such tools requires both technical and human resources which may not be cost-effective. These drawbacks as well as the potential distraction of the learning process can make such tools less practical in educational settings and classrooms. Researchers suggest speech is a very good source to identify emotion, however because of challenges such as environmental noise level it is not commonly practiced [23].

This paper focuses on emotion analysis as a subset of affect and measures the correlation between the emotions that students express in low-stake teams with their performance in the course. We identify students' emotions as they work in dyad teams by recording their verbal conversations and analyze how emotional states correlate with their performance. Identifying these relationships can help us in developing tools to monitor students' emotions and identify at-risk individuals at earlier stages of the semester.

In the next section, we discuss different dimensions of teamwork and methods of evaluating students' performance at both team and individual levels in educational settings and introduce emotion as an attitude construct that is associated with team performance. In section 3, we present our methodology, study design, and data collection protocol followed by the results of data analysis. Finally, we conclude by discussing the takeaways and the plan for future work.

2. Background

Evaluating students' involvement in teams and providing timely feedback are key components to a good team experience. Researchers have proposed different tools and insights on evaluating team performance over the past 30 years. The review of the literature of teamwork assessment shows there are four primary ways to collect data on team performance: 1) self-report, 2) peer assessment, 3) observation, and 4) objective outcomes. For optimal results, it is suggested to combine both qualitative and quantitative data collection [13]. For a comprehensive teamwork evaluation, we need to collect information from different sources [13]. Gathering the necessary information from all perspectives of teamwork requires a group of team observers [24]. In general, it is not easy for instructors to evaluate team dynamics inside and outside the classroom to verify a fair and successful team experience for every individual. For this reason, peer and self-evaluation are common methods of assessing individuals' contributions in teamwork. Surveys are employed for this kind of assessment and these surveys can include Likert scales, partner ranking, descriptive word matching, short answers about peers, and journaling about their effort and experiences [25]. Finally, the weight of the grade is either provided by the instructor in the form of a standard rubric, or the weight of each component is negotiated between the instructor and the students [26]. Teams have outcomes at both team and individual levels, therefore measuring teamwork at the individual level is important. The team-level outcome is the final product which is the result of the effort of all team members, while at the individual level the outcome can be the team member's attitude or contribution to the teamwork process [13].

In low-stake teams, no significant team-level final product is produced to be evaluated as a team performance metric. In these types of teams, students do not have assigned roles, and since teamwork activity has a low contribution to final grades, there is a high chance that students rely on their partners and don't participate in team activity as expected. This situation has severe consequences in collaborative active learning since students will not utilize the class time to learn course material through team activities. In such cases, the emphasis of the team evaluation should be more at the student level and their attitude towards teamwork, in order to provide timely feedback to them.

Lack of quantitative and objective measures of teamwork at the individual level is a barrier for evaluating teams' performance and assembling effective teams [27]. Most metrics to evaluate individuals in teams rely on experts who observe and rate teams by using rubrics or based on qualitative dimensions like leadership and team structure [27]. Here the question is what data needs to be collected and what factors should be used for evaluating individuals in low-stake teams. Recent research shows that attitudinal factors in teamwork can determine team effectiveness [28]. In this research, the first step is to identify the attitude that individuals pose in teams [29]. In the following section, we discuss how the affect (combination of emotion and mood) as attitudinal dimension is associated with performance in the educational setting and discuss the methods for measuring emotional states.

2.1. Affect and Emotion

Affect as an attitudinal attribute plays an important role in students' learning in educational settings [30]. The affective domain which includes both emotions and moods shows how much a person values the learning process, their willingness to contribute to learning new things, their ability to make a decision, and how generally one behaves in different situations [31], [32]. Emotional obstacles can negatively impact students' learning process while positive feelings of joy, happiness, and satisfaction positively influence students' learning [33]. Research shows that students who experience emotional difficulties that are not identified early may not receive appropriate feedback and can result in lower performance [34]. Furthermore, affective states not only impact the performance but also, influence the interpersonal relationships in teams in educational settings [32]. Therefore, considering affective states and emotional awareness is important for both students and instructors [35]. For students, emotional awareness empowers them with the required skills to manage their emotions and establish positive relationships with peers and handle challenging situations effectively [32]. This is critical in collaborative active learning in which the learning and cognitive

process are integrated into teamwork [35]. Identifying students' emotional state is important for instructors since it can lead to both cognitive and affective scaffolding, which can improve social knowledge construction [35].

Psychologists claim a person can be known based on three domains called the ABCs of attitude which is: Affect, Behavior, and Cognition [36]. Cognition can be measured based on the learning outcome and behavior can be observed, but affect is not directly observable and is more challenging to measure [31]. There is a wide range of methods for measuring affective states. These methods vary from psychological measures such as heart rate, diagnostic tasks, self-reports, facial expressions, and knowledge-based methods to derive affective state in a given context like time of day and length of the task and individual journals [23]. The most common form of measuring affect is using self-report surveys [13] which has certain drawbacks as discussed in the previous section. More advanced tools such as EEG technology have been applied to capture emotional states through brain signals or by observing facial expressions, gestures, posture, and periods of silence [37]-[27]. Although these tools provide better results on capturing emotion compared to individual surveys and self-reports, still they are not error-free, and deploying these tools and using experts as observers is demanding in the educational domain for a large number of students.

To date, most of the existing tools to identify learners' emotional states use self-report instruments such as the Achievement Emotions Questionnaire (AEQ) [35], [40]. Fewer researchers recognize emotional states by doing sentiment analysis on students' journals and learning diaries or from their conversations in forums or other asynchronous textual data [41]. These methods are applied to either identify the polarity of students' emotions or the expressions of six fundamental emotions of anger, trust, surprise, sadness, joy, fear, disgust, and anticipation [41]. Although sentiment analysis is more promising to evaluate emotion, it has not been widely applied in the educational domain compared to the social media or e-commerce review corpora due to data limitations [33], [41].

Researchers claim speech is one of the best tools to extract information about emotion, however because of challenges in data collection it is not widely used in the educational domain [23]. The selection of a suitable method for emotion analysis depends on different factors such as the type of emotions to be identified. the resources required to collect data, and the context in which the task is performed [23]. For example, in some contexts using self-reports may be more appropriate than using sensors, since sensors can cause interference with the task being performed by the students [23].

One of the innovations of this research is assessing students' emotions based on their verbal speech in teams. The advantage of this approach is that emotion is captured directly, it minimizes distraction during the learning activity, and data collection can be done at a large scale. In the next section, we introduce our methodology to operationalize emotional states and elaborate on how we design the study.

3. Methodology

In this research, we identify correlations between emotion and students' performance in low-stake teams. We further conduct aspect-based sentiment analysis to identify the topics in which the students expressed the most positive or negative emotion. The result of this analysis can help us identify the areas in which students have more challenges and apply pedagogical interventions. We formulate the following null hypothesis:

H_0: There is no correlation between a student's positive sentiment in low-stake teams and their individual performance in the course.

In our study, speech is the modality of data collection for measuring and analyzing students' emotions. We recorded students' speech as they talked in teams during the class activity throughout the semester. The recorded audio files were then transcribed for text mining and emotion analysis. The data processing phase includes the following steps: 1) noise reduction by applying filters on the audio to improve the quality, 2) audio transcription, 3) assigning IDs to each individual based on the voice recognition, and 4) removing the speech utterances from the third person (i.e., speakers other than the team members). As a result, the transcription of all speech utterances of every participant was stored as one data point. This resulted in 28 datasets to feed into the text mining and sentiment analysis algorithms. This process is illustrated in Figure 1.

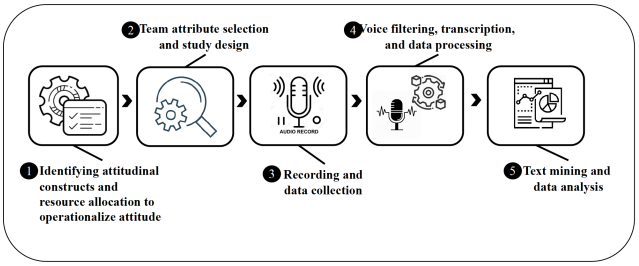


Figure 1: Data Collection and Analysis Process

We used different methods to analyze students' sentiments in speech and identified their correlation with performance. First, was measuring the polarity, frequency, and intensity of the sentiments in three classes of positive, negative, and neutral. Next, different classes of emotion were extracted.

Finally, we performed thematic analysis of the sentiments (aspect-based sentiment analysis) to explore the context and themes in which students expressed more positive sentiment as they spoke in the class. In the following, we explain how each method for sentiment analysis was conducted in this study.

3.1. Polarity Sentiment Analysis

The first step for polarity sentiment analysis was the segmentation of the datasets so that each vector is the speech initiation point, meaning each vector consists of the speech of a participant until it is finished or interrupted by the other teammate. As a result, the number of vectors in each dataset denotes the number of times a participant initiated the speech either in an active or reactive mode. The visualization of the text-mining algorithm for polarity and intensity sentiment analysis is presented in Figure 2.

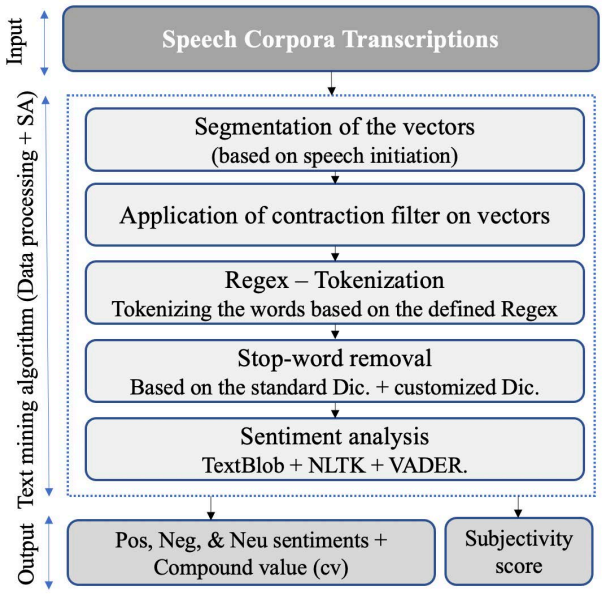


Figure 2: Algorithm for Polarity and Intensity Sentiment Analysis

After applying the contraction filtering on the vectors, we developed a regular expression for tokenizing the vectors and eliminating extra characters which did not impact the sentiment score. We used both the standard English dictionary and also created a custom dictionary to remove the stop words and as well as the redundant words that speakers used habitually and didn't impact the sentiment score. For developing the customized dictionary, we measured the frequency of unigrams in each dataset, and the most frequent common unigrams across all datasets were included in the dictionary to be removed from the corpora. Examples of such words are 'yeah', 'ok' that students uttered repeatedly.

Finally, we analyzed the sentiments from the processed data by applying lexicon-based and rule-based NLP python tools; TextBlob, NLTK, and VADER. The VADER algorithm was applied due to its higher precision and accuracy on short tokens at string-level compared to other sentiment analysis tools [42]-[44]. Moreover, many sentiment analysis tools are either polarity-based or valence-based but VADER measures both the polarity and valence of the input. VADER outputs four classes of 'Negative', 'Neutral', 'Positive', and 'Compound' with values between -1 to 1. The compound score is the adjusted normalized value of the sum of valence scores of each word in the lexicon. The equation of compound value is presented in Equation (1) [45]:

$$Compound\ value = \frac{sum_val}{\sqrt{(sum_val)^2 + 15}} \quad (1)$$

where sum_val is the sum of the sentiment arguments passed to the score_valence() function in the VADER algorithm.

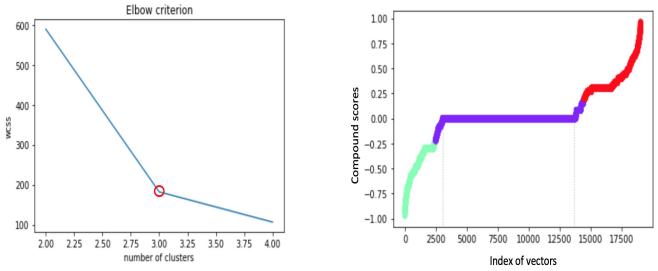
We used compound value (cv) as a metric for the unidimensional measure of sentiment. The threshold for classifying positive, negative, and neutral sentiments can vary based on the context. The typical thresholds are [12][46]:

Positive sentiment: cv >= 0.05

Neutral sentiment: -0.05 < cv < 0.05

Negative sentiment: cv <= - 0.05

In this study, we conducted the k-means clustering on the compound values to determine the right threshold for the data. The Elbow method was used to identify the optimum number of clusters. The Elbow method explains the percentage of the variance as a function of the number of clusters meaning the optimum number of clusters is defined such that adding one more cluster does not provide better modeling of the data [47]. The best number of clusters is chosen based on the 'Elbow criterion', which is shown as an angle in the graph (Figure 3.a) [47]. The Elbow criterion determines three as the optimum number of clusters. The 3-means clustering of data shows the most density of compound values in the cluster with cv= 0 as shown in Figure 3.b.



3.a. Elbow Criterion 3.b. 3-Means Clustering

Figure 3: Elbow Criterion and 3-Means Clustering of cv

As a result, we considered zero as the threshold for classifying cv into positive, neutral, and negative sentiments. This means any vector with cv < 0 was labeled as negative, vectors with cv > 0 were considered as positive, and vectors with cv = 0 were treated as neutral sentiment classes.

Next, we identified multiple classes of sentiments from speech corpora. The methodology for doing so is described in the following section.

3.2. Multi-class Sentiment Analysis

In addition to analyzing the polarity of the sentiments, we analyzed how different types of emotions contributed to students' performance. For extracting multiple classes of emotions, we applied the LIWC (Linguistic Inquiry and Word Count) which is an efficient text analysis tool for analyzing diverse emotional and cognitive components in verbal or written speech corpora [48]. The core of the LIWC is the LIWC2015 dictionary which includes almost 6400 word-stems and selected emotions. In this tool, a single word such as "cry" that carries emotion can belong to different categories such as sadness, negative emotion, and overall affect [48]. Each category of LIWC2015 includes a list of dictionary words that define the relevant scale. The complete list of scales and scale words can be found in the LIWC2015 Development manual [48].

In this study, we selected specific dictionary categories to identify which emotions more contribute to individuals' performance. The total number of 93 dimensions in all 4 sub-dictionaries exist in the LIWC2015, out of which we identified 63 dimensions to be most relevant to the context of this study. These

dimensions exist in the four sub-dictionaries of 1) psychological processes, 2) Linguistic processes 3) punctuation, 4) other grammar. The proportion of selected dimensions in each sub-dictionary were 67% psychological processes, 24% linguistic processes, 6% other grammar, and 3% punctuations. The scale score of each 63 dimensions is a numerical value measured by using the LIWC framework.

In the next step, we applied Principal Component Analysis (PCA) technique to conduct factor analysis. PCA is an unsupervised statistical method to reduce the dimensionality of large and complex data [49]. It identifies the interrelation between features and reduces the number of features by creating a new feature (component) based on the linear combination of initial features. The main steps of PCA are 1) removal of the target feature and class labeling by creating a d-dimensional matrix, 2) calculating the covariance matrix of the whole dataset, 3) computing eigenvectors (e1, e2, ...ed) and corresponding eigenvalues (λ1, λ2, ..., λd), 4) sorting the eigenvectors by decreasing eigenvalues and choose k eighteen vectors with the larger eigenvalues to form a d*k dimensional matrix (in this study, we considered the threshold of 1 and chose the number of components (k) with eigenvalues > 1), 5) using the d*k eigenvector matrix to transform the samples onto the new subset.

The PCA method is widely applied in developing predictive models. After dimensionality reduction, the dataset is divided to test and train datasets to fit the data into the training algorithm. In this study, the goal is to identify the correlation of different emotions with performance and identify the ones that most contribute to performance. However, the development of the predictive models is not the focus of this study and will be done in future research. The next step of analyzing emotions is aspect-based sentiment analysis, which is described in the following section.

3.3. Aspect-based Sentiment Analysis

The last step of sentiment analysis in this study was the aspect-based analysis and exploring the themes of the positive sentiments. We applied word frequency analysis and measured the frequency of the tokens both unigrams and bi-grams in the positive sentiment corpora. After identifying the most frequent word tokens we identified the proportion of course-related tokens among them.

3.4. Data Collection

The data for this study was collected from an active learning CS1 class in which students worked in low-stake teams to do the class activities. Based on the research [13] and our empirical evidence, in order to improve students' engagement, we decided to form teams of two students. Team formation was done subjectively, and they were mixed of both males and females with diverse backgrounds. We formed the teams at the beginning of the semester by a gamified activity in which students were paired based on their month of birth. Since students were mostly involved in the team formation game, we noticed high levels of satisfaction between teammates throughout the semester.

The class included 65 students among which 48 participated in this study. We collected the recorded speech of 48 participants throughout the whole semester, however, due to some technical issues or absence of students during some sessions, the recording of 28 individuals (14 teams of two students) was considered in this study due to their consistent pattern of attendance in the class.

The class was held for 75 minutes during two days of the week. In each class, the first 15-20 minutes was dedicated to resolving student's misconceptions from the previous session followed by a low-stake clicker quiz and a 15-minute mini-lecture. The remaining class time (about 40 minutes) was dedicated to peer learning class activities in the form of low-stake teamwork.

During data collection, we faced multiple environmental and technical issues or human errors which made the data unavailable for analysis. For example, sometimes students pressed the stop button on the recorder accidentally, or sometimes it was the battery or hardware issue which led to the loss of data. On top of that, environmental noise level and cross talk in class when all team members were talking at the same time was a challenge in data collection. We employed some protocols to overcome these challenges such as training specific TAs for the recording-related tasks or encouraging students to sit in certain locations to minimize the noise level and improving the quality of the recording. To overcome the hardware-related challenges, we identified recording devices that had features like bidirectional paired microphones, built-in noise cancellation, lasting battery, high-level user-friendliness, and low cost. In the next section, the result of the data analysis is presented, and the null hypothesis is evaluated.

4. Data Analysis

The main goal of data analysis is to identify the correlation between positive emotions and students' performance and identifying emotions that can serve as predictive metrics.

The performance metric is considered to be students' grades in the course. Students' performance was assessed in a formative style during the semester by three major assignments, three lecture tests, and three lab tests. Each milestone had a certain contribution to the final grade: assignment 20%, lecture test 30%, and lab test 30%. Participants' grade distribution is presented in Figure 4.

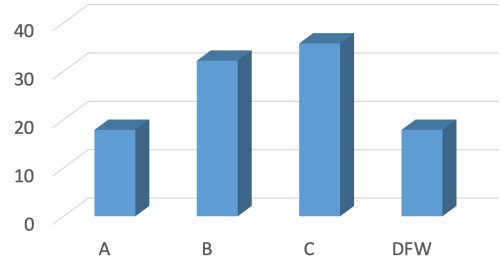


Figure 4: Grade Distribution of the Participants

The linear regression analysis of the performance scores over time determines the slope of trendlines (i.e., the ratio of change over time). Accordingly, the participants were categorized into three groups; upward trend (slope >0): 57%, downward trend (slope <0): 28%, no trend (slope = 0): 14%.

The data shows that most of the participants had an upward trend which means their performance increased during the semester. In the rest of this section, we present the analysis result.

4.1. Polarity and Intensity of Sentiments

To analyze the polarity of the sentiments NLTK and VADER algorithms were applied on the 28 datapoints to classify each vector into four classes of Positive (Pos), Neutral (Neu). Negative (Neg) and Compound (Comp). As described in the methodology section, the vectors with cv > 0 are considered to be positive sentiments. The intensity and frequency of positive sentiments were normalized based on the amount of speech. Frequency refers to the number of vectors with cv > 0 and intensity refers to the actual score of cv in each vector. Equations (2) and (3) show how the mean frequency and intensity scores are measured.

$$Frequency = \frac{m}{n} \quad (2)$$

$$Intensity = \frac{\sum_{x=1}^{m}(comp_value_x)}{n} \quad (3)$$

where:

n= total number of vectors in each dataset

Vectors with positive comp_value= {1,2, …m}

We identified a homogeneous pattern between frequency and intensity of positive sentiments (cv > 0), based on the linear regression analysis. This means the students with higher frequency of positive sentiments had higher levels of intensity in their positive emotions. The linear regression plot of the frequency and intensity of positive sentiments is presented in Figure 5.

The Spearman's rank correlation coefficient was applied to identify if there is a positive relationship between students' positive sentiments (intensity and frequency) and their performance. Spearman's rank correlation coefficient is a nonparametric (distribution-free) rank statistic for measuring the strength of an association between two variables [50]. It evaluates how the relationship between two variables can be described using a monotonic function, without making any assumptions about the frequency distribution of the variables [50]. The coefficient value (r_s) can be anywhere between -1 and 1, and the closer r_s is to +1 and -1, the two variables have a stronger monotonic relationship. The r_s value is calculated using equation (4) and the strength of the correlation is interpreted based on the absolute value of r_s; r_s = 00-.19 "very weak", r_s =.20-.39 "weak", r_s = .40-.59 "moderate", r_s = .60-.79 "strong", r_s = .80-1.0 "very strong" [51].

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \quad (4)$$

where d_i is the difference in ranks for variables and n is the number of cases. The calculated r_s values are presented in Table 1.

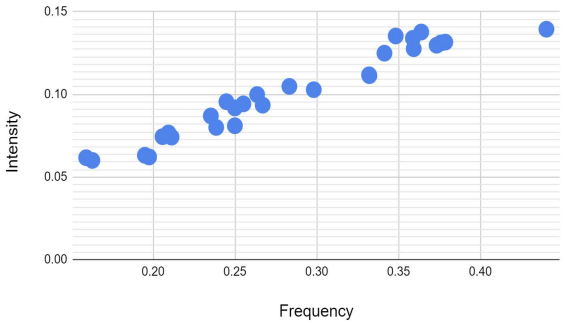


Figure 5: Linear Regression of Frequency and Intensity of Positive Sentiments

Table 1: Coefficient Values of Positive Sentiments and Performance

	Coefficient Value (r_s)	
Frequency and performance	.61	Strong positive correlation
Intensity and performance	.64	Strong positive correlation

The coefficient values of both intensity and frequency of positive sentiments show that both have a strong positive correlation with performance.

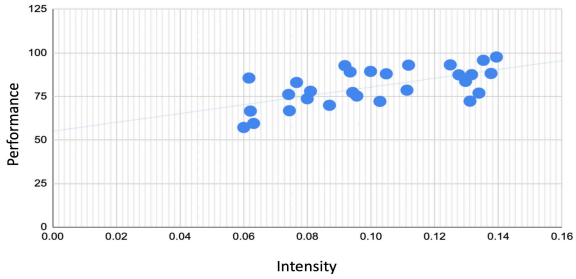


Figure 6: Regression plot of the Intensity of Positive Sentiment vs Performance

The linear regression plot of intensity and frequency of positive sentiments vs performance are presented in Figures 6 and 7. In these plots, the horizontal axis shows the intensity and frequency of positive sentiments, and the vertical axis shows the performance score.

To test the null hypothesis (H_0), we applied the chi-square test and measured the two-tailed p-value with the confidence level of 0.05. The Null hypothesis states: There is no correlation between students' positive sentiment in low-stake teams and their individual performance in the course.

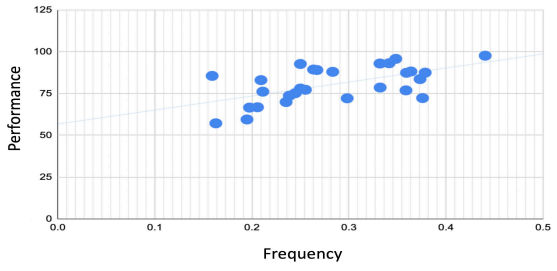


Figure 7: The Regression Plot of Frequency of Positive Sentiment vs Performance

The calculated p-value for the intensity of positive compound value is 0.002 and the p-value for the frequency of compound value is 0.001. The p-values for both the intensity and frequency of positive sentiments are less than the confidence level (0.05), therefore the null hypothesis is rejected which confirms that there is a statistically significant correlation between intensity and frequency of positive sentiment and students' performance. The calculated two-tailed p-values are presented in Table 2.

Table 2: Two-tailed p-values of Positive Sentiment and Performance

Tow-tailed p-value		
Frequency and performance	.001	Statistically significant correlation
Intensity and performance	.002	Statistically significant correlation

The next step in our analysis is classifying students' sentiments into multiple classes beyond positive, negative, and neutral to investigate which classes can be used as predictive metrics for performance as the target value. The result of multiclass sentiment analysis is presented in the following section.

4.2. Multi-class sentiment analysis

The LIWC text analysis tool was applied to extract sentiment class from the corpora. The selected 63 dimensions of LIWC are in psychological, affective, and cognitive domains. The output features of each domain scale differently, in other words, the range of features in one domain might be greater than the features in another one. For example, the word count (WC) features are greater than the affective features such as negative emotion (negemo). To standardize the distribution of the data we applied the StandardScaler() algorithm from the SKLEARN library. This algorithm standardizes the features by removing the mean and scaling to the unit variance as shown in equation (5) [52]:

$$z = (x-u)/s \quad (5)$$

where u is the mean and s is the standard deviation of the training sample.

The output features of LIWC have very high dimensionality, which makes it challenging to interpret the features and identify the main ones for developing predictive models. In order to reduce the dimension of the feature space to the most critical ones and preserving as much 'variability' (i.e., statistical information) as possible, we used the Principal Component Analysis (PCA) method. To determine the minimum number of principal components, we calculated the cumulative proportion of variances that the components explain. The acceptance level of cumulative variance can vary from 80% to 95% depending on the application. We identified 14 principal components with 92% variance which is an acceptable number to represent data. Figure 8 shows the Scree plot of the LIWC features.

The heatmap visualization of the 14 principal components and the 63 LIWC features is presented in Figure 9. The horizontal axis shows the total of 63 original features, the vertical axis denotes the identified 14 principal components and the color bar on the right side represents the correlation between the original features and the principal components. The lighter color indicates a more positive correlation exists between the features and the principal component and the darker colors show a negative correlation.

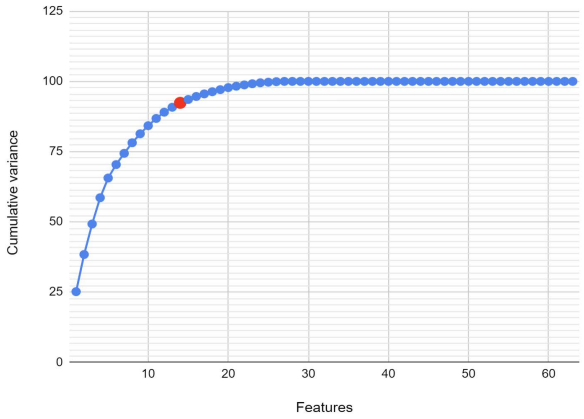


Figure 8: Scree Plot of LIWC Dimensions

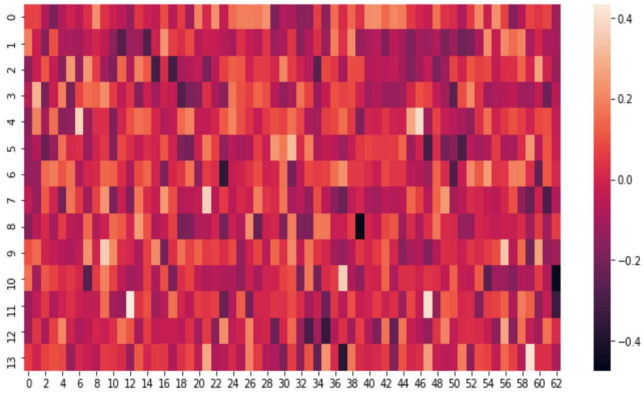


Figure 9: The Heatmap Visualization of Principal Components vs. Features

Table 3: Mapping of Principal Components (PCs) to the Original LIWC Features

PCs	LIWC Feature	PCS	LIWC Feature
PC_1	Anxiety	**PC_8**	Negative emotion
PC_2	Negations (e.g., no, not, never)	**PC_9**	Perceptual process (hear)
PC_3	Common verbs	**PC_10**	Personal pronounce (she/he)
PC_4	Analytical thinking	**PC_11**	Exclamation marks
PC_5	Drives (including affiliation, achievement, power, reward focus, risk focus)	**PC_12**	Auxiliary verbs
PC_6	Insight (e.g., think, know, consider)	**PC_13**	Certainty
PC_7	Anger	**PC_14**	Non-fluencies (e.g., Er, hm, umm)

To interpret each principal component in terms of the original variables, we measured the magnitude of the coefficients for the original variables such that larger absolute values of coefficients confirm corresponding variables have more importance in calculating the component. The maximum absolute values of the coefficients are mapped to the original LIWC features in Table 3.

The table 3 lists the most principal features that determine the target value which is students' performance.

The identified features can be applied to train the predictive algorithms to predict student's performance based on their psychological, affective, and cognitive behaviors. Fitting the sample data into predictive models is not the scope of this study and will be considered in future work. In the next section, the result of thematic analysis (aspect analysis) on students' positive sentiments is presented.

4.3. Aspect-based sentiment analysis

After identifying the correlation between the positive sentiments and the performance we explore the theme of the positive sentiments to identify whether students expressed a more positive emotion toward the course content or any other topic or experience. In order to identify the themes, we applied word frequencies algorithms to parse the positive sentiment vectors and extract unigrams and bigrams tokens. The frequency of unigrams and bigrams determines the dominant theme in students' positive sentiments.

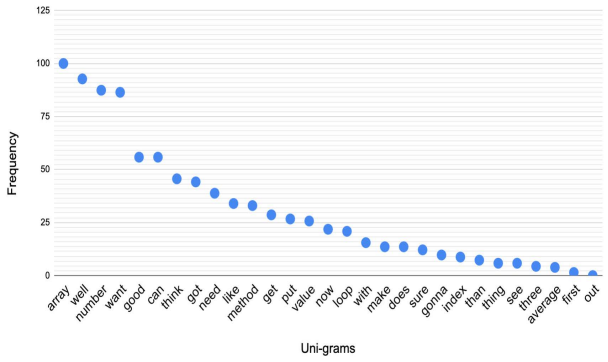


Figure 10: Frequency of Unigrams in Positive Sentiment Vectors (Comp>0)

Figures 10 and 11 plot the normalized frequency of the top 30 unigrams and bigrams of vectors with cv > 0. Based on these plots the dominant theme of word tokens that most frequently were uttered in students' speech is specifically about the course content (i.e., Array, random number).

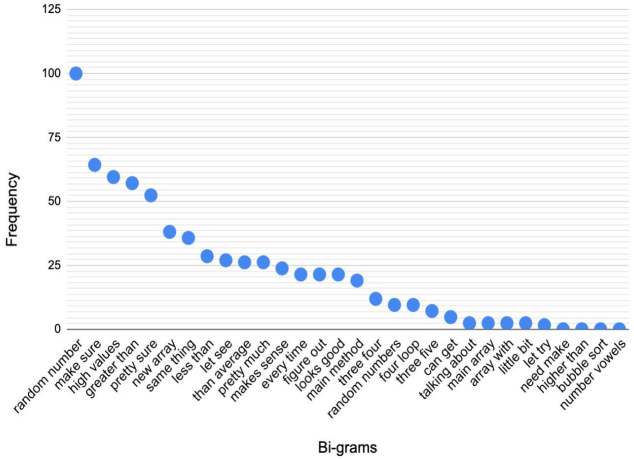


Figure 11: Frequency of Bi-Grams in Positive Sentiment Vectors (Comp>0)

The proportion of course-specific unigrams and bi-grams tokens in the vectors with cv>0 is presented in Figures 12 and 13.

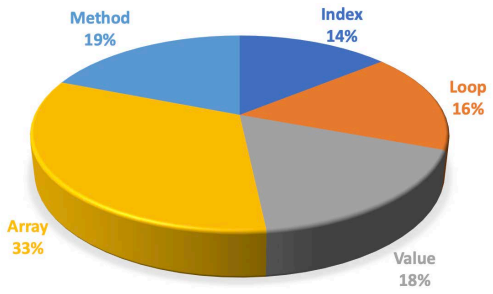


Figure 12: Course-Specific Unigrams in the Vectors with cv>0

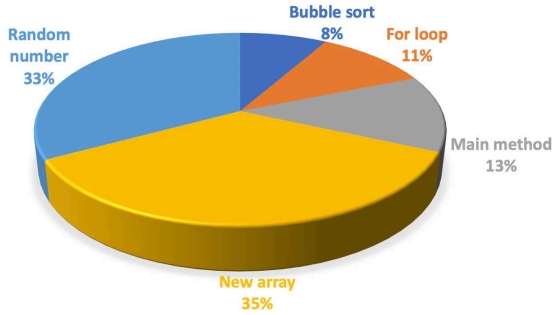


Figure 13: Course-specific Bi-Grams in the Vectors with cv>0

Based on the frequency analysis we conclude that students' positive sentiments were mainly centered on course-related topics. This finding also confirms that the course design was effective in engaging the students in the teamwork activity since they were timed activities that allowed the well-prepared students to finish them within 35-40 minutes and they did not have extra time to discuss the topics outside the course content.

5. Discussion

The k-means clustering of the output sentiments showed that most vectors fall into the neutral class which is the cluster where cv=0 (figure 3.b). We further analyzed students' speech to identify the correlation between the neutral sentiments (cv = 0) as well as the subjectivity level in their speech and their performance. We further conducted a thematic analysis on the negative sentiments' corpora (vectors with (cv<0) to identify the themes in which students expressed negative emotions.

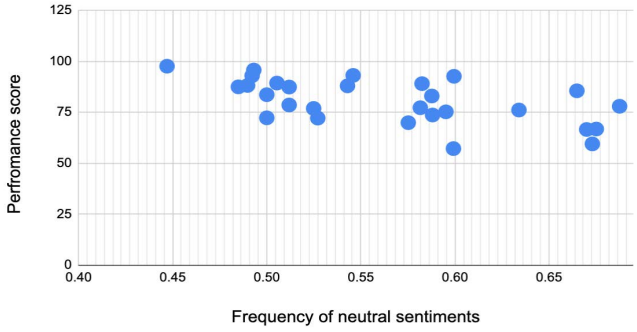


Figure 14: Frequency of Neural Sentiments vs Performance Score

To identify the correlation between neutral sentiments and performance we applied the VADER tool to extract neutral sentiments (cv =0). The regression plot of the frequency of neutral sentiments vs performance is presented in Figure 14. In this figure, the horizontal axis is the frequency of the neutral

sentiments, and the vertical axis is the performance. We observe a homogeneous pattern in which by decreasing the performance the frequency of neutral sentiments increases.

The Spearman's rank correlation coefficient was conducted and the calculated r_s value is -0.61. It confirms there is a strong negative correlation between these two variables of neutral sentiments and performance.

Next, we analyzed the speech to identify the correlation between the level of subjectivity and performance. We applied TextBlob which is a rule-based sentiment analysis tool to extract the subjectivity metric [53]. The output subjectivity score is a float number between 0.0 and 1.0 where 0.0 is very objective and 1.0 is very subjective [54]. We applied Equation (6) to calculate the subjectivity score of each participant.

$$\text{Subjectivity Score} = \sum_{x=1}^{n} Sx \qquad (6)$$

where, Sx is the subjectivity score of each vector (0<=Sx<=1), and n is the total number of vectors in each dataset.

The subjectivity level vs. performance score is plotted in Figure 15, where the vertical axis is the performance score, and the horizontal axis is the subjectivity score.

Based on Spearman's rank correlation coefficient the calculated r_s value is -.02, which shows a very weak negative correlation exists between the level of subjectivity and students' performance.

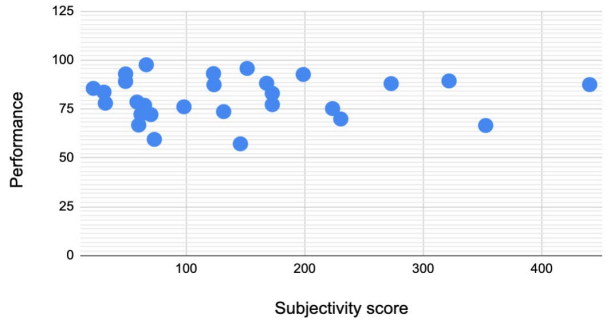


Figure 15: The Regression Plot of Subjectivity vs Performance Score

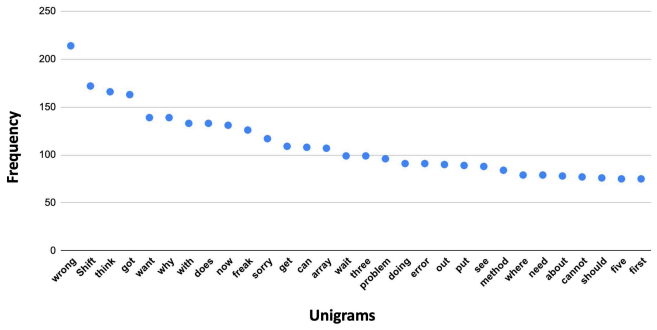


Figure 16: Frequency of Negative Sentiment Unigrams

Next, we conducted a thematic analysis on negative sentiment vectors (cv < 0) by calculating the frequency of unigrams. The result shows that students expressed more negative sentiments when they made mistakes and faced a problem in solving the problems in the class activities. For the purpose of this study, we have replaced the swear words with "shift" and "freak".

The scatter plot of the frequency of the unigrams in the negative sentiment vectors (cv<0) (Figure 16) shows the most frequent token in negative sentiments is the word "wrong".

And finally, it is worth mentioning that we did not find a statistically significant correlation between vectors with cv <0 and the performance score in this study.

6. Conclusion

Teamwork and collaboration are critical aspects of cognition and learning. In low-stake teams, students don't have assigned roles and tasks since they are more focused on formative learning tasks than final artifact production. This might lead to a lack of contribution from some team members which as a result may lead to poor peer learning and performance. Assessing student's contribution and performance in low stake teams is thus challenging for instructors. We posit that attitude constructs and individuals' emotions can be assessed as an indicator of performance.

In the existing literature, emotion is mainly operationalized in the form of surveys and self-reports [13] which have their own challenges and may not necessarily provide reliable results. A major drawback of surveys is the lack of commitment from participants to provide responses.

Text analysis methods are another way to operationalize sentiments and have been applied to students' textual conversations on discussion forums or blogs [41]. The drawback of these methods is that they are conducted on asynchronous conversations and don't provide sentiment analysis in a real-time context. Speech is an alternative to capture emotional states. Recording and speech analysis have their own challenges, such a being demanding in terms of time, effort, and resources. Due to the environmental noise level, this method of data collection is not often used in educational settings.

In this article, we proposed a novel approach to operationalize students' sentiments by recording their real-time speech in low-stake teams during class. Finding the correlation between different emotions and students' performance can help in developing models to predict students' performance at earlier stages of the semester to provide timely feedback to them and apply pedagogical interventions. The main goal of this research is to identify the correlation between students' positive emotions as they work in teams during class with their performance.

The novelty of our research is sentiment analysis based on speech. This method has minimum distraction for students during the class activity and does not require additional time to answer the surveys. We implemented our method in a CS1 class during one semester with 28 participants. The result of data analysis showed the students who had higher scores in frequency and intensity of positive sentiments performed better in the course. On the other hand, the students who had more frequency in their neutral sentiments had lower performance scores in the course. We did not identify a statistically significant correlation between negative sentiments and students' performance. The result of

thematic analysis on the positive sentiments confirms that the theme of the students' positive emotions were the topics related to the course content.

This study was conducted on 28 samples thus the results of the data analysis may not generalize to other courses or other populations of students. However, the results provide a basis for future research to analyze larger samples of data and for developing predictive models.

The result of this research can benefit both students and instructors. Instructors can identify at-risk students in earlier stages of the semester and help them develop their social skills and learning experience.

6.1. Future work

In the next phase of this research, we aim to collect more samples to study the correlation of positive emotion with performance. In this research, the principal components from multiple features (classes of emotion) were identified that determine the performance. In future work, we will fit these principal features to train machine-learning algorithms to predict performance.

Another future direction is to develop algorithms for automatic and real-time transcriptions of the speech to be able to provide real-time analysis as students work in teams. The development of such a system that provides emotional feedback to both instructors and students as they work in teams helps students to develop emotional awareness and adjust their behavior and interaction with peers. This system can also enable the instructor to observe the climate of different teams and cue them to provide timely feedback to students.

References

[1] N. Dehbozorgi, M. Lou Maher, M. Dorodchi, "Sentiment analysis on conversations in collaborative active learning as an early predictor of performance," in 2020 IEEE Frontiers in Education Conference (FIE), IEEE: 1–9, 2020. DOI: 10.1109/FIE44824.2020.9274119

[2] G. Salomon, D.N. Perkins, "Chapter 1: Individual and social aspects of learning," Review of Research in Education, **23**(1), 1–24, 1998. https://doi.org/10.3102/0091732X023001001

[3] C. Latulipe, N.B. Long, C.E. Seminario, "Structuring flipped classes with lightweight teams and gamification," in Proceedings of the 46th ACM Technical Symposium on Computer Science Education, 392–397, 2015. https://doi.org/10.1145/2676723.2677240

[4] N. Dehbozorgi, S. MacNeil, "Semi-automated Analysis of Reflections as a Continuous Course," in 2019 IEEE Frontiers in Education Conference (FIE), IEEE: 1–5, 2019. DOI: 10.1109/FIE43999.2019.9028636

[5] M. Prince, "Does active learning work? A review of the research," Journal of Engineering Education, **93**(3), 223–231, 2004. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

[6] M.K. Smith, W.B. Wood, W.K. Adams, C. Wieman, J.K. Knight, N. Guild, T.T. Su, "Why peer discussion improves student performance on in-class concept questions," Science, **323**(5910), 122–124, 2009. https://doi.org/10.1126/science.1165919

[7] B.J. Millis, P.G. Cottell Jr, Cooperative Learning for Higher Education Faculty. Series on Higher Education., ERIC, 1997.

[8] P.D. Feden, R.M. Vogel, Methods of teaching: Applying cognitive science to promote student learning, McGraw-Hill Humanities, Social Sciences & World Languages, 2003.

[9] N.H. Narayanan, C. Hundhausen, D. Hendrix, M. Crosby, "Transforming the CS classroom with studio-based learning," in Proceedings of the 43rd ACM technical symposium on Computer Science Education, 165–166, 2012. https://doi.org/10.1145/2157136.2157188

[10] N. Dehbozorgi, "Active learning design patterns for CS education," in ICER 2017 - Proceedings of the 2017 ACM Conference on International Computing Education Research, 2017, doi:10.1145/3105726.3105741.

[11] N. Dehbozorgi, S. MacNeil, M. Lou Maher, M. Dorodchi, "A comparison of lecture-based and active learning design patterns in CS education," in 2018 IEEE Frontiers in Education Conference (FIE), IEEE: 1–8, 2018. https://doi.org/10.1109/FIE.2018.8659339

[12] L. Liu, J. Hao, A.A. von Davier, P. Kyllonen, J.-D. Zapata-Rivera, A tough nut to crack: Measuring collaborative problem solving, IGI Global: 344–359, 2016. https://doi.org/10.4018/978-1-4666-9441-5.ch013

[13] E. Salas, D.L. Reyes, A.L. Woods, The assessment of team performance: Observations and needs, Springer: 21–36, 2017. https://doi.org/10.1007/978-3-319-33261-1_2

[14] D.R. Ilgen, J.R. Hollenbeck, M. Johnson, D. Jundt, "Teams in organizations: From input-process-output models to IMOI models," Annu. Rev. Psychol., 56,517–543,2005.https://doi.org/10.1146/annurev.psych.56.091103.070250

[15] S.W.J. Kozlowski, B.S. Bell, "Work groups and teams in organizations.," Handbook of psychology: Industrial and organizational psychology, London, England: Wiley, **12**, 333–375, 2013. https://doi.org/10.1002/9781118133880.hop212017

[16] S.W.J. Kozlowski, D.R. Ilgen, "Enhancing the effectiveness of work groups and teams," Psychological Science in the Public Interest, **7**(3), 77–124, 2006. https://doi.org/10.1111/j.1529-1006.2006.00030.x

[17] J. Mathieu, M.T. Maynard, T. Rapp, L. Gilson, "Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future," Journal of Management, **34**(3), 410–476, 2008. https://doi.org/10.1177/0149206308316061

[18] L.A. DeChurch, J.R. Mesmer-Magnus, "The cognitive underpinnings of effective teamwork: a meta-analysis.," Journal of Applied Psychology, **95**(1), 32, 2010. https://doi.org/10.1037/a0017328

[19] S. Anders, J. Heinzle, N. Weiskopf, T. Ethofer, J.-D. Haynes, "Flow of affective information between communicating brains," Neuroimage, **54**(1), 439–446, 2011. https://doi.org/10.1016/j.neuroimage.2010.07.004

[20] M.B. Schippers, A. Roebroeck, R. Renken, L. Nanetti, C. Keysers, "Mapping the information flow from one brain to another during gestural communication," Proceedings of the National Academy of Sciences, **107**(20), 9388–9393, 2010. https://doi.org/10.1073/pnas.1001791107

[21] K. Shockley, M.-V. Santana, C.A. Fowler, "Mutual interpersonal postural constraints are involved in cooperative conversation.," Journal of Experimental Psychology: Human Perception and Performance, **29**(2), 326, 2003. https://doi.org/10.1037/0096-1523.29.2.326

[22] R. Stevens, T. Galloway, J. Lamb, R. Steed, C. Lamb, Linking team neurodynamic organizations with observational ratings of team performance, Springer: 315–330, 2017. https://doi.org/10.1007/978-3-319-33261-1_20

[23] E. Hudlicka, "To feel or not to feel: The role of affect in human–computer interaction," International Journal of Human-Computer Studies, **59**(1–2), 1–32, 2003. https://doi.org/10.1016/S1071-5819(03)00047-8

[24] T.L. Dickinson, R.M. McIntyre, A conceptual framework for teamwork measurement, Psychology Press: 31–56, 1997.

[25] D.E. Wilkins, P.B. Lawhead, "Evaluating individuals in team projects," ACM SIGCSE Bulletin, **32**(1), 172–175, 2000. https://doi.org/10.1145/331795.331849

[26] L.K. Michaelsen, M. Sweet, "The essential elements of team-based learning," New Directions for Teaching and Learning, **2008**(116), 7–27, 2008. https://doi.org/10.1002/tl.330

[27] R. Stevens, T. Galloway, J. Lamb, R. Steed, C. Lamb, Linking team neurodynamic organizations with observational ratings of team performance, Springer: 315–330, 2017. https://doi.org/10.1007/978-3-319-33261-1_20

[28] E. Salas, C.S. Burke, J.E. Fowlkes, "Measuring team performance in the wild: Challenges and tips", Psychology Press: 263–290, 2014.

[29] J.E. Driskell, E. Salas, S. Hughes, "Collective orientation and team performance: Development of an individual differences measure," Human Factors, **52**(2), 316–328, 2010. https://doi.org/10.1177/0018720809359522

[30] P.A. Schutz, R. Pekrun, G.D. Phye, "Emotion in education", Elsevier, 2007.

[31] D.B. McLeod, "Research on affect in mathematics education: A reconceptualization," Handbook of Research on Mathematics Teaching and Learning, 1, 575–596, 1992.

[32] A.C. Araújo-Simões, S.M. Guedes-Gondim, "Performance and affects in group problem-solving," Revista de Psicología Del Trabajo y de Las Organizaciones, **32**(1), 47–54, 2016. https://doi.org/10.1016/j.rpto.2015.09.001

[33] M. Munezero, C.S. Montero, M. Mozgovoy, E. Sutinen, "Exploiting sentiment analysis to track emotions in students' learning diaries," in Proceedings of the 13th Koli Calling International Conference on Computing Education Research, 145–152, 2013.

https://doi.org/10.1145/2526968.2526984
[34] P. Maras, P. Kutnick, "Emotional and behavioural difficulties in schools: Consideration of relationships between theory and practice," Social Psychology of Education, **3**(3), 135–153, 1999. https://doi.org/10.1023/A:1009618701899
[35] M. Arguedas, T. Daradoumis, F. Xhafa, "Towards an emotion labeling model to detect emotions in educational discourse," in 2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems, IEEE: 72–78, 2014. https://doi.org/10.1109/CISIS.2014.11
[36] D.T. Mizokawa, N. Hansen-Krening, "The ABCs of attitudes toward reading: Inquiring about the reader's response," Journal of Adolescent & Adult Literacy, **44**(1), 72–79, 2000.
[37] S. Anders, J. Heinzle, N. Weiskopf, T. Ethofer, J.-D. Haynes, "Flow of affective information between communicating brains," Neuroimage, **54**(1), 439–446, 2011. https://doi.org/10.1016/j.neuroimage.2010.07.004
[38] M.B. Schippers, A. Roebroeck, R. Renken, L. Nanetti, C. Keysers, "Mapping the information flow from one brain to another during gestural communication," Proceedings of the National Academy of Sciences, **107**(20), 9388–9393, 2010.
[39] K. Shockley, M.-V. Santana, C.A. Fowler, "Mutual interpersonal postural constraints are involved in cooperative conversation.," Journal of Experimental Psychology: Human Perception and Performance, **29**(2), 326, 2003. https://doi.org/10.1037/0096-1523.29.2.326
[40] R. Pekrun, T. Goetz, A.C. Frenzel, P. Barchfeld, R.P. Perry, "Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ)," Contemporary Educational Psychology, **36**(1), 36–48, 2011. https://doi.org/10.1016/j.cedpsych.2010.10.002
[41] H. Tarmazdi, R. Vivian, C. Szabo, K. Falkner, N. Falkner, "Using learning analytics to visualise computer science teamwork," in Proceedings of the 2015 ACM Conference on Innovation and technology in computer science education, 165–170, 2015. https://doi.org/10.1145/2729094.2742613
[42] T. Ahmed, A. Bosu, A. Iqbal, S. Rahimi, "SentiCR: a customized sentiment analysis tool for code review interactions," in 2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE), IEEE: 106–111, 2017. https://doi.org/10.1109/ASE.2017.8115623
[43] K. Hood, P.K. Kuiper, "Improving student surveys with natural language processing," in 2018 Second IEEE International Conference on Robotic Computing (IRC), IEEE: 383–386, 2018. https://doi.org/10.1109/IRC.2018.00079
[44] C. Hutto, E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in Proceedings of the International AAAI Conference on Web and Social Media, 2014.
[45] A. Amin, I. Hossain, A. Akther, K.M. Alam, "Bengali vader: A sentiment analysis approach using modified vader," in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), IEEE: 1–6, 2019. https://doi.org/10.1109/ECACE.2019.8679144
[46] V. Bonta, N.K.N. Janardhan, "A comprehensive study on lexicon based approaches for sentiment analysis," Asian Journal of Computer Science and Technology, **8**(S2), 1–6, 2019. https://doi.org/10.51983/ajcst-2019.8.S2.2037
[47] P. Bholowalia, A. Kumar, "EBK-means: A clustering technique based on elbow method and k-means in WSN," International Journal of Computer Applications, **105**(9), 2014.
[48] Y.R. Tausczik, J.W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods," Journal of Language and Social Psychology, **29**(1), 24–54, 2010. https://doi.org/10.1177/0261927X09351676
[49] A.S. Hess, J.R. Hess, "Principal component analysis," Transfusion, **58**(7), 1580–1582, 2018. https://doi.org/10.1111/trf.14639
[50] J. Hauke, T. Kossowski, "Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data," 2011. https://doi.org/10.2478/v10117-011-0021-1
[51] M.M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," Malawi Medical Journal, **24**(3), 69–71, 2012.
[52] https://scikit-learn.org/
[53] A. Hasan, S. Moin, A. Karim, S. Shamshirband, "Machine learning-based sentiment analysis for twitter accounts," Mathematical and Computational Applications, **23**(1), 11, 2018. https://doi.org/10.3390/mca23010011
[54] S. Loria, "textblob Documentation," Release 0.15, 2, 269, 2018.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 98-111 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Improving of Heat Spreading in a SiC Propulsion Inverter using Graphene Assembled Films

Sepideh Amirpour*,[1], Torbjörn Thiringer[2], Yasin Sharifi[3], Marco Majid Kabiri Samani[4]

[1]China Euro Vehicle Technology AB, Department of Motion and Energy, Gothenburg, SE-41755, Sweden

[2]Chalmers University of Technology, Department of Electrical Engineering, Gothenburg, SE-41296, Sweden

[3]Fengco Real Time Control AB, Gothenburg, SE-41109, Sweden

[4]INFOTIV AB, Gothenburg, SE-41117, Sweden

ARTICLE INFO

Article history:
Received: 01 August, 2021
Accepted: 14 November, 2021
Online: 17 December, 2021

Keywords:
Silicon Carbide (SiC)
Voltage Source Inverters (VSI)
Thermal Feedback
Graphene Assembled Films
Thermal Simulations
Electrified Vehicles

ABSTRACT

The focus of this work is first to establish the effect of the chip temperature and thermal feedback on the determination of the power loss in a three-phase propulsion inverter, then to demonstrate the possibility of achieving an improved heat spreading through the different layers inside a SiC power module by using graphene assembled films in the packaging of the power module. The power loss analysis has been carried out for two Silicon Carbide (SiC) modules in a vehicle inverter, incorporating the MOSFET's reverse conduction as well as including the impact of blanking time on the inverter on-state losses. This data for calculating the losses is determined at an operating situation below the field weakening speed with a high torque for a permanent magnet synchronous machine (PMSM). The operating point is found to be the worst operating condition point when looking at the power loss point. First, it can be noted that not accounting for the thermal feedback, the power loss is considerably underrated, i.e.,11-15% on the on-state converter. Following, the analysis of utilizing the graphene layer in the SiC module reveals a reduction of 10°C per SiC chips in the junction temperature of the SiC MOSFET is achievable. The reduction is calculated based on an applied power loss per SiC chips in steady-state simulation. Furthermore, up to 15°C decrease in the transient computation over the Worldwide Harmonized Light Vehicles Test Cycle (WLTC) per SiC chip is noticed. Moreover, a reduction up to 50% for the junction to case thermal resistance ($R_{th,JC}$) is observed by adding the graphene layer in the power module.

1. Introduction

Increasing the efficiency of vehicle inverters for electrified powertrains is becoming a significant demand [1]. Improved Wide Band Gap silicon carbide (SiC) MOSFETs are promising replacements of the standard silicon insulated gate bipolar transistors (Si IGBTs) in EV applications [2–4]. This is a result of lower switching losses due to quicker switching transitions, as well as good thermal properties in the SiC MOSFETs, superior to those of Si IGBTs. Moreover, lower the on-state losses can be reduced through the utilization of the MOSFET's ability to also conduct current in the opposite direction [5, 6]. Therefore, improving the efficiency of the propulsion inverter utilizing SiC MOSFETs, can lead to reduced losses in the powertrain, which results in a higher power density, to some extent also through the reduction of cooling circuitry. However, the development of switching devices towards miniaturization, causes severe temperature stresses that significantly threaten the performance and lifetime of the semiconductor chips and dies in high power applications. A possibility to tackle this issue is enhanced thermal conduction in various directions inside the power module with effective cooling of the semiconductor chips. Integrating heat spreading materials with excellent mechanical and thermal properties for the power electronics packaging can efficiently move the heat away from the power devices, resulting in a lower operating temperature of the system [7].

Compared to metal materials, nano-scaled carbon materials, in especial, graphene, have a thermal conductivity up to 5300 W/m.K [8], ten times higher than that of the metals [7, 9] as well as an amazing lightweight property, 2.2 g/cm3 and good stability [7], [10, 11]. Particularly, compared with the single layer graphene,

*Corresponding Author: Sepideh Amirpour, Gothenburg, SE-41755, Sweden, +46721843949, sepideh.amirpour@cevt.se

https://dx.doi.org/10.25046/aj060614

there is the incentive in fabricating graphene assembled films (GFs) as new heat spreading materials [12–15] due to their promising electrical and thermal properties. Consequently, research community as well as manufacturers have put intensive efforts into the research on new high-thermally conducting materials like graphene as an effective heat spreader and these efforts have been reported in literature [16–21]. In [22–24] research on the usage of graphene as a spreader of heat in power modules packaging have been conducted. Further interesting research are in [25–27] and a procedure for studying graphene transistors using standard TCAD tools is brought forward in [28]. However, these applications have their own limited scope regarding the design and do not satisfy the comprehensive aspects of the combined electrical and thermal analysis of the SiC semiconductor power module from the chip level towards the different material layers.

To fill this knowledge gap, a research effort conducted is presented in this paper. The specific novel contribution of this article (which is an addition of the research originally provided at the IEEE Industrial Electronics Society (IECON) 2020 conference [29]) is: How can the temperature distribution through the SiC MOSFET's chips as well as the Direct Bonded Copper (DBC) layers in the power module be efficiently spread, through the utilization of graphene with its excellent thermal properties. The finite element method (FEM) in the COMSOL Multiphysics software is very useful when it comes to determining the thermal aspects and has accordingly been utilized. First, a comprehensive numerical analysis quantifies to what extent incorporating and omitting the thermal feedback will affect the losses in a SiC-based propulsion inverter when including the blanking time and the MOSFET's conduction in its opposite direction [29]. Then the graphene solution is proposed to boost the heat dissipation in the SiC power module.

The analysis is caried out for two SiC half-bridge modules, CAB450M12XM3, a recently introduced 3rd generation [30], and the previous module, CAS300M12BM2 also with 1200V rating [31] from Wolfspeed /Cree. Figure 1 shows the procedure used for the power loss determination.

2. Reverse Conduction and Blanking Time

In Si IGBTs the flow of the current in opposite direction is through an anti-parallel diode, while MOSFETs can conduct this current through their 'conduction channel'. In the inverter, when the current and voltage in one leg does not have the same signs, one of the upper or the lower diodes in the phase leg is conducting. When Vds of the corresponding MOSFET exceeds the diode's threshold voltage, both will conduct in parallel. The diode can either be the intrinsic inbuilt diode or a separate one. This ability affects the distribution of on-state losses in a SiC MOSFET module, with the consequence of lower losses.

In order to prevent a short-circuit of the dc-link, a time duration when the lower as well as upper transistor is off at the same time, is needed in a PWM-controlled inverter, a so-called blanking time. The result is that, the risk of a short circuit of the dc-link can be kept to a minimum. The total on-state loss of the modules is affected by the conduction of the diode as a consequence of the blanking time, increasing the losses slightly.

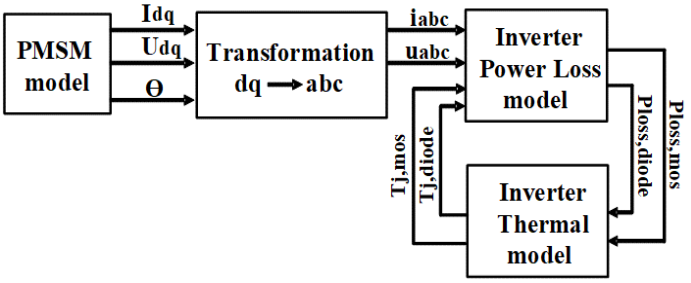


Figure 1: Overview of the system simulation model.

3. Conduction Losses including the MOSFET's Reverse Conduction and Blanking Time

3.1. Conduction Losses in the SiC Power Module

The on-state losses for the investigated inverter are determined numerically using MATLAB. The methodology used in this work is based on the procedure presented in [6] and [32] for calculating the SiC MOSFETs' conduction losses. In this procedure, the MOSFET on-state loss over a machine line period is determined according to

$$P_{cond,MOS} = \frac{1}{2\pi}\int_0^{2\pi} R_{on} I_M^2(\alpha)\tau(\alpha)\, d\alpha \quad (1)$$

here R_{on} is the MOSFET on-state resistance, I_M the MOSFET current, $\alpha = 2\pi f t$ where f is the machine fundamental frequency and τ is the duty cycle found as

$$\tau(\alpha) = \frac{1}{2}(1 + m\sin\alpha) \quad (2)$$

here m is the modulation index [33]. Similarly, the on-state losses of the diode can be formulated as

$$P_{cond,D} = \frac{1}{2\pi}\int_0^{2\pi} (R_d I_D^2(\alpha) + V_d I_D(\alpha))\tau(\alpha)\, d\alpha \quad (3)$$

From the data sheet, V_d, the voltage drop and, R_d, the on-state resistance can be found, I_D is the current through the diode.

The resulting expressions for the current through the MOSFET and the diode thus becomes

$$I_M = \frac{R_d I_p \sin(\alpha-\varphi) - V_d}{R_d + R_{on}} \quad (4)$$

$$I_D = \frac{R_{on} I_p \sin(\alpha-\varphi) + V_d}{R_d + R_{on}} \quad (5)$$

where φ is the angle of fundamental power factor and I_p is the peak value of the phase current [32].

The losses when incorporating and not incorporating the MOSFET reverse conduction are displayed in Figure 2 for the upper diode /MOSFET combination in a phase leg of a CAS300, SiC inverter. A substantial decrement of up to 83% in the diodes' total on-state losses originating from the parallel conduction was found. In Table 1 the operating condition used for calculating the

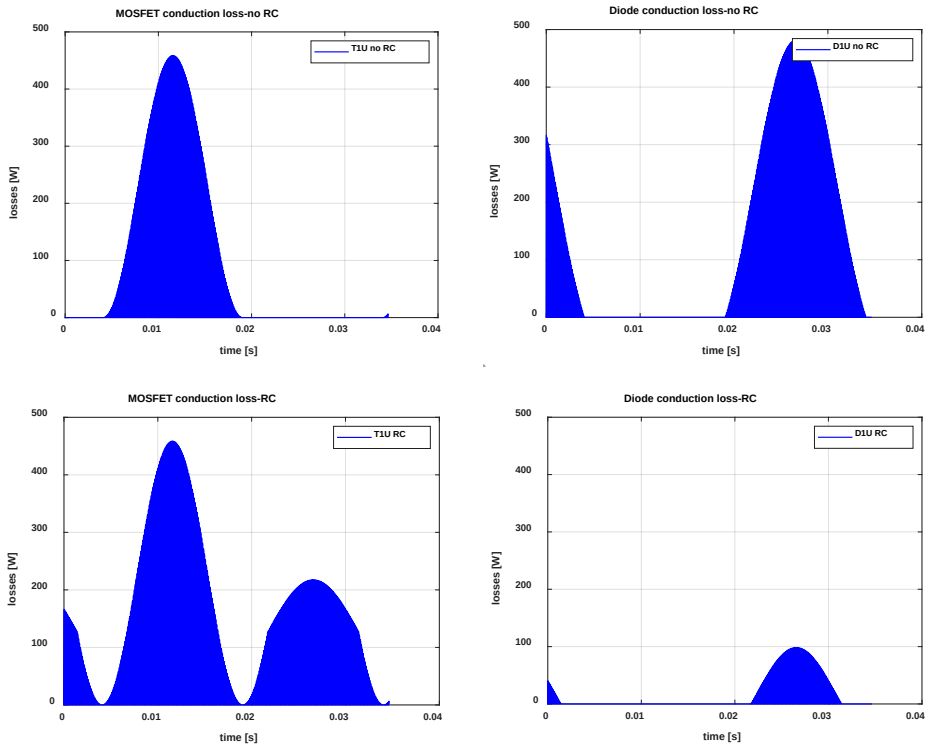


Figure 2: MOSFET and diode conduction losses with (RC), and without MOSFET reverse conduction (no RC), in CAS300 SiC inverter for upper MOSFET and diode in a phase leg.

Table 1: Chosen Operating Point of PMSM for Power Loss Calculations

Variable	Value	Unit
Current magnitude	565	[A]
DC voltage	300	[V]
Blanking time	0.5	[µs]
Switching Frequency	10	[kHz]
Modulation index	0.084	[-]
Torque	160	[Nm]
Mechanical speed	500	[rpm]

losses is given. Likewise, a reduction up to 97% is noticed in the total diodes' conduction losses of the CAB450 SiC inverter.

3.2. Blanking Time

The influence of blanking time is incorporated in the calculation of the on-state losses by formulating a representative duty cycle. The resulting formulation becomes,

$$\tau_{eq}(\alpha) = \tau(\alpha) - t_{blanking} f_{sw} = \frac{1}{2}(1 - 2t_{blanking} f_{sw} + m sin\alpha) \quad (6)$$

where f_{sw} is the switching frequency [32].

The calculated currents of the diodes and MOSFETs in a phase leg of the CAS300 inverter are presented in Figure 3, without (top one) and with (bottom one) utilizing blanking time, also incorporating the MOSFET reverse conduction. As can be seen in the bottom figure, there is current only in the diode for the blanking time intervals. This leads to an increase of 20 W (12%) in the diodes' conduction losses for the CAS300 inverter which is

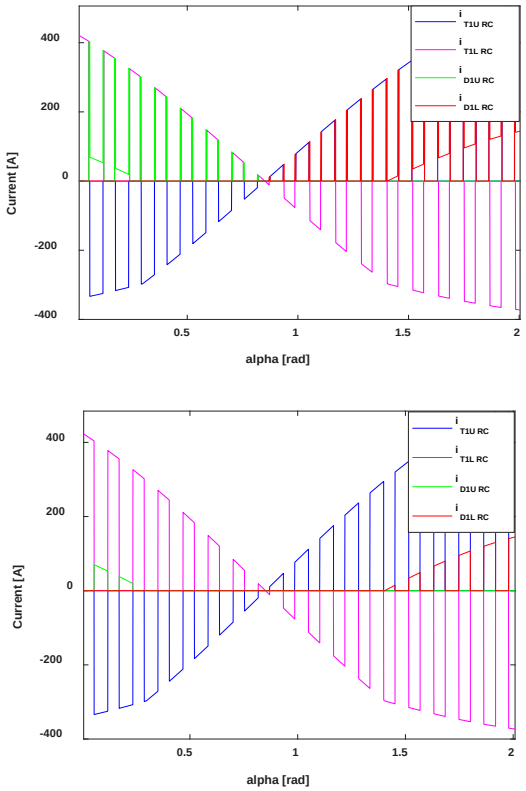


Figure 3: Influence of blanking time on the diode and MOSFET currents as a function of $\alpha = 2\pi f t$ in one phase leg of a CAS300 equipped inverter. Both when incorporating the MOSFET's reverse conduction and not. Upper figure, no blanking time, lower figure with blanking time.

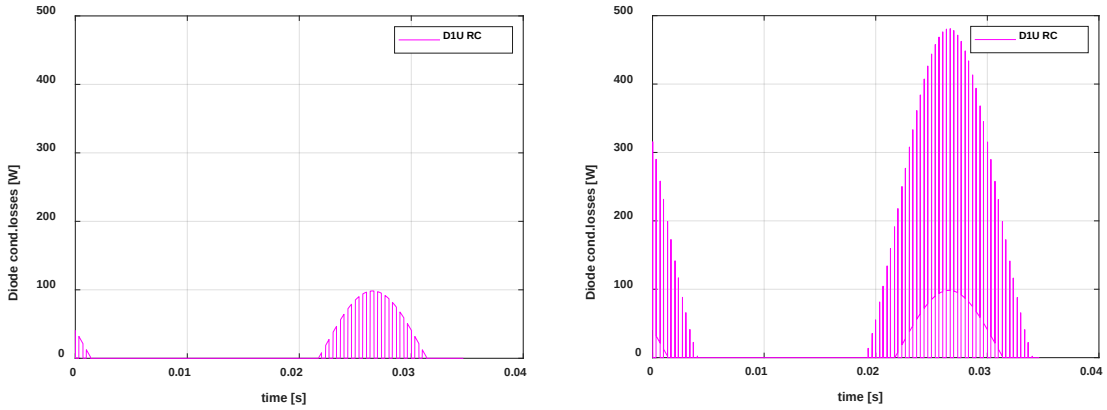


Figure 4: Diode conduction loss, without blanking time (left figure), with blanking time (right figure) in a phase leg of CAS300 SiC inverter.

illustrated in Figure 4 (right one). The impact on the MOSFETs' conduction losses is very low. Likewise, for the CAB450, a rise of up to 70 W, of the diodes' conduction losses is noted. Figure 4 presents the upper diode conduction loss in a phase leg of the CAS300, with and without blanking time, when incorporating the MOSFET reverse conduction. In Table 1 the operating condition used for the determination of the conduction losses are provided.

4. Switching Losses in SiC MOSFET

In the MOSFETs and its anti-parallel/body-diode, a resulting loss is created from each turn off and turn on. The switching losses in a MOSFET and diode can be obtained analytically by the expression as

$$P_{sw.MOSFET,Diode} = f_{sw} \cdot E_{sw(@I_{nom},V_{nom})} \cdot \left(\frac{1}{\pi}\frac{I_p}{I_{nom}}\right)^{k_i} \cdot \left(\frac{V_{dc}}{V_{nom}}\right)^{k_v} \quad (7)$$

In the expression E_{sw} represents the switching energy loss, I_p the peak phase current, I_{nom} and V_{nom} the nominal current and voltage values and k_i , and k_v the current and voltage exponents [34]. In this study, in the numerical implementation, the switching loss is established at each switch-off and switch-on occasion of the module as

$$P_{sw} = \frac{\sum E_{sw}}{t} \quad (8)$$

with t being the simulation time.

5. Electro-Thermal Calculation Network

Due to that several parameters in the modules are temperature-dependent, for instance the on-state resistances, forward voltage drops, switching and reverse recovery energies, the inverters' power losses are determined through the usage of a thermal model, that is given in Figure 5. T_j, are the junction temperatures of the MOSFET and the diodes, while T_c is the case temperature, T_s and T_f are the heatsink and fluid temperatures. In order to have a representative comparison, the same heatsink is utilized for both inverters. Noteworthy is that, for the inverters' design, three SiC half-bridge modules were needed. All set-ups are normalized to a

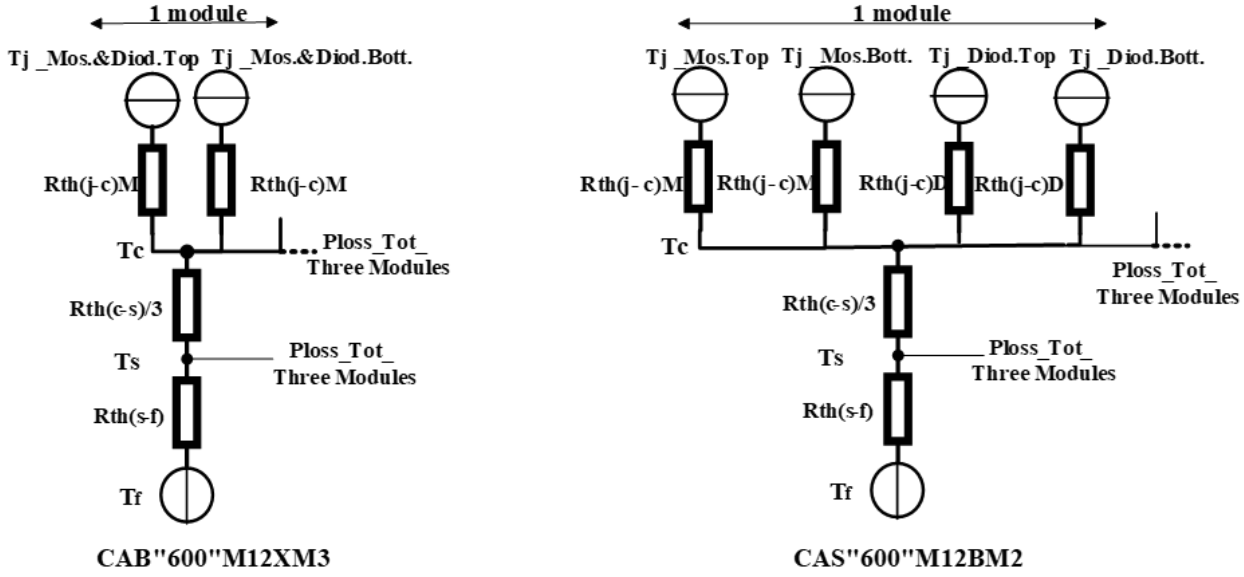


Figure 5: Thermal calculation model to find the power losses and temperatures.

Table 2: PMSM Rating Parameters

Variable	Value	Unit
DC-Link voltage	300	[V]
Rated current	400 RMS	[A]
Pole Pairs	4	[-]
Max. speed	12000	[rpm]
Max. torque	160	[Nm]

rated current of 600 A following the specification of the inverters. The temperature and flow rate of the coolant is 65 ℃ and 10 L/min. Worth mentioning is that the latest 3rd GEN MOSFET dies have been used in the CAB450 module which have an intrinsic body diode, and accordingly the module does not need an antiparallel diode, as can be observed in Figure 5. In comparison, the CAS300 module is equipped with an anti-parallel Schottky diode.

6. Power Losses Analysis in SiC Inverters with and without Accounting for Thermal Feedback

As mentioned before, an operating point of a PMSM at high torque, low speed and high current magnitude is used to analyze the power losses of the two investigated SiC power modules. The chosen operating point can be noticed as the toughest operating condition in the urban driving cycles. MPTA/MTPV field-oriented control is utilized to generate the inverter currents. Based on this control strategy, the machine operates in the constant torque speed range and in the partial and full field weakening speed ranges [35,36]. Table 1 gives the chosen operating point and key machine parameters are given in Table 2. As depicted in Figure 1, an iterative approach is used to find the losses based on the temperature in which, the devices' steady-state and transient parameters (V_T, V_d, R_f, R_d, R_{on}, E_{sw} and E_{rr}) are interpolated using the feedback of the junction temperature. The losses and the temperatures are looped until they have converged. Worth mentioning is that the switching energies of the devices are interpolated as a function of junction temperature and current to fulfill both the current and temperature dependency of the switching energies. The results of the investigations are given in Table 3. It is worth to note that, for a fair comparison, a power loss scaling factor is utilized on the SiC modules in order to form them into 600 A modules.

As shown in Table 3, a significant reduction in diode conduction losses is observed for the two SiC modules. As was explained in sections 2 and 3, this is due to the MOSFET's reverse conduction ability to utilize the diode and the MOSFET to share the current. In the SiC modules, the increase of on-state losses for the built-in diodes as a function of an increase in junction temperature is substantially lower compared to that of the MOSFETs' channel. This is due to a more moderate rise in the diode dynamic resistance with respect to the temperature increase. Moreover, for the two investigated modules, in SiC-MOSFETs, a temperature increase has almost zero effect on the switching losses.

For the comparison of the two SiC modules, a substantial decrease is noticed in the on-state losses in the CAB450's diodes. This is due to the fact that, the CAB450 module is using the latest 3rd GEN MOSFET dies with a very robust intrinsic body diode and reliable operation in the 3rd quadrant. This feature is weaker in the CAS 300 Module.

All in all, the analysis shows that the thermal feedback has a considerable impact on the losses and leads to an increase of up to 11% for CAS300 and 15.6% for CAB450 on the inverters' total conduction losses at the chosen operating condition. This leads to that the demand for effective cooling strategies of the semiconductor power modules is increasing continuously, especially for traction applications which can typically have a high ambient temperature, particularly in SiC-MOSFETs where the temperature variations have larger magnitudes compared to those of Si-IGBTs [37]. Hence, in the following sections, utilizing graphene, a novel thermal conductive, cost-effective, and eco-friendly material is proposed in the power module packaging and

Table 3: Calculated Average Values of Conduction and Switching Losses of the SiC Inverters utilizing and not utilizing Thermal Feedback, Considering Blanking Time and using the MOSFET's Reverse Conduction at 160 Nm, 500 rpm Mechanical Speed and $T_f = 65$℃.

CAB "600"M12XM3 (CAB450)									
	Cond. [W]	Cond. Thermal Feedback [W]	Diff. [W]	Diff.%	Sw. [W]	Sw. Thermal Feedback [W]	Diff. [W]	Diff.%	Tjunc. Thermal Feedback [℃]
MOSFETs	977.8	1140.5	162.7	16.6	206.6	206.8	0.2	0.09	99.8
Diodes	74.4	75.4	1	1.35	3.51	3.52	0.01	0.28	99.8
Tot. Inverter	1052.2	1215.9	163.7	15.6	210.1	210.3	0.22	0.1	
CAS "600"M12BM2 (CAS300)									
	Cond. [W]	Cond. Thermal Feedback [W]	Diff. [W]	Diff.%	Sw. [W]	Sw. Thermal Feedback [W]	Diff. [W]	Diff.%	Tjunc. Thermal Feedback [℃]
MOSFETs	1072	1164	92	8.6	196.3	196.5	0.2	0.1	95.6
Diodes	140.1	182.1	42	29.9	0	0	0	0	88.2
Tot. Inverter	1212.1	1346.1	134	11	196.3	196.5	0.2	0.1	

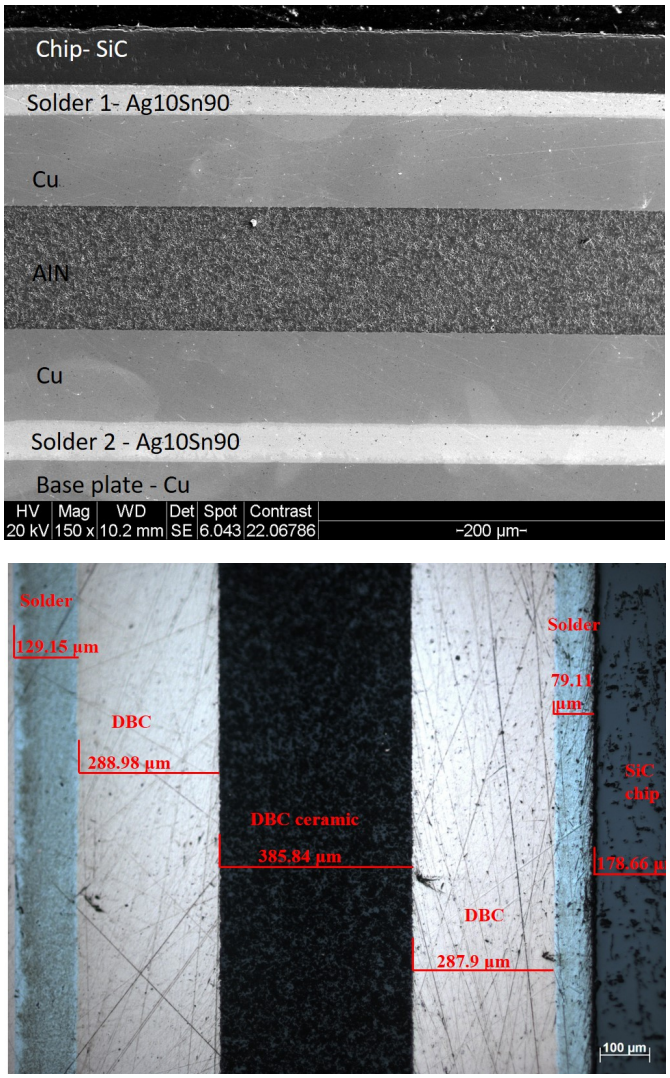


Figure 6: The Layer materials of the reference SiC module CAS300 from SEM and ESD analyses (top) and the materials respective thickness of CAS300 from SEM and ESD analyses (bottom) (given in Table 4).

analyzed by FEM in the COMSOL Multiphysics software intended to improve the heat spreading inside the module.

7. 3D-CAD Model and Thermal Modeling Set-up of the Power Module using SEM and EDS Analyses

As was mentioned before, to enhance the heat spreading and cooling of the semiconductor chips in the chosen half-bridge SiC power module, CAB450M12XM3, a thermal model is built up as close as possible to the real module starting from CAD and then FEM simulation in COMSOL Multiphysics. Even though the thermal packaging aspects of the module was mostly unknown due to confidentiality. However, the key features were known from the reverse engineering effort, which showed that the base plate and insulator have been made of copper and silicon nitride, respectively. Therefore, a Scanning Electron Microscopy (SEM) together with an Energy Dispersive X-Ray Spectroscopy (EDS) investigation were conducted on the available similar voltage class SiC power module, CAS300M12BM2, as a reference to approximate further material information. EDS is a chemical micro-analysis technique, which identifies the x-rays released from the sample through a bombardment process with electrons, so that the element arrangement of the analyzed volume can be found. The above-mentioned analyses have been conducted on a cross-sectional cut off piece of the module. As illustrated in Figure 6 (top), the chips are all silicon-carbide, the base plate is made of copper and an aluminum nitride (AlN) insulator has been used as one of the DBC layers. The other two DBC layers are made of copper. Moreover, the solder $Sn_{90}Ag_{10}$ is used for both solder layers. The thickness of the layers which is shown in Figure 6 (bottom), is given in Table 4.

In the next step, the CAD geometry of the CAB450 SiC power module including the thermal aspects is meshed in COMSOL Multiphysics based on the SEM and EDS analyses data from the reference CAS300 module as well as the CAB450 datasheet information. Instead of AlN, silicon nitride (Si_3N_4) is used as the ceramic material in the DBC substrate based on the CAB450 datasheet. Figure 7 (left) presents a 3D-geometry of the CAB450 half-bridge power module including the chips in which each group of five chips is representing one SiC switch place. The dimensions are implemented according to the datasheet information.

8. Automized FEM Computation in COMSOL for Steady- state/Transient Heat Dissipation

As mentioned before, to be able to capture all the heat dissipation surface interactions, a 3D FEM simulation in COMSOL Multiphysics is performed. Each of the power module materials and their thicknesses were accounted for in the meshing procedure as separate computational domains including the solders, copper, DBC, etc. In order to take the thermal coupling and heat spreading effects into consideration, the contact resistances among the material layers were also modelled using the provided realistic values by material datasets. As the focus of

Table 4: Power Module Layers' Thicknesses and Thermal Properties

Layer	Material	Thickness (µm)	Thermal conductivity, K (W/m.K)	Density, ρ (kg/m^3)	Heat capacity, C (J/kg.K)
Chip	SiC	178	490	3216	690
DBC	Cu	288	400	8940	385
DBC ceramic	AlN/Si_3N_4	386	20	3100	700
Solder (below the chip)/ (below the DBC layer)	$Sn_{90}Ag_{10}$	79/129	50	9000	150
Baseplate	Cu	3	400	8940	385

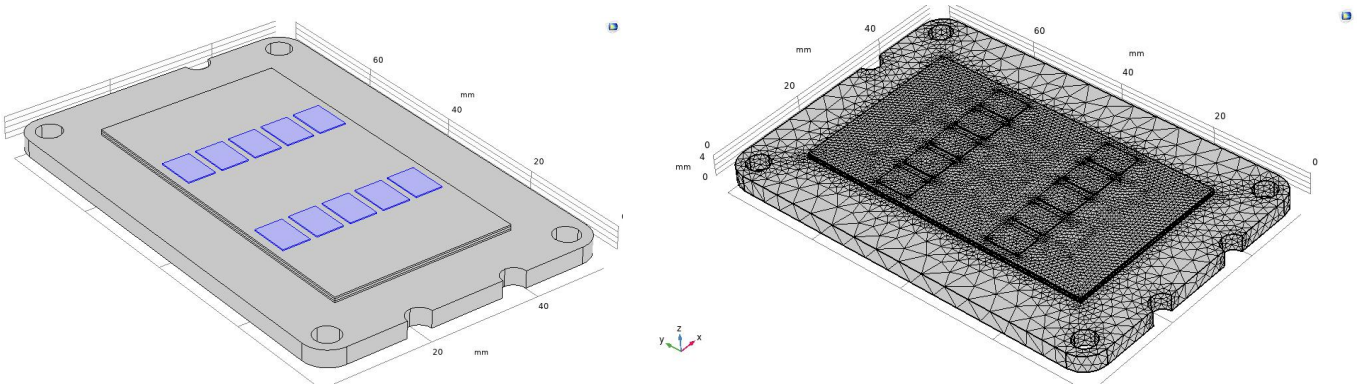


Figure 7: 3D-geometry of CAB450 SiC module built for simulation (left) and the mesh visualization of the module (right).

the analysis is on the thermal distribution through the chips and DBC layers, a manually refined mesh is applied in these domains. Overall, a uniform mesh distribution is built up as depicted in Figure 7 (right). The boundary conditions are caried out by considering the chips as heat sources and a selected power losses of 1262 W, i.e., 126.2 W per chip from the previous power loss calculations of the investigated module is applied to these heat sources.

To mimic a liquid cooling solution, the bottom surface of the baseplate is given a heat transfer coefficient of 3000 W/(m^2.K) to represent the convective heat flux as a boundary condition with a liquid flow of water with a rate of 1 m/s (one meter per second) and a temperature of 65°C. All other boundaries are set to have a heat transfer coefficient of 10 W/(m^2.K), which is the convection of low-speed air flow above a surface [38].

In general, the transportation of heat can be described as the transportation of the thermal energy as a result of a gradient in temperature. Then the heat flux can be expressed by Fourier's law as

$$q = -k\,\nabla T \tag{9}$$

which defines the theory behind the heat conduction. The equation shows that the thermal conductivity, k in (W/m.K), is proportional to the magnitude of the temperature gradient. q is the heat flux measured in (W/m^2) and ∇T is the temperature gradient.

The heat convection, which can either be natural or forced, is added to the boundaries of the system, and is dependent on the geometry and its length. When natural convection is implemented, the system is cooled naturally through air. The steady-state heat flux density is calculated as the following equation

$$q_s = h\,(T_s - T_m) \tag{10}$$

when the convective cooling/heating is involved. h is the heat transfer coefficient in W/(m^2.K), T_s represents the surface temperature and T_m is the media temperature.

Since there are different materials involved in this study, then the heat conduction or heat diffusion must be taken into account by the expression as

$$\rho C \frac{\delta T}{\delta t} = \nabla K(T)\nabla \mathrm{T} + q_v \tag{11}$$

where ρ is the density, C is the heat capacity, $\frac{\delta T}{\delta t}$ represents the difference of temperature over time, and ∇T is the temperature gradient.

Finally, the model is ready for analyzing the time variations of the temperature inside the different layers of the power module and the temperature profiles across it, within the boundary layers. The first results will illustrate the thermal performance of the investigated power module which is built up as close as possible to the real model and in the next phase an assembled graphene film layer is placed under the first copper layer as close as possible to the chips as heat sources to demonstrate the improvement of the heat distribution inside the module from the chip level towards the different layers.

9. Results of Steady-state Thermal Computation of the Power Module

The stationary solver in COMSOL is used to evaluate the steady-state temperature distribution across the power module from the chips towards the bottom layers. The heat distribution is measured by applying a 3D-cutline from the middle chip to the baseplate as depicted in Figure 8. The reason for choosing the middle chip is since this one will experience the highest thermal stress. As it can be seen in Figure 9 (top), the result shows an average maximum temperature of around 99°C for the chips for the applied power loss of 126.2 W per chip which is approximately close to the junction temperature calculated by the electro-thermal model presented in section 6, Table 3. The lowest temperature of 70°C is also noticed around the edges of the module. The temperature profile across all the layers is illustrated in Figure 9 (bottom).

In addition, the thermal resistance of the junction to case (R_{thJC}) of 0.2 K/W has been obtained from the simulation.

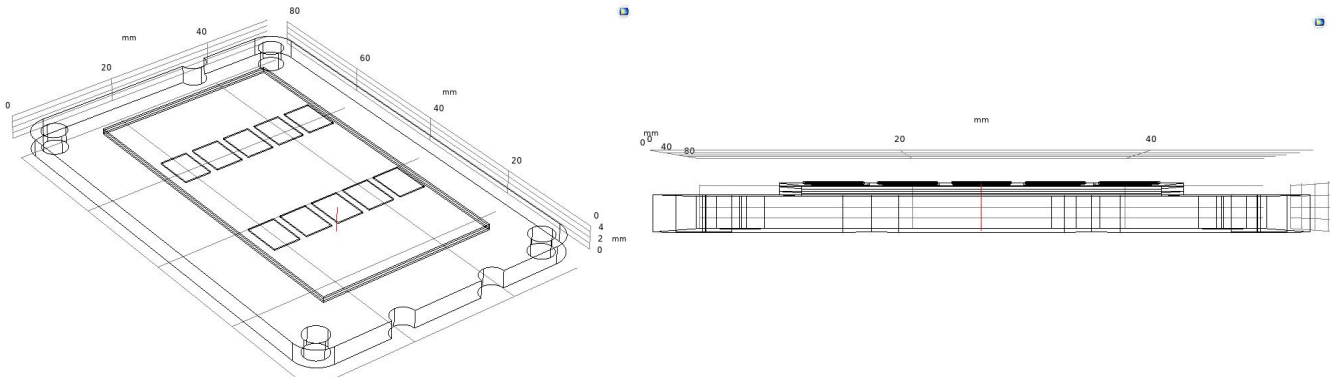


Figure 8: 3D-cutline across the module for temperature measurements, top view (left) and side view (right).

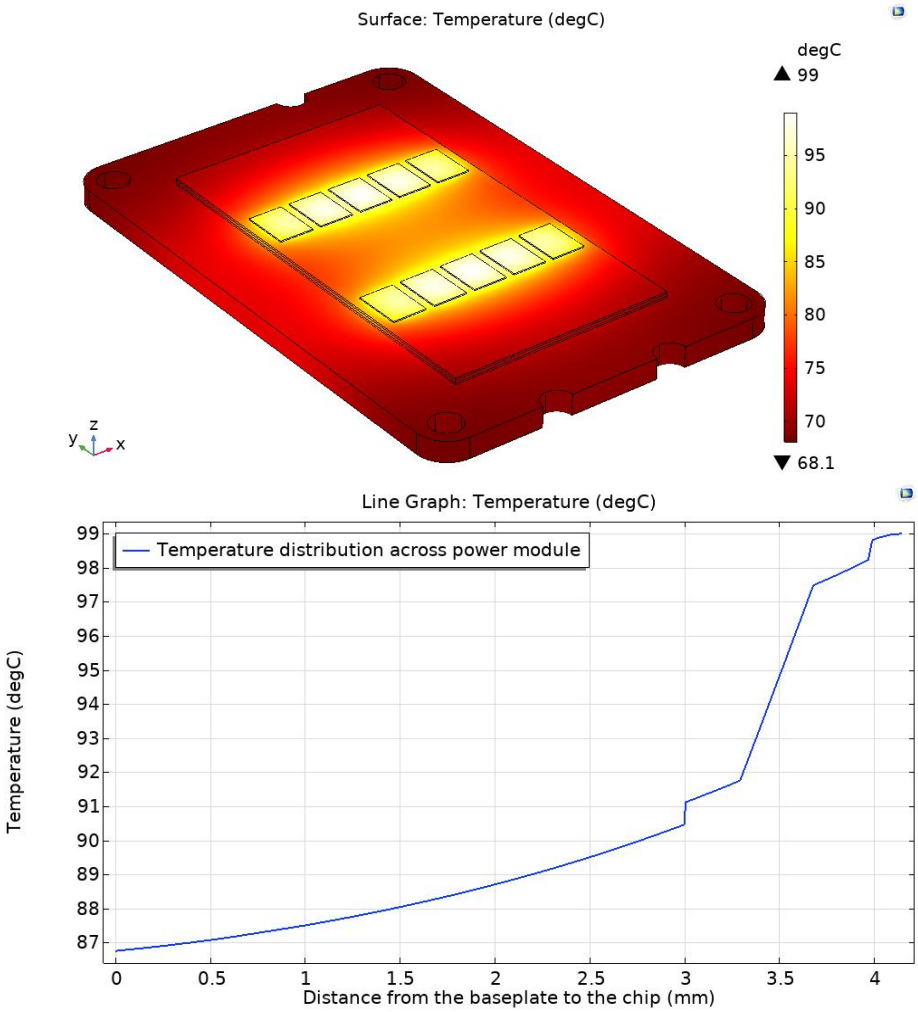


Figure 9: Temperature distribution across the power module (top) and temperature profile across all the layers (bottom).

10. Using the Graphene Assembled Film in the module Packaging

As mentioned before, a novel solution to efficiently transport the heat away from the power devices, especially in high power applications can be achieved by integrating ultra-high thermal conducting as well as flexible and robust materials into the cooling path [7]. Therefore, graphene which is one of the most promising materials that can provide the above-mentioned properties, has been used in this study to optimize the thermal impedance of the power module as well as to make a uniform and fast heat spreading over the module.

To prove this goal, the graphene layer has been placed as close as possible to the chips which are the heat sources, with an in-plane thermal conductivity of 2900 W/m.K in x and y directions and a cross-plane thermal conductivity of 14 W/m.K in the z direction. The directions as the design parameters should be chosen in such a way that they conduct and spread the heat flow as fast and

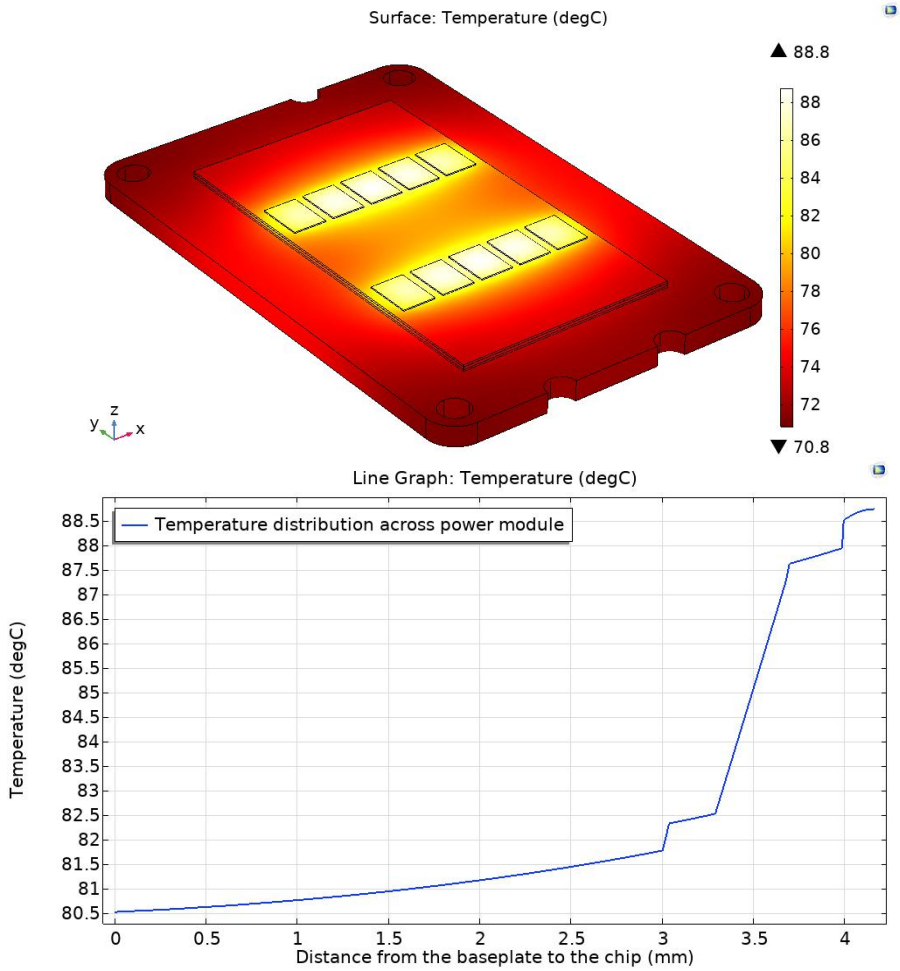


Figure 10: Temperature distribution across the power module (top) and temperature profile across all the layers (bottom) with graphene.

Table 5: Graphene Thermal Properties

Material	Thermal conductivity, K (W/m.K)	Density ρ (kg/m^3)	Heat capacity, C (J/kg.K)
Graphene	(x,y,z) = (2900,2900,14)	2267	720

efficient as possible on the surface of the module and then towards the bottom layers.

Worth mentioning is that, to reach high performing thermal and mechanical properties, the fabricating freestanding graphene films (GFs) have shown a promising potential compared to using single/few layer graphene for the thermal management of high-power electronics [7]. Hence, in this work, 1 μm thick graphene films are utilized to make an appropriate layer thickness in respect to the thermal conductivity which is challenging since the thermal conductivity levels are very much dependent on the thickness of the graphene layer [7]. The layer is placed under the first copper layer as a realistic placement in order to not weaken the electrical conductivity through the module which otherwise might happen as a result of a poor insulating ability of the graphene material.

In addition, in the fabrication processes, in order to minimize the thermal contact resistance between the graphene layer and other different substrates, a molecular functionalization method is used in which the small molecules are utilized to functionalize the surface of the graphene film which is in contact with the device surfaces. The small molecule, as the functional agent, gives as result a formation of molecular bridges between the device substrates and the grapheme-based surface and in this way the thermal resistance between these layers is strongly reduced [39].

Therefore, based on this method, the contact resistance is set to 1.1×10^{-8} $K.m^2/W$ in this work to achieve the appropriate thermal coupling between the layers [39]. Table 5 gives the material properties of graphene.

11. Results of Steady-state Thermal Computation of the Power Module with Graphene Layer

To investigate the thermal behavior in the presence of the graphene layer in the power module, a 3D-cutline from the middle chip to the baseplate, similar as in section 9 is applied to the module.

The result reveals a reduction of 10°C in the average temperature per chip for the applied power loss of 126.2 W per chip, as depicted in Figure 10, and Figure 10 (bottom) which illustrates the temperature profile across all the layers. Moreover, as can be seen in Figure 11, a comparison between the two cases,

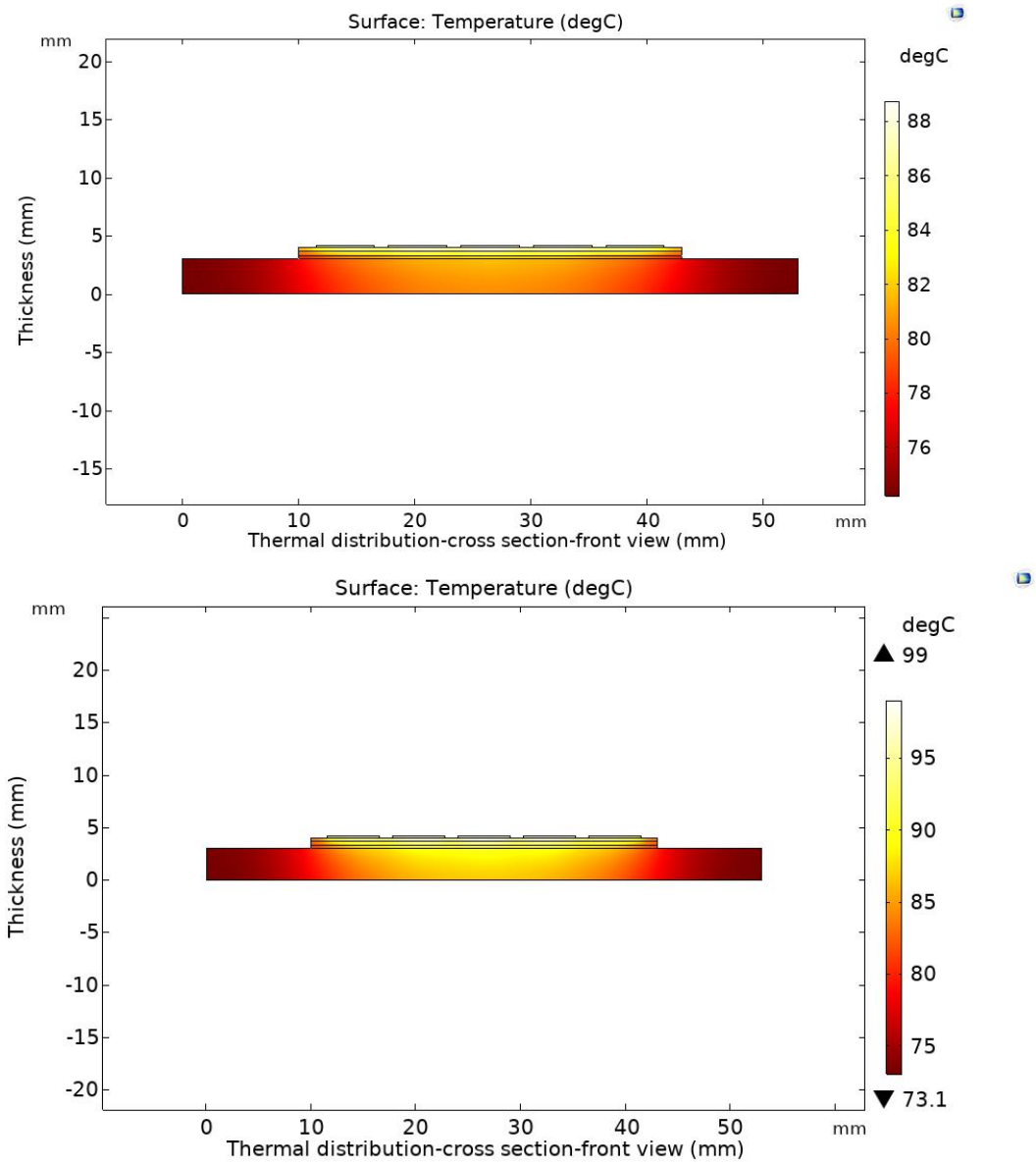


Figure 11: Temperature distribution with graphene (top) and without graphene (bottom).

with and without using graphene, confirms the expected results, i.e., a noticeable difference in the temperature distribution in the module. With graphene (top figure), the temperature is spreading flatter and more homogeneously in each layer, which means that each layer is absorbing the heat uniformly. Consequently, the heat is more distributed across the baseplate compared to that of the case without graphene (bottom figure) in which the heat is mostly concentrated in the middle where the chips are located. Therefore, this temperature uniformity leads to a significant improvement, 54.3%. The junction to case thermal resistance (R_{thJC}) equals to 0.0913 K/W, using the graphene layer compared to that of 0.2 K/W without graphene.

Moreover, Figure 12 depicts the temperature distribution across each layer which clearly shows the temperature difference of more than 10°C between the two cases for the applied power loss per chips.

12. Transient Thermal Computation of the Power Module, with/without Graphene layer using WLTC Driving Cycle

Since the load of the inverter is constantly varying, hence the real-time thermal computations are necessary to more accurately quantify the thermal behavior and reliability of the inverters.

Therefore, the time-dependent solver in COMSOL is used to evaluate the transient temperature distribution across the power module from the chips towards the bottom layers over the WLTC driving schedule consisting of both the urban and highway phases. Currently in many countries including the EU, WLTC is the main drive test procedure for light-duty vehicles. First based on the vehicle model used in this study, the required torque and speed values for the PMSM model are determined. The values are then utilized for the power loss calculation over the WLTC drive test

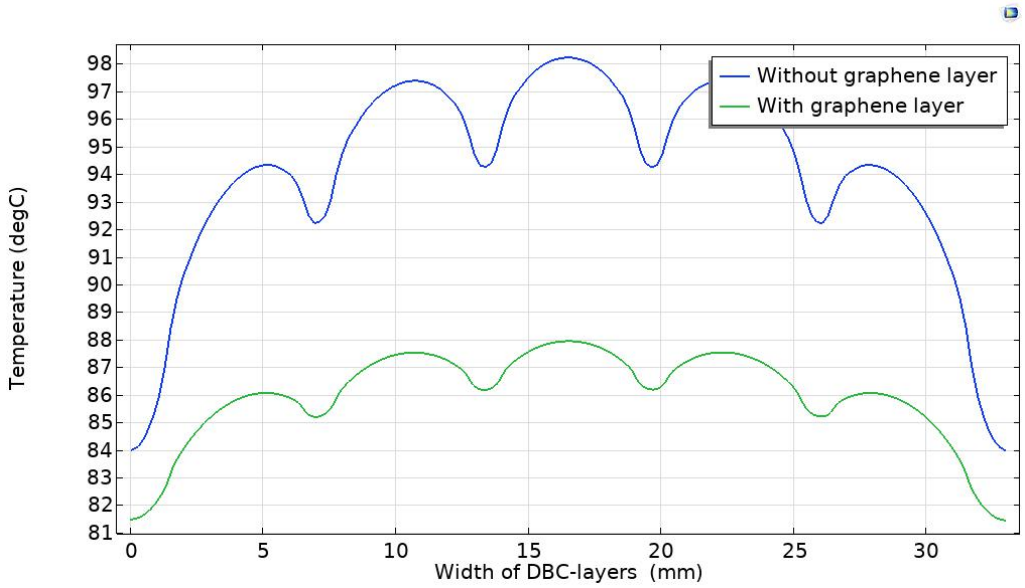


Figure 12: Temperature distribution across each layer with/without graphene.

Table 6: Vehicle Parameters

Vehicle Parameter	Value	Unit
Total mass of the vehicle	1700	[kg]
Aerodynamic drag coefficient	0.35	[-]
Vehicle cross section area	2	[m^2]
Rolling friction coefficient	0.007	[-]
Wheel radius	0.3	[m]

and thereafter for the thermal loading in COMSOL. Table 6 gives the vehicle model parameters.

In order to measure the transient heat distribution, a 3D-cutline from the middle chip to the baseplate is used similarly as done in the steady-state simulation and then the calculated power losses over WLTC are applied to the chips as the heat sources.

Figure 13 illustrates the calculated power losses in the WLTC, and Figure 14 shows the temperature distribution across the power module for the two cases (with/without) graphene. The thermal profile comparison per chip between the case of the power module with the assembled graphene film and without the graphene film is depicted in Figure 15. As can be seen in Figure 13, the highest calculated power losses are observed around 1000 second for the related speeds and torques of the PMSM based on the used vehicle model parameters. Therefore, the temperature distributions for the two cases, with/without graphene are considered at 1033 s.

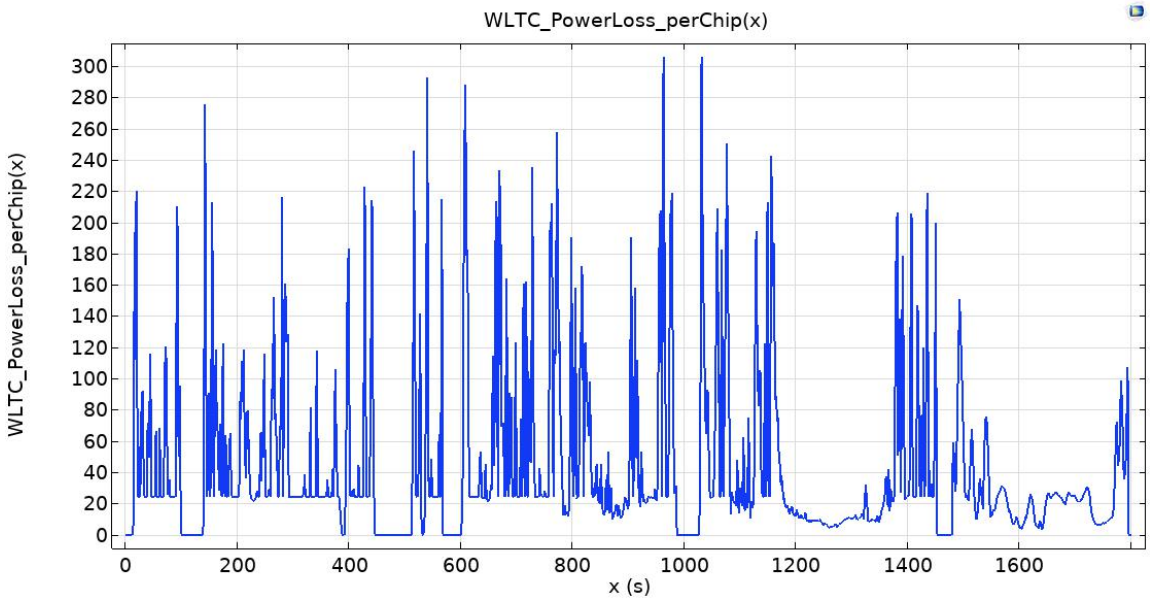


Figure 13: Calculated power losses over WLTC

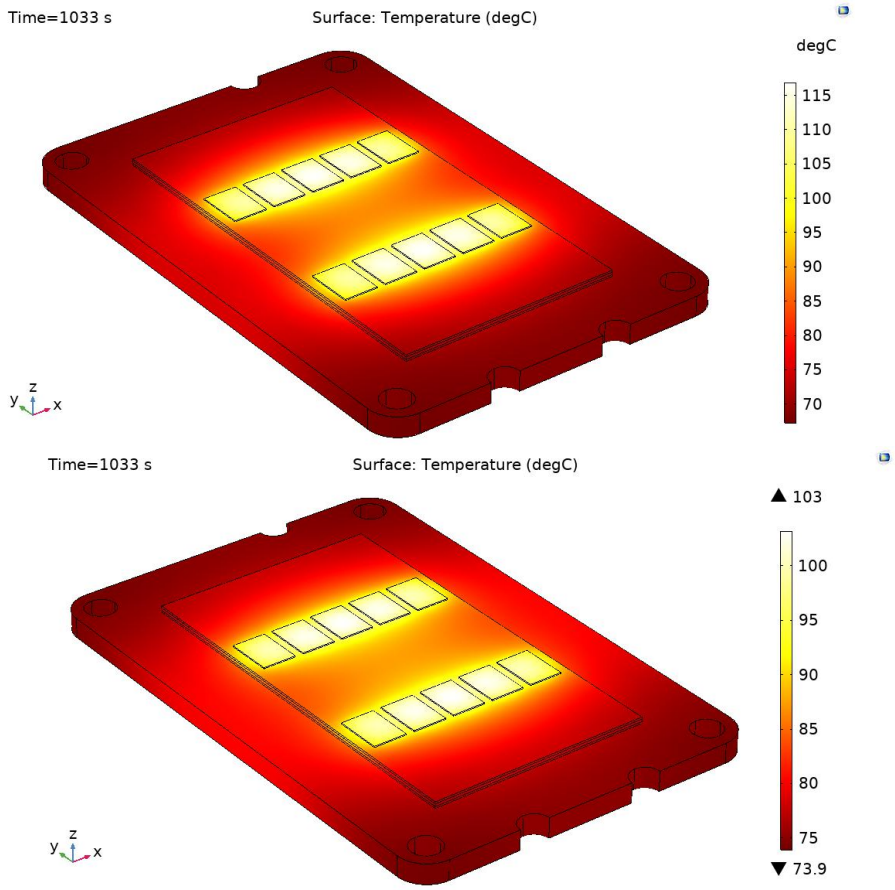


Figure 14: Temperature distribution across the power module without the graphene layer (top) and with the graphene layer (bottom) at 1033 s.

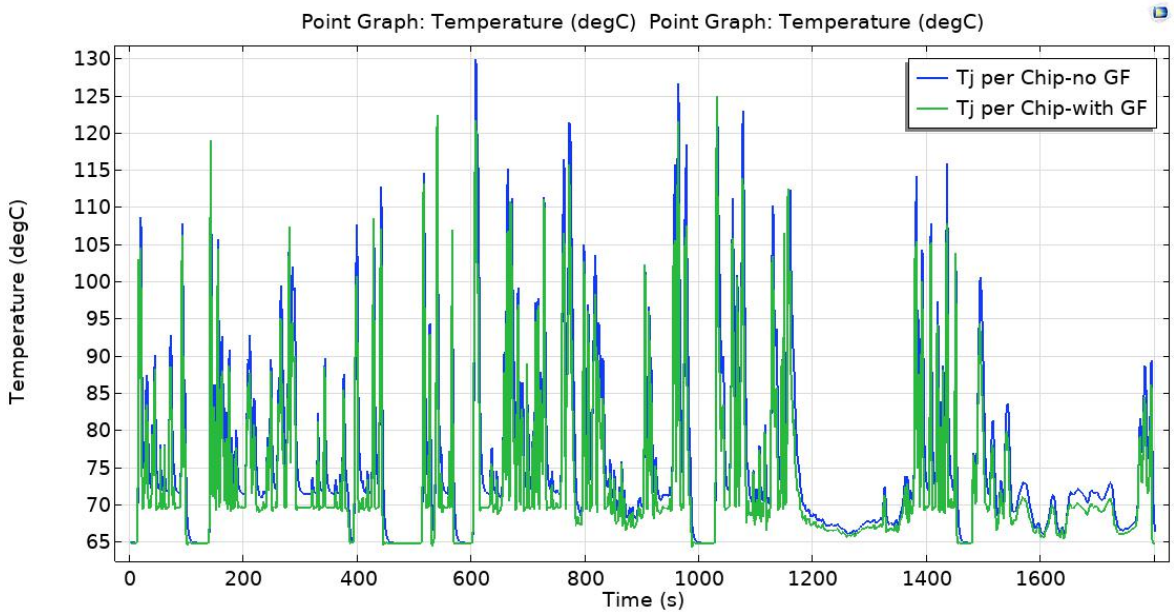


Figure 15: Comparison of the average junction temperature profile of each chip over the WLTC drive cycle.

According to Figure 14, a temperature reduction of about 15°C per chip at 1033 s is observed in case of using the graphene layer in the power module. Moreover, a close look into the heat distribution over the module shows a flatter distribution in case of using graphene. Also, with graphene the heat is more distributed across the baseplate compared to the case without graphene in which the heat is more concentrated in the middle, where the chips are located. Quite clearly proven, graphene as a heat conductor is improving the heat spreading across the module and this leads to an effectivization of the cooling of the semiconductor power module.

Finally, Figure 15 illustrates the thermal profile comparison over the WLTC driving schedule per chip between the two cases, with and without the graphene layer. As depicted, the graphene layer gives a lower average junction temperature per chip and this can lead to a lower operating temperature of the semiconductor power module in general. Worth mentioning is that this effect can be more apparent in high-power applications where due to the high voltage and current levels used, a high number of chips are needed. Also, as observed before in the steady-state simulation, boosting the heat dissipation performance in the chips and power module by using graphene, leads to a noticeable reduction in the junction to case thermal resistance of the power module and decreasing the operating temperature accordingly.

13. Conclusion

In this paper, the impact of neglecting the thermal feedback on the conduction losses for two SiC inverters were determined. The effect is demonstrated for a chosen heavy operating point of a PMSM, high torque and low speed, which is a difficult operating point from a power loss perspective. Since accounting for the thermal feedback shows a considerable increase in the inverter total conduction losses, therefore, graphene as a novel thermal management material is utilized as a layer in the power module packaging to improve the thermal dissipation and cooling of the SiC chips in the power module. A steady- state as well as a transient thermal computation in COMSOL Multiphysics are carried out on the investigated SiC power module which is created in a way to be as close as possible to a real module in terms of layers material, electrical and thermal aspects.

The thermal computation results imply a noticeable improvement in thermal spreading over the module which leads to a lower thermal resistance from junction to case of the module and a lower operating temperature in both the steady-state and transient computations. The transient computations have been applied over a WLTC drive cycle to investigate the thermal variations of the SiC chips when the load varies in a hybrid vehicle powertrain.

Worth mentioning is that, inserting the graphene assembled films in the SiC power module and functionalizing the films to reduce the contact thermal resistances in a standard electronics lab is very challenging due to a very thin and flexible texture of the graphene material. In addition, different variables including the inner temperatures, the heat distribution, and the streamlines, etc., within the multiple layers from the SiC chips down to the whole bottom layers cannot be easily measured on the hardware implementation even with many sensors. Therefore, using FEM in COMSOL Multiphysics is found to be a reliable method to simulate the thermal aspects precisely and prove the concepts of this paper.

To sum up, with this study, the application of using graphene in the building of the power module demonstrates a promising potential to boost the heat spreading in the semiconductor power modules especially in the high-power applications where also lightweight is required as priority in the light-duty vehicles.

As most important future work, measurement verification is the next step.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

References

[1] J. Reimers, L. Dorn-Gomba, C. Mak, A. Emadi, "Automotive traction inverters: Current status and future trends," IEEE Transactions on Vehicular Technology, **68**(4), 3337–3350, 2019.

[2] K. Kumar, M. Bertoluzzo, G. Buja, "Impact of SiC MOSFET traction inverters on compact-class electric car range," in 2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), 1–6, 2014, doi:10.1109/PEDES.2014.7042156.

[3] J. Biela, M. Schweizer, S. Waffler, J.W. Kolar, "SiC versus Si—Evaluation of Potentials for Performance Improvement of Inverter and DC–DC Converter Systems by SiC Power Semiconductors," IEEE Transactions on Industrial Electronics, **58**(7), 2872–2882, 2011, doi:10.1109/TIE.2010.2072896.

[4] K. Olejniczak, T. McNutt, D. Simco, A. Wijenayake, T. Flint, B. Passmore, R. Shaw, D. Martin, A. Curbow, J. Casady, B. Hull, "A 200 kVA electric vehicle traction drive inverter having enhanced performance over its entire operating region," in 2017 IEEE 5th Workshop on Wide Bandgap Power Devices and Applications (WiPDA), 335–341, 2017, doi:10.1109/WiPDA.2017.8170569.

[5] S.E. Schulz, "Exploring the High-Power Inverter: Reviewing critical design elements for electric vehicle applications," IEEE Electrification Magazine, **5**(1), 28–35, 2017, doi:10.1109/MELE.2016.2644281.

[6] J. Rąbkowski, T. Płatek, "Comparison of the power losses in 1700V Si IGBT and SiC MOSFET modules including reverse conduction," in 2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe), 1–10, 2015, doi:10.1109/EPE.2015.7309444.

[7] J.L. Nan Wang, Majid Kabiri Samani, Hu Li, Lan Dong, Zhongwei Zhang, Peng Su, Shujing Chen, Jie Chen, Shirong Huang, Guangjie Yuan, Xiangfan Xu, Baowen Li, Klaus Leifer, Lilei Ye, Tailoring the Thermal and Mechanical Properties of Graphene Film by Structural Engineering, Small, **14**(29), 1801346, doi:10.1002/smll.201801346.

[8] T.M. Tritt, Thermal conductivity : theory, properties, and applications, Springer, 2010.

[9] L. Peng, Z. Xu, Z. Liu, Y. Guo, P. Li, C. Gao, "Ultrahigh Thermal Conductive yet Superflexible Graphene Films," Advanced Materials, **29**(27), 2017, doi:10.1002/ADMA.201700589.

[10] A.A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, C.N. Lau, "Superior thermal conductivity of single-layer graphene," Nano Letters, **8**(3), 902–907, 2008, doi:10.1021/NL0731872.

[11] Y. Liu, S. Chen, Y. Fu, N. Wang, D. Mencarelli, L. Pierantoni, H. Lu, J. Liu, "A lightweight and high thermal performance graphene heat pipe," Nano Select, **2**, 2020, doi:10.1002/nano.202000195.

[12] N.J. Song, C.M. Chen, C. Lu, Z. Liu, Q.Q. Kong, R. Cai, "Thermally reduced graphene oxide films as flexible lateral heat spreaders," Journal of Materials Chemistry A, **2**(39), 16563–16568, 2014, doi:10.1039/C4TA02693D.

[13] G. Xin, H. Sun, T. Hu, H.R. Fard, X. Sun, N. Koratkar, T. Borca-Tasciuc, J. Lian, "Large-area freestanding graphene paper for superior thermal management," Advanced Materials, **26**(26), 4521–4526, 2014, doi:10.1002/ADMA.201400951.

[14] Z. Liu, Z. Li, Z. Xu, Z. Xia, X. Hu, L. Kou, L. Peng, Y. Wei, C. Gao, "Wet-spun continuous graphene films," Chemistry of Materials, **26**(23), 6786–6795, 2014, doi:10.1021/CM5033089.

[15] J. Li, F. Ye, S. Vaziri, M. Muhammed, M.C. Lemme, M. Östling, "Efficient inkjet printing of graphene," Advanced Materials, **25**(29), 3985–3992, 2013, doi:10.1002/ADMA.201300361.

[16] Y. Fu, J. Hansson, Y. Liu, S. Chen, A. Zehri, M.K. Samani, N. Wang, Y. Ni, Y. Zhang, Z. Bin Zhang, Q. Wang, M. Li, H. Lu, M. Sledzinska, C.M.S. Torres, S. Volz, A.A. Balandin, X. Xu, J. Liu, "Graphene related materials for thermal management," 2D Materials, **7**(1), 2020, doi:10.1088/2053-1583/AB48D9.

[17] N. Wang, Y. Liu, S. Chen, L. Ye, J. Liu, "Highly Thermal Conductive and Electrically Insulated Graphene Based Thermal Interface Material with Long-Term Reliability," in 2019 IEEE 69th Electronic Components and

Technology Conference (ECTC), 1564–1568, 2019, doi:10.1109/ECTC.2019.00240.
[18] M.F. Abdullah, S.A. Mohamad Badaruddin, M.R. Mat Hussin, M.I. Syono, "EVALUATION ON THE REDUCED GRAPHENE OXIDE THERMAL INTERFACE MATERIAL AND HEAT SPREADER FOR THERMAL MANAGEMENT IN HIGH-TEMPERATURE POWER DEVICE," Jurnal Teknologi, **83**(3 SE-), 53–59, 2021, doi:10.11113/jurnalteknologi.v83.15102.
[19] J.S. Lewis, Reduction of Device Operating Temperatures with Graphene-Filled Thermal Interface Materials, C , **7**(3), 2021, doi:10.3390/c7030053.
[20] Y. Liu, T. Thiringer, N. Wang, Y. Fu, H. Lu, J. Liu, "Graphene based thermal management system for battery cooling in electric vehicles," in 2020 IEEE 8th Electronics System-Integration Technology Conference (ESTC), 1–4, 2020, doi:10.1109/ESTC48849.2020.9229856.
[21] N. Wang, Y. Liu, L. Ye, J. Li, "Highly Thermal Conductive and Light-weight Graphene-based Heatsink," in 2019 22nd European Microelectronics and Packaging Conference Exhibition (EMPC), 1–4, 2019, doi:10.23919/EMPC44848.2019.8951839.
[22] S. Huang, Y. Zhang, S. Sun, X. Fan, L. Wang, Y. Fu, Y. Zhang, J. Liu, "Graphene based heat spreader for high power chip cooling using flip-chip technology," in 2013 IEEE 15th Electronics Packaging Technology Conference (EPTC 2013), 347–352, 2013, doi:10.1109/EPTC.2013.6745740.
[23] S. Huang, Y. Zhang, N. Wang, N. Wang, Y. Fu, L. Ye, J. Liu, "Reliability of graphene-based films used for high power electronics packaging," in 2015 16th International Conference on Electronic Packaging Technology (ICEPT), 852–855, 2015, doi:10.1109/ICEPT.2015.7236714.
[24] S. Huang, J. Bao, H. Ye, N. Wang, G. Yuan, W. Ke, D. Zhang, W. Yue, Y. Fu, L. Ye, K. Jeppson, J. Liu, "The effects of graphene-based films as heat spreaders for thermal management in electronic packaging," in 2016 17th International Conference on Electronic Packaging Technology (ICEPT), 889–892, 2016, doi:10.1109/ICEPT.2016.7583272.
[25] P. Zhang, N. Wang, C. Zandén, L. Ye, Y. Fu, J. Liu, "Thermal characterization of power devices using graphene-based film," in 2014 IEEE 64th Electronic Components and Technology Conference (ECTC), 459–463, 2014, doi:10.1109/ECTC.2014.6897324.
[26] M.R.M. Hussin, M.M. Ramli, S.K.W. Sabli, I.M. Nasir, M.I. Syono, H.Y. Wong, M. Zama, "Fabrication and characterization of graphene-on-silicon schottky diode for advanced power electronic design," Sains Malaysiana, **46**(7), 1147–1154, 2017, doi:10.17576/JSM-2017-4607-18.
[27] J. Wang, L. Shi, Y. Li, L. Jin, Y. Xu, H. Zhang, Y. Zou, Y. Lan, X. Ma, "Thermal management of graphene-induced high-power semiconductor laser package with bidirectional conduction structure," Optics and Laser Technology, **139**, 2021, doi:10.1016/J.OPTLASTEC.2021.106927.
[28] K. Hemanjaneyulu, M. Khaneja, A. Meersha, H.B. Variar, M. Shrivastava, "Comprehensive Computational Modelling Approach for Graphene FETs," in 2018 4th IEEE International Conference on Emerging Electronics (ICEE), 1–5, 2018, doi:10.1109/ICEE44586.2018.8937909.
[29] S. Amirpour, T. Thiringer, D. Hagstedt, "Power Loss Analysis in a SiC/IGBT Propulsion Inverter Including Blanking Time, MOSFET's Reverse Conduction and the Effect of Thermal Feedback Using a PMSM Model," in IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, 1424–1430, 2020, doi:10.1109/IECON43393.2020.9254297.
[30] Wolfspeed, CAB450M12XM3, Nov. 2021.
[31] Wolfspeed, CAS300M12BM2, Nov. 2021.
[32] A. Acquaviva, T. Thiringer, "Energy efficiency of a SiC MOSFET propulsion inverter accounting for the MOSFET's reverse conduction and the blanking time," in 2017 19th European Conference on Power Electronics and Applications (EPE'17 ECCE Europe), P.1-P.9, 2017, doi:10.23919/EPE17ECCEEurope.2017.8099052.
[33] J.W. Kolar, H. Ertl, F.C. Zach, "Influence of the modulation method on the conduction and switching losses of a PWM converter system," IEEE Transactions on Industry Applications, **27**(6), 1063–1075, 1991, doi:10.1109/28.108456.
[34] SEMIKRON, Application Manual Power Semiconductors, Ilmenau, 2015.
[35] M. Meyer, J. Bocker, "Optimum Control for Interior Permanent Magnet Synchronous Motors (IPMSM) in Constant Torque and Flux Weakening Range," in 2006 12th International Power Electronics and Motion Control Conference, 282–286, 2006, doi:10.1109/EPEPEMC.2006.4778413.
[36] M. Meyer, T. Grote, J. Bocker, "Direct torque control for interior permanent magnet synchronous motors with respect to optimal efficiency," in 2007 European Conference on Power Electronics and Applications, 1–9, 2007, doi:10.1109/EPE.2007.4417370.
[37] L. Ceccarelli, R.M. Kotecha, A.S. Bahman, F. Iannuzzo, H.A. Mantooth, "Mission-Profile-Based Lifetime Prediction for a SiC mosfet Power Module Using a Multi-Step Condition-Mapping Simulation Strategy," IEEE Transactions on Power Electronics, **34**(10), 9698–9708, 2019, doi:10.1109/TPEL.2019.2893636.
[38] E. EDGE, Convective heat transfer coefficients table chart, Nov. 2021.
[39] H. Han, Y. Zhang, N. Wang, M.K. Samani, Y. Ni, Z.Y. Mijbil, M. Edwards, S. Xiong, K. Sääskilahti, M. Murugesan, Y. Fu, L. Ye, H. Sadeghi, S. Bailey, Y.A. Kosevich, C.J. Lambert, J. Liu, S. Volz, "Functionalization mediates heat transport in graphene nanoflakes," Nature Communications, **7**(1), 11281, 2016, doi:10.1038/ncomms11281.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 101-106 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Monthly Rainfall Forecasting from Sea Surface Temperature Spatial Pattern

Prattana Deeprasertkul*, Royol Chitradon

Hydro - Informatics Institute (Public Organization), Bangkok, 10900, Thailand

ARTICLE INFO

Article history:
Received: 20 July, 2021
Accepted: 08 September, 2021
Online: 16 September, 2021

Keywords:
Rainfall Forecast
Sea Surface Temperature
Pattern Matching

ABSTRACT

The ocean surface temperatures or sea surface temperatures have a significant influence on local and global weather. The change in sea surface temperatures will lead to the change in rainfall patterns. In this paper, the long-term rainfall forecasting is developed for planning and decision making in water resource management. The similarity of sea surface temperature images pattern that was applied to analyze and develop the monthly rainfall forecasting model will be proposed. In this work, the convolutional neural network and autoencoder techniques are applied to retrieve the similar sea surface temperature images in database store. The accuracy values of the monthly rainfall forecasting model which is the long-term forecasting were evaluated as well. The average value of the model accuracies was around 82.514%.

1. Introduction

The rainfall forecast modeling systems are available to monitor and warn flood and drought events in Thailand. There are two fundamental approaches which are dynamical and empirical methods for rainfall forecasting. The dynamical prediction is generated by physical model based on numerical analysis that represents the evolution in seasonal forecasting by using observed Sea Surface Temperature anomalies as a boundary condition. The empirical approach relies on the observed and measured data. It is used to analyse the atmospheric and oceanic variables in the real world [1].

Tropical rainfall is influenced strongly by nearby sea temperatures in the real world [1]. Recently, there are some climate predictions used empirical relations between Sea Surface Temperature or SST and atmospheric variables [2]. The empirical approaches is the most widely used to predict the climate using linear regression, machine learning, artificial neural network, and the methods of data-driven handling [3, 4].

In this paper, the monthly rainfall forecasting model is implemented by using the convolutional autoencoder technique for efficient unsupervised encoding of image features. We used 7 years (2014-2020) datasets of the global climate data such as SST and the Global Satellite Mapping of Precipitation project or GSMaP datasets that both are the satellite images. This model has predicted the monthly rainfall amount in image format. This is the improvement of the data-driven analytics which is a project of Hydro – Informatics Institute (Public Organization) [5].

The monthly rainfall forecasting amount with 6 months ahead were predicted. The experimental results show the accuracy that was calculate between the prediction rainfall amount and the actual rainfall amount.

The remaining of this paper is organized as follows: Section 2 describes the background knowledge related to this work. Section 3 explains the analysis of SST matching. In section 4, the experimental results of monthly rainfall forecasting model are presented. Finally, this work is concluded in section 5.

2. Background

The datasets and image processing technologies using in this monthly rainfall forecasting model are described in this section.

2.1. Data

The set of data using in this work consist of two datasets that were SST as displayed in Figure 1 and the satellite precipitation or GSMaP as displayed in Figure 2. Firstly, SST is the ocean temperature that closes to the ocean's surface. The sea surface is between 1 millimetre and 20 meters below the sea surface [6]. The weather satellite sensors are available to measure and define sea surface temperatures. Secondly, the GSMaP project [6,7,8,9] is a high resolution global precipitation map production applying satellite data. This project is supported by Core Research for Evolutional Science and Technology (CREST) of the Japan

*Corresponding Author: Prattana Deeprasertkul, prattana@hii.or.th

Science and Technology Agency (JST) during 2002-2007. The GSMaP project activities are promoted by the JAXA Precipitation Measuring Mission (PMM) Science Team since 2007.

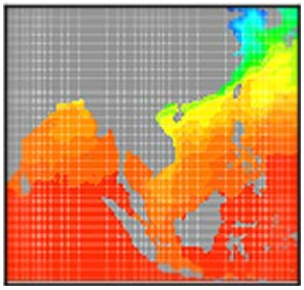


Figure 1: The example of SST data.

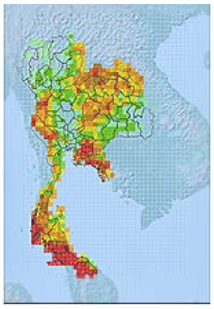


Figure 2: The example of GSMaP data.

2.2. Type and Format of Data

Because SST and GSMaP are the satellite images in the digital formats, to display both on Map and to have a good performance, both are converted from images (JPEG) format to text (ESRI ASCII raster) format. The ESRI ASCII raster file format has the basic structure as illustrated in Table 1 [10]. The header fields at the top of the file displayed in "Parameter" column in Table 1 [10]. The example of ESRI ASCII raster file is shown in Figure 3.

Table 1: The description of the header fields of ESRI ASCII raster file.

Parameter	Description	Requirements
NCOLS	Number of cells in a column.	Integer
NROWS	Number of cells in a row.	Integer
XLLCENTER	Coordinate X on the origin point (at lower left corner of the cell).	Match with Y coordinate type.
YLLCENTER	Coordinate Y on the origin point (at lower left corner of the cell).	Match with X coordinate type.
CELLSIZE	Cell size.	Decimal
NODATA_VALUE	The value means NoData in the raster file.	Optional

```
ncols 480
nrows 450
xllcorner 378923
yllcorner 4072345
cellsize 30
nodata_value -32768
43 2 45 7 3 56 2 5 23 65 34 6 32 54 57 34 2 2 54 6
35 45 65 34 2 6 78 4 2 6 89 3 2 7 45 23 5 8 4 1 62 ...
```

Figure 3: The example of ESRI ASCII raster file.

2.3. Convolutional Autoencoder

A Convolutional Autoencoder or ConvAE shown in Figure 4 is a type of Convolutional Neural Network (CNN) designed for unsupervised deep learning. CNN is widely used in the field of pattern recognition. CNN architecture is constructed from various layers in the neural network. The convoluted layers are used for automatic extraction of an image feature hierarchy. The size of the input layer is the same as output layer. An autoencoder consists of image encoder and image decoder. The training of the autoencoder encodes image features into compressed images. The encoded image features are then clustered using k-means clustering. a small sub volume set is characterized by the decoded cluster center [11]. This technique is applied to reduce the storage usage, speeding up the computation time, improving performance by removing redundant variables [12].

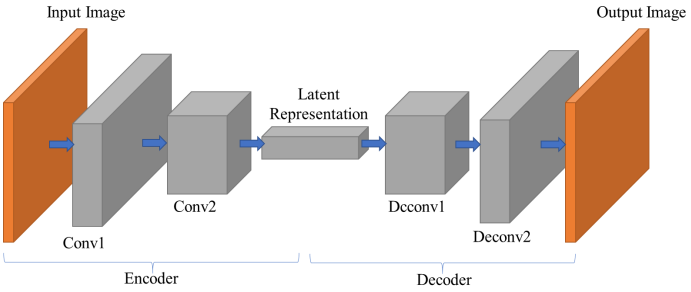


Figure 4: Convolutional Autoencoder.

- An encoder as shown in Figure 4 is an algorithm used to transfer one format to another format. In this work, the encoder is applied to the original image. Then, the original image is represented in smaller feature vectors image or it is compressed [13].
- The decoder as shown in Figure 4 is an algorithm used for image decompression. The important features of the original image are not lost after the decoding process is applied in the encoded image. The compressed image is returned to its original image [13].

2.4. Image Retrieval

To retrieve the similar images stored in the images database and to get the top-k most similar database images, K-Nearest Neighbours or KNN method with cosine similarity as the distance metric. KNN is applied to verify the nearest neighbours and to get k nearest points as the similar images from a query image.

The Content Based Image Retrieval method was applied in this work as the process flow shown in Figure 5. In training process, the encoded representation of all images from training database is calculated using ConvAE and saved in a file. In testing process, the encoded representation of query image from testing database is calculated using ConvAE and saved in a file. The distance between the encoded representations of query image to encoded representation of all the images in training database is calculated to find nearest neighbors. Finally, KNN retrieves the

top-k images in descending order of distance to the query image [14].

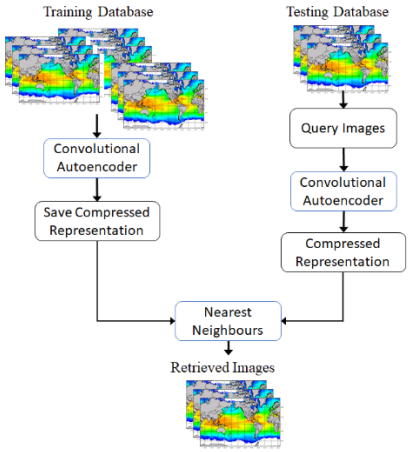


Figure 5: The Content Based Image Retrieval process flow.

2.5. Image Matching

Image matching is method to evaluate the maximum similarity values between two digital images. The method is divided into two categories: 1) gray-based matching 2) feature-based matching. The gray-based matching is a matching algorithm of the grayscale feature applied in the image pixels. The feature-based matching is a matching algorithm of the feature vector points [15]. Image similarity is the similarity measurement between two images with the metrics. Therefore, some similarity criteria rules are used to verify the corresponding points of two images. The similarity quality is evaluated by the distance. Recently, there are many metrics of the distance measurement such as Euclidean distance, humming distance, cosine distance, and etc. the Hamming distance is a metric for measuring the distance between two binary strings.

2.6. Image Hashing

Image hashing has various methods which can be the mapping of the similar images with similar hashing codes. The well-known image hashing methods are the followings [16]:

- Average image hashing is the simplest hashing method. The images are converted to the greyscales which are transformed to binary codes with the average values.
- Perceptual image hashing extracts the real image features to hash values with discrete cosine transformation. This image hashing is the popular hashing method in signal processing.
- Difference image hashing uses the relative gradient technique to calculate the difference between adjacent pixels.

Hamming distance is the metric to compare and calculate the distance between two hashing strings. From comparing between the equal length of two hashing strings, hamming distance computes the number of bits which are different [17].

3. Sea Surface Temperature Matching

This section has described the convolutional autoencoder layers when the process was implemented and the image matching process for monthly rainfall forecast application.

3.1. Convolutional Autoencoder Layers

The convolutional autoencoder aims to predict monthly rainfall from the Sea Surface Temperature images. The rainfall predictor is implemented by adding the following layers to the neural network as described in Table 2 [13]. There is a procedure as the following:

- Step 1: A convoluted layer with shape of (287, 309, 16) is added and collected separately in the output layer 1. It is extracted significant features from the original image before reducing the shape in smaller features.
- Step 2: A maximum layer selected and pooled with the shape of (144, 155, 16), is associated to the first step convoluted layer and it resizes the layer to smaller features.
- Step 3: A dropout layer, with the shape of (144, 155, 8), is regularization technique to avoid overfitting of neural network.
- Step 4: A maximum layer pooled with the shape of (72, 78, 8), is associated to the previous step convoluted layer.
- Step 5: The previous maximum layer is pooled to the shape of (72, 78, 8) before a convoluted layer is applied.
- Step 6: A maximum layer selected and pooled with the shape of (36, 39, 8), is associated to the previous step convoluted layer.
- Step 7: The previous maximum layer is pooled to the shape of (36, 39, 8) before a convoluted layer is applied.
- Step 8: An upsampled layer with the shape of (72, 78, 8), is applied to increase the image resolution to the larger image shape.
- Step 9: The previous convoluted layer is performed by the additional operation on the output layer 2 with the shape of (72, 78, 8).
- Step 10: An upsampled layer with the shape of (144, 156, 8), is applied to increase the image resolution to the larger image shape.
- Step 11: The previous convoluted layer is performed by the additional operation on the output layer 1 with the shape of (142, 154, 16).
- Step 12: An upsampled layer with the shape of (284, 308, 16), is applied to increase the image resolution to the larger image shape.
- Step 13: The final layer of neural networks is applied to the convoluted layer with the shape of (284, 308, 4) to receive the output.

Table 2: The layers of Autoencoder method [13].

Layer (type)	Output Shape
Input Layer	(287, 309, 4)
Conv	(287, 309, 16)
Maximum Pooling	(144, 155, 16)
Conv	(144, 155, 8)

Maximum Pooling	(72, 78, 8)
Conv	(72, 78, 8)
Maximum Pooling	(36, 39, 8)
Conv	(36, 39, 8)
Upsampling	(72, 78, 8)
Conv	(72, 78, 8)
Upsampling	(144, 156, 8)
Conv	(142, 154, 16)
Upsampling	(284, 308, 16)
Conv	(284, 308, 4)

3.2. Image Retrieval Process

This application is the SST pattern data analysis for 6 months rainfall forecasting. In this work, the Content Based Images Retrieval technique to estimate the similarity values was applied in monthly rainfall forecasting of Thailand. Figure 6 illustrates the SST image matching process. The current SST has been detected and matched with the historical SST dataset stored in the database. In this project, MongoDB database application is used to collect the images which are the unstructured data. The program code is developed to retrieve the similar database images with comparing to the current SST images. The top three highest values of image similarities are the results of monthly rainfall forecasting.

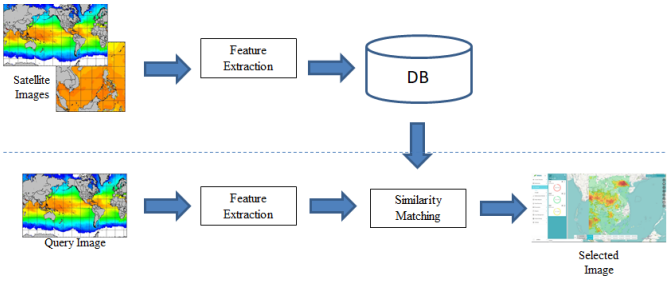


Figure 6: The image matching process flow [5].

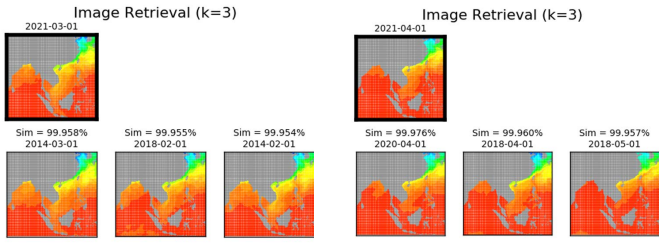


Figure 7: The results of the top three highest SST similarities.

4. Experimental Results

The program has processed and generated the results of the top three highest SST similarities written by Python. These results are shown in Figure 7. The rainfall amounts using in this work are the accumulated rainfall or the accumulated GSMaP rainfall at the same months of top three highest SST matching. These months results are used to define the predicted months of rainfall data and they are shown on the application screen in Figure 8. In addition, the SST and GSMaP images can be simultaneously overlaid on application screen as illustrated in Figure 9.

The accurate measurement of this monthly rainfall forecasting analysis is described the detail of process in this section as well. The process flow is displayed in Figure 10. For this measurement, the image hashing method was applied to retrieve the GSMaP image similarities. The hashing codes are used to be the index to discover the most similar database images. The image hashing codes are used to compare the images which are easier and faster method than using the whole of original images [18].

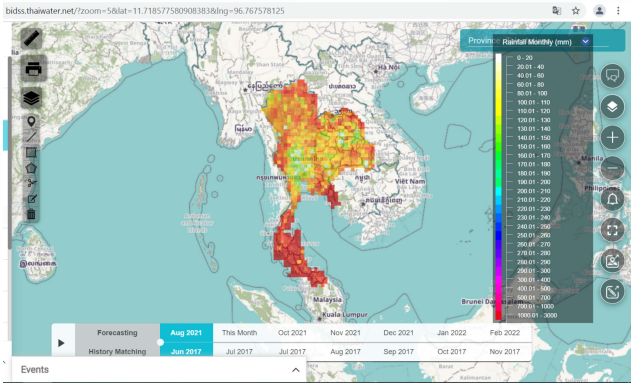


Figure 8: The example of monthly GSMaP data overlaid on map.

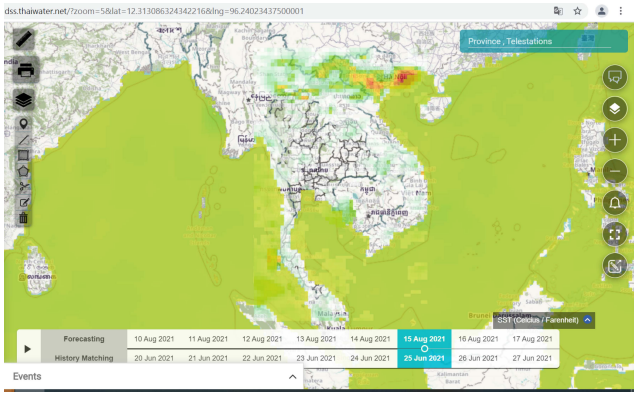


Figure 9: SST and GSMaP data were overlaid simultaneously on the application screen.

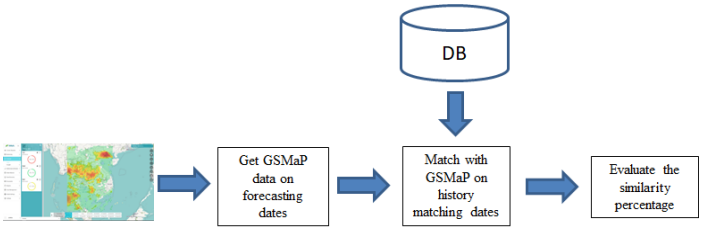


Figure 10: The process flow of the accurate measurement of rainfall forecasting model [5].

The program code of accuracy evaluation of this monthly rainfall forecasting model by using image pattern similarity was written by Python Jupyter as shown in Figure 11. The current monthly GSMaP rainfall datasets were matched with the historical GSMaP datasets collected in MongoDB. Both of the current and the historical GSMaP hashing codes were created and then the similarity values between two images were computed with their hashing method. The example between the n^{th} months

GSMaP rainfall forecasting and the n^{th} months of the historical matching GSMaP as shown in Table 3 and Table 4.

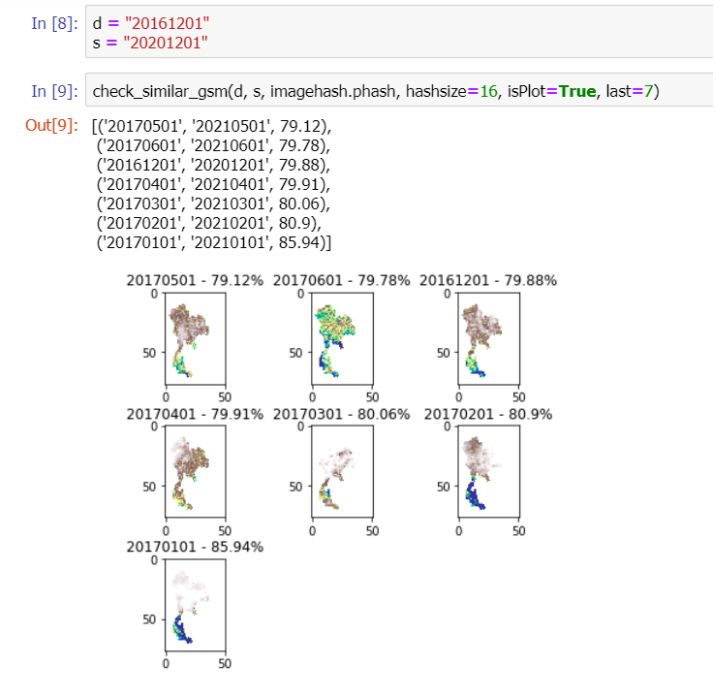


Figure 11: The example of Python Jupyter displaying for GSMaP comparison.

Table 3: The example I of the rainfall forecasting date and the historical matching date of GSMaP data.

Forecasting date	History matching date
Dec. 2020	Dec. 2016
Jan. 2021	Jan. 2017
Feb. 2021	Feb. 2017
Mar. 2021	Mar. 2017
Apr. 2021	Apr. 2017
May 2021	May 2017
Jun. 2021	Jun. 2017

Table 4: The example II of the rainfall forecasting date and the historical matching date of GSMaP data.

Forecasting date	History matching date
Nov. 2020	Nov. 2016
Dec. 2020	Dec. 2016
Jan. 2021	Jan. 2017
Feb. 2021	Feb. 2017
Mar. 2021	Mar. 2017
Apr. 2021	Apr. 2017
May 2021	May 2017

The GSMaP rainfall images on the forecasting months were computed the similarity values with the historical GSMaP matching months. The example image of GSMaP similarity (%) case is shown in Figure 12. In addition, the example of the similarity values of monthly GSMaP rainfall forecasting and their historical matching were calculated as shown in Table 5. Table 5 is the example 1 which shows the GSMaP data matching between forecasting months, 2020/12 – 2021/06, and history matching months, 2016/12 – 2017/06. The similarity values of example 1 are 81.21%, 93.18%, 84.14%, 83.75%, 82.3%, 79.53%, and 80.52%, respectively and the average of similarity values is 83.52%. For the example 2 in Table 6, the forecasting months are 2020/11 – 2021/05 that match with the history matching months, 2016/11 – 2017/05 and the similarity values are 81.46%, 81.21%, 93.18%, 84.14%, 83.75%, 82.3%, and 79.53%, respectively. The average of similarity values in the example 2 is 83.65%.

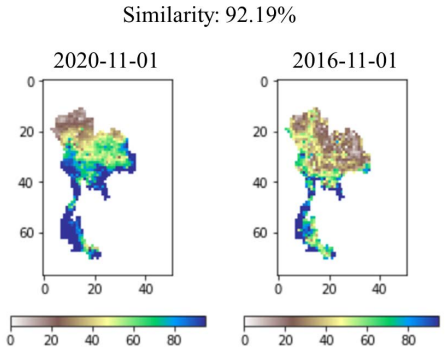


Figure 12: The example images of GSMaP similarity (%)

Table 5: The example I of the average of monthly GSMaP forecasting and the historical matching similarity

Forecasting months	History matching months	Similarity (%)
2020/12	2016/12	81.21
2021/01	2017/01	93.18
2021/02	2017/02	84.14
2021/03	2017/03	83.75
2021/04	2017/04	82.3
2021/05	2017/05	79.53
2021/06	2017/06	80.52
	Avg.	83.52

Table 6: The example II of the average of monthly GSMaP forecasting and the historical matching similarity

Forecasting months	History matching months	Similarity (%)
2020/11	2016/11	81.46
2020/12	2016/12	81.21
2021/01	2017/01	93.18
2021/02	2017/02	84.14
2021/03	2017/03	83.75
2021/04	2017/04	82.3
2021/05	2017/05	79.53
	Avg.	83.65

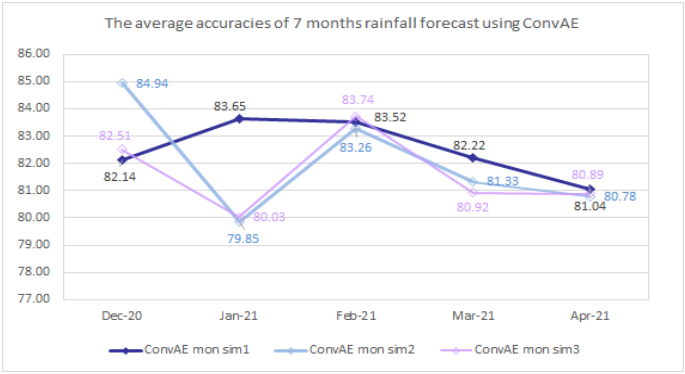


Figure 13: The graph of the monthly rainfall forecast similarity since December 2020 to April 2021

Figure 13 is the comparison between the accurate values of top three highest similarities or top three rainfall forecasting. The graph shows the accurate values of 5 months GSMaP rainfall forecast since December 2020, January 2021, February 2021, March 2021, to April 2021. The accurate values of the highest similarity (dark blue line) are 82.14%, 83.65%, 83.52%, 82.22%, and 81.04%, respectively. The average of the first accuracy is 82.51%. The average of the second and the third accurate values are 82.03% and 81.62%, respectively. Therefore, it can be found that the highest GSMaP rainfall forecast is better than others.

5. Conclusion

The monthly rainfall forecasting was the SST image patterns analysis by computing the average similarity between the recent monthly SST image datasets and historical monthly SST image datasets and applying convolutional autoencoder method to retrieve the top three highest of SST similarities. Moreover, the accurate value of this monthly rainfall forecasting was evaluated by applying the GSMaP image matching method as well. Then the result of this monthly rainfall forecasting was 82.514% which was an average value in matching GSMaP testing datasets.

However, this monthly rainfall forecasting model will be improved the accuracy, in the near future, by considering more atmospheric variables related to rainfall in Thailand. In addition, the weekly rainfall forecasting will be added in the application.

References

[1] M. Kannan et al., “Rainfall Forecasting Using Data Mining Technique,” International Journal of Engineering and Technology, **2**(6):397-401, 2010.

[2] University of Reading, "Tropical Rainfall and Sea Surface Temperature Link Could Improves Forecast," https://www.reading.ac.uk/news-and-events/releases/PR849878.aspx, 26 October 2020.

[3] M. Haghbin et al., "Applications of soft computing models for predicting sea surface temperature: a comprehensive review and assessment," Progress in Earth Planetary Science, **8:**4, 2021, doi:10.1186/s40645-020-00400-9.

[4] O. Ieremeiev, V. Lukin, K. Okarma, and K. Egiazarian, "Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images," Remote Sensing, **12**(15):2349, 2020, doi:10.3390/rs12152349.

[5] P. Deeprasertkul, "An Assessment of Rainfall Forecast using Image Similarity Processing," Proceedings of the 2020 3rd International Conference on Image and Graphics Processing, 141-145, 2020, doi:10.1145/3383812.3383821.

[6] K. Okamoto et al., "The Global Satellite Mapping of Precipitation (GSMaP) project," 25th IGARSS Proceedings, 3414-3416, 2005, doi: 10.1109/IGARSS.2005.1526575.

[7] K. Aonashi, J. Awaka, M. Hirose, T. Kozu, T. Kubota, G. Liu, S. Shige, S. Kida, S. Seto, N. Takahashi, and Y. N. Takayabu, "GSMaP passive, microwave precipitation retrieval algorithm: Algorithm description and validation," J. Meteor. Soc. Japan, 87A, 119-136, 2009.

[8] T. Kubota, S. Shige, H. Hashizume, K. Aonashi, N. Takahashi, S. Seto, M. Hirose, Y. N. Takayabu, K. Nakagawa, K. Iwanami, T. Ushio, M. Kachi, and K. Okamoto, "Global Precipitation Map using Satelliteborne Microwave Radiometers by the GSMaP Project : Production and Validation," IEEE MicroRad, 290-295, 2006, doi: 10.1109/MICRAD.2006.1677106.

[9] T. Ushio, T. Kubota, S. Shige, K. Okamoto, K. Aonashi, T. Inoue, N. Takahashi, T. Iguchi, M. Kachi, R. Oki, T. Morimoto, and Z. Kawasaki, "A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data," J. Meteor. Soc. Japan, 87A, 137-151, 2009, doi:10.2151/jmsj.87A.137.

[10] ESRI Inc., "ESRI ASCII raster format," http://resources.esri.com/help/9.3/arcgisengine/java/GPToolRef/spatial_analyst_tools/esri_ascii_raster_format.htm.

[11] X. Zeng, M. Ricardo Leung, T. Zeev-Ben-Mordehai, and M. Xu, "A convolutional autoencoder approach for mining features in cellular electron cryo-tomograms and weakly supervised coarse segmentation," Journal of Structural Biology, **202**(2):150-160, 2018, doi:10.1016/j.jsb.2017.12.015.

[12] Y. Zhang, “A Better Autoencoder for Image: Convolutional Autoencoder.” In ICONIP17-DCEC, 2018.

[13] P. Kumar and N. Goel, "Image Resolution Enhancement Using Convolutional Autoencoders." 7th Electronic Conference on Sensors and Applications Proceedings, 8259, 2020, doi:10.3390/ecsa-7-08259.

[14] P. Deore, V. S. Patil, and V. B. Patil, "Content Based Image Retrieval using Deep Convolutional Variational Autoencoder," Proteus Journal, **11** (10), 2020, ISSN/eISSN: 0889-6348.

[15] G. Jianhua, Y. Fan, T. Hai, W. Jing-xue, and L. Zhiheng, "Image matching using structural similarity and geometric constraint approaches on remote sensing images," Journal of Applied Remote Sensing, **10**(4), 2016, doi:10.1117/1.JRS.10.045007.

[16] T. Zhenjun, W. Shuozhong, Z. Xinpeng, and W. Weimin, "Structural Feature-Based Image Hashing and Similarity Metric for Tampering Detection," Fundamenta Informticae, **106**(1):75-91, 2011, doi:10.3233/FI-2011-377.

[17] N. Raut, "What is Hamming Distance?," TutorialPoint, 2018.

[18] Study.com, "Discuss the advantages of hashing," 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 5, 107-124 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

A Summary of Canonical Multivariate Permutation Entropies on Multivariate Fractional Brownian Motion

Marisa Mohr[*,1,2], Ralf Möller[1,3]

[1] *University of Lübeck, Institute of Information Systems, Lübeck, 23562, Germany*

[2] *inovex GmbH, Data Management & Analytics, Hamburg, 22763, Germany*

[3] *German Research Center for Artificial Intelligence, Stochastic Relational AI in Healthcare, Lübeck, 23562, Germany*

ARTICLE INFO

Article history:
Received: 14 July, 2021
Accepted: 08 September, 2021
Online: 16 September, 2021

Keywords:
Fractional Brownian Motion
Multivariate Time Series
Ordinal Pattern
Permutation Entropy
Hurst Parameter Estimation
Feature Engineering

ABSTRACT

Real-world applications modelled by time-dependent dynamical systems with specific properties such as long-range dependence or self-similarity are usually described by fractional Brownian motion. The investigation of the qualitative behaviour of its realisations is an important topic. For this purpose, efficient mappings from realisations of the dynamical system, i.e., time series, to a set of scalar-valued representations that capture certain properties are considered. Permutation entropy is a well-known measure to quantify the complexity of univariate time series in a scalar-valued representation, for example, to derive estimates for self-similarity or as features or representations in learning tasks. However, since many real-world problems involve multivariate time series, permutation entropy needs to be extended to the multivariate case. This work summarises the behaviour of pooled permutation entropy (PPE), multivariate multi-scale permutation entropy (MMSPE), and multivariate weighted permutation entropy (MWPE) on multivariate fractional Brownian motion, and this work fills the gaps in existing research. In addition, we provide a new study of multivariate ordinal pattern permutation entropy (MOPPE) on multivariate fractional Brownian motion. We conclude with a detailed experimental evaluation and comparison between all multivariate extensions, for example, demonstrating identical behaviour of PPE and MMSPE or uncovering different aspects such as amplitude and cross-correlations by using MWPE and MOPPE, respectively.

1 Introduction

This paper is an extension of two works originally presented in the *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2020 (APSIPA ASC)* [1] and in the *International Florida Artificial Intelligence Society Conference Proceedings (FLAIRS-34)* [2]. Both papers study the qualitative behaviour of multivariate fractional Brownian motion by analysing the distribution of up and down encodings in terms of ordinal patterns and permutation entropy.

Time series data is part of many real-world applications, e.g., weather records, patient health evolution metrics, (industrial) sensor data, stock prices, website clicks, server metrics, network data, political or sociological indicators, video or voice messages – just to name a few. Time series can be understood as a series of (real-valued) random variables of a time-dependent system that changes over time, also called a dynamical system. For many real-world applications, modelling time-dependent dynamical systems require specific properties such as *long-range dependence* or *self-similarity*. For example, the long memory is fundamental in internet traffic [3] or for financial data [4]. If financial return series exhibit a long-range dependence, this indicates that observed returns over time are not independent. When returns are not independent, past returns may help predict future returns [5]. This work focuses on a special class of dynamical systems or stochastic processes that describe long-range dependence and self-similarity properties, called fractional Brownian motion (fBm) and denoted as $B_H(t)$, where the so-called Hurst parameter H describes the self-similarity.

While classical time series analysis examines the values of the time series themselves, symbolic time series analysis considers a non-parametric mapping or encoding into a sequence of symbols. This is particularly promising in the case of dynamical systems such as fBm, as the overall dynamics of the generating system is uncovered [6]. To measure the qualitative behaviour of a generating system, efficient measures of the sequence of symbols to a set of scalar-valued representations, also called features, that capture

[*]Corresponding Author: Marisa Mohr, mohr@ifis.uni-luebeck.de

specific properties are considered. Entropy, a measure that has its origins in information theory, is promising through an encoding that preserves the information content of an object [7]. PE is a robust, scalar-valued measure for determining the degree of complexity of dynamical systems or observed time series by analysing the distribution of *symbolic sequences of ordinal patterns* [8]. While the associated entropy is low for a deterministic time series, it approaches its maximum value in case of uncorrelated randomness or high complexity. To distinguish between randomness and complexity, an additional measure is necessary. As a complement to permutation entropy (PE), multi-scale permutation entropy (MSPE) captures the complexity of time series on different time scales [9]. On higher time scales, random noise tends to cancel out, resulting in a low entropy value, where complex signals retain a high entropy value. In this manner, it is possible to gain a deeper insight into the randomness and complexity of a system. Another limitation of PE is the inability to distinguish between different patterns of a given motif in amplitudes. In [10], the author introduce weighted permutation entropy (WPE), which takes into account patterns that differ in amplitudes by assigning weights to each extracted pattern.

The behaviour of PE on fBm in the univariate case is well understood [8, 11, 12]. Not only the PE but also the distributions of ordinal patterns of certain lengths yield interesting properties. The distributions of ordinal patterns of lengths (also called order) two and three yield parameter functions that can be used for descriptive purposes such as autocorrelation. The distribution of order four has irrational values and depends on the noise distribution. The resulting closed formulas for computation are used in this work. The distributions of ordinal patterns of lengths greater than four do not yield closed formulas. Thus, a relationship between autocorrelation or spectrum on the one hand and ordinal patterns, on the other hand, does no longer exist [8]. In [13], the author investigate the behaviour of MSPE on fBm, denoted as $B_H(t)$, and show that MSPE of the fractional Gaussian noise, the increment process of fBm, i.e., $X(t) = B_H(t+1) - B_H(t)$, is time-scale invariant. In [2] we investigate the behaviour of WPE on fBm and show that the distribution of ordinal patterns of order $d = 2$ and thus also WPE are independent of the weighting.

Moreover, in many fields of applications, multivariate measurements are performed. Examples of multivariate fractional Brownian motion (mfBm) can be found in economic time series [14], or functional Magnetic Resonance Imaging of several brain regions [15]. This work summarises the behaviour of several multivariate extensions of PE on mfBm. The first original conference paper [1] studies the behaviour of two canonical multivariate extensions of PE, known as pooled permutation entropy (PPE) and multivariate multiscale permutation entropy (MMSPE) of orders $d = 2$ and $d = 3$, on mfBm. The second original conference paper [2] investigates the behaviour of multivariate weighted permutation entropy (MWPE) of orders $d = 2$ and $d = 3$ on mfBm. We extend original works [1] and [2] by

- a **comprehensive review** of all existing extensions to multivariate permutation entropy including code for computing on Github[1] and Python Package Index (PyPI)[2],
- a **complement** in existing theoretical and experimental investigations of PPE, MMSPE and MWPE on mfBm in a variation of its Hurst parameter H, i.e., we fill the gaps by investigating the behaviour of additional orders $d = 4$ and $d = 5$ as well as other parameters such as in particular different time delays, time scales and number of variables of mfBms,
- a **new study** on the behaviour of multivariate ordinal pattern permutation entropy (MOPPE) on mfBm, and
- a **comparison** of all considered multivariate extensions.

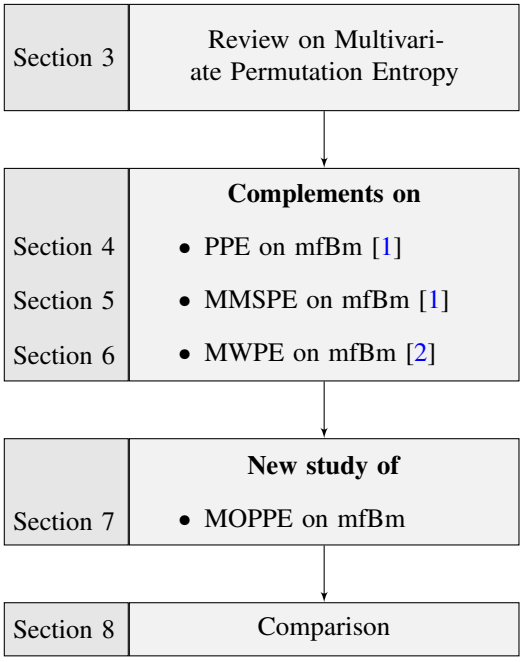


Figure 1: Structure of this research.

This paper is structured as follows and as visualised in Figure 1. After presenting background on (m)fBm and univariate PE in Section 2, needed for the paper to be self-consistent, we provide in Section 3 a review of multivariate extensions of PE and categorise them according to their approach. Since this paper focuses on canonical extensions, we present all those extensions in detail and provide algorithms for their computations. In Section 4, Section 5, and Section 6 we investigate the behaviour of PPE, MMSPE and MWPE, respectively, on mfBm in variation of the Hurst parameter H, and we complement the original work with previously mentioned investigations. In Section 7 we present a new study on the behaviour of MOPPE on mfBm in variation to the Hurst parameter H. In Section 8 we compare our findings and show that PPE and MMSPE behave identically for all investigated orders, delays, scales, and the number of variables. In addition, MWPE and MOPPE of orders $d > 2$ uncover different aspects of mfBm, i.e., amplitudes and cross-correlations, respectively. In Section 9 we conclude our work with mentioning future work.

[1] https://github.com/marisamohr/mpePy
[2] https://pypi.org/project/mpePy

2 Background

We shortly formalise fBm and subsequently key concepts of ordinal patterns and PE. We then summarise known results from the application of PE on fBm in the univariate case.

2.1 Multivariate Fractional Brownian Motion

A stochastic process, or more generally a mathematical object that is similar to itself at all scales, is called a fractal. When you zoom in on a fractal, it resembles or looks exactly like the original shape. The mathematical property is called self-similarity and is expressed by the so-called Hurst parameter $H \in (0, 1)$ in honour of Harold Edwin Hurst [16]. The work described here focuses on a class of special stochastic processes, namely *fractional Brownian motion* (fBm), involving the properties of self-similarity and long-range dependency.

Definition 2.1 (Fractional Brownian Motion) A stochastic process is called *fractional Brownian motion* (fBm) with Hurst parameter $H \in (0, 1)$ and denoted as $B_H(t)$ if it is

1. Gaussian,
2. self-similar, i.e., there exists $H \in (0, 1)$ such that for any $a > 0$ it is
$$B_H(at) \sim a^H B_H(t), \tag{1}$$
where $\sim$ denotes the equality of probability distributions, and it has
3. stationary increments, i.e., $B_H(t) - B_H(s) \sim B_H(t - s)$.

In case $H = 1/2$, fBm corresponds to the ordinary Brownian motion. In case $H > 1/2$, the process has a persistence property and positively correlated increments, i.e., an upward jump is more likely followed by another upward jump and vice versa, and the process exhibits long-range dependence. For $H \to 1$, the process becomes smoother, less irregular and more trendy. In case $H < 1/2$, the process has negatively correlated increments and an anti-persistence property. Figure 2 shows three different realisations of fBms of length $T = 5000$ with different Hurst parameters $H \in \{0.25, 0.5, 0.75\}$ illustrating the properties mentioned above.

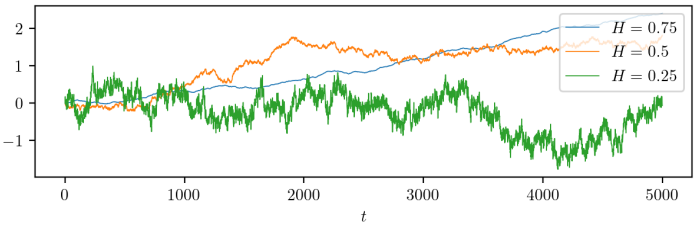


Figure 2: Three different realisations of fBm of length $T = 5000$ with different H.

Stationarity refers to the fact that the distribution of the process doesn't change in time, which has important consequences. In particular, the $B_H(t)$ are identically distributed, i.e., the expectation values and variances of components do not depend on time t. Furthermore, the distribution $B_H(t) - B_H(s)$ depends only on $t - s$, so the correlations of the components also depend only on $t - s$.

Multivariate generalisations of fBms are introduced in [17, 18] and require an extension of self-similarity to multivariate processes first. While the definition given in [19] concerns a far-reaching generalisation of fBm, where self-similarity becomes operator self-similarity for the multivariate case, in [18] the authors restrict themselves to joint self-similarity, where joint self-similarity can be seen as a special case of operator self-similarity where the operator is diagonal. For the sake of simplicity, we focus on the definition introduced in [18].

Definition 2.2 (Multivariate Self-Similar Process) A multivariate process $((X^i(t))_{i=1}^m)_{t\in\mathbb{R}} = (X^1(t), \dots, X^m(t))_{t\in\mathbb{R}}$ with variable-dimension $m \in \mathbb{N}$ is called *self-similar*, if there exists a vector $H = (H_1, \dots, H_m)$ with $H_i \in (0, 1)$ for $i = 1, ..., m$ such that for any $a > 0$ it holds that

$$(X^1(at), \dots, X^m(at))_{t\in\mathbb{R}} \sim (a^{H_1} X^1(t), \dots, a^{H_m} X^m(t))_{t\in\mathbb{R}},$$

where $\sim$ denotes the equality of finite-dimensional distributions.

With Definition 2.2, mfBm is defined as follows.

Definition 2.3 (Multivariate Fractional Brownian Motion [20]) An m-multivariate process $((X^i(t))_{i=1}^m)_{t\in\mathbb{R}}$ is called *multivariate fractional Brownian motion (mfBm)* with self-similarity or Hurst parameter $H = (H_1, \dots, H_m)$ with $H_i \in (0, 1)$ for $i = 1, ..., m$, and is denoted as $\mathbf{B}_H^m(t)$, if it is

1. Gaussian distributed,
2. self-similar with Hurst parameter H, and it has
3. stationary increments, i.e., $\mathbf{B}_H^m(t) - \mathbf{B}_H^m(s) \sim \mathbf{B}_H^m(t - s)$.

Multivariate self-similarity imposes some constraints on the covariance structure of mfBm [18]. The covariance structure is characterised by three parameters $\sigma_i > 0$, $\rho_{ij} \in (-1, 1)$ and $\eta_{ij} \in \mathbb{R}$ for $i, j = 1, ..., m$, which allow two components to be more or less correlated and the process to be reversible in time or not. Parameter $\sigma_i > 0$ is the standard deviation of the i-th variable at time $t = 1$. Parameter $\rho_{ij} = \rho_{ji}$ is the correlation coefficient between the variables i and j at time $t = 1$. Parameters $\eta_{ij} = -\eta_{ji}$ are antisymmetric and linked with the time-reversibility of mfBm. Multivariate fractional Brownian motion can be characterised by its covariances and cross-covariances of its variables as follows.

Lemma 2.1 (Covariance Function of mfBm [21]) *The mfBm $\mathbf{B}_H^m(t)$ is marginally an fBm, such that the covariance function of the i-th variable $B_{H_i}^i$ of mBfm is as in the univariate case, i.e.,*

$$Cov(B_{H_i}^i(s), B_{H_i}^i(t)) = \frac{\sigma_i^2}{2}(|s|^{2H_i} + |t|^{2H_i} - |t - s|^{2H_i}). \tag{2}$$

where $\sigma_i^2 = Var(B_{H_i}^i(1))$. The cross-covariances of mfBm for all $(i, j) \in \{1, ..., m\}^2$ and $i \neq j$ are given by

$$Cov(B_{H_i}^i(s), B_{H_j}^j(t)) = \frac{\sigma_i \sigma_j}{2}(w_{ij}(-s) + w_{ij}(t) - w_{ij}(t - s)) \tag{3}$$

where the function $w_{ij}(h)$ is defined by

$$w_{ij}(h) = \begin{cases} (\rho_{ij} - \eta_{ij}\,\mathrm{sign}(h))|h|^{H_i+H_j} & \text{if } H_i + H_j \neq 1, \\ \rho_{ij}|h| + \eta_{ij} h \log|h| & \text{if } H_i + H_j = 1. \end{cases} \tag{4}$$

Moreover, the setting of $\rho_{ij} = 1$ and $\eta_{ij} = 0$ in (3) and (4) matches with the definition in the univariate case in (2). For $m = 1$, Definition 2.3 matches with the univariate Definition 2.1 of fBm. Figure 3 shows three realisations of mfBm of length $T = 2000$, $\rho_{i,j} = 0.3$, $\eta_{i,j} = 0.1/(1 - H_i - H_j)$ with different Hurst parameters.

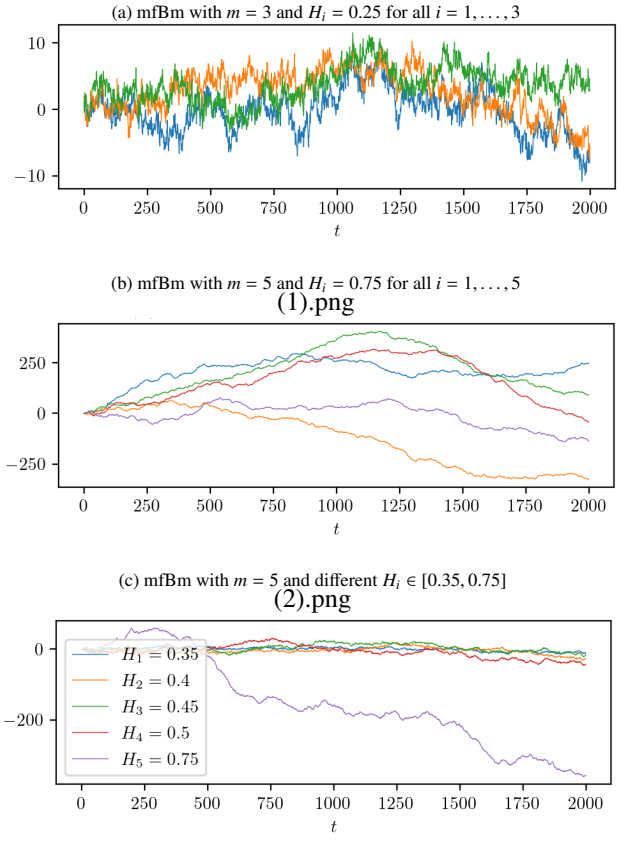


Figure 3: Three different realisations of mfBm with different Hurst parameters.

2.2 Ordinal Pattern Representations

To investigate the qualitative behaviour of mfBm, we use ordinal pattern symbols. Before we get into the details of multivariate ordinal patterns, we introduce univariate concepts.

2.2.1 Ordinal Pattern Symbolisation

In the mathematical theory of symbolic dynamics, a dynamical system is modelled by a discrete space consisting of infinite (sequences of) abstract symbols. These sequences of symbols are the subject of advanced analysis such as entropies. As far as current research is concerned, there are two general approaches to encode the sequence of real-valued measurements into a sequence of symbols as visualised in Figure 4.

On the one hand, classical symbolisation approaches partition the data range according to specified mapping rules in order to then encode a numerical time series into a sequence of discrete symbols from a predefined alphabet Σ. An example of the partitioning of a data range can be found in Figure 4a, while the assignment of the symbols (encoding) to the time series is visualised in Figure 4b. A corresponding and well-known algorithm for determining data range partitions is Symbolic Aggregate ApproXimation (SAX) presented in [22]. On the other hand, ordinal pattern symbolisation is another approach that, independent of the data range of the time series, encodes the total order between two or more neighbours ($x < y$ or $x > y$) into symbols. This ordinal pattern approach is based on the idea of Bandt and Pompe [23] and visualised in Figures 4c and 4d. Since we use the ordinal symbolisation scheme in the remainder of this work, we present its formalism and advantages in detail in the following.

Ordinal patterns describe the total order between two or more neighbours encoded by permutations.

Definition 2.4 (Univariate Ordinal Pattern)
A vector $(x_1, \ldots, x_d) \in \mathbb{R}^d$ has *ordinal pattern* $(r_1, \ldots, r_d) \in \mathbb{N}^d$ of order $d \in \mathbb{N}$ if $x_{r_1} \geq \cdots \geq x_{r_d}$ and $r_{l-1} > r_l$ in the case $x_{r_{l-1}} = x_{r_l}$.

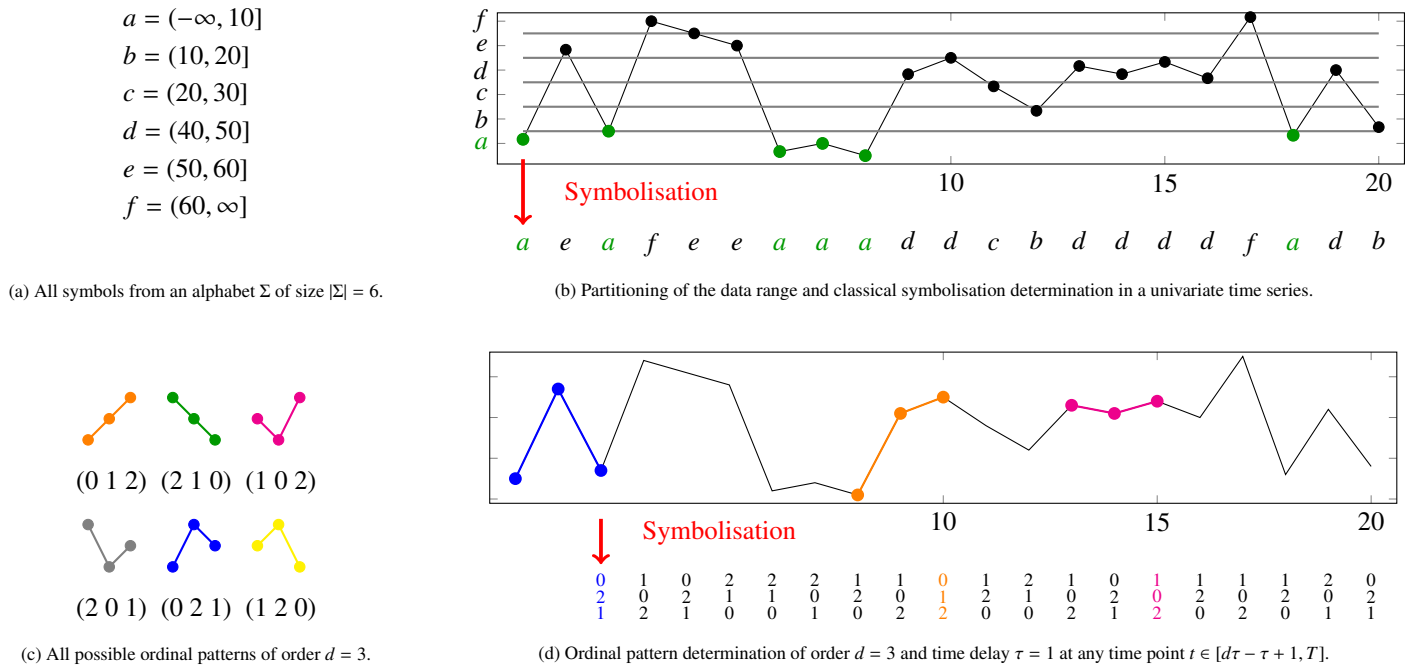
Note that equality of two values within a pattern is not allowed. In this case, without loss of generality, the newer value is replaced with a smaller value. Figure 4c shows all possible ordinal patterns of order $d = 3$ of a vector (x_1, x_2, x_3). To symbolise a time series $(x_1, x_2, ..., x_T) \in \mathbb{R}^T$ each point in time $t \in \{d, ..., T\}$ is assigned its ordinal pattern of order d. The order d is chosen to be much smaller than the total length T of the time series to look at smaller ordinal pattern windows within the series and their distributions of "up and down" movements. To assess the overarching trend, delayed behaviour is of interest. The time delay $\tau \in \mathbb{N}$ is the delay between successive points in the symbol sequences. Different delays show different details of the structure of a time series. Figure 4d visualises ordinal pattern determination of order $d = 3$ and time delay $\tau = 1$ of three different time points in a time series marked in blue, orange and magenta. Note that ordinal patterns are determined at any arbitrary point in time.

The ordinal approach has notable advantages in its practical application. First of all, the method is conceptually simple as it reflects man's natural thought of up and down movements and is, therefore, open to interpretation. Second, prior knowledge of the data range or the type of time series is not necessary. The concept can be applied to any time series as long as the range of values is ordered, e.g. $x_t \in \mathbb{R}$. Third, the ordinal approach supports robust and fast implementations [24, 25]. Fourth, it allows for an easier estimation of a good symbolisation scheme compared to the classical symbolisation approaches [26, 27].

2.2.2 Permutation Entropy

Not the ordinal patterns themselves, but their distributions in a *univariate time series* $(x_t)_{t=1}^T$ are of interest. Thus, each pattern is identified with exactly one of the *ordinal pattern symbols* $j = 1, 2, ..., d!$. To assess the (dis)order of the identified symbols in the system or in the time series, we use a measure of the dispersion from the field of statistics, namely entropy. The greater the disorder, the higher the entropy. The statistical interpretation of entropy corresponds to Shannon's information entropy used in information theory or computer science. Applying the well-known formula for (Shannon) entropy

$$I(Z) = -\sum_{z \in Z} p_z \ln p_z, \tag{5}$$



(a) All symbols from an alphabet Σ of size $|\Sigma| = 6$.

(b) Partitioning of the data range and classical symbolisation determination in a univariate time series.

(c) All possible ordinal patterns of order $d = 3$.

(d) Ordinal pattern determination of order $d = 3$ and time delay $\tau = 1$ at any time point $t \in [d\tau - \tau + 1, T]$.

Figure 4: Two approaches for symbolising a univariate time series: (a) and (b) classical symbolisation and (c) and (d) ordinal pattern symbolisation. Best viewed in colour.

where Z is an alphabet of symbols $Z = \{z_1, z_2, \ldots, z_n\}$ to ordinal pattern symbols leads directly to the definition of permutation entropy.

Definition 2.5 (Permutation Entropy [23])
Permutation entropy (PE) of order $d \in \mathbb{N}$ and delay $\tau \in \mathbb{N}$ of a univariate time series $\mathbf{x} = (x_t)_{t=1}^{T}$, $T \in \mathbb{N}$ is defined by

$$\mathrm{PE}_{d,\tau}(\mathbf{x}) = -\sum_{j=1}^{d!} p_j^{\tau,d} \ln p_j^{\tau,d}, \tag{6}$$

where $p_j^{\tau,d}$ is the frequency of occurrence of ordinal pattern j in the time series.

In time series with maximally random ordinal pattern symbols, the ordinal patterns are equally distributed so that PE is $\ln(d!)$. For a time series with a regular pattern, e.g., in the case of strict monotony, PE is equal to zero [7].

2.2.3 *Multi-Scale-Permutation Entropy*

For example, biological and physiological time series, as well as time series from other fields, often contain complex correlations between both temporal and spatial levels (scales) that need to be uncovered. PE achieves a maximum entropy value both on completely random time series and on series with complex correlations based on a determinism. To uncover multi-scale complex correlations, multi-scale entropy (MSE) proposed in [28] provides a systematic procedure to assign small complexity values to both fully predictable and fully random uncorrelated time series. In contrast, correlated processes across different scales have a high complexity value. In [9], the author extend the concept of MSE to univariate ordinal patterns.

For the consideration of different scales of a time series and an associated definition of MSPE, a coarse-grained procedure is used: From the original time series, several consecutive time data points are averaged within a non-overlapping time window of scaling length ϵ, also called *scaling factor*. Each element of the coarse-grained time series $\mathbf{y} = (y_l^{(\epsilon)})_{l=1}^{T/\epsilon}$ is calculated as:

$$y_l^{(\epsilon)} = \frac{1}{\epsilon} \sum_{t=(l-1)\epsilon+1}^{l\epsilon} x_t \tag{7}$$

for $1 \leq l \leq \frac{T}{\epsilon}$.

Definition 2.6 (Multi-Scale Permutation Entropy [9])
Multi-scale permutation entropy (MSPE) of order $d \in \mathbb{N}$ and delay $\tau \in \mathbb{N}$ of a univariate time series $\mathbf{x} = (x_t)_{t=1}^{T}$, $T \in \mathbb{N}$ is defined as PE of its coarse-grained time series $\mathbf{y} = (y_l^{(\epsilon)})_{l=1}^{T/\epsilon}$, that is

$$\mathrm{MSPE}_{d,\tau,\epsilon}(\mathbf{x}) = \mathrm{PE}_{d,\tau}(\mathbf{y}). \tag{8}$$

2.2.4 *Weighted Permutation Entropy*

Another shortcoming in the above Definition 2.5 of PE is that when the ordinal patterns are extracted, no information other than the order structure is preserved. In particular, information about the amplitude in a time series is lost. However, ordinal patterns with large differences in amplitude should contribute in different ways to the computation of PE. The weighted permutation entropy introduced in [10] allows for the weighting of ordinal patterns by exploiting amplitude information resulting from small fluctuations in the time series due to the effect of noise to be weighted less than ordinal patterns with large amplitudes.

Definition 2.7 (Weighted Permutation Entropy [10])
Weighted permutation entropy (WPE) of order $d \in \mathbb{N}$ and delay

$\tau \in \mathbb{N}$ of a univariate time series $\mathbf{x} = (x_t)_{t=1}^T$, $T \in \mathbb{N}$ is defined as

$$\mathrm{WPE}_{d,\tau}(\mathbf{x}) = -\sum_{j}^{d!} p_{w_j}^{\tau,d} \ln p_{w_j}^{\tau,d} \tag{9}$$

with

$$p_{w_j}^{\tau,d} = \frac{\sum_{t\leq T} w_t \cdot [(x_{t-(d-1)\tau}, ..., x_{t-\tau}, x_t) \text{ has pattern } j]}{\sum_{t\leq T} w_t \cdot [(x_{t-(d-1)\tau}, ..., x_{t-\tau}, x_t)]} \tag{10}$$

where $[x] = 1$ if x is true, and 0 otherwise, and

$$w_t = \frac{1}{d}\sum_{k=1}^{d}(x_{t-(k-1)\tau} - \overline{x}_t^{d,\tau})^2 \tag{11}$$

is the empirical variance of the sub-sequence and $\overline{x}_t^{d,\tau}$ denotes its arithmetic mean, i.e., $\overline{x}_t^{d,\tau} = \frac{1}{d}\sum_{k=1}^{d} x_{t-(k-1)\tau}$.

2.3 On the Behaviour of Permutation Entropy on Fractional Brownian Motion

This work investigates the behaviour of PE on fBm in a multivariate setting. To this purpose, we recapitulate univariate results from [8]. They investigate the distribution of ordinal patterns in fBm of different orders and, if possible, provide closed formulas for computation of pattern distributions as follows:

Order $d = 2$. The ordinal patterns are equally distributed, more specifically

$$p_{01}^{\tau} = p_{10}^{\tau} = 1/2, \tag{12}$$

such that $\mathrm{PE}_{2,\tau}(B_H(t)) = -\ln(1/2)$ for all τ.

Order $d = 3$. The distribution of ordinal patterns is given by

$$p_{012}^{\tau} = \frac{1}{\pi}\arcsin 2^{H-1} =: u \tag{13}$$

for all τ. Furthermore, for a Gaussian process with stationary increments (as fBm), it holds that

$$p_j^{\tau} = \begin{cases} u & \text{if } j = (012), (210), \\ (1-2u)/4 & \text{otherwise,} \end{cases} \tag{14}$$

for all τ. In particular, $\mathrm{PE}_{3,\tau}(B_H(t))$ is independent of the delay τ but monotonically dependent on the Hurst parameter H.

Order $d = 4$. The distribution of ordinal patterns can also be expressed, albeit in a more complex formula

$$p_j^{\tau} = \frac{1}{8} + \frac{1}{4\pi}\cdot v_j \tag{15}$$

for all delays τ, where

$$v_j = \begin{cases} \arcsin\alpha_1 + 2\arcsin\alpha_2 & \text{if } j = (0123), (3210), \\ \arcsin\alpha_4 - 2\arcsin\alpha_5 & \text{if } j = (3120), (0213), \\ 2\arcsin\alpha_6 + \arcsin\alpha_1 & \text{if } j = (1032), (2301), \\ \arcsin\alpha_7 - \arcsin\alpha_1 - \arcsin\alpha_5 & \text{if } j = (0132), (1023), \\ & \quad (2310), (3201), \\ \arcsin\alpha_7 - \arcsin\alpha_4 - \arcsin\alpha_5 & \text{if } j = (0312), (3021), \\ & \quad (2130), (1203), \\ \arcsin\alpha_3 + \arcsin\alpha_8 - \arcsin\alpha_5 & \text{if } j = (2013), (0231), \\ & \quad (3102), (1320), \\ \arcsin\alpha_6 - \arcsin\alpha_8 + \arcsin\alpha_2 & \text{if } j = (0321), (3012), \\ & \quad (1230), (2103), \end{cases} \tag{16}$$

with

$$\alpha_1 = \frac{1+3^{2H}-2^{2H+1}}{2}, \quad \alpha_2 = 2^{2H-1}-1, \quad \alpha_3 = \frac{1-3^{2H}-2^{2H}}{2\cdot 6^H},$$
$$\alpha_4 = \frac{3^{2H}-1}{2^{2H+1}}, \quad \alpha_5 = 2^{2H-1}, \quad \alpha_6 = \frac{-1-3^{2H}+2^{2H}}{2\cdot 3^H},$$
$$\alpha_7 = \frac{3^{2H}-2^{2H}-1}{2^{2H+1}}, \quad \text{and} \quad \alpha_8 = \frac{2^{2H-1}}{3^H}.$$

In particular, $\mathrm{PE}_{4,\tau}(B_H(t))$ is independent of the delay τ but monotonically dependent on the Hurst parameter H.

Order $d \geq 5$. There are no closed formulas.

The formulas for the distribution of univariate order patterns of different orders lead to the following behaviour of PE and its extensions MSPE and WPE on fBm.

Theorem 2.2 *For order $d = 2$, it holds*

$$\mathrm{PE}(B_H(t)) = -\ln(1/2) \tag{17}$$
$$= \mathrm{MSPE}(B_H(t)) \tag{18}$$
$$= \mathrm{WPE}(B_H(t)). \tag{19}$$

for all delays $\tau \in \mathbb{N}$ and all scales ϵ.

Proof. Equation (17) follows directly from (12). Equations (18) or (19) are shown in [1] and [2], respectively. The independence of MSPE from the scale ϵ follows from the invariance of the coarse-grained procedure of fBm. □

Theorem 2.3 *For orders $d = 3$ and $d = 4$, $\mathrm{PE}(B_H(t))$ and $\mathrm{MSPE}(B_H(t))$ are independent of all delays $\tau \in \mathbb{N}$ but monotonically dependent on the Hurst parameter H. Moreover, $\mathrm{MSPE}(B_H(t))$ is independent of all scales ϵ, and thus $\mathrm{PE}(B_H(t))$ and $\mathrm{MSPE}(B_H(t))$ are identical.*

Proof. The independence of MSPE from the scale ϵ follows from the invariance of the coarse grain procedure of fBm and is shown in [1]. Thus, independence of the delay τ and the dependence on the Hurst parameter H follows directly from (13). □

In [2] we observe that for WPE of order $d = 3$, certain ordinal patterns, namely the strictly ascending ordinal pattern (012) and the strictly descending pattern (210), have higher weights and thus have more impact on the computation of WPE than the other four ordinal patterns. As a result, WPE decreases faster than PE for increasing Hurst parameter H. We elaborate on this experimental discovery in Section 6.

3 A Review of Multivariate Permutation Entropy

This section provides a comprehensive overview of existing extensions of permutation entropy to the multivariate case. We categorise existing multivariate extensions according to their procedure before presenting them in detail.

3.1 Different Approaches for Multivariate Permutation Entropy

A *multivariate time series* $((x_t^i)_{i=1}^m)_{t=1}^T$ such as a certain realisation of mfBm has more than one time dependent variable. Each variable x^i for $i \in 1, ..., m$ not only depends on the respective past values in time but also has some dependence on other variables in phase space. Considering two time points $(x_t^i)_{i=1}^m$ and $(x_{t+1}^i)_{i=1}^m$ with m variables, simply put two vectors, it is not possible to establish a total order between them. A total order is only possible if $x_t^i > x_{t+1}^i$ or $x_t^i < x_{t+1}^i$ for all $i \in 1, ..., m$. Therefore, there is no trivial generalisation of the PE algorithm to the multivariate case.

Nevertheless, numerous studies deal with the multivariate extensions of PE. We classify proposed extensions into four procedures. The multivariate time series is

a) projected into univariate ordinal space by determining univariate ordinal pattern *between neighbouring values in time space* (row marked in red in Figure 5a) in each single variable i , and then pooled over all m variables for multidimensionality, or

b) projected into univariate ordinal space by determining univariate ordinal pattern *between values of all m variables* (column marked in red in Figure 5b) for each single time point t, and then pooled over all T time points for multidimensionality, or

c) projected into multivariate ordinal space by determining the multivariate ordinal pattern *between d vectors of variable dimension m* (box marked in blue in Figure 5c for $d = 2$), or

d) projected onto a single-dimensional reduction first (box marked in green in Figure 5d), and then transformed into ordinal space by determining univariate ordinal pattern.

The first procedure a) is a canonical extension of the univariate definition and is presented in [29] as pooled permutation entropy (PPE). PPE is probably the most widely used multivariate extension. While PPE measures the complexity of each variable in time space, procedure b) analyses the complexity of the variables in phase space. A variant of procedure b) is introduced in [30] as Multivariate Permutation Entropy (MvPE). The third procedure c) first requires an extension of univariate ordinal pattern to multidimensionality. Some theoretical basis for this is set in [31, 32]. A detailed discussion and challenge on real-world data can be found in [33]. The idea of the fourth procedure d) is to first reduce the number of variables m to a single dimension, i.e., $m = 1$, by applying a distance measure or dimension reduction algorithm. Consequently, Definition 2.5 can then be used for PE computation directly. Rayan, i n[34], the author propose several distance measures to reduce the dimensionality, that is Euclidian distance with reference point $(x_{t=0}^i)_{i=1}^m$, Manhattan distance with reference point $(x_{t=0}^i)_{i=1}^m$, and Euclidian distance with reference point $\mathbf{0}$. To consider correlations between variables in space, in [33], the author propose to reduce the dimension using principal component analysis.

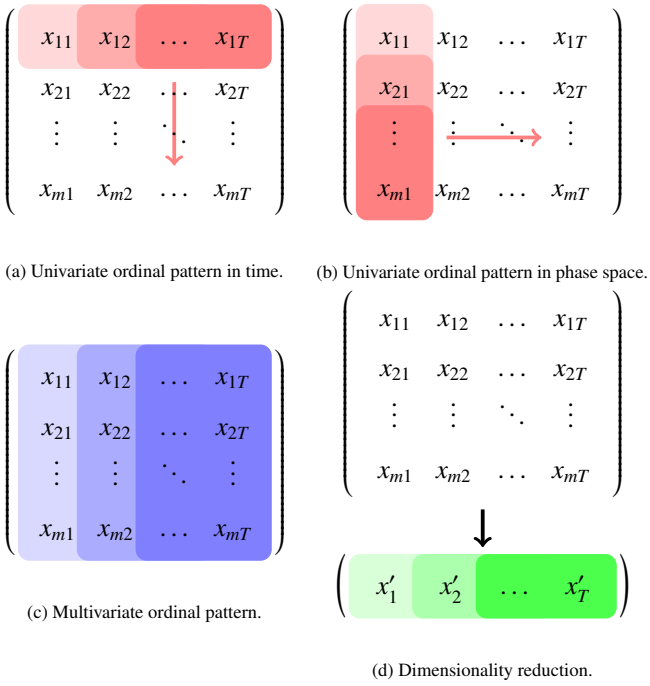


Figure 5: Four procedures of ordinal pattern determination in a multivariate time series.

Since the determination of ordinal patterns in both procedures b) and d) requires transformations of the structure of mfBms, we restrict ourselves in the following to extensions a) and c) that are some kind of canonical extensions.

3.2 Pooled Permutation Entropy and Adaptions

The procedure in Figure 5a is a canonical extension of PE to the multivariate case, i.e., univariate ordinal patterns are determined for each variable of the multivariate time series before each pattern of all variables is pooled. This idea is introduced in [29] and referred to as *pooled permutation entropy* (PPE). In addition, we introduce two adaptations based on PPE that take into account different aspects of a time series, i.e., different scales or amplitudes, called *multivariate multi-scale permutation entropy* (MMSPE) or *multivariate weighted permutation entropy* (MWPE), respectively.

3.2.1 Pooled Permutation Entropy

The idea of PPE is to use marginal frequencies of $d!$ ordinal patterns regarding all m variables as input for entropy computation. For the determination of marginal frequencies, an auxiliary matrix has to be established first as follows.

1. For each variable $i = 1, .., m$ and for each ordinal pattern $j = 1, ..., d!$, count all time steps $s \in [\tau(d-1)+1, T]$, for which the variable-time pair (i, s) has the ordinal pattern j.

2. Divide the counts by $m \cdot \delta$, where $\delta := T - \tau(d-1)$ is the total count of ordinal patterns each variable has.

3. Store the results, i.e., frequencies $p_{ij}^{\tau,d}$ in a so-called *pooling matrix* $P \in (0,1)^{m \times d!}$, which reflects the distribution of ordinal patterns in the multivariate time series across its m variables.

It holds $\sum_{i=1}^{m} \sum_{j=1}^{d!} p_{ij}^{\tau,d} = 1$. For computational reasons, marginal frequencies $p_{.j}^{\tau,d} = \sum_{i=1}^{m} p_{ij}^{\tau,d}$ must not vanish for $j = 1, ..., d!$. If they vanish, set the value close to zero.

Definition 3.1 (Pooled Permutation Entropy [29])
Pooled permutation entropy (PPE) of a multivariate time series $\mathbf{X} = ((x_t^i)_{i=1}^m)_{t=1}^T$ is defined as PE of the marginal frequencies $p_{.j}^{\tau,d} = \sum_{i=1}^{m} p_{ij}^{\tau,d}$ for $j = 1, ..., d!$ describing the distribution of the ordinal pattern and can be calculated as

$$\mathrm{PPE}_{d,\tau}(\mathbf{X}) = -\sum_{j}^{d!} p_{.j}^{\tau,d} \ln p_{.j}^{\tau,d}. \tag{20}$$

In the literature, PPE is often referred to as Multivariate Permutation Entropy. To avoid confusion with other multivariate versions, we use the original naming. A major advantage of PPE is that within the procedure also univariate PE can be derived for each variable by calculating the PE based on the frequencies p_{ij} of all ordinal patterns $j = 1, \ldots, d!$ for corresponding $i = 1, \ldots, m$, i.e. row by row on the pooling matrix. Algorithm 1 provides pseudocode for computing PPE.

Algorithm 1: Computation of PPE

```
Input: Multivariate Time Series X^{m×T}, Order d, Delay τ
Function pooling(X,d,τ):
    P^{m×d!} ← pooling matrix initialised with zeros
    for every time series variable i = 1,...,m in X do
        for each ordinal pattern j = 1,...,d! do
            c ← number of time steps s ∈ [τ(d−1)+1, T]
              with pattern j
            P_ij ← divide c by the number of total time steps
              δ · m in the multivariate time series X
    return P
Function marginalisation(P):
    p^{1×d!} ← vector of marginalisation initialised with zeros
    for every column j = 1,...,d! in P^{m×d!} do
        p_j ← sum up p_ij
    return p
return PE(p)
```

For example, PPE is successfully used in the analysis of electroencephalography (EEG) signals, as cross-channel regularities between spatially distant variables, i.e., on different hemispheres or in different areas, can be extracted by long-range spatial nonlinear correlations [29]. Furthermore, in [24], the author successfully used PPE to study the 19-channel scalp EEG reflecting changes in brain dynamics of a boy with lesions predominantly in the left temporal lobe caused by connatal toxoplasmosis. In [35], the author use PPE to characterise sleep EEG signals from more than 80 hours of nocturnal sleep recordings and classify sleep stages. They find that each sleep stage is characterised by statistically different PPE values and that the observed pattern of PPE is consistent with the physiological properties of the EEG in each sleep stage.

3.2.2 Multivariate Multi-Scale Permutation Entropy

In addition to PPE and by analogy with MSPE, Morabito et al. [9] provide a canonical definition of multivariate multi-scale permutation entropy. MSPE and MMSPE are calculated on different time scales by processing the coarse-grained time series as a function of the scale factor ϵ. Per variable i in a multivariate time series, several consecutive time data points are averaged within a non-overlapping time window of the scaling length ϵ. Each element of the coarse-grained time series $\mathbf{Y} = ((y_{i,l}^{(\epsilon)})_{l=1}^{T/\epsilon})_{i=1}^{m}$ is calculated as:

$$y_{i,l}^{(\epsilon)} = \frac{1}{\epsilon} \sum_{t=(l-1)\epsilon+1}^{l\epsilon} x_{i,t} \tag{21}$$

for all $i = 1, ..., m$ and $1 \leq l \leq \frac{T}{\epsilon}$.

Definition 3.2 (Multivariate Multi-Scale PE [9])
Multivariate multi-scale permutation entropy (MMSPE) of order $d \in \mathbb{N}$ and delay $\tau \in \mathbb{N}$ of a multivariate time series $\mathbf{X}$ is defined as PPE of its coarse-grained time series $\mathbf{Y}$, that is

$$\mathrm{MMSPE}_{d,\tau,\epsilon}(\mathbf{X}) = \mathrm{PPE}_{d,\tau}(\mathbf{Y}). \tag{22}$$

The simultaneous utilisation of a multi-scale approach and the consideration of multiple variables of the time series facilitates assessing the complexity of the underlying dynamical system. Algorithm 2 provides pseudocode for computing MMSPE.

Algorithm 2: Computation of MMSPE

```
Input: Multivariate Time Series X^{m×T}, Order d, Delay τ,
       Scale Factor ε
Y^{m×(T/ε)} ← coarse-grained multivariate time series
  initialized with zeros
for every time series variable i = 1,...,m in X do
    y_i ← calculate coarse-grained univariate time series
                                        // see Eq.(21)
return PPE(Y, d, τ)
```

3.2.3 Multivariate Weighted Permutation Entropy

In addition to PPE and by analogy with WPE, in [2] we provide a canonical definition of multivariate weighted permutation entropy. Again, for the determination of MWPE, an auxiliary matrix has to be established first:

1. For each variable $i = 1, .., m$ and for each ordinal pattern $j = 1, ..., d!$, select all time steps $s \in [\tau(d-1)+1, T]$ for which the variable-time pair (i, s) has the ordinal pattern j.

2. Add up the weights w_t, i.e., $\mathbf{w_{ij}} = \sum_{t=d\tau-\tau+1}^{T} w_t$ for all selected ordinal pattern vectors j and for each variable $i = 1, ..., m$. Note that the total count of weights w_t^i for each variable i is $\delta := T - (d\tau - \tau)$.

3. Divide the weighted sum $\mathbf{w_{ij}}$ by the total sum of all $m \cdot \delta$ weights to obtain the weighted frequencies for every ordinal pattern j.
4. Store the results, i.e., weighted frequencies $p_{w_{ij}}^{\tau,d}$ in a so-called *weighted pooling matrix* $P_w^{\tau,d} \in \mathbb{R}^{m \times d!}$, which reflects the weighted distribution of ordinal patterns in the multivariate time series across its m variables.

Based on the weighted pooling matrix $P_w^{\tau,d}$, multivariate weighted permutation entropy can be calculated as follows.

Definition 3.3 (Multivariate Weighted PE [2])
Multivariate weighted permutation entropy (MWPE) of a multivariate time series $\mathbf{X} = ((x_t^i)_{i=1}^m)_{t=1}^T$ is defined as PE of the marginal weighted frequencies $p_{w_{\cdot j}}^{\tau,d} = \sum_{i=1}^m p_{w_{ij}}^{\tau,d}$ for $j = 1, ..., d!$ describing the distribution of the weighted ordinal pattern and can be calculated as

$$\mathrm{MWPE}_{d,\tau}(\mathbf{X}) = -\sum_{j}^{d!} p_{w_{\cdot j}}^{\tau,d} \ln p_{w_{\cdot j}}^{\tau,d}. \tag{23}$$

Algorithm 3 provides pseudocode for computing MWPE.

Algorithm 3: Computation of MWPE

Input: Multivariate Time Series $X^{m \times T}$, Order d, Delay τ
Function `weightedPooling(X,d,τ):`
  $P^{m \times d!} \leftarrow$ weighted pooling matrix initialised with zeros
  for *every time series variable* $i = 1, \ldots, m$ *in* X **do**
    for *each ordinal pattern* $j = 1, \ldots, d!$ **do**
      $w_{ij} \leftarrow$ add up weights w_t `// see Eq.(11)`
      $p_{ij} \leftarrow$ divide w_{ij} by the total sum of all $m \cdot \delta$ weights
  return P
$p \leftarrow$ `marginalisation(P)`
return PE(p)

3.3 Multivariate Ordinal Pattern Permutation Entropy

A *multivariate* time series $((x_t^i)_{i=1}^m)_{t=1}^T$ and its corresponding data matrix $X \in \mathbb{R}^{m \times T}$ has more than one time-dependent variable. Each variable x^i for $i \in 1, ..., m$ not only depends on the respective past values in time but also has some dependence on other variables in phase space. Since all procedures from the previous section consider univariate ordinal patterns per variable i separately, interdependence is not considered. To involve interdependence, an intuitive idea introduced in [33] is to store the univariate ordinal patterns of all variables at a time point t together into one common multivariate ordinal pattern.

Definition 3.4 (Multivariate Ordinal Pattern) A data matrix $(x_1, ..., x_d) \in \mathbb{R}^{m \times d}$ has *multivariate ordinal pattern*

$$\begin{pmatrix} r_{11} & \cdots & r_{1d} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{md} \end{pmatrix} \in \mathbb{N}^{m \times d} \tag{24}$$

of order $d \in \mathbb{N}$ if $x_{r_{i1}} \geq ... \geq x_{r_{id}}$ for all $i = 1, ..., m$ and $r_{il-1} > r_{il}$ in the case $x_{r_{il-1}} = x_{r_{il}}$.

Figure 6 shows all possible multivariate ordinal patterns of order $d = 3$ and number of variables $m = 2$.

$$\begin{matrix} \binom{0\,1\,2}{0\,1\,2} & \binom{0\,1\,2}{2\,1\,0} & \binom{0\,1\,2}{1\,0\,2} & \cdots & \binom{0\,1\,2}{1\,2\,0} \\ \binom{2\,1\,0}{0\,1\,2} & \binom{2\,1\,0}{2\,1\,0} & \binom{2\,1\,0}{1\,0\,2} & \cdots & \binom{2\,1\,0}{1\,2\,0} \\ \binom{1\,0\,2}{0\,1\,2} & \binom{1\,0\,2}{2\,1\,0} & \binom{1\,0\,2}{1\,0\,2} & \cdots & \binom{1\,0\,2}{1\,2\,0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{1\,2\,0}{0\,1\,2} & \binom{1\,2\,0}{2\,1\,0} & \binom{1\,2\,0}{1\,0\,2} & \cdots & \binom{1\,2\,0}{1\,2\,0} \end{matrix}$$

Figure 6: All possible multivariate ordinal patterns of order $d = 3$ with $m = 2$ variables. Best viewed in colour.

With the natural extension of univariate ordinal pattern, which leads to Definition 3.4, it is possible to apply the PE algorithm from Definition 2.5 to multivariate time series in its original form.

Definition 3.5 (Multivariate Ordinal Pattern PE [33])
Multivariate ordinal pattern permutation entropy (MOPPE) of order $d \in \mathbb{N}$ and delay $\tau \in \mathbb{N}$ of a multivariate time series $\mathbf{X} = ((x_t^i)_{i=1}^m)_{t=1}^T$ is defined by

$$\mathrm{MOPPE}_{d,\tau}(\mathbf{X}) = -\sum_{j=1}^{d!} p_j^{\tau,d} \ln p_j^{\tau,d}, \tag{25}$$

where $p_j^{\tau,d}$ is the frequency of multivariate ordinal pattern j in the multivariate time series.

Algorithms 1, 2, 3 and 4 for computing PPE,MMSPE, MWPE and MOPPE, respectively, can also be found on Github[1] and PyPI[2].

Algorithm 4: Computation of MOPPE

Input: Multivariate Time Series $X^{m \times T}$, Order d, Delay τ
$p^{1 \times d!} \leftarrow$ frequencies initialised with zeros
for *every time step* t *in* X **do**
  for *each ordinal pattern* $j = 1, \ldots, d!$ **do**
    $c \leftarrow$ count all time points t with multivariate ordinal pattern j `// see Def. 3.4`
    $p_j \leftarrow$ divide c by the number of total time steps $T - \tau(d-1)$ in the multivariate time series X
return PE(p)

The number of the possible multivariate ordinal pattern increases exponentially with the number of variables m, i.e., $(d!)^m$. Therefore, if d and m are too large, depending on the application, each pattern occurs only rarely or some not at all, resulting in a uniform distribution of ordinal patterns. This has the consequence that subsequent learning procedures may fail. Nevertheless, for small order d and sufficiently large length T of the time series, the use of multivariate ordinal patterns can lead to higher accuracy in learning tasks, e.g., classification [33], because they incorporate interdependence.

4 PPE Applied to mfBm

In this section, we investigate the behaviour of PPE of different orders and delays on mfBm in the variation of its Hurst parameter from both theoretical and experimental points of view. For the sake of completeness, we recapitulate results from previous work and fill in gaps with further results.

4.1 Theoretical Analysis

Orders 2 **and** 3**.** In previous work, [1] we show that for order $d = 2$ the PPE of mfBm is constant for all delays τ as well as any number of variables m, in particular, it holds that

$$\mathrm{PPE}_{2,\tau}(\mathbf{B}_H^m(t)) = -\ln(1/2) \tag{26}$$

for all τ, m. Furthermore, we show that for order $d = 3$, the PPE of mfBm stays also independent of the delay τ, but depends monotonically on the number of variables m as well as the Hurst parameter H. In the case $H_i = H_j$ for all variables i, j, PPE is also independent of the number of variables m. Details and corresponding formulas for entropy calculation can be found in the original paper [1].

Orders greater than 3**.** In this work, we show that a behaviour analogous to that of PPE with order $d = 3$ can be transferred to orders $d > 3$.

Theorem 4.1 $\mathrm{PPE}(\mathbf{B}_H^m(t))$ *of order* $d = 4$ *is independent of all delays* τ *but monotonically dependent on the number of variables* m *and the Hurst parameter* H *of mfBm. In the case* $H_i = H_j$ *for all* i, j*, it is independent on the number of variables* m*.*

Proof. Let $d = 4$. Since mfBm $\mathbf{B}_H^m(t)$ is marginally an fBm $B_H(t)$, we conclude from Definition 2.3 and (15), that

$$p_{ij}^{\tau,4} = \frac{1}{m}\left(\frac{1}{8} + \frac{1}{4\pi}\cdot v_j\right) \tag{27}$$

for all $i = 1, ..., m$ and v_j as in (16).

The marginal frequencies for $j = 1, ..., d!$ are given by

$$p_{.j}^{\tau,4} = \sum_{i=1}^{m} \frac{1}{m}\left(\frac{1}{8} + \frac{1}{4\pi}\cdot v_j\right), \tag{28}$$

where v_j is as in (16), while

$$\alpha_1 = \frac{1 + 3^{2H_i} - 2^{2H_i+1}}{2}, \quad \alpha_2 = 2^{2H_i-1} - 1, \quad \alpha_3 = \frac{1 - 3^{2H_i} - 2^{2H_i}}{2\cdot 6^{H_i}},$$
$$\alpha_4 = \frac{3^{2H_i} - 1}{2^{2H_i+1}}, \quad \alpha_5 = 2^{2H_i-1}, \quad \alpha_6 = \frac{-1 - 3^{2H_i} + 2^{2H_i}}{2\cdot 3^{H_i}},$$
$$\alpha_7 = \frac{3^{2H_i} - 2^{2H_i} - 1}{2^{2H_i+1}}, \quad \text{and} \quad \alpha_8 = \frac{2^{2H_i-1}}{3^{H_i}}.$$

Considering the distribution of all possible realisations of fBm, we see from the use of (28) in Definition 3.1 that the number of variables m in the case of $H_i = H_j$ for all i, j cancels out. Thus, $\mathrm{PPE}(\mathbf{B}_H^m(t))$ is independent of the number of variables m. In addition, $\mathrm{PPE}(\mathbf{B}_H^m(t))$ stays independent of all delays τ. Monotonic dependence on the Hurst parameter H is preserved since the transformations carried out, i.e. additions, do not change the monotonicity. □

For order $d > 4$, no closed formulas for the distribution of ordinal patterns in fBm exist (see Section 2.3). Nevertheless, analogous behaviour can also be observed in this case, and we we evaluate that in an experimental setting in the next section.

4.2 Experimental Evaluation

In this experimental evaluation, we investigate the behaviour of PPE on mfBm in the variation of the Hurst parameter H, the number of variables m and the delay τ. All experimental computations are based on simulations of mfBms using an algorithm described in [20] with parameters $H_i = H_j$ for all i, j, $\rho_{i,j} = 0.0$ and $\eta_{i,j} = 0.1/(1 - H_i - H_j)$ for Lemma 2.1. The lengths $T = 10,000$ of mfBms are assumed to be large. The results of computations are visualised in Figure 7.

In Figure 7a we show the independence of PPE of all orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm from number of variables m. For this, the delay is fixed to $\tau = 1$. All green lines, i.e., for $d = 2$, are constant, namely $-\ln(1/2)$, for each m, confirming the independence of numbers of variables m in (26). All other lines, i.e., for $d > 2$, are also the same for any m, but monotonically dependent on the Hurst parameter H, i.e., in particular confirming Theorem 4.1. In Figure 7b we show the independence of PPE on mfBm from delay τ. For this, the number of variables is fixed to $m = 3$. Again, all green lines, i.e., for $d = 2$, are constant $-\ln(1/2)$ for each τ, confirming the independence of delay τ in (26). All other lines, i.e., for $d > 2$, are also the same for any τ, but monotonically dependent on the Hurst parameter H, in particular confirming Theorem 4.1.

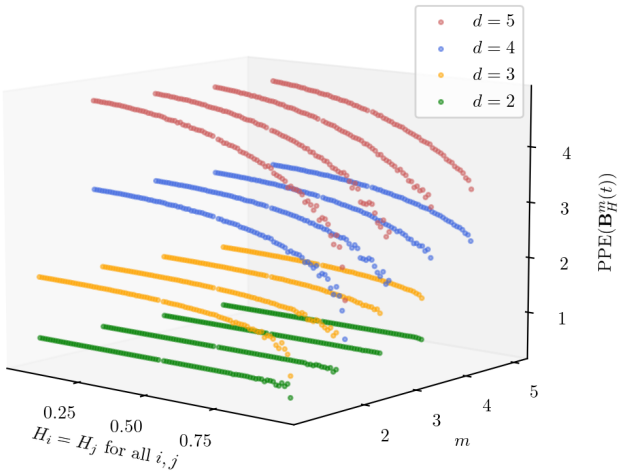
The deviations of the values of PPE in Figure 7 with increasing Hurst parameter H thus result from the length restriction in the simulation, i.e., the distributions as given in the theoretical part are only to be expected if $T \to \infty$ converges [13]. Thus, for a small length T, the estimates of the probabilities for the ordinal patterns differ from the true values of a hypothetical mfBm of infinite length. All in all, the experiment underpins our theoretical results. The results are not surprising since PPE pools over the univariate patterns so that mfBm can be understood by the computational logic as a kind of extended univariate fBm. This kind of extension ensures that the deviations for greater Hurst parameter H and larger number of variables become smaller and can also be observed in Figure 7a.

5 MMSPE Applied to mfBm

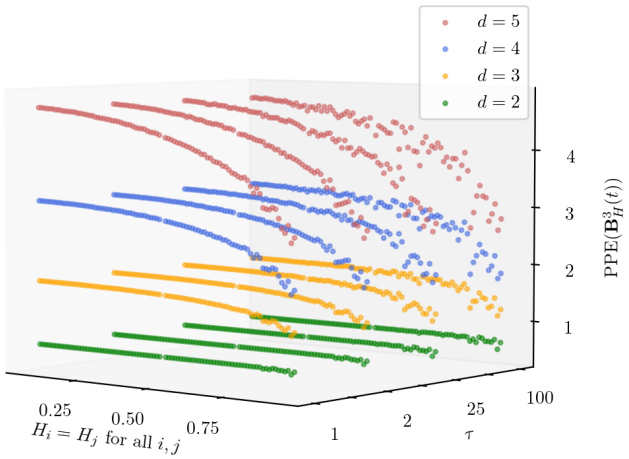
In this section, we investigate the behaviour of MMSPE of different orders, delays, and scales on mfBm in the variation of its Hurst parameter from both theoretical and experimental points of view. For the sake of completeness, we recapitulate results from previous work and fill in remaining results.

5.1 Theoretical Analysis

To investigate the behaviour of MMSPE on mfBm, coarse-grained fractional Brownian motion has to be considered. By using (7)



(a) PPE of fixed delay $\tau = 1$ on mfBm for different numbers of variables $m = 3$.



(b) PPE of different delays τ on mfBm for fixed variable dimension $m = 3$.

Figure 7: Experimental computations of PPE of orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm in variation of its Hurst parameters H.

or (21) on fBm, coarse-grained fractional Brownian motion (cfBm) is defined as

$$B_{H_i}^{i,(\epsilon)}(l) = \frac{1}{\epsilon} \sum_{j=1}^{\epsilon} B_{H_i}^{i}((l-1)\epsilon + j) \tag{29}$$

for $l = 1, ..., T/\epsilon$ and all $i = 1, ..., m$.

In [1] we show that the structure of the covariance function is the same as the original fBm for all variables $i = 1, ..., m$, but with additional information of the scale factor ϵ. Therefore the covariance function of the original fBm or mfBm is invariant to the coarse-grained procedure. For this reason, the results from the previous Section 4 are directly transferable to MMSPE.

Orders 2 and 3. For order $d = 2$ the MMSPE on mfBm is constant for all delays τ, number of variables m, as well as any scales ϵ. In particular holds

$$\mathrm{MMSPE}_{2,\tau,\epsilon}(\mathbf{B}_H^m(t)) = -\ln(1/2) \tag{30}$$

for all τ, m, ϵ. Furthermore, for order $d = 3$, the MMSPE on mfBm is also independent of delays τ and scales ϵ but depends monotonically on the number of variables m as well as the Hurst parameter H of mfBm. In the case $H_i = H_j$ for all variables i, j, MMSPE is also independent of the number of variables m. Details and corresponding formulas for entropy calculation can be found in the original work [1].

Orders greater than 3. With previous insights, the transfer of results from MMSPE of order $d = 3$ to MMSPE of order $d = 4$ is straightforward. For the sake of completeness, we formulate the following theorem.

Theorem 5.1 $\mathrm{MMSPE}(\mathbf{B}_H^m(t))$ *of order* $d = 4$ *is independent of all delays* τ *and scales* ϵ*, but monotonically dependent on the number of variables* m *and Hurst parameter* H *of mfBm. In the case* $H_i = H_j$ *for all* i, j*, it is independent on the number of variables* m*.*

Proof. For mfBm, each variable $i = 1, \ldots, m$ is marginally an fBm. Since the covariance function of fBm is invariant to the coarse-grained procedure in (29), the independence of scales ϵ follows [1]. Then Theorem 4.1 is directly applicable, which implies independence from the delay τ, the number of variables m and monotonic dependence on the Hurst parameter H. □

For order $d > 4$, no closed formulas for the distribution of ordinal patterns in fBm exist (see Section 2.3). Nevertheless, analogous behaviour can also be observed in this case. We evaluate this behaviour in an experimental setting in the next section.

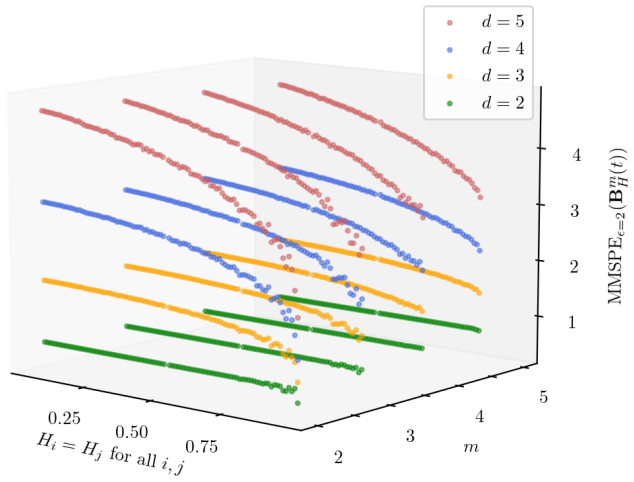
Summarising the results of this subsection, the inclusion of different scales in mfBm has no influence. This is particularly related to or reflects the fractal property because it resembles the original shape when a fractal is zoomed in or scaled. Thus, MMSPE and PPE show identical behaviour, which we investigate and discuss in detail in Section 8.

5.2 Experimental Evaluation

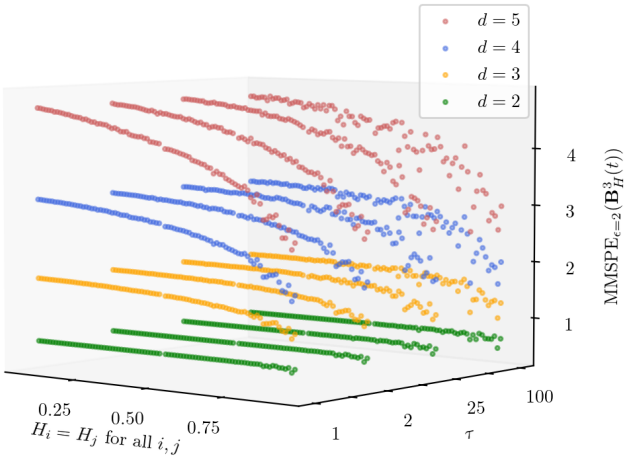
In this experimental evaluation, we investigate the behaviour of MMSPE on mfBm in the variation of the Hurst parameter H, delay τ, number of variables m, as well as scale ϵ. The experimental computations are based on the same simulations of mfBms as in Section 4.2.

In Figure 8a we show the independence of MMSPE of all orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm from number of variables m. For this, the delay is fixed to $\tau = 1$, and the scale is fixed to $\epsilon = 2$. All green lines, i.e., for $d = 2$, are constant, namely $-\ln(1/2)$, for each m, confirming the independence of numbers of variables m in (30). All other lines, i.e., for $d > 2$, are also the same for any m, but monotonically dependent on the Hurst parameter H, in particular confirming Theorem 5.1. In Figure 8b we show the independence of MMSPE on mfBm from delay τ. For this, the number of variables is fixed to $m = 3$, and the scale is fixed to $\epsilon = 2$. Again, all green lines, i.e., for $d = 2$, are constant with value $-\ln(1/2)$ for each τ, confirming the independence of delay τ in (30). All other lines, i.e., for $d > 2$, are also the same for any τ, but monotonically dependent on the Hurst parameter H,

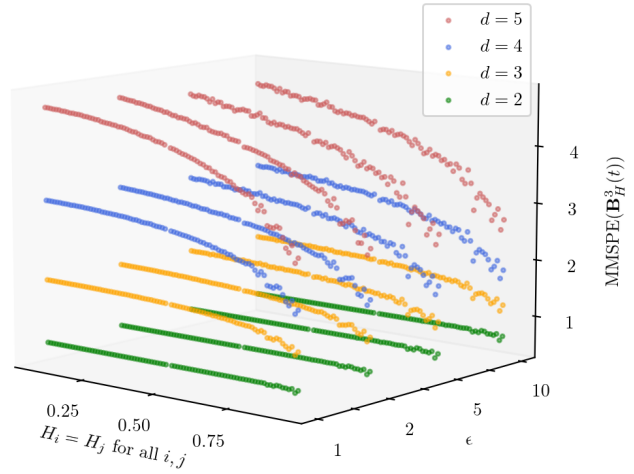
in particular confirming Theorem 5.1.



(a) MMSPE of fixed delay $\tau = 1$ and fixed scale $\epsilon = 2$ on mfBm for different numbers of variables $m = 3$.



(b) MMSPE of different delays τ and fixed scale $\epsilon = 2$ on mfBm for fixed variable dimension $m = 3$.



(c) MMSPE of fixed delay $\tau = 1$ and different scales ϵ on mfBm for fixed variable dimension $m = 3$.

Figure 8: Experimental computations of MMSPE of orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm in variation of its Hurst parameters H.

In Figure 8c we show the independence of MMSPE on mfBm from scale ϵ. For this, the number of variables is fixed to $m = 3$ and the delay to $\tau = 1$. Again, all green lines, i.e., for $d = 2$, are constant with value $-\ln(1/2)$ for each ϵ, confirming the independence of scale ϵ in (30). All other lines, i.e., for $d > 2$, are also the same for any scale ϵ, but monotonically dependent on the Hurst parameter H, in particular confirming Theorem 5.1. Again, the reason for the deviations of the values of MMSPE in Figure 8 with increasing Hurst parameter H is the same as in the previous Section 4, namely from length restriction $T < \infty$ of mfBm.

All in all, the experiment underpins our theoretical results from Since MMSPE is based on PPE and scaling has no influence on mfBm, the behaviour of MMSPE on mfBm in Figure 8 is to be expected the same as that of PPE in Figure 7. A direct comparison is discussed in Section 8.

6 MWPE Applied to mfBm

In this section, we investigate the behaviour of MWPE of different orders and delays on mfBm in the variation of its Hurst parameter from both theoretical and experimental points of view. For the sake of completeness, we recap results from past work and add some further results to fill in some of the remaining research gaps.

6.1 Theoretical Analysis

Order 2. In previous work, [2] we show that for order $d = 2$, PPE of mfBm is constant for all delays τ as well as numbers of variables m.

Theorem 6.1 *MWPE of order $d = 2$ on mfBm $B_H(t)$ is given by*

$$\mathrm{MWPE}_{2,\tau}(\mathbf{B}_H^m(t)) = -\ln(1/2) \tag{31}$$

for all τ, m.

Proof. From [2]: WPE or MWPE differs from PE or PPE, respectively, in that the ordinal patterns are weighted depending on their position t according to (11). For a weight w_t of order $d = 2$, i.e., of two time steps $x_{t-\tau}$ and x_t, we have

$$w_t = \frac{1}{2}\sum_{k=1}^{2}(x_{t-(k-1)\tau} - \overline{x}_t^{2,\tau})^2 \tag{32}$$

$$= \frac{1}{2}(x_t - x_{t-\tau})^2. \tag{33}$$

Since $x_t \sim B_H(t)$, we conclude from Definition 2.3, property (3), i.e., fBm has stationary increments, that

$$\frac{1}{2}(B_H(t) - B_H(t-\tau))^2 \sim \frac{1}{2}(B_H(\tau))^2 \tag{34}$$

with $\mathrm{Var}(B_H(\tau)) = \sigma^2\tau^{2H}$ and $\sigma^2 = \mathrm{Var}(B_H(1)) = s^{2H}$ as in (2). Consequently, the weights w_t are independently distributed from t, i.e.,

$$w_t \sim \mathcal{N}(0, \frac{1}{2}(s\tau)^{2H}). \tag{35}$$

Considering the distribution of all possible realisations of fBm, we see from the use of the weights w_t from (11) in the computation of WPE that the weights cancel out for a constant delay $\tau \in \mathbb{N}$. With (12) it follows that $\mathrm{WPE}(B_H(t)) = -\ln(1/2)$.

Calculating the frequencies of ordinal patterns for the weighted pooling matrix $P_w^{\tau,d} \in \mathbb{R}^{m\times d!}$ results in an equal distribution with

$$P^{\tau,2}_{w_{i,(1,0)}} = P^{\tau,2}_{w_{i,(0,1)}} = \frac{1}{2m} \tag{36}$$

for all $i = 1, .., m$. By calculating the marginal relative frequencies, the number of variables m is reduced, i.e.,

$$P^{\tau,2}_{w_{\cdot,j}} = \frac{1}{2} \tag{37}$$

for $j = \{(0,1),(1,0)\}$. Hence the claim follows. □

Orders greater than 2. For orders $d = 3$ and $d = 4$, the following theorem can be derived by inspecting corresponding weights of ordinal patterns in combination with results from PPE.

Theorem 6.2 MWPE($\mathbf{B}_H^m(t)$) *of orders $d = 3$ and $d = 4$ are independent of all delays τ but monotonically dependent on the number of variables m and Hurst parameter H of mfBm. In the case $H_i = H_j$ for all i, j, it is independent on the number of variables m.*

Proof. The weights for ordinal patterns of order $d = 3$ are given by

$$w_t = \frac{2}{9} \cdot (x_t^2 - x_t(x_{t-\tau} - x_{t-2\tau}) + x_{t-\tau}^2 - x_{t-\tau}x_{t-2\tau} + x_{t-2\tau}^2). \tag{38}$$

In contrast to $d = 2$, in particular in contrast to (33), due to the lack of symmetry, the weights depend on the time t, and therefore do not cancel and thus influence the value of MWPE. Since the distribution of the ordinal pattern is independent of the weights w_t, the independence of the delay τ is retained. In addition, the multiplication of weights w_t and counts do not affect the monotony that follows from (13) and (14) of PPE. The independence of MWPE from the number of variables m for $H_i = H_j$ for all variables i, j follows with the same line of reasoning as from the end of the proof of Theorem 4.1 when calculating the marginal frequencies.

The weights for ordinal patterns of order $d = 4$ are given by

$$\begin{aligned} w_t \quad = \tfrac{1}{16} \cdot \quad & (3x_t^2 - 2x_t(x_{t-\tau} + x_{t-2\tau} + x_{t-3\tau}) \\ & +3x_{t-\tau}^2 - 2x_{t-\tau}(x_{t-2\tau} + x_{t-3\tau}) \\ & +3x_{t-2\tau}^2 - 2x_{t-2\tau}x_{t-3\tau} + 3x_{t-3\tau}^2). \end{aligned} \tag{39}$$

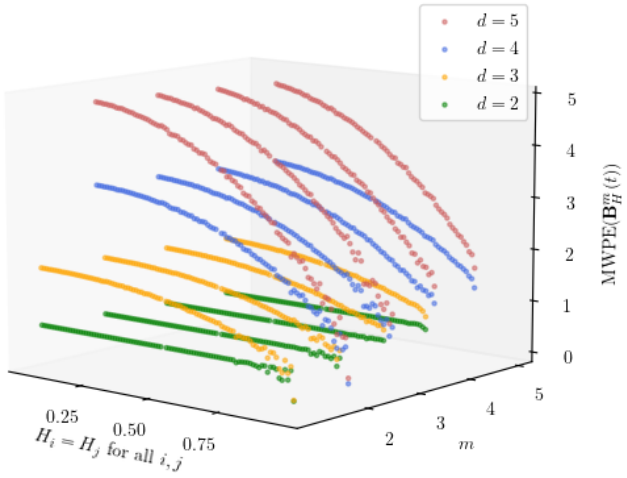
Using the same reasoning as in the case of $d = 3$ and with (15), the independence of the delay τ, the number of variables m and the monotonic dependence of the Hurst parameter H for MWPE are preserved. □

For order $d > 4$, no closed formulas for the distribution of ordinal patterns in fBm exist (see Section 2.3). Nevertheless, analogous behaviour can also be observed in this case, which we evaluate in an experimental setting in the next section.

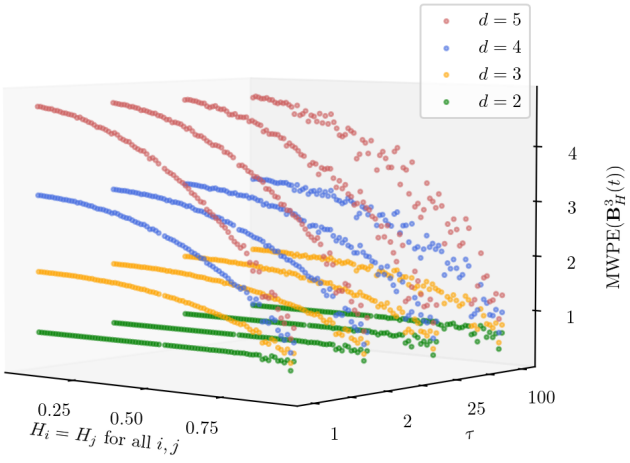
For increasing Hurst parameter $H \to 1$, m(fBm) is more positively correlated, i.e., after an upward jump, a further upward jump is more likely to follow and vice versa. Consequently, it is more likely that the strictly ascending ordinal pattern $(0, 1, 2)$ or the strictly descending ordinal pattern $(2, 1, 0)$ occurs than the other four ordinal patterns $(0, 2, 1)$, $(1, 0, 2)$, $(2, 0, 1)$ and $(1, 2, 0)$. Since for $H \to 1$ and $t \to \infty$ m(fBm) becomes trendier, the weights of $(0, 1, 2)$ and $(2, 1, 0)$ respectively must be larger on average. Thus, by weighting the ordinal patterns of orders $d > 2$, the values of MWPE are expected to decrease faster than the values of PPE.

6.2 Experimental Evaluation

In this experimental evaluation, we investigate the behaviour of MWPE on mfBm in the variation of Hurst parameter H, delay τ, and the number of variables m. The experimental computations are based on the same simulations of mfBms as in Section 4.2.



(a) MWPE of fixed delay $\tau = 1$ on mfBm for different numbers of variables $m = 3$.



(b) MWPE of different delays τ on mfBm for fixed variable dimension $m = 3$.

Figure 9: Experimental computations of MWPE of orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm in variation of its Hurst parameters H.

In Figure 9a we show the independence of MWPE of all orders $d = 2$ (green), $d = 3$ (orange), $d = 4$ (blue), and $d = 5$ (red) on mfBm from number of variables m. For this experiment, the delay is fixed to $\tau = 1$. All green lines, i.e., for $d = 2$, are constant, namely $-\ln(1/2)$, for each m, confirming the independence of numbers of variables m in (31). All other lines, i.e., for $d > 2$, are the same for any number of variables m, but monotonically dependent on the Hurst parameter H. In Figure 9b we show the independence of MWPE on mfBm from delay τ. For this experiment, the number of variables is fixed to $m = 3$. Again, all green lines, i.e., for $d = 2$, are constant with value $-\ln(1/2)$ for each delay τ, confirming the independence of delay τ in (31). All other lines, i.e., for $d > 2$, are

also the same for any delay τ, but monotonically dependent on the Hurst parameter H. The reason for the deviations of the values of MWPE in Figure 9 with increasing Hurst parameter H is the same as in the previous Section 4, i.e., the deviations are caused by length constraint $T < \infty$ of mfBm.

All in all, the experiment underpins our theoretical results statet in Theorem 6.1 and Theorem 6.2. In particular, the behaviour of MWPE in Figure 9 is similar to that of PPE in Figure 7 in terms of the independence of the parameters τ and m, as well as the constant or monotonic progression at orders $d = 2$ or $d > 2$, respectively.

In the theoretical analysis, we discussed the influence of the weighting of ordinal patterns of different orders on the values of MWPE. For $d = 2$, we show that weighting has no influence (Theorem 6.1). For orders $d > 2$, we assume that certain ordinal patterns (strictly ascending and strictly descending) are weighted more than others. Thus the values of MWPE fall faster than the values of PPE as the Hurst parameter H increases. Exemplarily, Table 1 shows the frequencies of (weighted) ordinal patterns $(p_{w_j})\ p_j$ of order $d = 3$ on a simulated mfBm of length $T = 10000$ with number of variables $m = 3$, correlation coefficient $\rho = 0.0$, and Hurst parameter $H = (0.8, 0.8, 0.8)$. While for PPE it holds that $w_t = 1$ for every pattern j, we see that the weights w_t of ordinal patterns $(0, 1, 2)$ and $(2, 1, 0)$ on average are larger than those of the other four ordinal patterns. Consequently, the frequencies of these ordinal patterns are also higher than in the unweighted case of PPE. For more details on the difference between MWPE and PPE, see Section 8.

j	counts	w_t	p_{w_j}	p_j
$(0, 1, 2)$	10539	0.76	0.48	0.35
$(0, 2, 1)$	2470	0.19	0.03	0.08
$(1, 0, 2)$	2482	0.19	0.03	0.08
$(1, 2, 0)$	2450	0.19	0.03	0.08
$(2, 0, 1)$	2462	0.19	0.03	0.08
$(2, 1, 0)$	9588	0.72	0.40	0.33

Table 1: Frequencies of (weighted) ordinal patterns $(p_{w_j})\ p_j$ of order $d = 3$.

7 MOPPE Applied to mfBm

In this section, for the first time, the behaviour of MOPPE of different orders and delays on mfBm is investigated in the variation of its Hurst parameter, both from a theoretical and experimental point of view.

7.1 Theoretical Analysis

Since multivariate ordinal patterns are defined as univariate ordinal patterns combined in a matrix, joint probabilities are considered to investigate the behaviour of MOPPE on mfBm. Let p_j^τ be a probability for a univariate ordinal pattern j of arbitrary order d in fBm given in Section 2.3. Since any univariate ordinal pattern of any order d is independent of the delay τ, we write p_j in the following.

Order 2. In contrast to the previous sections, the behaviour of MOPPE on mfBm is obviously dependent on the number of variables m but remains independent of the delays τ.

Lemma 7.1 *Let $B^k_{H_k}(s)$ and $B^l_{H_l}(t)$ for every $k, l = 1, \dots, m$ conditionally independent, then it holds that*

$$\mathrm{MOPPE}_{2,\tau}(\mathbf{B}^m_H(t)) = -\ln\Big(\frac{1}{2^m}\Big) \tag{40}$$

for all τ.

Proof. The independence from the delay τ follows directly from Definition 3.4 and the distribution of the univariate ordinal patterns of order $d = 2$ in (12). Let $j \in \{(0, 1), (1, 0)\}$ be a univariate ordinal patterns of order $d = 2$. Let $X_k = B^k_{H_k}(t)$ and $X_l = B^l_{H_l}(t)$ for every $k, l = 1, \dots, m$ conditionally independent, then the joint probability function satisfies

$$\begin{aligned} P(X_1 = j, \dots, X_m = j) &= P(X_1 = j) \cdot \dots \cdot P(X_m = j) && (41)\\ &= {p_j}^m && (42) \end{aligned}$$

With (12) it is $p_j = 1/2$ for every j, so that the joint distribution of every m-fold combination of ordinal patterns $j \in \{(0, 1), (1, 0)\}$ stays ${p_j}^m$. For the number of variables m, there exist 2^m multivariate ordinal patterns as combinations from univariate ordinal patterns so that

$$\mathrm{MOPPE}_{2,\tau}(\mathbf{B}^m_H(t)) = -\sum_{j=1}^{2^m} \frac{1}{2}^m \cdot \ln\Big(\frac{1}{2}^m\Big). \tag{43}$$

□

Orders greater than 2. For orders $d = 3$ and $d = 4$, the following theorem can be derived by considering joint probabilities of univariate ordinal pattern distributions introduced in Section 2.3.

Lemma 7.2 $\mathrm{MOPPE}(\mathbf{B}^m_H(t))$ *of orders $d = 3$ and $d = 4$ are independent of all delays τ, but dependent on number of variables m, and monotonically dependent on the Hurst parameter H.*

Proof. The independence from the delay τ follows directly from Definition 3.4 and the distribution of the univariate ordinal patterns or orders $d = 3$ or $d = 4$ in (13) or (15), respectively. As introduced in Section 2.3, the distribution of ordinal patterns of order $d = 3$ is twofold (see (14)). Therefore, let $A = \{(0, 1, 2), (2, 1, 0)\}$ and $B = \{(0, 2, 1), (1, 2, 0), (1, 0, 2), (2, 0, 1)\}$. Since multivariate ordinal patterns are combinations of univariate ordinal patterns, in the following, we consider three cases where the ordinal patterns of all m variables are either from A, from B or from A and B. Again, let $B^k_{H_k}(s)$ and $B^l_{H_l}(t)$ be conditionally independent for each $k, l = 1, \dots, m$.

1. Let the univariate pattern a_i of the i-th variable be $a_i \in A$ for all $i = 1, \dots, m$. Using (13) and (41) the joint distributions of all 2^m combinations of $a_1, \dots, a_m \in A$ are given by

$$\prod_{i=1}^{m} \frac{1}{\pi^m} \arcsin 2^{H_i - 1} =: \prod_{i=1}^{m} \frac{1}{\pi^m} u_i. \tag{44}$$

2. Let the univariate pattern b_i of the i-th variable be $b_i \in B$ for all $i = 1, \dots, m$. Using (14) and (41) the joint distributions of all 4^m combinations of $b_1, \dots, b_m \in B$ are given by

$$\prod_{i=1}^{m} \frac{1}{4}(1 - \frac{1}{\pi^m} u_i). \tag{45}$$

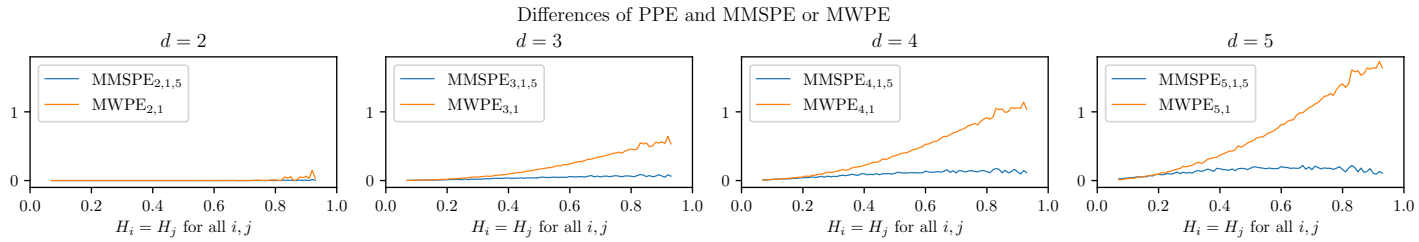


Figure 10: Differences of PPE and MMSPE or MWPE of different orders (from left to right) and fixed delay $\tau = 1$ on simulations of mfBm with fixed number of variables $m = 3$ in variation of Hurst parameters $H_i = H_j$ for all i, j.

3. Let the univariate pattern c_i of the i-th variable be, where $c_1, \dots, c_k \in A$ and $c_{k+1}, \dots, c_m \in B$. Then using (13), (14), and (41) the joint distributions of all $6^m - 4^m - 2^m$ remaining combinations $c_1, \dots, c_m$ are given by

$$\prod_{i=1}^{k} \frac{1}{\pi^k} u_i \cdot \prod_{i=k+1}^{m} \frac{1}{4}(1 - \frac{1}{\pi^{m-k}} u_i). \tag{46}$$

In particular, the joint distributions or the distributions of the multivariate ordinal patterns remain monotonically dependent on the Hurst parameter H after the monotonic transformations.

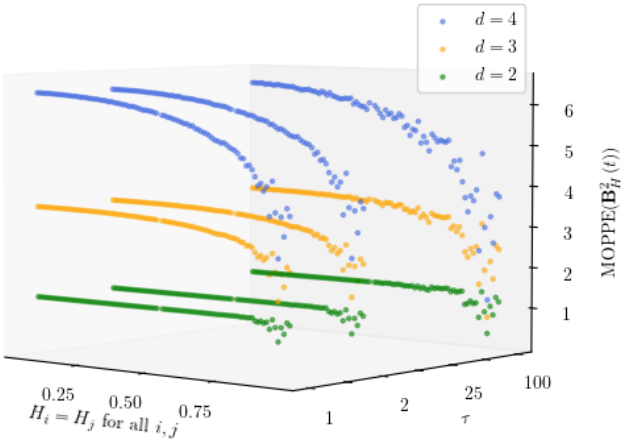
For order $d = 4$, formulas for joint distributions corresponding to multivariate ordinal patterns are derivable with (15) and (41) in the same way as for order $d = 3$. □

For orders $d > 4$, no closed formulas for the distribution of ordinal patterns in fBm exist. Nevertheless, analogous behaviour of MOPPE of orders $d > 4$ can also be observed in this case, which we evaluate in an experimental setting in the next section.

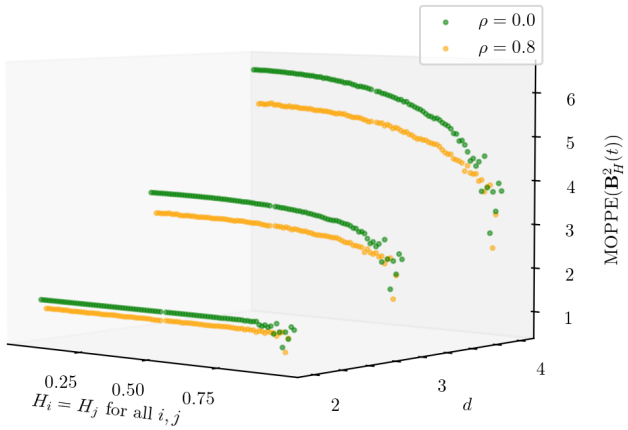
7.2 Experimental Evaluation

In this experimental evaluation, we investigate the behaviour of MOPPE on mfBm in the variation of the Hurst parameter H and delay τ. Since the number of multivariate ordinal patterns depends on the number of variables m and thus also on MOPPE, we refrain from an experimental evaluation of different m, i.e., we assume $m = 2$ small for computational reasons. The experimental computations are based on the same simulations of mfBms as in Section 4.2 with the difference that in this experiment, we include more than one correlation parameter as MOPPE takes into account the interdependence of variables.

In Figure 11a we show the independence of MOPPE on mfBm from delay τ. For this, the number of variables is fixed to $m = 2$. Again, all green lines, i.e., for $d = 2$, are constant $-\ln(1/4)$ for each τ, confirming the independence of delay τ in (40). All other lines, i.e., for $d > 2$, are also the same for any τ, but monotonically dependent on the Hurst parameter H, i.e., for increasing Hurst parameter H the value of MOPPE decreases. The reason for the deviations of the values of MOPPE in Figure 11 with increasing Hurst parameter H is the same as in the previous Section 4. All in all, the experiment underpins our theoretical results from Theorem 7.1 and Theorem 7.2.



(a) MOPPE of different delays τ on mfBm with fixed number of variables $m = 2$.



(b) MOPPE of fixed delay $\tau = 1$ on mfBm with fixed number of variables $m = 2$.

Figure 11: Experimental computations of MOPPE for different delays τ on mfBm with different correlations $\rho = \{0.0, 0.8\}$ in variation of its Hurst parameters H.

Although we have focused on independent variables in the theoretical part for simplicity as well as due to the lack of results in the literature, we examine the dependence on the correlations of variables in space in Figure 11b, since the definition of multivariate ordinal patterns involves the dependence on several variables. A high cross-covariance or correlation of the variables causes the behaviour of the individual variables (for different Hurst parameters) to converge. Another influences the behaviour of one variable, and so is the occurrence of the higher-order ordinal patterns that are dependent on the Hurst parameter. For the investigation of correlations using MOPPE, the delay or the number of variables are fixed to $\tau = 1$ or $m = 2$, respectively. In Figure 11b, the yellow

lines, corresponding to a correlation coefficient of 0.8, are constantly below the green lines corresponding to a correlation coefficient of 0.0. This reflects the fact that due to a higher correlation of the variables, there is a greater fit of the individual univariate ordinal pattern order in the multivariate combination, resulting in a smaller MOPPE value. Advantages are discussed in Section 8.

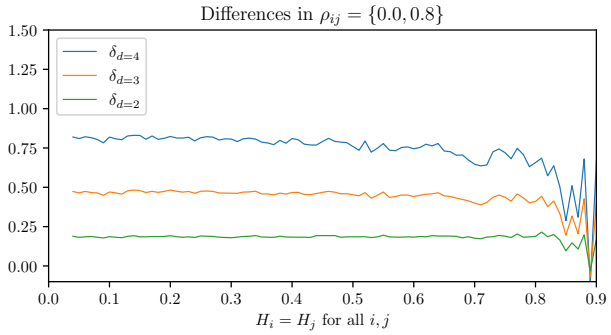


Figure 12: Differences of MOPPE values on mfBm with different correlations $\rho_{ij} = \{0.0, 0.8\}$ in variation of its Hurst parameter H.

8 Comparison

Up to this point, we have examined the behaviour of numerous multivariate representations, namely pooled permutation entropy (PPE), multivariate multi-scale permutation entropy (MMSPE), multivariate weighted permutation entropy (MWPE), and multivariate ordinal pattern permutation entropy (MOPPE). In this section, we summarise both the differences and the similarities between the various representations. This leads to different possible applications as well as to different recommendations.

PPE can be understood as a canonical extension of univariate ordinal patterns to a multivariate variant that pools the univariate ordinal patterns over the variables. MMSPE and MWPE belong to the same family as PPE. They are all based on a pooled matrix while addressing different aspects of the time series and ordinal patterns, namely scaling and amplitudes. One advantage for these approaches is that, in contrast to MOPPE, they also allow univariate analyses on the individual variables within the algorithm. Figure 10 compares PPE with its extensions MMSPE and MWPE with different orders $d = 1, \dots, 5$ from left to right. Since, as evaluated in the individual sections, all measures are independent of the delay and the number of variables, we set $\tau = 1$ and $m = 3$. Similarly, MMSPE is independent of the scale parameter ϵ, so we set the scale factor $\epsilon = 5$. The left sub-figure shows the equality of PPE, MMSPE and MWPE for order $d = 2$, which again underpin (26), (30) and (31). Thus, scaling or weighting has no effect. The three right sub-figures in Figure 10 show the equality of PPE and MMSPE (blue lines) for orders $d = 3, 4, 5$. Similarly, they show a difference between PPE and MWPE as the orange lines increase in variation of increasing H, indicating a faster decrease of MWPE.

MOPPE can be understood as a canonical extension of univariate ordinal patterns to multivariate ordinal patterns by conceiving univariate patterns as multidimensional patterns in a matrix. The advantage of this approach lies in consideration of the interdependencies of several variables at a single time point. A disadvantage is the exponentially increasing number of possible ordinal patterns, which increase the complexity of the computation and result in a uniform distribution and thus maximum entropy for $T \ll \infty$. Furthermore, MOPPE considers the multivariate time series as a whole, i.e., not each variable individually, which means that an analysis of the individual variables using MOPPE within the algorithm is not possible, in contrast to PPE.

Figure 12 visualises the differences of MOPPE values of different orders $d = 2$ (green), $d = 3$ (orange), and $d = 4$ (blue) and fixed delay τ on mfBm with fixed variable dimension $m = 2$ and different correlations $\rho_{ij} = \{0.0, 0.8\}$ in variation of its Hurst parameter $H = (H_1, H_2)$, where $H_1 = H_2$. As the visualisation shows, the differences for any $d \in \{2, 3, 4\}$ are almost constant for all Hurst parameters H. Increasing differences for increasing order d result from increasing maximum entropy values per order d.

The differences in the multivariate extensions of PE are important because they may be relevant to the use of different applications. In the field of inverse problems, the aim is to estimate the parameters, e.g., the Hurst parameter, of a generating dynamic system from observed realisations [36]. As all multivariate variants for orders $d > 2$ involve a direct relationship between the distribution of univariate ordinal patterns and the Hurst parameter H, PPE, MWPE and MOPPE are suitable candidates to solve the inverse problem regarding H. Given a realisation of mfBm, for example, the calculation of entropies that decrease as H increases can provide conclusions about the value of H. In addition to estimating H, MOPPE has the potential to uncover additional correlations by taking interdependencies into account, and it does so from one source. However, MOPPE is only practical for a small number of variables m. In machine learning, a high accuracy of the prediction depends not only on the model used but also on the representations of the data. Therefore, extracting good or expressive representations, also called features, from the data is essential. Especially in classification tasks, it is necessary to extract features that separate well [37]. In this case, MWPE may promise advantages over PPE because MWPE favours certain ordinal patterns (strictly rising and strictly falling) and thus decreases faster compared to PPE.

9 Conclusion and Future Work

Fractional Brownian motion (fBm) is usually used for modelling real-world applications with specific properties such as long-range dependence or self-similarity. Since many real-world applications also contain multivariate interdependencies between several variables, we focus on multivariate fractional Brownian motion (mfBm). This paper provides a comprehensive study of several multivariate measures that study the qualitative behaviour of multivariate fractional Brownian motion (mfBm), namely pooled permutation entropy (PPE), multivariate multi-scale permutation entropy

(MMSPE), multivariate weighted permutation entropy (MWPE), and multivariate ordinal pattern permutation entropy (MOPPE). All measures are understood as multivariate extensions of univariate permutation entropy (PE), which is based on the distribution of ordinal patterns, i.e., on the ups and downs of different orders $d \in \mathbb{N}_{>1}$. The main contributions of this paper are the extensions of the theoretical and experimental investigations of PPE, MMSPE and MWPE on mfBm in the original papers [1, 2] to additional orders $d = 4$ and $d = 5$. Furthermore, this paper provides a new study on the behaviour of MOPPE on mfBm, and for the first time, a comparison of all multivariate extensions considered.

Our theoretical and experimental analyses show the following results on the behaviour of PPE, MMSPE, MWPE and MOPPE on mfBm. For order $d = 2$, for all four cases the multivariate permutation entropy is constant $-\ln(1/2)$ regardless of number of variables, Hurst parameters, delays, scales, or weights (or $-\ln(1/2^m)$ in case of MOPPE). Because all measures are equal, in particular constant and independent of all parameters, no characteristics can be determined via these measures with order $d = 2$. Therefore, usage is not reasonable in applications. Since scaling does not change the structure of mfBm, MMSPE of any scale ϵ is equal to PPE and analysis with MMSPE of higher orders $d > 2$ does not provide any additional insight than PPE.

However, for orders $d > 2$, the use of PPE, MWPE and MOPPE provide interesting insights and possible applications. The distribution of ordinal patterns, and thus also PPE, MWPE and MOPPE, are directly related to the Hurst parameter H. For example, considering the estimation of H is an inverse problem, PPE, MWPE and MOPPE can be used since they all depend monotonically on H, i.e., the entropy decreases as H increases. In contrast to PPE and MOPPE, MWPE is substantially influenced by strictly ascending and descending ordinal patterns. For this reason, MWPE decreases faster than PPE as H increases, providing more expressive representations that may promise better discriminability. Since multivariate ordinal patterns involve the dependence of several variables on mfBm, MOPPE is promising for estimating the Hurst parameter H and correlations from one hand, which the other measures cannot. To prove the exact theoretical relationship, conditional probabilities of univariate ordinal patterns must first be derived.

In the literature, there are additional approaches to multivariate permutation entropy determination, specifically Multivariate Permutation Entropy (MvPE), introduced in [30], or multivariate permutation entropy based on principal component analysis (MPE-PCA), introduced in [33]. Since their application requires a prior transformation of mfBm and thus is not a canonical extension of the univariate definition, its study is not part of this work. Nevertheless, both have the potential to uncover (additional) structures in mfBm but are left for future work. In addition, the theoretical results of the measures investigated in this work motivate various potential applications, e.g., for estimating the Hurst parameter or as a feature in a learning task in the field of machine learning. Similarly, the application of these measures is of great interest not only in the context of mfBm but also for any real-world data set. Since this work focuses on the theoretic contexts, we leave concrete applications for future work.

References

[1] M. Mohr, N. Finke, R. Möller, "On the Behaviour of Permutation Entropy on Fractional Brownian Motion in a Multivariate Setting," in Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2020 (APSIPA-ASC), 2020.

[2] M. Mohr, F. Wilhelm, R. Möller, "On the Behaviour of Weighted Permutation Entropy on Fractional Brownian Motion in the Univariate and Multivariate Setting," The International FLAIRS Conference Proceedings, **34**, 2021, doi: 10.32473/flairs.v34i1.128464.

[3] T. Konstantopoulos, S.-J. Lin, "Fractional Brownian Approximations of Queueing Networks," in P. Glasserman, K. Sigman, D. D. Yao, editors, Stochastic Networks, 257–273, Springer New York, New York, NY, 1996.

[4] R. J. Elliott, J. van der Hoek, "Fractional Brownian Motion and Financial Modelling," in M. Kohlmann, S. Tang, editors, Mathematical Finance, 140–151, Birkhäuser Basel, 2001.

[5] B. S. Nath, B. Mousumi, "Long Memory in Stock Returns: A Study of Emerging Markets," Iranian Journal of Management Studies, **5**(2), 67–88, 2012.

[6] F. Traversaro, F. O. Redelico, M. R. Risk, A. C. Frery, O. A. Rosso, "Bandt-Pompe symbolization dynamics for time series with tied values: A data-driven approach," Chaos: An Interdisciplinary Journal of Nonlinear Science, **28**(7), 2018, doi:10.1063/1.5022021.

[7] J. Amigó, Permutation Complexity in Dynamical Systems: Ordinal Patterns, Permutation Entropy and All That, Springer Series in Synergetics, Springer-Verlag, Berlin Heidelberg, 2010.

[8] C. Bandt, F. Shiha, "Order Patterns in Time Series," Journal of Time Series Analysis, **28**(5), 646–665, 2007.

[9] F. C. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito, I. Palamara, "Multivariate Multi-Scale Permutation Entropy for Complexity Analysis of Alzheimer's Disease EEG," Entropy, **14**(7), 1186–1202, 2012, doi: 10.3390/e14071186.

[10] B. Fadlallah, B. Chen, A. Keil, J. Principe, "Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information," Physical review. E, Statistical, nonlinear, and soft matter physics, **87**, 022911, 2013, doi:10.1103/PhysRevE.87.022911.

[11] M. Sinn, K. Keller, "Estimation of Ordinal Pattern Probabilities in Gaussian Processes with Stationary Increments," Comput. Stat. Data Anal., **55**(4), 1781–1790, 2011, doi:10.1016/j.csda.2010.11.009.

[12] L. Zunino, D. G. Pérez, M. T. Martín, M. Garavaglia, A. Plastino, O. A. Rosso, "Permutation entropy of fractional Brownian motion and fractional Gaussian noise," Physics Letters A, **372**(27), 4768–4774, 2008.

[13] A. Dávalos, M. Jabloun, P. Ravier, O. Buttelli, "Theoretical Study of Multiscale Permutation Entropy on Finite-Length Fractional Gaussian Noise," in 2018 26th European Signal Processing Conference (EUSIPCO), 1087–1091, 2018, doi:10.23919/EUSIPCO.2018.8553070, iSSN: 2076-1465.

[14] L. A. Gil-Alana, "A fractional multivariate long memory model for the US and the Canadian real output," Economics Letters, **81**(3), 355–359, 2003, doi:10.1016/S0165-1765(03)00217-9.

[15] S. Achard, D. S. Bassett, A. Meyer-Lindenberg, E. Bullmore, "Fractal connectivity of long-memory networks," Physical Review E, **77**(3), 2008.

[16] B. B. Mandelbrot, J. R. Wallis, "Noah, Joseph, and Operational Hydrology," Water Resources Research, **4**(5), 909–918, 1968, doi:10.1029/WR004i005p00909.

[17] G. Didier, V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," Bernoulli, **17**(1), 1–33, 2011, doi:10.3150/10-BEJ259.

[18] F. Lavancier, A. Philippe, D. Surgailis, "Covariance function of vector self-similar processes," Statistics & Probability Letters, **79**(23), 2415–2421, 2009, doi:10.1016/j.spl.2009.08.015.

[19] G. Didier, V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," Bernoulli, **17**(1), 1–33, 2011.

[20] P.-O. Amblard, J.-F. Coeurjolly, F. Lavancier, A. Philippe, "Basic properties of the Multivariate Fractional Brownian Motion," Séminaires et congrès, **28**, 65–87, 2013.

[21] P.-O. Amblard, J.-F. Coeurjolly, "Identification of the Multivariate Fractional Brownian Motion," IEEE Transactions on Signal Processing, **59**(11), 5152–5168, 2011, doi:10.1109/TSP.2011.2162835.

[22] B. Chiu, E. Keogh, S. Lonardi, "Probabilistic Discovery of Time Series Motifs," in Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03, 493–498, ACM, 2003.

[23] C. Bandt, B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," Physical Review Letters, **88**(17), 174102, 2002, publisher: American Physical Society.

[24] K. Keller, T. Mangold, I. Stolz, J. Werner, "Permutation Entropy: New Ideas and Challenges," Entropy, **19**(3), 134, 2017, doi:10.3390/e19030134.

[25] A. B. Piek, I. Stolz, K. Keller, "Algorithmics, Possibilities and Limits of Ordinal Pattern Based Entropies," Entropy, **21**(6), 547, 2019, doi:10.3390/e21060547.

[26] K. Keller, S. Maksymenko, I. Stolz, "Entropy determination based on the ordinal structure of a dynamical system," Discrete and Continuous Dynamical Systems - Series B, **20**(10), 3507–3524, 2015, doi:10.3934/dcdsb.2015.20.3507.

[27] I. Stolz, K. Keller, "A General Symbolic Approach to Kolmogorov-Sinai Entropy," Entropy, **19**(12), 675, 2017, doi:10.3390/e19120675.

[28] M. Costa, A. L. Goldberger, C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," Physical Review Letters, **89**(6), 068102, 2002.

[29] K. Keller, H. Lauffer, "Symbolic Analysis of High-Dimensional Time Series," International Journal of Bifurcation and Chaos, **13**(09), 2657–2668, 2003, doi:10.1142/S0218127403008168.

[30] S. He, K. Sun, H. Wang, "Multivariate Permutation Entropy and Its Application for Complexity Analysis of Chaotic Systems," Physica A: Statistical Mechanics and its Applications, **461**, 812–823, 2016, doi:10.1016/j.physa.2016.06.012.

[31] K. Keller, "Permutations and the Kolmogorov-Sinai Entropy," Discrete & Continuous Dynamical Systems - A, **32**(3), 891, 2012, doi:10.3934/dcds.2012.32.891.

[32] A. Antoniouk, K. Keller, S. Maksymenko, "Kolmogorov-Sinai Entropy via Separation Properties of Order-Generated Sigma-Algebras," Discrete and Continuous Dynamical Systems, **34**(5), 1793–1809, 2013, doi:10.3934/dcds.2014.34.1793.

[33] M. Mohr, F. Wilhelm, M. Hartwig, R. Möller, K. Keller, "New Approaches in Ordinal Pattern Representations for Multivariate Time Series," in Proceedings of the 33rd International Florida Artificial Intelligence Research Society Conference (FLAIRS-33), 124–129, AAAI Press, 2020.

[34] Y. Rayan, Y. Mohammad, S. A. Ali, "Multidimensional Permutation Entropy for Constrained Motif Discovery," in N. T. Nguyen, F. L. Gaol, T.-P. Hong, B. Trawiński, editors, Intelligent Information and Database Systems, Lecture Notes in Computer Science, 231–243, Springer International Publishing, Cham, 2019, doi:10.1007/978-3-030-14799-0_20.

[35] N. Nicolaou, J. Georgiou, "The Use of Permutation Entropy to Characterize Sleep Electroencephalograms," Clinical EEG and Neuroscience, **42**(1), 24–28, 2011, doi:10.1177/155005941104200107.

[36] B. D'Ambrogi-Ola, "Inverse Problem for Fractional Brownian Motion with Discrete Data," 2009.

[37] Y. Bengio, A. Courville, P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, **35**(8), 1798–1828, 2013, doi:10.1109/TPAMI.2013.50.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 6, 119-129 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Associated Risk Factors for the Development of Extensively Resistant Pulmonary Tuberculosis in the First Level of Health Care: From a Public Health Perspective

Mauricio Mamani[*,1], Mario Chauca[2], Edward Huamani[3], Richard Gonzales[3]

[1]*Universidad Privada San Juan Bautista, Lima, 15067, Perú*

[2]*Universidad Ricardo Palma, Lima, 15039, Perú*

[3]*Federico Villarreal National University, Lima, 15088, Perú*

ARTICLE INFO

Article history:
Received: 09 July, 2021
Accepted: 11 November, 2021
Online: 17 December, 2021

Keywords:
Risk factor
Extensively drug-resistant
Tuberculosis

ABSTRACT

To determine the risk factors associated with the development of extensively resistant pulmonary tuberculosis in the DIRIS Lima Sur, 2017. The type of research was observational, quantitative, analytical, retrospective case-control approach. The population consisted of 158 patients with MDR-TB treatment schedule between 2016 and 2017. The sample consisted of 24 cases diagnosed with extensively drug-resistant pulmonary tuberculosis (XDR-TB) and 48 controls with MDR-TB. The survey was used for both groups (cases - controls) as a data collection technique and a questionnaire as an instrument. 5 risk factors associated with XDR-TB were identified. Among the internal factors associated at the bivariate level were: drug use, previous history of MDR-TB/TB treatment, having taken MDR-TB treatment for less than one year, history of failure to primary and individualized scheme for MDR-TB/TB ($p<0.05$); and the multivariate analysis confirmed the influence of the factor "history of failure to primary and individualized scheme for MDR-TB/TB". Among the external factors most associated at the multivariate level were: having a history of family contact deceased by XDR-TB/MDR-TB and extra domiciliary contact in the neighborhood deceased by TB ($p<0.05$) It was concluded that the factors associated with the development of extensively resistant tuberculosis are "failure to follow the primary and individualized MDR-TB regimen" and "history of deceased family contact with XDR-TB-MDR-TB".

1. Introduction

Tuberculosis (TB) is a disease caused by bacteria corresponding to the Mycobacterium tuberculosis complex, which usually attacks the lungs, but are not exclusive to them, is considered infectious to have the ability to be transmitted from a sick person to a healthy one, when the first expectora flugge drops in the form of aerosols [1].

TB, according to the World Health Organization (WHO), is among the ten leading causes of death worldwide, being in 2016 an approximate of 1.7 million deaths, of which 0.4 million had association with HIV, likewise, its statistics reveal that about 10.4 million individuals became ill with TB; as for its resistant forms, in the same year, an estimated 480,000 (360,000 to 600,000) incident cases of MDRTB, where approximately 190,000 (120,000 to 260,000) would have died from this cause.

Only 123,000 cases were diagnosed, of which 12,240 developed XDR-TB (9% of MDR-TB cases), where countries such as India, the People's Republic of China and the Russian Federation account for 50% of all extremely resistant cases; likewise, in Latin America, Peru, is among eight countries reporting presence of XDR-TB in a total of 117 cases, followed by Brazil, Mexico, Haiti, Argentina, Ecuador, Colombia and Bolivia [2], [3].

Likewise, in 2017 our country detected 31,518 incident cases of TB in all its forms, 27 578 incident cases of pulmonary TB, 16 877 cases of smear-positive TB (FP), 1335 of MDR TB and 83 of XDRTB, representing TB morbidity statistics of 99.0 per 100 thousand inhabitants, an incidence of 86.7 per 100 thousand inhabitants and a smear-positive incidence of 53.0 per 100 thousand inhabitants.

[*]Corresponding Author: Mauricio Mamani, maurim130@gmail.com

Peru accounts for 14% of the estimated cases of TB in the Region of the Americas; Metropolitan Lima and Callao report the highest percentage of TB patients in all its forms in the country with 64%, in relation to multidrug-resistant TB (MDR-TB) it reaches 79% and extensively drug-resistant TB (XDR-TB) 70%.

The Metropolitan Lima region is divided into four Integrated Health Network Directorates (DIRIS), which are responsible for the overall health of the population. The Lima City DIRIS is one of the main TB hotspots, as it has the highest concentration of TB cases in all its forms and resistances [4], [5].

The growing incidence of primary resistance and acquired resistance to XDR-TB in the DIRIS Lima Sur, which categorizes it as an epidemiological scenario of high risk of TB transmission, leads us to the particular analysis of the following factors: Prolonged contact with bacilliferous patients and substance abuse (tobacco, alcohol, drugs), whose development is carried out throughout our study and lead us to the decision to execute it and achieve the identification of those risk factors that are associated with the emergence of XDR-TB in MDR-TB patients, and that this useful and accurate information will support the approach of appropriate health policies in order to decrease the morbidity and mortality of the population vulnerable to MDR-TB and XDR-TB and contribute to the control and eradication of TB.

The study of risk factors is of great importance because they will allow to predict or prevent the development of TB disease, which are present in each individual and their variability is according to the socioeconomic development of their nation and the social status they have, so they can be classified into factors attributable to the individual, environmental factors or environment and socioeconomic factors [6].

Within the typology of risk factors they found two: the risk factors of the external environment or associated with the disease and the risk factors of the internal environment or predictors of a disease [7].

The consumption of harmful substances such as tobacco and alcohol are considered health risks and influence the therapeutic regimen of TB, the former being a common risk factor and of great importance in TB, with negative repercussions on public health [8].

The previous treatment of TB in its resistant forms as an incurable disease, together with the inadequate management of the follow-up of the cases, and an untimely and ineffective treatment, generated the mutation and appearance of new strains with greater resistance, thus the emergence of XDR-TB [9].

The deficiencies in the therapeutic regimen and its non-compliance, bring as a consequence the appearance of more resistant forms of TB (XDR-TB), which circulate freely in the community, increasing the transmission of these strains and their potential increase in morbidity and mortality [10]. XDR-TB statistics do not provide accurate official figures, but each country reports high and increasing figures for the lethality of this type of TB [11].

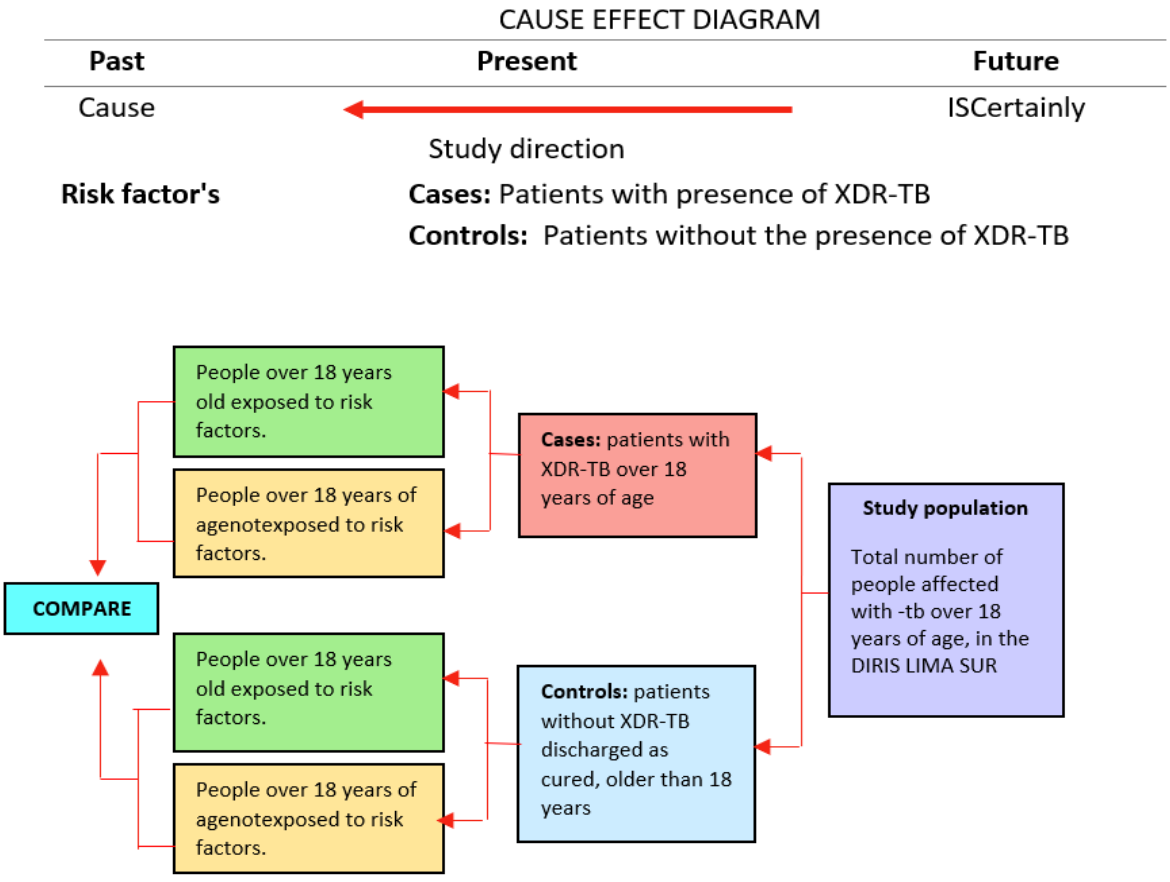


Figure 1: Cause Effect Diagram

The progressive emergence of XDR-TB is of great concern because of its contagiousness and the difficulty to combat and control, which classifies it as potentially irreversible, where the vulnerable population is the most affected, followed by the economically active, in addition to conditions such as previous history of having received treatment and epidemiological background that allow a mechanism of transmission to the person, family and community, exacerbating the deterioration of the health of the entire population; this situation leads to great repercussions not only at the individual level but economic and social for the country, considering it as a public health problem.

To date, XDR-TB remains in the new millennium, the most dangerous human infectious disease, generating high numbers of deaths in the world, however, the production of resistant strains, is not yet defined to what extent would constitute a health problem today, so this article represents the extension of the work originally presented at the 4th International Conference on Medical and Health Informatics [12].

2. Analysis y Method

2.1. Type of research

A retrospective case-control study was developed, with quantitative approach, observational and analytical level, where we identified the risk factors of extensively drug-resistant TB (XDR-TB), applying the epidemiological method [13], [14].

2.2. Variables

The factors are events or phenomena from different sources that predispose people, belonging to a certain region, whose effects could generate the initiation of a disease, that is, that individual who is exposed to a greater number of risk factors will have a greater chance of developing a disease in relation to another less exposed, therefore, will have a high health risk in contrast to one who is away from those factors. [7] A risk factor is an agent that interacts with others; it is a series of detectable particularities in an individual or group way in people, related to the possibility that they are potentiated in influencing the state of health of people. These factors are usually accompanied by three elements and an exposure that predisposes to the development of a certain result (a disease) [15].

Risk factors associated with XDR-TB: These are those characteristics discernible in individuals or groups that are associated with an increased likelihood of developing an unfavorable effect on their health. That is, the exposed individual contracts and develops extensively resistant tuberculosis (XDRTB), this has extended the pattern of resistance to at least two of the most effective drugs of each second line group: the fluoroquinolone group and the second-line injectable group; having among them ciprofloxacin, levofloxacin and moxifloxacin for the first group and amikacin, kanamycin or capreomycin for the second group respectively. [16]

Presence of XDR TB: patients diagnosed with XDR TB, with positive smear or culture results and with a susceptibility panel showing resistance to the most potent first-line antituberculotic drugs and to a fluoroquinolone and aminoglycoside as second-line drugs. TB in this form of presentation has virtually no cure, reflects a high mortality rate and keeps the same transmission mechanism as in its simple form, being its potential affected also their family and close environment [17].

2.3. Population, sample and sampling

The population confirmation for this study was 158 MDR-TB patients diagnosed between 2016 and 2017; and 24 cases who were diagnosed as XDR-TB under treatment. The study area covered health facilities belonging to the Directorate of Integrated Health Networks Lima South (DIRIS). This total was subdivided into two groups of cases and controls. Using Open Epi software, a sample of 24 cases and 48 controls was obtained for unpaired case-controls [18], [19]. For the control group, simple random probability sampling was used, because as this is a subgroup of the population, each constituent element has the same possibility of choice.

2.4. Instruments

The survey was used as a technique and the questionnaire as an instrument, which was applied to both groups: control and cases, in some cases support was sought from medical records, follow-up records, treatment cards and home visits for verification of the information. The validity of the instrument was obtained through expert judgment made up of thematic, methodological and statistical specialists with academic degrees of Doctorate, likewise, it was validated by the technical standard and in other research studies carried out in Peru [20] (The instrument has a reliability of 0.730 (acceptable higher than 0.600) for all items.

2.5. Data collection and analysis procedure

The study began by reviewing the MDR-TB and XDR-TB case follow-up books of thirty-five (35) health facilities of the DIRIS Lima Sur that have a National Health Strategy for the Prevention and Control of Tuberculosis (ESN-PCT), and then selected those patients who met the inclusion criteria and through home visits proceeded to the application of the instruments after informed consent of the participants, having knowledge of the objectives and aims of the study. The data collected were processed by means of the statistical software STATA version 15, and presented in tables, parameters or measures for a better description and analysis of each study variable. In the hypothesis testing, the Odds Ratio (OR) was used in both simple logistic regression and multivariate regression, with a confidence interval (CI) of 95%.

3. Result

3.1. Descriptive analysis

Table 1: Distribution of sociodemographic factors of patients with XDR-TB and MDR-TB in DIRIS Lima Sur, 2017

		Cases		Controls		Total	
		n_1	%	n_2	%	n	%
Age	10 – 19	1	4.2	4	8.3	5	6.9
	20 – 25	5	20.8	13	27.1	18	25.0
	26 – 60	17	70.8	25	52.1	42	58.3
	61 to more	1	4.2	6	12.5	7	9.7

Genre	Male	14	58.3	23	47.9	37	51.4
	Female	10	41.7	25	52.1	35	48.6
Level of education	Primary	5	20.8	17	35.4	22	30.6
	Secondary	15	62.5	18	37.5	33	45.8
	Technical superior	4	16.7	8	16.7	12	16.7
	Superior Univ.	0	0.0	5	10.4	5	6.9
Occupation	Does not work or study	14	58.3	19	39.6	33	45.8
	Work	7	29.2	16	33.3	23	31.9
	Study	2	8.3	8	16.7	10	13.9
	Other	1	4.2	5	10.4	6	8.3

Source: DIRIS Lima Sur, 2017

Table 1 describes the sociodemographic factors of the patients enrolled in the study. It can be seen that the age from 26 to 60 years is found in 70.8% of the cases and 52.1% of the controls, 58.3% of the cases were males and 52.1% of the controls were females, the secondary education level is found in 62.5% of the cases and 37.5% of the controls and the condition of not working or studying is found in 58.3% of the cases and 39.6% of the controls.

Table 2: Distribution of clinical factors of patients with XDR-TB and MDR-TB in DIRIS Lima Sur, 2017.

		Cases		Controls		Total	
		n_1	%	n_2	%	n	%
Body Mass Index	Normal	17	70,8	32	66,7	49	68,1
	Overweight	7	29,2	12	25,0	19	26,4
	Obesity	0	0,0	4	8,3	4	5,5
Smear test result (BK)	+	9	37.5	25	52.1	34	47.2
	++	7	29.2	13	27.1	20	27.8
	+++	8	33.3	10	20.8	18	25.0
Cultivation result	+	8	33.3	23	47.9	31	43.1
	++	13	54.2	13	27.1	26	36.1
	+++	3	12.5	12	25.0	15	20.8
Condition of entry	Previously	14	58.3	20	41.7	34	47.2
	Never treated	10	41.7	28	58.3	38	52.8
Hematocrit result	Under	7	29,2	8	16,7	15	20,8
	Normal	17	70,8	40	83,3	57	79,2

Source: DIRIS Lima Sur, 2017.

Table 2 shows that normal BMI was present in 70.8% of the cases and 66.7% of the controls, smear microscopy (+) was present in 37.5% of the cases and 52.1% of the controls, on the other hand, the culture result for 54.2% of the cases was (++) and 47.9% of the controls, (+). As for the admission condition, 58.3% of the cases were previously treated and 41.7% of the controls were never treated. Normal hematocrit was present in 70.8% of cases and 83.3% of controls.

Table 3: Distribution of internal risk factors associated with MDR-TB in DIRIS Lima Sur, 2017.

			Cases		Controls		Total	
			n_1	%	n_2	%	n	%
Harmful Habits	Alcohol consumption	Yes	5	20.8	6	12.5	11	15.3
		No	19	79.2	42	87.5	61	84.7
	Drug Use	Yes	7	29.2	4	8.3	11	15.3
		No	17	70.8	44	91.7	61	84.7
	Tabaco use	Yes	4	16.7	4	8.3	8	11.1
		No	20	83.3	44	91.7	64	88.9
History of comorbidities	Positive HIV serology	Yes	2	8.3	1	2.1	3	4.2
		No	22	91.7	47	97.9	69	95.8
	Diagnosis of Diabetes Mellitus	Yes	5	20.8	4	8.3	9	12.5
		No	19	79.2	44	91.7	63	87.5
	Cortico prior therapy	Yes	2	8.3	0	0.0	2	2.8
		No	22	91.7	48	100.0	70	97.2

Source: DIRIS Lima Sur, 2017.

Table 3 shows the distribution of internal risk factors for the appearance of XDR-TB. Regarding harmful habits, 79.2%, 70.8% and 83.3% of the cases and 87.5%, 91.7% and 91.7% of the controls did not consume alcohol, drugs or tobacco, respectively.

The history of comorbidity such as a history of positive HIV serology, diagnosis of diabetes and use of corticotherapy were not present in the majority of cases or controls.

Table 4: Distribution of risk factors related to treatment history associated with MDR-TB in DIRIS Lima Sur, 2017.

			Cases		Controls		Total	
			n_1	%	n_2	%	n	%
Treatment History	Previous TB/ MDR-TB treatment	Yes	17	70.8	19	39.6	36	50.0
		No	7	29.2	29	60.4	36	50.0
	Poor adherence to TB/ MDR-TB treatment	Yes	9	37.5	10	20.8	19	26.4
		No	15	62.5	38	79.2	53	73.6
	TB/ MDR-TB treatment less than one year	Yes	11	45.8	6	12.5	17	23.6
		No	13	54.2	42	87.5	55	76.4
	Failure to adhere to primary and individualized TB/TB-MDR-TB régimen	Yes	12	50.0	4	8.3	16	22.2
		No	12	50.0	44	91.7	56	77.8
	TB/MDR-TB treatment dropout	Yes	4	16.7	7	14.6	11	15.3
		No	20	83.3	41	85.4	61	84.7
	Early relapse in less than 6 months at discharge as cured of TB/MDR-TB	Yes	2	8.3	4	8.3	6	8.3
		No	22	91.7	44	91.7	66	91.7
	Previous Fluoroquinolone treatment	Yes	1	4.2	0	0.0	1	1.4
		No	23	95.8	48	100.0	71	98.6
	Self-administered home treatment for TB/TB-MDR.	Yes	2	8.3	0	.0	2	2.8
		No	22	91.7	48	100.0	70	97.2
	Adverse reactions to antituberculosis treatmen	Yes	1	4.2	1	2.1	2	2.8
		No	23	95.8	47	97.9	70	97.2

Source: DIRIS Lima Sur, 2017.

Table 4 shows that 70.8% of the cases were previously treated for TB/MDR-TB, 79.2% of the controls did not have poor adherence to TB/MDR-TB treatment. In 54.2% of the cases and 87.5% of the controls, no less than one year of previous treatment was required; 83.3% of the cases and 85.4% of the controls did not have a history of abandonment of MDR-TB/MDR-TB treatment; no early relapse within 6 months of discharge as cured was present in 91.7% of cases and controls; self-administration of treatment was not present in 91.7% of cases and 100% of controls; and the presence of RAFAS was not present in the majority of cases and controls (>95%).

Table 5: Distribution of external risk factors associated with MDR-TB in DIRIS Lima Sur, 2017.

			Cases		Controls		Total	
			n_1	%	n_2	%	n	%
Contact history	Family member at home with XDR-TB/MDR-	Yes	14	58.3	27	56.3	41	56.9
		No	10	41.7	21	43.8	31	43.1
	Had family members who died of XDR-TB/MDR-TB	Yes	10	41.7	7	14.6	17	23.6
		No	14	58.3	41	85.4	55	76.4
	Have you had neighbors or friends in the neighborhood with tuberculosis?	Yes	11	45.8	25	52.1	36	50.0
		No	13	54.2	23	47.9	36	50.0
	Have you had neighbors or friends in the neighborhood who died from TB?	Yes	5	20.8	2	4.2	7	9.7
		No	19	79.2	46	95.8	65	90.3
	Prison history for the past two years.	Yes	5	20.8	4	8.3	9	12.5
		No	19	79.2	44	91.7	63	87.5
	Previously worked in hospital or was a health science student	Yes	0	0.0	3	6.3	3	4.2
		No	24	100.0	45	93.8	69	95.8
	Hospitalized in the last two years before becoming ill with TB, for more than 15 days.	Yes	1	4.2	1	2.1	2	2.8
		No	23	95.8	47	97.9	70	97.2

Source: DIRIS Lima Sur, 2017.

Table 5 presents the external factors associated with XDR-TB. Intradomiciliary contact with XDRTB/MDR-TB was present in 58.3% of the cases and 56.3% of the controls; history of a deceased family member with XDR-TB/MDR-TB was not present in 58.3% of the cases and 85.4% of the controls. The majority of cases and controls did not have out-of-home contact with TB, were not former prisoners did not work in health facilities, were not students of health sciences, and had not been hospitalized in the last two years for TB.

3.2. Hypothesis testing

General hypothesis

H0: Risk factors are not significantly associated with extensively resistant tuberculosis in multidrug-resistant patients. in the Health Directorate Lima Sur, 2017.

H1: Risk factors are significantly associated with extensively resistant tuberculosis in multidrug-resistant patients. in Health Directorate Lima Sur, 2017.

The association of risk factors with the development of XDR-TB was contrasted from the sub-hypotheses, to which each factor separately was applied simple logistic regression (bivariate) and multivariate, for subsequent determination of the existence or non-existence of significant association; using the Odds Ratio (OR) with a confidence interval (CI) of 95% and considered as a statistically significant result at a level □=0.05

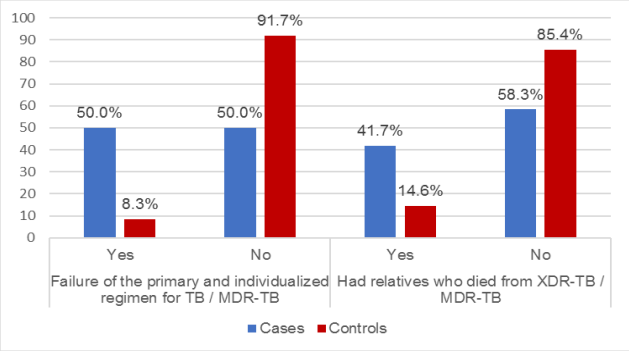


Figure 2: Factores de riesgo internos y externos asociados significativamente a la TB-XDR

According to the figure, 50% of the cases if they failed the primary and individualized scheme for TB / MDR-TB and 41.7% of the cases if they had relatives who died from XDR-TB / MDR-TB; in controls, 91.7% did not fail the primary and individualized regimen for MDR-TB / TB and 85.4% had no relatives who died from XDR-TB / MDR-TB.

Specific Hypothesis 1

H0: Internal risk factors are not significantly associated with the presence of extensively drug-resistant tuberculosis in multidrug-resistant patients.

H1: Internal risk factors are significantly associated with the presence of extensively drug-resistant tuberculosis in multidrug-resistant patients

Table 6 shows that the factors associated with XDR-TB in the DIRIS Lima Sur were: drug use, previous history of having used MDR-TB/TB treatment, having received MDR-TB treatment for less than one year and history of treatment failure for MDR-TB/TB, since their p-value is < 0.05, rejecting the H0 and concluding that these factors are significant. The multivariate analysis reaffirmed the influence of the variable: history of failure to primary and individualized treatment for MDR-TB/TB. In this sense: The odds of presenting XDR-TB in the group of cases with failure to the primary and individualized scheme for MDR-TB/TB is 6.81 times the odds of presenting XDR-TB. In the controls with primary and individualized TB/TB-MDR-TB scheme failure.

Table 6: Statistical test of internal risk factors associated with XDR-TB in DIRIS Lima Sur, 2017.

		p value	Bivariate OR	Bivariate IC95%	Multivariate OR	Multivariate IC95%
Harmful Habits	Alcohol consumption	0,354	1,84	0,49-6,79		
	Drug use	**0,021**	4.53	1,17-17,47		
	Tabaco use	0,289	2,2	0,49-9,69		
History of comorbidities	Positive serology for HIV	0,211	4,27	0,37-49,68		
	Diagnosis de Diabetes Mellitus	0,131	2,89	0,69-11,98		
	Corticosteroid therapy	0,108*	-	-		
Treatment History	Previous TB/ MDR-TB treatment	**0,012**	3,71	1,29-10,63		
	Poor adherence to TB/ MDR-TB treatment	0,130	2,28	0,77-6,72		
	TB/ MDR-TB treatment less than one year	**0,002**	5,92	1,83-19,15		
	Failure to adhere to primary and individualized TB/TB-MDR-TB régimen	**0,000**	11	3,00-40,33	6,81	1,39-33,27
	TB/MDR-TB treatment dropout	0,817	1,17	0,31-4,47		
	Early relapse in less than 6 months at discharge as cured of TB/MDR-TB.	0,108*	1	0,17-5,89		
	Previous Fluoroquinolone treatment	0,333*	-	-		
	Self-administered home treatment for TB/TB-MDR.	0,108*	-	-		
	Adverse reactions to antituberculosis treatment	0,559*	2,04	0,12-34,16		

Source: DIRIS Lima Sur, 2017

Specific Hypothesis 2

H0: External risk factors are not significantly associated with the presence of extensively drug-resistant tuberculosis in multidrug-resistant patients.

H1: External risk factors are significantly associated with the presence of tuberculosis. extensively resistant in multidrug-resistant patients.

Table 6: Statistical test of external risk factors associated with XDR-TB in DIRIS Lima Sur, 2017.

		p-value	Bivariate		Multivariate	
			OR	IC95%	OR	IC95%
Contact history	Family member at home with XDR-TB/MDR-TB	0,866	1,09	0,40-2,94		
	Had family members who died of XDR-TB/MDR-TB	**0,011**	4,18	1,34-13,09	9,73	2,00-47,26
	Have you had neighbors or friends in the neighborhood with tuberculosis?	0,617	0,78	0,29-2,08		
	Have you had neighbors or friends in the neighborhood who died from TB?	**0,024**	6,05	1,08-33,96		
	Prison history for the past two years.	0,131	2,89	0,69-11,98		
	Previously worked in hospital or was a health science student	0,546*	-	-		
	Hospitalized in the last two years before becoming ill with TB, for more than 15 days.	0,559*	2,04	0,12-34,16		

Source: DIRIS Lima Sur, 2017.

Table 7 shows that the external factors associated with XDR-TB in the DIRIS Lima Sur were: history of a family member who died of XDR-TB/MDR-TB and having friends or neighbors in the neighborhood who died of TB, due to their p-value < 0.05, rejecting the H0 and concluding that these factors are significant. The multivariate analysis reaffirmed the influence of the variable: history of a family member who died from XDR-TB/MDR-TB. In this sense: The odds of presenting XDR-TB in the group of cases who had relatives who had died of XDR-TB/MDR-TB is 9.73 times the odds of presenting XDR-TB in the controls who had relatives who had died of XDR-TB/MDR-TB.

4. Discussion

When analyzing our results, we found that according to sociodemographic factors of the 72 patients with XDR-TB and MDR-TB in the DIRIS Lima Sur; taking into account the majority of cases, 70.8% are between 26 to 60 years old, 58.3% are male, 62.5% have secondary school education and 58.3% neither work nor study. As for the majority of controls, 52.1% were between 26 and 60 years of age, 52.1% were women, 45.8% had secondary education and 45.8% did not work or study. The results are similar in some respects to those of other studies in which the variables age, age, sex, sex, and age related variables were crossed. [21] Where the variables age, sex and occupation were crossed. Likewise with the study of [22] where the variables age, sex and occupation were considered as predictors of the appearance of XDR-TB and MDR-TB. It is also worth mentioning the study of [23] where they show age greater than or equal to 40 years as a significant protective factor. Another study by [24] Another study of the study of the study of the age of the cases shows an average age of 37 years for the case group and 38 years for the controls, as well as 56.2% and 66.7% for the male gender for cases and controls respectively.

The socioeconomic level of a person has an impact on morbidity and mortality in vulnerable population groups and therefore in their country. Studies have shown that a lower socioeconomic status has an impact on the mother's education and that a poor educational level is associated with increasing morbidity rates [25].

One of the great changes in the epidemiology of TB worldwide are the resistant forms of this disease, from mono-resistant TB to XDR-TB, constituting from a great challenge for public health to an incurable disease with very high lethality, especially if it is associated with HIV/AIDS.

The existence of factors such as those found in our study, for the group of cases with a history of a family member who died of XDR-TB, is 9.73 times more likely to develop XDR-TB and the history of having received previous anti-tuberculosis treatment is 6.81 times more likely to develop XDR-TB than in the control group. These epidemiological factors predispose morbidity and mortality in the economically active population.

The internal factors significantly associated with the development of XDR-TB were drug use, previous history of having received MDR-TB/TB treatment, having received previous MDR-TB treatment for less than one year, and multivariate analysis confirmed the significant influence of the variable failure to primary and individualized MDR-TB/TB treatment. The study of [20] study highlights that the failure of an antituberculosis treatment scheme consisting of drugs such as fluoroquinolone and a second-line injectable has a strong association with XDR-TB. Another study by [22] concludes that one of the factors associated with multidrug resistant TB is a history of previous TB treatment. Also in the research [26] also stated that problems or failure to adhere to treatment is more strongly associated with XDR-TB as well as inadequate screening, among others. Another study [21] highlights a history of anti-TB treatment as a risk factor for the development of MDR TB and XDR TB, and that failure to adhere to treatment regimens is strongly associated with XDR TB [9] is strongly associated with XDR-TB.

Regarding harmful habits, patients with a history of drug use have a higher risk of developing XDR-TB, results that are related to the study of [24] where he reaffirms drug addiction along with deterioration in adherence to treatment as risk factors for the

presence of MDR-TB. The success of a TB control program and the health of the affected person will depend on the degree of adherence to the therapeutic regimen, which is one of the most important aspects in the management of resistant TB, for which the interdisciplinary team of health institutions should formulate patient-centered and oriented strategies in order to seek adherence to supervised treatment and avoid irregularities, in partnership and involvement of grassroots organizations, social actors in the community, and community health workers [17].

The external factors with significant association to the development of XDR-TB found were extradomiciliary contact with someone who died of TB; and through multivariate analysis, the influence of the variable: family member who died of XDR-TB/MDR-TB was reaffirmed, results that are coherent with the study of [23] study in relation to the occurrence of primary MDR-TB in patients from Callao, ratifying as a risk factor having had contact with someone with MDR/XDR-TB and as a protective factor being younger than 40 years old. Other studies, such as [20] highlights as factors with a strong association to the development of XDR-TB the failure of a treatment scheme with second-line drugs and having a close relationship with someone affected by documented XDR-TB, [21] also mentions contact with MDR TB as a risk factor, and in the case-control study, the study also [27] also considers previous contact with MDRTB cases in both groups as a risk factor.

The contact is the person who establishes a close relationship with the affected person; therefore, an extra-domiciliary contact is comprised of partners, friends, neighbors and family members who have a regular relationship with the person with TB and who frequently visit the home. The primary practices for environmental control of TB and its resistant forms such as XDR is to ensure and maximize natural and cross ventilation in the environment, free of obstacles in the entrances, to ensure an exchange of air flow with a minimum of twelve air changes per hour and to ensure natural lighting, these strategies comprise primary prevention [17].

The next level of prevention is secondary, through early detection and timely treatment will ensure the cure and reduction of serious consequences of the disease, so in the case of TB, this is not aimed at reducing the incident cases but to reduce its severity (duration and complications) and lethality. In [28], Likewise, tertiary prevention aimed at rehabilitation seeks to reduce the complications of a disease through therapy and rehabilitative medicine, guaranteeing the reduction of sequelae, disabilities, and suffering, and promoting the adaptation of affected persons and their families to their environment. All these levels of prevention also include good quality medical care, which together will allow the individual to reintegrate into productive activities without being re-exposed to the same risk factors that contributed to getting sick [28].

Our findings showed that, despite the existence of a health strategy for the prevention and control of tuberculosis with defined objectives and treatment schemes approved by technical standards, there are risk factors that are associated with the emergence of extensively resistant TB, which demonstrates deficiencies in the approach to people affected by tuberculosis, which goes beyond an exclusively health approach, but also demographic and social. This situation does not contribute to an adequate isolation of the index case, generating chains of multiplication and transmission of TB and MDR/XDR TB bacilli in the family nucleus and community. In addition, overcrowded public transportation puts at risk the other passengers who share the same reduced and unventilated space, under the same mechanism.

There are not enough human resources for the comprehensive care of TB patients, and the health personnel at the first level of care are responsible for the approach and management of TB, MDR-TB and XDR-TB cases, making early diagnosis and timely treatment of this complex disease difficult. These patients with MDR-TB and XDR-TB receive treatment divided into two doses, both in the morning and in the afternoon; however, most first-level health care institutions only provide care in the mornings, a situation that contributes to the large-scale increase in the morbidity of tuberculosis in the country.

5. Conclusions

There are risk factors that are significantly associated with the development of extensively resistant tuberculosis in multidrug-resistant patients in the Directorate of Integrated Health Networks of Lima Sur, demonstrating that these factors were determinant for its appearance.

Multivariate analysis showed that the association of the internal risk factor: having failed the primary and individualized regimen for MDR-TB/TB is 6.81 times higher risk of developing XDR-TB/MDRTB than in those who did not fail.

The external risk factor: having a deceased family member with XDR-TB/MDR-TB is 9.73 times more likely to develop XDR-TB/MDR-TB than those without a deceased family history; this association was confirmed by multivariate analysis.

5.1. Recommendations

It is suggested to implement and train health professionals in the approach and comprehensive care of the person affected by tuberculosis, as well as in the sustainable work of adherence to supervised treatment.

Implement the family and community health program to reduce the epidemiological profile of tuberculosis as a key strategy in the Directorate of Integrated Health Networks in southern Lima. Also, promote research to develop new diagnostic and susceptibility panel tests and prioritize research with an epidemiological approach

References

[1] M. Raviglione, Tuberculosis: The Essentials, Fourth Edition. Boca Raton: CRC Press 2016.
[2] WHO. Implementation of the end of TB strategy. Genova: WHO 2016.
[3] OPS. Tuberculosis in the Region of the Americas. Retrieved February 28, 2015.
[4] OPS. Tuberculosis. Retrieved from https://bit.ly/2rf0xlG.
[5] MINSA. Tuberculosis intervention and control plan 2018 - 2020. Lima. Retrieved on May 10, 2018.
[6] J. Garcia, Tuberculosis and risk factors. Emerging diseases, 153, 2016.
[7] K. Colimon, Foundations of Epidemiology. Madrid: Diaz de Santos S.A 2008.
[8] WHO. World report on tuberculosis. Retrieved June 12, 2014, from http://www.who.int/mediacentre/factsheets/fs104/es/.N. Cao, M. Štěpnička, M. Burda, "Fuzzy Quantifiers and Compositions of Partial Fuzzy Relations Employing Dragonfly Algebras," in 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 1–6, 2019, doi:10.1109/FUZZ-

IEEE.2019.8858832.
[9] Y. Llamas, Extensively antibiotic-resistant tuberculosis (XDR-TB): therapies used successfully clinically to cure the disease. Journal of Clinical Research, **65**(3), 255-262, 2013. DOI: 10.1016/j.ijantimicag.2009.12.019
[10] J. Robledo, Control of multidrug-resistant tuberculosis: a possible goal. Rev Biomedica, **39**(3), 431-433, 2019. DOI: 10.2165/11595340-000000000-00000
[11] M. Contrera, Drug-resistant tuberculosis: diagnosis and treatment. National Guide for e; tuberculosis management, 147, 2017.
[12] M. Mamani et al., Characterization of risk factors associated with tuberculosis. ICMHI 2020: Proceedings of the 4th International Conference on Health and Medical Informatics (230-234). Ibaraki-Japan: Association for Computing Machinery, 2014
[13] R. Bonita, Basic Epidemiology. Washington, D.C: PAHO, 2017.
[14] R. Hernandez et al., Research methodology (Sixth ed.). Mexico: McGraw Hill, 2018.
[15] CDC. Risk factors for tuberculosis. Obtained from the Center for the Control and Prevention of Nurses: https://www.cdc.gov/tb/esp/topic/basics/risk.htm.
[16] C. Daley, An Issue of Clinics in Chest Medicine. Pennsylvania: Elsevier Health Sciences, 2019.
[17] MINSA. Technical Health Standard for the comprehensive care of people affected by tuberculosis. Lima: MINSA, 2013.
[18] J. Kelsey, Methods in Observational Epidemiology. New York: Oxford University Press, 1986.
[19] J. Fleiss, J. Statistical Methods for Relationships and Proportions. New York: John Wiley and sons, 1981.
[20] M. Ticona, Extremely resistant tuberculosis (XDR-TB), history and current situation. Acta Med Per, 4, 25, 2008.
[21] Y. Libón, Risk factors for the development of multidrug-resistant tuberculosis in patients of the Hospital Nacional "Dos de Mayo. Ricardo Palma University, Lima, Lima, 2017.
[22] D. Asenjo, Factors associated with multidrug resistance in. patients diagnosed with tuberculosis Guillermo Almenara Irigoyen National Hospital. San Martin de Porres University, Lima, 2016.
[23] A. Avalos, Factors associated with primary multidrug-resistant tuberculosis in. An Fac med, 75, 233-246, 2016.
[24] Rodriguez. Risk Factors for Multidrug-resistant Pulmonary Tuberculosis in the La Libertad Region, Peru, 2013.
[25] J. Gofin, Global Community Health, principles and programs in the world. Madrid, 2012.
[26] Y. Gonzales, Extensively antibiotic-resistant tuberculosis (XDR-TB): therapies used successfully clinically to cure the disease. Clinical Research Rev., **60**(3), 2011. DOI: 10.1016/j.ijantimicag.2009.12.019
[27] T. Chang et al., Risk factors in the appearance of multidrug resistance in patients with pulmonary tuberculosis. Rev medic panacea, **4**(2), 31-35, 2014.
[28] OPS. Module of principles of epidemiology for the control of diseases (MOPECE) (Second ed.). Washington D.C .: OPC, 2011.

Annexes

QUESTIONNAIRE TO DETECT RISK FACTORS ASSOCIATED WITH THE PRESENCE OF TB-XDR IN A FIRST LEVEL HEALTH CARE INSTITUTION - DIRIS LIMA SUR, *2017*

	Cases	Controls

I. Socio-demographic characteristics (Intervening variable)

A. Age (in years)

1) 10 – 19		1
2) 20 – 25		2
3) 26 – 60		3
4) 61 plus		4

B. Genero

1) Male		1
2) Female		2

C. Level of education

1) Primary		1	3) Technical superior		3
2) Secondary		2	4) Superior Univ.		4

D. Occupation

1) Not working, not studying		1	3) Study		3
2) Works		2	4) Other		4

E. BMI category.

a) Weight _____ b) Size _____

c) BMI result _____

1) Low weight <18.5		1

2) Normal [18.5 – 25〉		2
3) Overweight ≥25		3

F. Results of smear microscopy

a) BK (+)		1	b) BK (++)		2	c) BK (+++)		3

G. Cultivation result

a) Cultivation (+)		1	b) Cultivation (++)		2	c) Cultivation (+++)		3

H. CONDITION OF ENTRY:

	a) BEFORE TREATED		b) NEVER TREATED

I. Result of number of drugs with resistance IN VITRIO (PS)

1) Resistance (R) 2) Sensitive (S) *(Mark with X)*	R	H	Km	Am	Cm	Lfx	Mfx	Eto	Cs	PAS	Cfz	Amx/Clv	Cpx

J. Haematocrit result

a) Result _____

b) Men: Normal: <41% - 53%>		1
Low: < 41%		2
c) Women: Normal: <36% - 46%>		1
Low: < 36%		2

II. Risk factor associated with TB- XDR-TB-MDR cases and controls

2.1.	Internal case factors and controls		
	A. Harmful Habits	Yes(1)	No (0)
1	Consumed alcohol more than 3 times per week		
2	Used drugs in the past year		
3	Consumed more than 5 cigarettes per week		
	B. History of comorbidities		
4	Showed positive serology for HIV		
5	Diabetes Mellitus		
6	Prior use of corticosteroid therapy		
	C. History of Tuberculosis Treatment		
7	You were previously treated for tuberculosis / MDR-TB		
8	Showed poor adherence to TB/MDR-TB treatment		
9	You were on TB/ MDR-TB treatment for less than 1 year		
10	You failed the primary, individualized TB/TB-MDR-TB regimen		
11	You have previously stopped treatment for TB/MDR-TB.		
12	Had early relapse within 6 months of being discharged as cured of TB/MDR-TB		
13	Previous use of Fluoroquinolones		
14	You received self-administered home treatment for TB- /TB- MDR.		
15	Had adverse reactions to anti-tuberculous treatment (RAFA).		
2.2.	External case-control factors		
	D. History of intra- and extra-home contact	Yes(1)	No (0)
16	Had a family member at home with TB -XDR/TB - MDR		
17	Had family members who died of XDR-TB/MDR-TB		
18	Have you had neighbors or friends in the neighborhood with tuberculosis?		
19	Have you had neighbors or friends in the neighborhood who died from TB?		
20	History of incarceration in a correctional facility within the last two years		
21	Previously worked in a hospital or was a health science student		
22	He was hospitalized in the last two years before getting sick with TB, for more than 15 days.		

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 125-140 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Machine Learning Algorithms for Real Time Blind Audio Source Separation with Natural Language Detection

Arwa Alghamdi[1], Graham Healy[2], Hoda Abdelhafez[*,1,3]

[1]*College of Computer & Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 12211, Saudi Arabia*

[2]*School of Computing, Dublin City University, Dublin, D09 NA55, Ireland*

[3]*Faculty of Computer & Informatics, Suez Canal University, Ismailia, 41611, Egypt*

ARTICLE INFO

Article history:
Received: 30 April, 2021
Accepted: 10 September, 2021
Online: 27 September, 2021

Keywords:
Machine learning
Audio source separation
Conv-TasNet
Demucs algorithm
Natural Languages

ABSTRACT

The Conv-TasNet and Demucs algorithms, can differentiate between two mixed signals, such as music and speech, the mixing operation proceed without any support information. The network of convolutional time-domain audio separations is used in Conv-TasNet algorithm, while there is a new waveform-to-waveform model in Demucs algorithm. The Demucs algorithm utilizes a procedure like the audio generation model and sizable decoder capacity. The algorithms are not pretrained; so, the process of separation is blindly without any function of three Natural Languages (NL) detection. This research study evaluated the quality and execution time of the separation output signals. It focused on studying the effectiveness of NL in Both algorithms based on four sound signal experiments: (music & male), (music &female), (music & conversation), and finally (music & child). In addition, this research studies three NL, which are English, Arabic and Chinese. The results are evaluated based on R square and mir_eval libraries, mean absolute Error (MAE) scores and root mean square error (RMSE). Conv-TasNet has the highest Signal-to-distortion-Ratio (SDR) score is 9.21 of music at (music & female) experiment, and the highest SDR value of child signal is 8.14. The SDR score of music at (music & female) experiment is 7.8 during the Demucs algorithm, whereas child output signal has the highest SDR score 8.15. However, the average execution time of English experiment of Conv-TasNet is seven times faster than Demucs. For accuracy measure, RMSE indicates absolute values, and MAE handles the errors between observations and prediction signals. Both algorithms show high accuracy and excellent results in the separation process.

1. Introduction

This paper is an extension of work originally presented in conference Novel Intelligent and Leading Emerging Sciences Conference (NILES) [1]. As everybody knows that a human is intelligent and has an aural source separation that helps to identify singular sounds within a mixed sound. The type of mixed sound could be a speech or music, which can be detected via the human aural system; because the brain nerves are trained since birth to differentiate between human speech and music. Artificial intelligence, machine learning algorithms, and other Technological advancements have contributed to the revolution of machines to simulate the human intelligence.

Machine learning algorithms contributed to the development in multiple fields such as signal processing. An approach called Blind Source Separation (BSS) can separate unknown set of source signals and the mixing methodology. Furthermore, BSS recovers all individual sources simultaneously as output from the mixture, which helps in performing a real-time technique. Thus, BSS can process sounds in a real-time using many algorithms such as independent component analysis (ICA), Principal Component Analysis (PCA), Degenerate Unmixing Estimation Technique (DUET), Short Time Fourier Transform (STFT) [2], [3]. Some studies applied this methodology to improve other algorithms such as the Deep Neural Network (DNN) algorithm [4], the Independent Deeply Learned Matrix Analysis (IDLMA)

[*]Corresponding Author: Hoda Abdelhafez, hodaabdelhafez@gmail.com

algorithm [5], the Minimum-Volume Beta-Divergence (NMF) algorithm [6], and others.

The real-time approach can be applied simultaneously when using BSS. This enhances separation by using a synchronization procedure during processing. One of the most significant evaluation criteria for comparison is real-time processing because it prefers the algorithm that has a minimum execution time during the separation process. The execution time value computes the actual time through the separation process.

This paper compares the efficiency of Conv-TasNet and Demucs algorithms which use the BSS approach to separate a mixture of two sounds (human speech and music). The biggest challenges are training and testing the algorithms to distinguish between original sounds with high-quality output signals in real-time, with a small execution time value. Both algorithms implement a synchronization method, which substantially reduces the execution time and applies a multi-channel input audio source as well.

2. Related Work

In [7], the authors applied time-domain methodology with the BSS approach. They presented a real-time approach using a Deep Learning System (DLS), which is a robust speech processing during multi-talker environments. Moreover, the authors used signals on nonnegative encoder outputs through the encoder-decoder framework that implemented a model called Time-domain Audio separation Network (TasNet) [8]. The Wall Street Journal (WSJ0-mix) dataset was used to study speech separation problems using two speakers. The dataset was separated into testing and training sets randomly produced mixture. The evaluation appeared that the DLS was 0.23 milliseconds which is six times faster than the state-of-art based system [9]. Nevertheless, the DLS system is not applicable in this study, because DLS supports a short latency like telecommunication applications and listening devices. Furthermore, the experiment is limited only on speech for robust speech processing environment, which represents different object in this study.

Another study implemented Degenerate Unmixing Estimation Technique (DUET) with the Levenberg-Marquardt algorithm for improving the synchronization and signal ratio [10]. This study aimed to build a new DUET with a real-time application while enhancing the quality of the separation process. They used the Bach10 audio dataset [11], which has ten excerpts of classical music, and the single tracks are obtainable. The random mixing with distinct parameters was used to compute the simulated stereophonic mixtures. Using Signal-to-Interference Ratio (SIR), according to the evaluation, a higher SIR means a higher separation of signals quality where they found that SIR was 0.06. Improving the disjoint orthogonality between prediction and observations sources using a recursive implementation in Levenberg-Marquardt algorithm. This result added value to our research study; in terms of enhancing the quality of separation of the BSS in real-time.

In [12], the authors proposed Time-Frequency (TF) method with BSS by estimating the quantity of TF masks and sources with a current spatial aliasing problem. They used DUET and Observation Vector Clustering (OVC). Also, they developed offline and online approach to deal with spatial aliasing. The online separation algorithm operated on gradient ascent and single frame basis. In addition, it optimized using on gradient ascent search with fixed known of the maximum quantity of sources N. The audio dataset was the (CHAINS corpus) story [13]; the authors chose the six-voice signals of the story. Their experiments were implemented simulating a cocktail party situation. Higher distance-miking is the best for the separation performance. The online algorithm was able to track moving sources by estimating of a fixed known of the maximum quantity of sources. The authors mentioned that in mildly reverberant environments, both offline and online algorithms show a good separation performance. However, they do not mention the processing speed of online or offline algorithms.

BSS approaches can be implemented in real-time on hearing aid systems [14]. It is possible to extract and reproduce the target component from mixed binaural sounds using a pocket-size real-time BSS module. The author compared Single-Input Multiple-Output (SIMO) of ICA module with SIMO-of binary masking on convolutive mixtures of speech. Then, produce binaural sounds from the decomposition of mixed binaural sounds. Their experiment evaluated both objectively and subjectively, using 9 users (8 male and 1 female). The distance between the sound source and user was 1.5 m. Furthermore, the speech sample was limited to 9 s and the sampling frequency was 8 kHz. The performance of the modules was comparatively high because of ICA value estimation. Besides, noise reduction rating NRR proposed as a better method rather than conventional methods. However, the authors did not mention more details about their datasets.

In [15], the authors used DUET algorithm of BSS in their research study. They proposed random number of sources while using only two mixtures. Tracking mixing parameters via an online algorithm was applied in a Maximum Likelihood based gradient search method. They generated a mixture dataset using different noises and voices, and they recorded varying angles in echoic and anechoic rooms. Anechoic and echoic rooms of 1500 tested mixtures, anechoic and echoic were 15 dB and 5 dB respectively, of office mixtures. The evaluation environment was a 750 MHz laptop computer, which run more than five times faster than real-time. They noticed that the separation results improved as the angle difference increased. The dataset of authors is completely different from our paper.

The Support Vector Machine (SVM) algorithm is also the solution of BSS under-determined convolutive problem. A study in [16], the authors applied an audio signals sparsity property as an alternative to the independence or stationery of the sources' expectation. Their test was based on the vocal of male and female. They recorded the audio set in a room using microphones with

0.2 seconds of reverberation time. The rate of sampling was 16KHz, and an interval of recording was 28 seconds. The authors presented visualization of signals, but without including a consistent description. The use of the SVM algorithm proved that the BSS could be enhancing via classification using the sparsity property during the separation procedure.

In [17]. the authors proposed optimization approaches about (DUET) algorithm. They developed and implemented DUET for distant speech recognition engine. The authors used two standard techniques for BSS real time: omni-directional microphones and soundcard at reverberation environments. Also, they had several speakers and noise sources, such as a TV. Their experiment focused on two human speakers in a room. They studied the Maximum Microphone Distance (MMD) and became MMD >1m or 2m, but this was a negative effect on the separation quality. However, there are some factors should be considered in a real-time approach when there are multiple sources of a streaming algorithm using position tracking, correlation, arbitrary statistics, condensed metrics. In addition to peak classier processed in 0.55s of execution time and running audio was 9.7s through a single thread with a five-intel core processor. They applied a huge and complex operations to come up with the result. Those techniques will be useful for creating new algorithms which has BSS in a real-time.

On the other hand, in [18], the authors introduced a BSS in real-time oriented to voice separation. They used low-power technique of a processor and synchronized signal for enhancing a real-time methodology. they focused on a cocktail party problem. The sampling rate was 8KHz and the size of data was 16 bit/sample. The audio sets were music, drum, male voice, and more recorded in an audio studio with a professional equipment environment. The authors realized that the algorithm was unable to increase the separation performance using a real-time approach; because of the high quality of separation procedure which requires five times level up iteration. Thus, their research study would be helpful in medical aspects. On the other hand, this research could be implemented in many signals field, mainly when there is a need of a real-time separation.

Demucs implements through following the methodology of WAVE-U-NET algorithm with some disagreements. WAVE-U-NET focuses on manipulating the feature on different period intervals for resampling feature maps frequently. Furthermore, Demucs convert the mixture audio signal into individual original audio signals. Also, Demucs has a technique which is a transferred convolution, that it requires low cost of computation and low memory [19] [20]. Conv-TasNet is a state-of-the-art approach for the separation procedure of speech and music, which helps waveform domain to predict a mask and to improve the whole spectrogram. In [20], the authors speed up the Conv-TasNet during source separation to enhance the execution time which supports a real-time methodology [21].

This paper is an extension of work [1] to study the effectiveness of the Natural Language on the separation process of Demucs and Conv-TasNet algorithms. The experiments' inputs were English, Arabic and Chinese Audioset. This is for studding the ability of the algorithm to separate speech from music blindly, without any previous known of the language type.

In [22], the authors presented the graph-clustering algorithm of Chinese Whispers. The evaluation of their research focused on Natural Language Processing (NLP) issues such as language separation, acquisition of word sense disambiguation, and syntactic word classes. This concludes to employee many graphs in NLP to optimize the separation.

In [23], the authors introduced continuous Arabic speech segmentation. They aimed to use high segmentation accuracies in Concatenative Speech Synthesis and Continuous Automatic Speech Recognition Systems. A tonal language such as Arabic has multiple allophonic phonemes; that will increase the segments computational cost. There are several factors such as ambiguous boundaries between phonemes and allophonic variations, makes recognition more challenging and affect the speech segmentation accuracy. FFT algorithm uses two main methodology frequency and time domains. The evaluation signals are recorded in a constricted environment. The audio sets were Ayat from Qur'an with a duration time 45s. Authors selected 10 audio speeches of 10 readers recorded with no interfering sounds such as noise. The samples file are 128 samples with mono- channel, bit rate 128kbps, type was .wav, audio format is PCM and audio sample size are 16-bit. In segmenting, they had a high segmentation result of intensity value and continuous signals. In the Arabic vowels, the process of the algorithm was accurate at segmentation. This research cannot be compared with the our study results because the differences in algorithms type and propose. The dataset of the experiment is a Small World graphs (SW-graph) [24]. PCW algorithm can process a huge number of graphs in a standard time, however its nondeterministic produces indecisive results.

3. Research Methodology

In this section, we focus on two algorithms for signal separation blindly on a real-time that were used in this research. These algorithms are the Demucs and Conv-TasNet algorithms to compare. The two algorithms are supported and updated in year 2020 by Alexandre Défossez. Moreover, both algorithms are efficient at computational cost and memory process. They are comparable; because they are based on state-of-the-art approaches [20], However, they have an MIT license that can be used and edited. Both Conv-TasNet and Demucs employ machine-learning algorithms to build the models. Also, discuss the evaluation method via the performance of the algorithm and the input and output signals. This research study was extended [1] to apply the Natural Language (NL) of three different languages English, Arabic and Chinese.

TasNet time-domain audio separation network reproduced with the convolution evaluation and encoder-decoder convolutional architecture to come up with Conv-TasNet. The encoder-decoder architecture is like an ICA algorithm operation. The input-signals of the algorithm are mixed sound. The encoder

builds a convolutional block with a particular kernel amount. In addition, the output-signals are a decoders' parameter produced in a non-negativity constraint form. Thus, the computational cost has a significant effect using long short-term memory (LSTM) network in the separation module process. This produces a limitation while training the original TasNet algorithm; because of determining the shorter length of the segment's waveform. Furthermore, LSTM requires a large quantity of parameters which has an immediate effect on expanding the cost of the operational procedure [21].

Conv-TasNet uses entirely time-domain convolution network for sound separation. It contains three stages of memory process: encoding, manipulating and decoding. At the first stage, the encoder module encodes the input signal into overlapping segments with a length (L). The selection depends upon the shorter L of a segment. The output of encoder becomes an input to the manipulating or separation stage, for estimating the mask of a waveform segments. The last stage is decoder, to transform the masks. It reverses the output of separation stage linearly to formulate the number of blocks. Figure 1 shows mathematically the Conv-TasNet process.

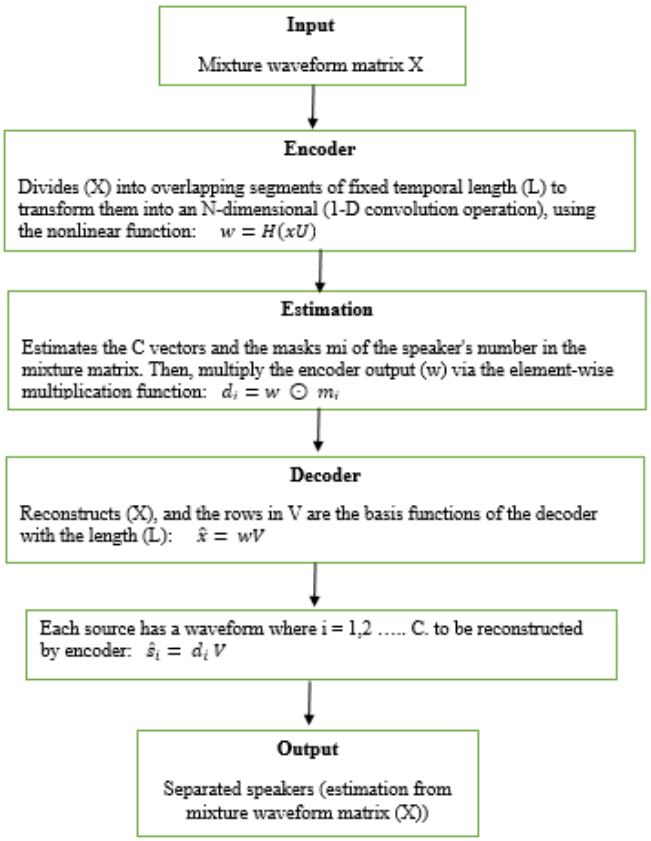


Figure 1: The Stages of Conv-TasNet Algorithm

Conv-TasNet algorithm is the evolution of TasNet algorithm, with application of entirely time-domain convolution network. As shown in figure 1, with the addition of the Temporal Convolutional Network (TCN) methodology. TCN assists in reducing the complexity of the model and the capacity and also reducing the computational cost. It can be applied with or without a real-time approach, minimize the execution time, and the precision of a separation process. In view of all these reasons, Conv-TasNet is suitable for our study. However, the algorithm has some limitation, it is incapable to denoise the signal or remove the artifacts of a signal such as reverberation or vibrations. Nevertheless, using a multi-channel could solve this problem. Additional limitation is that the constant temporal length can failure with lengthy tracking. Accordingly, this algorithm needs to develop in many aspects, such as accuracy and execution time.

Basic approach of Conv-TasNet algorithm is (1-D convolution separation model):

$$D - conv(Y, K) = concat(y_j \circledast k_j), j = 1, \ldots\ldots, N$$

$$S - conv(Y, K, L) = D - conv(Y, K) \circledast L$$

Algorithm Notation:

Y: is the input matrix
K: is the convolution kernel
P: is the kernel size
y, and k: are the rows on matrix Y and K, respectively.
L: is the convolution kernel with size 1.
$\circledast$: is the convolution operation.
D-conv(.): each row of the input Y convolves with the corresponding row of matrix K.
S-conv(.): standard convolution which is a separable convolution.

Figure 2 clarify how the separation process of Conv-TasNet algorithm is in signals form. It shows the separation results on a signal form of Conv-TasNet of (Music &Female) category

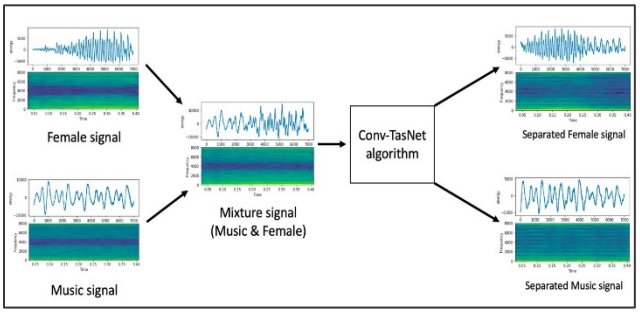


Figure 2: Conv-TasNet Algorithm

On the other hand, Demucs algorithm applies the convolution models, which are encoders, decoders, and LSTM models. Demucs is based on a U-network structure which is identical to the Wave-U-Net structure with differences. However, U-network performs a skip connection between the models, encoder, and decoder to supply a straight connection to the initial signal. Although, U-network does not use linear interpolation but uses a transposed convolution. That helps reducing computational cost and the size of the memory, which enhances the utilization of memory and process four times less in general. Nevertheless, running out of memory is the limitation of the Demucs multi-channel. Figure 3 shows the Demucs process.

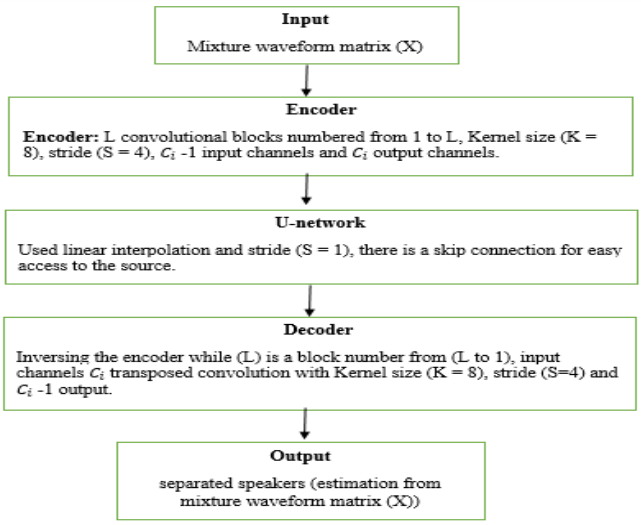


Figure 3: Demucs Stages

There is discrepancy between Conv-TasNet and Demucs. Demucs has a powerful decoder, that increments the kernel size 1 to 3, and the capacity as shown in table 1. Which gives advantage of preserve the lost information while mixing the instruments and unable to retain from the masking. Consequently, the disadvantage of retain information is execution time; because of the subsequent operations which could happen. Ultimately, time-domain and accuracy needs development.

3.1. Demucs General Equation

Basic approach of Demucs algorithm is (U-network structure), it uses Loss function to skip the association between encoder and decoder blocks to get straight transform to the initial input signal phase.

$$L_1(\hat{X}_S, X_S) = \frac{1}{T}\sum_{t=1}^{T} |\hat{X}_{s,t} - X_{s,t}|$$

$$L_2(\hat{X}_S, X_S) = \frac{1}{T}\sum_{t=1}^{T} (\hat{X}_{s,t} - X_{s,t})^2$$

Algorithm Notation:

$X_{s,t}$: *Waveform of source s, t is a subscript which is the* t^{th} *sample*

$\hat{X}_{s,t}$: *Predicted waveform, t is a subscript which is the* t^{th} *sample*

$\frac{1}{T}$: *Number of Sampling*

$L(\hat{X}_s, X_s)$: *is a reconstruction loose uses the average meant squrare error of average absolute error between waveforms.*

Figure 4 clarify how the separation process of Demucs algorithm is in signals form. shows the separation results on a signal form of Demucs of (Music &Female) category.

Table 1 shows the comparison between both algorithms depend on four features. These features are channel, kernel, time-domain, and speed.

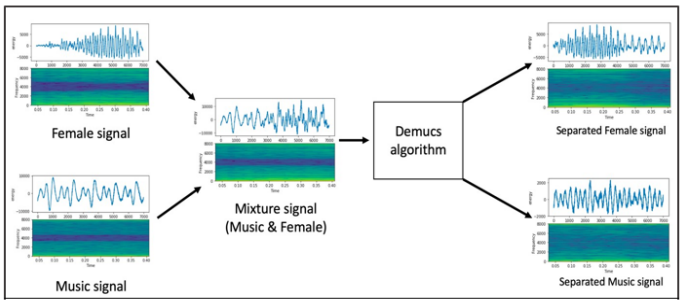


Figure 4: Demucs Algorithm

Table 1: Conv-TasNet & Demucs Coperasion

Features	Conv-TasNet	Demucs
channel	2 multiple channels	(2 or more) Multiple channels
kernel	(K = 1) low capacity	Bigger capacity kernel size (K = 3)
time-domain	offline/online	offline/online
speed	high speed	low speed

The comparison is based on some measurements. First, the SciPy Python library has been used to compute the standard deviation (SD) for the reference source (mixture matrix) and the estimation source (separated matrix). Additionally, the signal-to-noise-ratio (SNR) is a measurement tool used to measure the noise ratio of the signals. Second, the Scikit Learn Python library has been used to calculate the coefficient of determination for constant input arrays instead of accuracy. Moreover, , there is a root mean square error (RMSE) which is a management tool used to measure the difference between predicted and actual values. The mean absolute error tool (MAE) has been used to measure the prediction error for the testing and training datasets. Finally, the actual execution time has been calculated using the Python Time package. Since, this study is based on the BSS approach, it uses a performance evaluation package (mir_eval). It includes evaluation functions for music or audio signals. It has 4 measurement tools: (1) Signal-to-distortion-Ratio (SDR): it computes how is the similarity between the original signal and the produced signals, (2) Signal-to-Artifacts-Ratio (SAR): using artefacts helps in differentiate between the assessment errors, (3) Signal-to-Interference-Ratio (SIR) is the best measure of Gaussian noise and (4) Spatial-Distortion-Ratio (ISR) is A measurement tool usually used in image signals not in audio signal processing field.

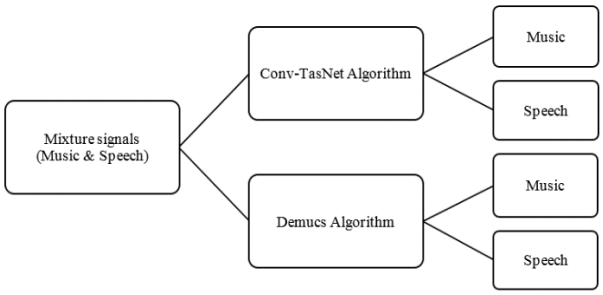


Figure 5: Data Processing for two algorithms

The mixture sound signal could be a mix of Drum, Vocals, Bass, and others. Demucs and Conv-TasNet can be tested on this mixture to get the estimation sound signal. The objective of this research is to study two estimated sources: music and vocals. Thus, the methodologies of these algorithms have changed to conform to the needs, as shown in figure 5. Then, the file of input and output stored the quantitative data as digital metrics at txt file. The digital metric files reduce the execution time and memory process; because they reduce the cost to process digital NumPy array rather than the audio signal.

4. Dataset and Pre-Processing

The dataset in this research project was obtained from youtube.com to compare and evaluate the Demucs and Conv-TasNet algorithms [25]. The data contained in a file with a WAV format. The dataset grouped within five classification experiments: Music with Male, Female, Child, and Conversation. The total size of experiment files is 10GB which contains a duration time of each experiment. The duration time is 3 hours, 28 minutes, and 20 seconds equal to 2GB for every single group. Data selection depends on the negativity of noise ratio, the clarity, and the suitability for instance, pure classic music such as Jazz, Piano, and more. Moreover, pure male sounds such as speech in international conferences like TED talk AND Apple. These sounds have no background noise or interfering sounds. They are a continuous speech extracted from Youtube.com. The factors (SD, SNR) of the pre-processing data were applied on all three languages mentioned before.

4.1. Data Types and signals

- Music (e.g., classical, piano, jazz): Figure 6 shows the jazz music acoustic vibrations.

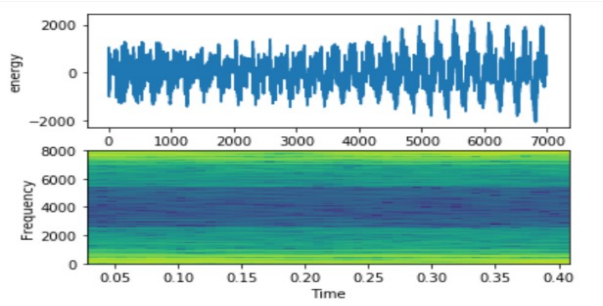


Figure 6: Music Signal on top energy on bottom frequency

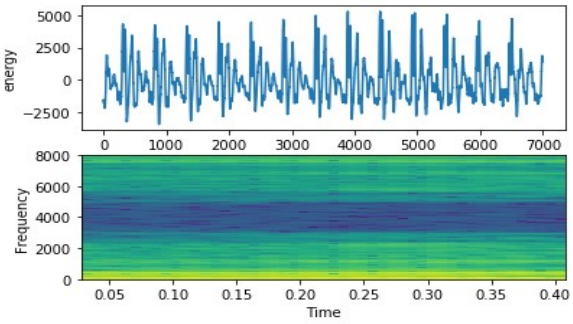


Figure 7: English Female Signal on top energy on bottom frequency

- Female (e.g. Michelle Obama speech): Figure 7 shows a high frequency of female signal like seasonal indices.
- Female (e.g., Balqees Fathi speech in women power conference): Figure 8 shows a high English female frequency of signal like seasonal but sequence indices.

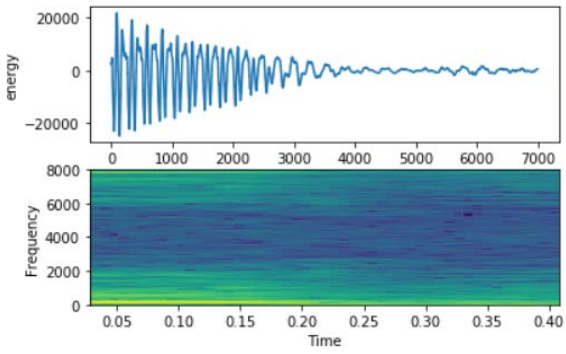


Figure 8: Arabic Female Signal on top energy on bottom frequency

- Chinese Female Figure 9 shows a mid-frequency of signal comparing with other signals.

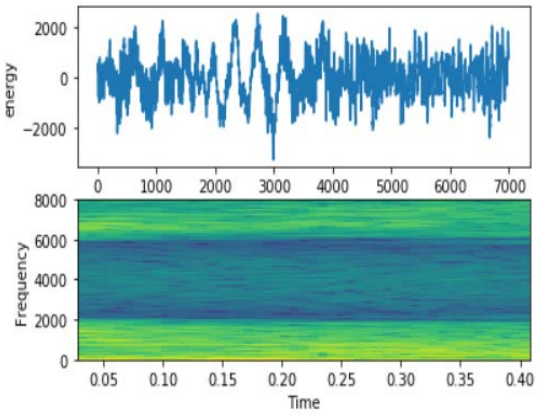


Figure 9: Chinese Female Signal on top energy on bottom frequency

- Male (e.g., male speech from Apple conference): Figure 10 shows a low frequency of male signal. Males' voices have lower intensity than females' voices. This figure shows acoustic vibrations likes a screaming sound.

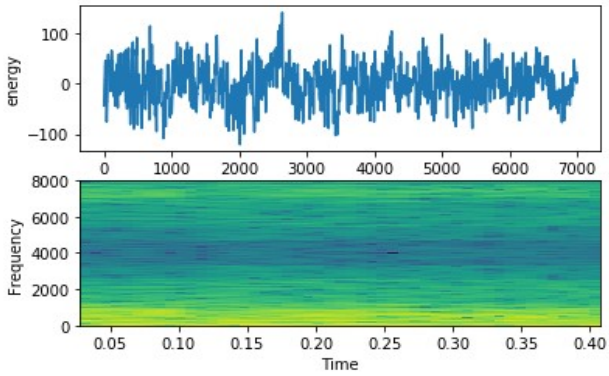


Figure 10: English Male Signal on top energy on bottom frequency

- Male (e.g., male speech from TEDX speech): Figure 11 shows the lowest frequency of male signal. Males' voices. This figure shows acoustic vibrations likes a screaming sound.

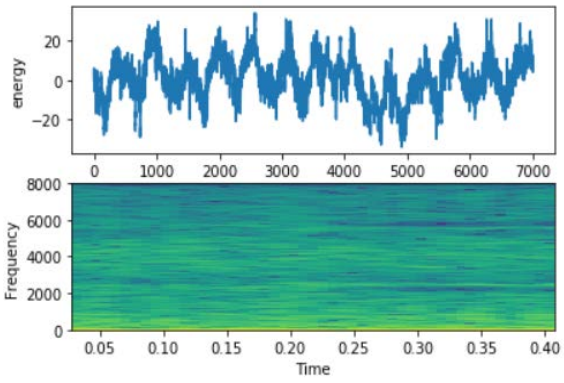


Figure 11: Arabic Male Signal on top energy on bottom frequency

- Chinese Male: Figure 12 shows the highest frequency of male signal. Males' voices. This figure shows more organized vibrations likes a seasonal.

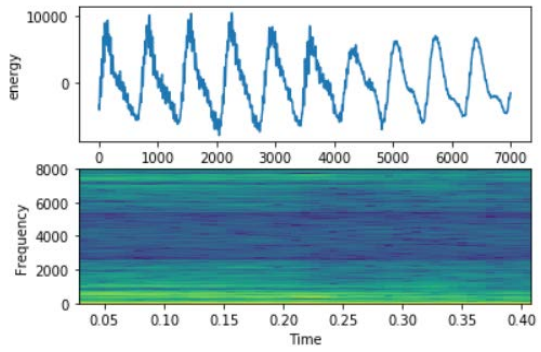


Figure 12: Chinese Male Signal on top energy on bottom frequency

- Conversation (e.g., male and female discussion): Figure 13 shows a mix between male and female voices. The first part is a male vocal while the other part is a female vocal.

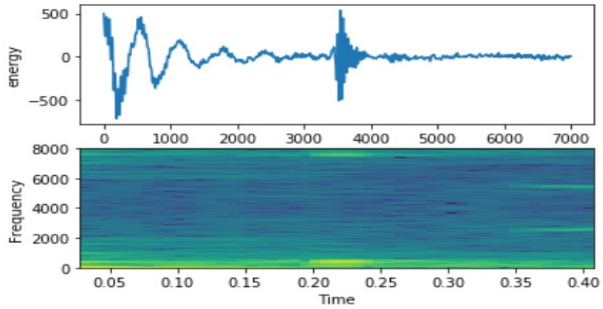


Figure 13: English Conversation Signal on top energy on bottom frequency

- Conversation (e.g., male and female discussion): Figure 14 shows a mix between male and female voices. The first part is a male vocal while the other part is a female vocal. It looks unorganized signal.
- Conversation (e.g., male and female discussion): Figure 15 shows a mix between male and female voices. The first part is a male vocal while the other part is a female vocal. It looks unorganized signal.

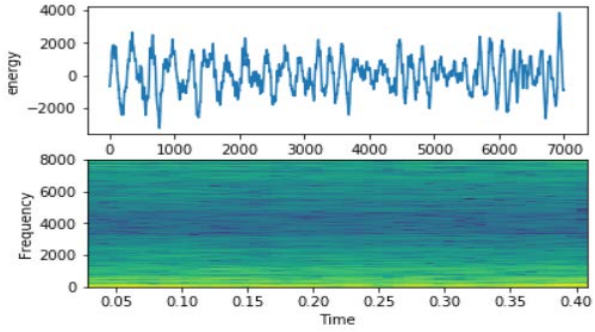


Figure 14: Arabic Conversation Signal on top energy on bottom frequency

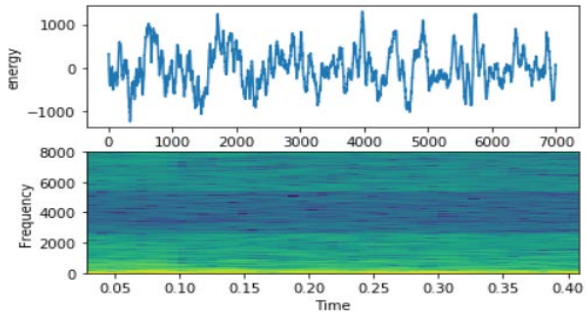


Figure 15: Chinese Conversation Signal on top energy on bottom frequency

- Child (e.g., child stories): Figure 16 shows a high vibration at the beginning, but slightly decreasing at the end.

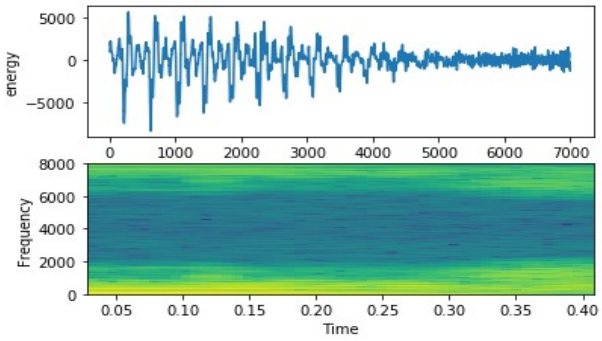


Figure 16: English Child Signal on top energy on bottom frequency

- Child (e.g., child stories): Figure 17 shows a high frequency.

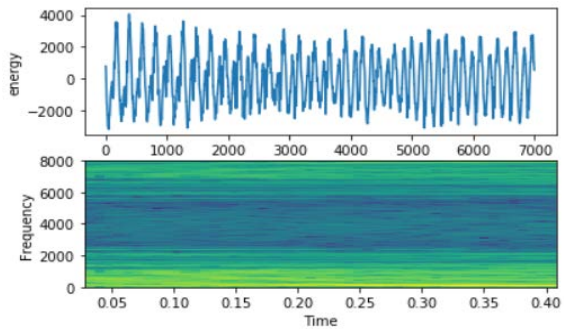


Figure 17: Arabic Child Signal on top energy on bottom frequency

- Child (e.g., child stories): Figure 18 shows a high frequency and heigh vibration

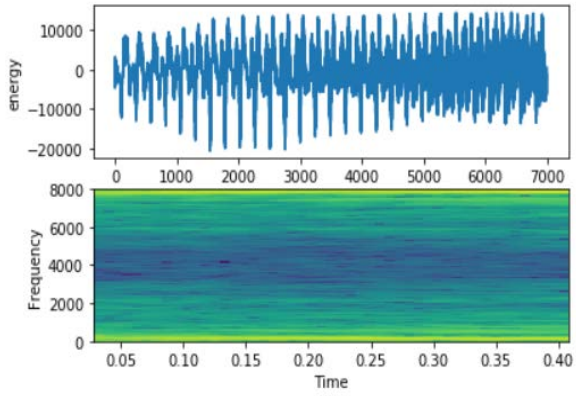


Figure 18: Chinese Child Signal on top energy on bottom frequency

4.2. Data pre-processing

Jupyter and Python libraries have helped produce the input segments of the separation process by a pre-processing, which contains two operations split and merge. The first operation is splitting, which splits the 3 hours audio file of each category into 1250 segments. The second operation is merging, which merges 1250 Music segments with each category one by one. Figure 19 shows the steps of pre-processing data segments. The mixture segment resources have 10 seconds of duration time per segment. Each file consists of two original segment resources (music & male), (music & female), (music & child) or/and (music & conversation).

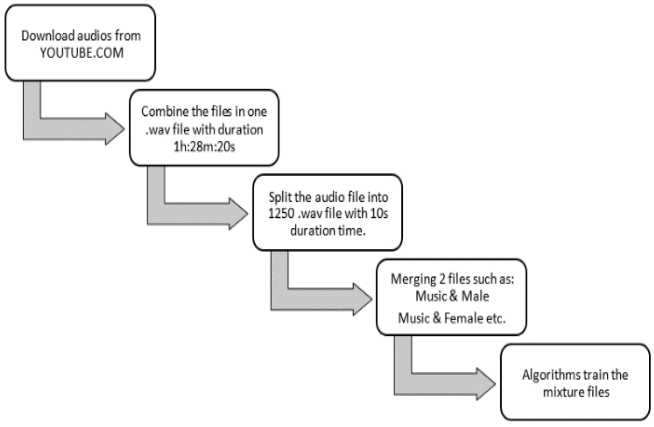


Figure 19: Steps of pre-processing data

The experiments are based on six dataset files. The main file is an Excel sheet contains four columns: Youtube URL (object), start time (float), end time (float), and label (object). There are five files for each category, one of them is a music file contains 1250 Records with IDs and music names [26]. Moreover, Male, Female, Child, and Conversation have files with IDs and their names. This research study evaluated an audio dataset of ten gigabytes. each category had two gigabytes of mixture segments with a music dataset. In addition, these files related to the original files by IDs. For each category of the dataset 70% was used for training and 30% for testing. Numpy is a Python library used for randomizing splitting by randomize function. The experiment started with training data and then testing data of each mixture dataset. Finally, measurement tools were used to evaluate the separation results.

4.3. Audio signals description in terms of SD and SNR

Standard Deviation (SD) is a measurement tool used to compute the data spreads (distribution) via computing the mean of squared deviations, then square root of the result. [25].

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Xi - \mu)^2}$$

Table 2: The mean of standard deviation of all data

Mean	Music	Male	Female	Child	Conversation
English	0.111309	0.052772335	0.0523693	0.078172052	0.0620758
Arabic	0.111309	0.046679	0.022318	0.232547	0.083778
Chinese	0.111309	0.103396	0.107852	0.134417	0.086935

In table 2, the value of all columns = 0.1 SD. The results are closing to zero value, which indicates all the estimation values are near to the mean.

Signal-Noise-Ratio (SNR): is a measuring tool that calculates the ratio between the original signal with the artifact signal such as a noise. It computes the mean of a signal, then dividing the result by the SD [27], [28].

SNR = SD(MEAN(X))

Table 3: Signal-Noise-Ratio

Mean	Music	Male	Female	Child	Conversation
English	0.000257	-2.16667E-06	-4.53405E-6	-9.37E-06	-9.2014E-06
Arabic	0.000257	-3.80E-05	-9.94E-05	0.000168	-3.26E-05
Chinese	0.000257	0.000142	7.59E-06	-9.47E-05	0.000594

Table 3 shows the noise ratio of all categories. All SNR values are negative; this signifies a minimum noise ratio and considers as a pure signal. The negativity of SNR indicates as a high-quality score.

5. Experiments and Results

The audio datasets run on personal computer OMEN hp, i7 Intel core and 16GB of RAM. Using Python language for data processing via Jupyter and Pycharm. Using multiple packages such as: mir_eval Version: 0.5, Sklearn Version: 0.0, Scipy Version: 1.4.1, Pandas Version: 0.25.1, NumPy Version: 1.18.2, youtube_dl Version: 2020.3.24, Pydub Version: 0.23.1 and more. The audio files downloaded from Youtube.com via youtube_dl libraries. This downloading process took more than 2 hours per label. Merging Music files with all labels such as Male files via Overlay function this process takes 1 hour per label. After that, applying the data on Conv-TasNet and Demucs algorithms. For Demucs the process takes 6 hours and for Conv-TasNet the process takes 3 hours per label. The output of each algorithm

produced 12.3 GB augments on the memory process, 1250 folders composed of 3750 files produced from each algorithm. Evaluation of each output for each algorithm takes 2 hours. this calculation for each 1250 WAV mixture files. Four experiments were applied for training and testing datasets, to study the effectiveness or speech classification for comparing the algorithms. Four different data types were used in these algorithms: (Music & Male), (Music & Female), (Music & Conversation) and (Music & Child).

5.1. Evaluation Results

The comparison between Conv-TasNet and Demucs algorithms was based on seven measurement tools using three packages Scikit-Learn, mir_eval and time. Scikit-learn package supports three accuracy metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R square (R2). Mir_eval package affords four tools to evaluate the performance: Spatial-Distortion-Ratio (ISR), Signal-to-Artefacts-Ratio (SAR), Signal-to-Interference-Ratio (SIR), and Signal-to-distortion-Ratio (SDR). Ultimately, the time package was used to calculate the execution time of the separation process.

The three tools were implemented to calculate the accuracy of the Conv-TasNet and Demucs algorithms. They support the quality score evaluation pertaining for observation inputs and estimation outputs. The observations cannot be calculated directly with accuracy because continuous metrics cannot apply to accuracy tools. RMSE, R2 and MAE were used to describe the results of algorithms.

a- R Square (R^2)

It is a statistical tool used to measure the proportion of how close the data points to the fitted regression line. Via subtract the observed data with the data fitted on the line regression. In other words, it tells how data goodness fit the regression model [29] [30].

$$R^2 = \frac{Sum\ of\ Squares_{regression}}{Sum\ of\ Squares_{total}}$$

Training experiments:

Table 4: the experiments results pertaining to the (Music & Male and Music & Female) training dataset of all three languages.

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	*English*	0.594750505	0.554857091	0.652751885	0.706623375
	Arabic	0.906899	0.730184	0.817224	0.252617
	Chinese	0.911638	0.924542	0.890167	0.924055
Conv-TasNet	*English*	0.673642922	0.611274307	0.725606045	0.782790036
	Arabic	0.939895	0.913543	0.91162	0.155181
	Chinese	0.937934	0.936259	0.898883	0.925415

Table 4 shows that Music has an R-squared (R^2) value of 59% and (Male) has an R^2 value of 55%. The R^2 value can be more accurate if the noise ratio is low. For the R^2 values described between 0% and 100%, 0% indicates that the observed data are far from the mean and have a low error estimation, while 100% indicates that the observed data are close to the mean [31].

As it shows that (Music & Male) experiments of both Arabic and Chinese are close to 100% which indicates the observed data are close to the mean. In another experiment, (Music & Female) of Arabic language is close to 0% which indicates the observed data are far from the mean and have a low error estimation.

Table 5: R Square of Conversation and child experiments of Training Dataset of all three languages.

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.641354444	0.682678489	0.624809687	0.763154352
	Arabic	0.938469	0.91497	0.67352	0.922591
	Chinese	0.865062	0.799423	0.884608	0.937521
Conv-TasNet	*English*	0.724467042	0.753158555	0.662892685	0.759285571
	Arabic	0.658099	0.000153	0.857881	1.16E-05
	Chinese	0.875329	0.822089	0.962708	0.477527

Tables 4 and 5 show the English language experiment results of Demucs algorithm, pertaining to the (Music & Female), (Music & Conversation) and (Music & Child) training datasets. The R^2 values of observation and prediction signals of each algorithm, are (Music) 65% and (Female) 70%, (Music) 64% and (Conversation) 68% and (Music) 62% and (Child) 76%, respectively. These results show that the (Music & Female) experiments achieved high R^2 values and that (Child) also achieved a higher value of Demucs of training experiment. While (Music & Male) experiment of Demucs has the lowest value 59% and 55% respectively. On the other hand, Conv-TasNet achieved the highest value of R^2, where (Female) R^2 = 78%, that means Female signals are the most signals close to the mean. In this evaluation it is noticeable that the Conv-TasNet always has the higher value compared to Demucs. Demucs training experiment during Arabic language has the highest value of R square which is 93% at Music (Music & Male) experiment, while it has the lowest speech value 25% at Female (Music & Female) experiment. On another side, Conv-TasNet experiment during Chinese language has the highest value 96% at Music (Music & Child) experiment and Arabic language has lowest value at speech 0.0001% at (Music & Conversation) and (Music & Child) experiments.

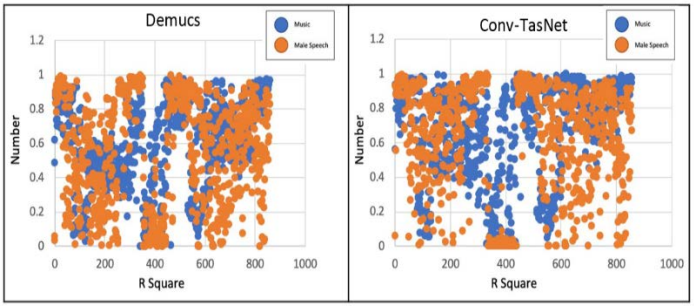


Figure 20: $R^2$of (Music & Male) experiment of Demucs & Conv-TasNet of English language.

On another side, Arabic language with Demucs algorithm achieved the highest value of R-square at speech side (Male, Conversation and Child). Also, Chinese language usually has high value of R-square near to 100% which means better performance, close to the mean and low estimation errors on signals.

The right plot in figure 20 represents the R square of observation and prediction data of (Music & Male) of Demucs. The left plot shows the spread points of observation and prediction data of (Music & Male) of Conv-TasNet of training data. The blue points show the music data and orange points show the males speech data. It explains the spread of data point of observation and prediction data. X-axis represent the R square values of (Music & Male) experiment. The sample data is not the reason for the increased R^2 value because all the experiments used the same number of segments. However, R^2 result may be an effect of the lower noise ratio of observation.

Testing experiments:

Like the training experiment, all values in tables 6 and 7 are above 50%. The highest value of R^2 is 77% of Conv-TasNet pertaining to the (Female) experiment and 67% of (Female) of Demucs algorithm. In training and testing data experiments, training data has the same R^2 with testing data which inference that data size sample has no affection of R^2 value.

Table 6: R^2 of male and female experiments of testing dataset

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	*English*	0.598482775	0.527509406	0.651269443	0.672236996
	Arabic	0.929927	0.90909	0.872887	0.54528
	Chinese	0.917021	0.908275	0.888142	0.911523
Conv-TasNet	*English*	0.677168022	0.659364504	0.712894379	0.774478377
	Arabic	0.927393	0.796538	0.958407	0.882355
	Chinese	0.966593	0.479971	0.85849	0.885474

Table 7: R^2 of Conversation and Child experiments of Testing Dataset

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.645854253	0.681211663	0.624127646	0.756858895
	Arabic	0.8902	0.866712	0.744405	0.669279
	Chinese	0.865062	0.799423	0.893974	0.94875
Conv-TasNet	*English*	0.718706834	0.767224614	0.670057154	0.75645193
	Arabic	0.899566	8.36E-06	0.676389	0.000152
	Chinese	0.510896	0.393926	0.812849	0.910011

Demucs testing experiment during Arabic language has the highest value of R square which is 93% at Music (Music & Male) experiment, while English language has the lowest speech value 53% at speech (Music & Male) experiment. Also, at Arabic language with the value of speech 54% at (Music & Female) experiment. On another side, Conv-TasNet experiment during Chinese language has highest value 97% at Music (Music & Male) experiment; and Arabic language has lowest value at speech 0.000008% at (Music & Conversation) experiment and 0.0001 at (Music & Child) experiment.

b- Root Mean Square Error (RMSE)

where R^2 is the relative measure of the fit of the model, RMSE is the absolute measure of the fit. It is also a tool to measure how accurately the algorithm predicts the response. It can be used instead of accuracy tools because the accuracy tool does not support the continuous array of reference resources [32] [33].

$$RMSE = \sqrt{1 - r^2 SD_y}$$

Training experiments:

Tables 8 and 9 show the RMSE values for the (Music & Male), (Music & Female), (Music & Conversation) and (Music & Child) training datasets. The RMSE values are (Music) 0.06 and (Male) 0.04, (Music) 0.06 and (Female) 0.03, (Music) 0.06 and (Conversation) 0.03 and (Music) 0.06 and (Child) 0.03, respectively of Demucs. Lower values of RMSE indicate absolute fit. Overall, both algorithms have positive values from (0.3 to 0.6) which means that the observations and prediction data are close.

Arabic and Chinese language have lower value relatively with English language. Mostly, the value of RMSE is around 0.3 which indicates the absolute fit of the output signals comparing with English separation of both algorithms.

Table 8: Root Mean Square Error of Male and Female experiments

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Mal	Music	Female
Demucs	*English*	0.060387951	0.035588642	0.055773658	0.026737967
	Arabic	0.028117	0.028008	0.039305	0.031757
	Chinese	0.030426	0.028518	0.034205	0.029987
Conv-TasNet	*English*	0.052309023	0.031107757	0.048070573	0.021558262
	Arabic	0.021331	0.021132	0.025871	0.025568
	Chinese	0.026394	0.025694	0.031611	0.030064

Table 9: Root Mean Square Error of conversation and child experiments.

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.058033883	0.027021274	0.057891011	0.032158433
	Arabic	0.027798	0.024634	0.062237	0.064626
	Chinese	0.039827	0.038917	0.036068	0.033564
Conv-TasNet	*English*	0.04975772	0.020641371	0.053277448	0.032526892
	Arabic	0.061912	0.247073	0.042143	0.173703
	Chinese	0.033942	0.036129	0.019014	0.015994

Testing experiments:

Tables 10 and 11 show the RMSE values for the (Music & Male), (Music & Female), (Music & Conversation) and (Music & Child) testing datasets. Similarly, with training dataset, the values of Conv-TasNet and Demucs are almost the same. Overall, both algorithms have positive values from (0.3 to 0.6) which indicates that observations and prediction data are almost fit.

Like training data, the testing experiment also shows better results of RMSE of both languages Arabic and Chinese.

Table 10: Root Mean Square Error of Male and Female experiments

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	*English*	0.061387858	0.038965244	0.058082848	0.02998228
	Arabic	0.024837	0.018296	0.03204	0.031994
	Chinese	0.024996	0.027324	0.034895	0.029049
Conv-TasNet	*English*	0.052586705	0.029218817	0.05128763	0.020946208
	Arabic	0.023592	0.023737	0.01534	0.014384
	Chinese	0.017425	0.013958	0.032243	0.031023

Table 11: Root Mean Square Error of conversation and child experiments.

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.05703156	0.026870191	0.059313099	0.032708388
	Arabic	0.033467	0.03027	0.041961	0.047127
	Chinese	0.039827	0.038917	0.030169	0.029733
Conv-TasNet	*English*	0.049186272	0.021074537	0.0546344	0.033307789
	Arabic	0.029296	0.156674	0.054295	0.24063
	Chinese	0.071935	0.080743	0.040955	0.040211

c- Mean Absolute Error (MAE)

MAE is a tool for measuring the accuracy of training algorithms. It handles continuous NumPy arrays and can be used to enhance the results of RMSE when measuring the algorithms' predictions. It computes the average magnitude of the errors that occur between predictions and actual observations.

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

Training experiments:

Notably, tables 12 and 13 show (Female) and (conversation) of training dataset experiments, via Conv-TasNet get the smallest value of error 0.01 between the reference and estimation resources. Demucs has 0.04 of (Music) of all experiments and 0.02 of each speech (Male, Female, Conversation and child).

Table 12: Mean Absolute Error of male and female experiments.

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	*English*	0.043319834	0.023187058	0.039918792	0.016335086
	Arabic	0.015769	0.015549	0.02708	0.022367
	Chinese	0.018334	0.016945	0.021714	0.019184
Conv-TasNet	*English*	0.036686502	0.020443172	0.034022512	0.01306448
	Arabic	0.012075	0.011286	0.015174	0.014449
	Chinese	0.016413	0.015721	0.019738	0.018671

Table 13: Mean Absolute Error of conversation and child experiments.

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.041430066	0.016731485	0.040630954	0.018055779
	Arabic	0.017914	0.015582	0.045522	0.047743
	Chinese	36.54131796	14.75716936	35.2270372	15.65436042
Conv-TasNet	*English*	0.034715515	0.012663916	0.035056178	0.016750811
	Arabic	0.045405	0.18224	0.029647	0.121688
	Chinese	30.34135987	11.05559904	0.012242	0.010555

Notably Demucs algorithm at training experiment shows the highest performance of Music separation 0.02 MAE in Arabic language of both (Music & Male) and (Music & Conversation) experiments as well as Chinese language at (Music & Male) experiment. Chinese language has significant values - highest value of error - through (Music & Conversation) and (Music & Child) experiments, these values are (36.5 & 14.8) and (35.2 & 15.7), respectively. Additionally, Conv-TasNet algorithm in training experiment shows the highest performance of Music separation 0.01 MAE in Arabic language at (Music & Male) experiment. However, Chinese language has significant values (30.3 & 11.1) at (Music & Conversation) sequentially which considered as the lowest performance in separation process.

Testing experiments:

Comparably, tables 14 and 15 of testing dataset experiment are like training dataset experiment of MAE values.

Table 14: Mean Absolute Error of male and female experiments.

Algorithm	Mean	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	*English*	0.044283635	0.025880494	0.041179648	0.018042131
	Arabic	0.013289	0.009229	0.020964	0.020834
	Chinese	0.015598	0.016635	15.89534399	6.964262591
Conv-TasNet	*English*	0.0370836	0.019214275	0.035862557	0.012616786
	Arabic	0.013402	0.013194	0.008488	0.007984
	Chinese	46.01418709	25.07100812	13.37673393	4.706061233

Table 15: Mean Absolute Error of Conversation and Child experiments

Algorithm	Mean	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	*English*	0.040552242	0.016646969	0.041642397	0.018309848
	Arabic	0.020308	0.017798	0.029497	0.033534
	Chinese	0.010246	0.007809	0.019657	0.019533
Conv-TasNet	*English*	0.034867332	0.013045045	0.036287197	0.017151002
	Arabic	0.019533	0.102707	0.038631	0.174129
	Chinese	0.056936	0.064269	0.024952	0.024837

In Demucs testing experiment, Arabic language has 0.01 at Music (Music & Male) experiment. Chinese language has 0.01 at Music (Music & Conversation) experiment. Nevertheless, Chinese language has highest value (15.9 & 6.10) at (Music & Female) experiment that indicates a high error rate through separation process. Conv-TasNet testing experiment has lowest error 0.01 in Chinese language at Music (Music & Conversation) experiment. On another side, Chinese language has a high error after separation process; the values are 46.0 & 25.1 at (Music & Male) experiment and 13.4 & 4.7 at (Music & Female) experiment.

5.2. Mir_eval of (SDR, SAR, SIR, ISR)

It is a Python package used to evaluate the results, which retravel machine learning algorithms. It extracts the music separation information pertaining to reference and estimation resources. Mir_eval package quantitatively compares the signals to algorithm implementations. The original mir_eval is a bss_eval package, which has been developed to evaluate the audio and image signals [34].

Training experiments:

Table 16: mir_eval of male and female experiments

Algorithm	SDR	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	English	6.365845562	2.261560323	7.724350679	5.891797598
	Arabic	14.34474	4.531938	11.95328	-4.01775
	Chinese	12.42069	11.31496	12.1474	11.05148
Conv-TasNet	English	7.768486164	2.785485859	9.208170136	7.873549345
	Arabic	14.66376	11.28408	14.29149	-5.51784
	Chinese	13.49518	12.34095	12.57517	11.18431
SIR					
Demucs	English	6.365845562	2.261560323	7.724350679	5.891797598
	Arabic	14.34474	4.531938	11.95328	-4.01775
	Chinese	12.42069	11.31496	12.1474	11.05148
Conv-TasNet	English	7.768486164	2.785485859	9.208170136	7.873549345
	Arabic	14.66376	11.28408	14.29149	-5.51784
	Chinese	13.49518	12.34095	12.57517	11.18431

SDR is the signal distortion ratio. It describes how similar the estimation resources are to the reference resources. SDR provides a global performance measure; but the other three measures are also important. Tables 16 and 17 show the average SDR values, taken from all the tracks. In these tables, Conv-TasNet produces a higher SDR score for the (Female & Music) experiment, and for the (Child) training data set where (Child) = 8.3 SDR'. In turn, Conv-TasNet produces a lower SDR score for the (Child & Music) training data set where (Music) = 6.6 SDR for the (Music) training dataset and a smaller SDR value where (Male) = 2.8 SDR for the (Male & Music). On the other side, Demucs produces a higher value where (Music) = 7.7 SDR for the (Female) training dataset and where (Child) = 8.1 SDR for the (Music & Child) experiment. However, it produces a lower SDR score where (Music) = 6.4 SDR and 6.5 SDR, respectively, for the (Male) and (Child) training datasets and where (Male) = 2.3 SDR for the (Male) training dataset.

Comparing between the language experiments at Table 19, The highest SDR is in an Arabic experiment. Where Music values around 11 to 14 SDR and Male speech, but Female speech has the lowest value -4 and -5 SDR in both algorithms. While Chinese experiment has a heigh value 11 to 12 at Music and Male or Female speech.

Table17: mir_eval of conversation and child experiments

Algorithm	SDR	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	English	7.332337909	4.889571299	6.460674567	8.145132574
	Arabic	13.08034	10.58806	3.800309	11.54092
	Chinese	10.14399	6.673121	9.924395	12.20906
Conv-TasNet	English	8.941261981	7.486695151	6.626825047	8.292722159
	Arabic	3.972427	3.972427	8.91058	-26.6596
	Chinese	10.63976	7.20355	16.42239	-8.05199
SIR					
Demucs	English	7.332337909	4.889571299	6.460674567	8.145132574
	Arabic	13.08034	10.58806	3.800309	11.54092
	Chinese	10.14399	6.673121	9.924395	12.20906
Conv-TasNet	English	8.941261981	7.486695151	6.626825047	8.292722159
	Arabic	3.972427	3.972427	8.91058	-26.6596
	Chinese	10.63976	7.20355	16.42239	-8.05199

At Table 17, there are more significant results of SDR affected by NL. Arabic experiment of Demucs algorithm has 13 at Music of conversation speech separation, but it has 3.8 SDR at Music of Child speech separation. On another side, Conv-TasNet algorithm during Arabic separation of conversation experiment is 4 SDR. While the lowest value in training experiment is -26.7 SDR.

Testing experiments:

One the other side, Conv-TastNet and Demucs have the higher values of (Child) testing dataset experiment, both values are 8.02 SDR. Then, (Conversation), (Female) and (Male) of testing dataset experiments respectively. SIR and SDR have the same values, which means a high correlation between these two measures. This is further indicated by the fact that they dominated in terms of estimation error. However, SIR is the best

measure of Gaussian noise, while SDR focuses more on interference-level intrusiveness and separation [35]-[37]. The experiments in this research study illustrated that Conv-TasNet is a preferable algorithm from the perspective of SDR and SIR. The reason is Demucs using a larger window-frame size'. Increasing weighting for noise power can cause potential modification of the SDR measure.

Table 18 and 19, in the context of comparing the Conv-TasNet algorithm with other research in terms of SDR, in [2], the authors have trained four models using the MusDB dataset [38], with 100 songs through Demucx, Conv-TasNet, Open-Unmix, MMDnseLSTM and Wave-U-Net. The average SDR of the vocals was 6.81 for the Conv-TasNet algorithm (the higher score) and 7.61 for the MMDnseLSTM algorithm. Their research study focused on music separation, while our research focused on music and speech separation in four categories. The highest value of music SDR separation for the (Music) and Female datasets was in Conv-TasNet. Moreover, there are some classification experiments that produced scores even higher than those found in [9], such as the Child, Female and Conversation training datasets. For Demucs algorithm, the score of SDR was 8.15 in the Child dataset; this was higher than the reported in [9], (6.08 for MusDB and 7.08 for 150 extra song tracks).

Table 18: mir_eval of Male and Female experiments

Algorithm	SDR	Music & Male Experiment		Music & Female Experiment	
		Music	Male	Music	Female
Demucs	English	6.303056998	1.461799732	7.607939697	4.703374514
	Arabic	12.92352	10.95636	11.55566	1.264586
	Chinese	11.78847	10.37697	10.47468	10.54035
Conv-TasNet	English	6.303056998	1.461799732	7.607939697	4.703374514
	Arabic	15.76899	6.375536	14.3024	9.198612
	Chinese	17.21237	-7.92302	9.808399	10.60297272
SIR					
Demucs	English	6.303056998	1.461799732	7.607939697	4.703374514
	Arabic	12.92352	10.95636	11.55566	1.264586
	Chinese	11.78847	10.37697	10.47468	10.54035
ConvTasNet	English	6.303056998	1.461799732	7.607939697	4.703374514
	Arabic	15.76899	6.375536	14.3024	9.198612
	Chinese	17.21237	-7.92302	9.808399	10.60297272

Table 19: mir_eval of conversation and child experiments

Algorithm	SDR	Music & Conversation Experiment		Music & Child Experiment	
		Music	Conversation	Music	Child
Demucs	English	7.390943771	4.908684372	6.583362429	8.02266854
	Arabic	11.81469	8.604822	6.909017	6.11558
	Chinese	10.14399	6.673121471	10.3703	13.03493
Conv-TasNet	English	8.912469304	7.531863904	6.66059257	8.02425869
	Arabic	10.93189	-26.1081	4.572872	-26.5985
	Chinese	1.261381	-0.95342	8.354678	10.88152
SIR					
Demucs	English	7.390943771	4.908684372	6.583362429	8.02266854
	Arabic	11.81469	8.604822	6.909017	6.11558
	Chinese	10.14399	6.673121471	10.3703	13.03493
Conv-TasNet	English	8.912469304	7.531863904	6.66059257	8.02425869
	Arabic	10.93189	-26.1081	4.572872	-26.5985
	Chinese	1.261381	-0.95342	8.354678	10.88152

In [8], the author studied the application of the Conv-TasNet algorithm to the WSJ0-3mix dataset for speech separation. Their study focused on how causality for Conv-TasNet that is a causal configuration due to causal convolution and/or layer normalization operations, leads to drops in Conv-TasNet performance. The SDR of causal Conv-TasNet value is 8.2, while the corresponding value is 13.1 for non-causal Conv-TasNet. The high SDR score for non-causal Conv-TasNet refers to the high signal noise ratio that is equal to 12.7, while -0.00001 SNR of music tracks of this research.

Notably, the signal-to-artifacts ratio (SAR) value has no correlation to the SIR and SNR values because it calculates the quantization and interferences of signals. It assists in distinguishing between the estimation errors using artefacts. For all the experiments of different languages and algorithms, SAR has zeros value. Spatial-Distortion-Ratio (ISR) is more often used for image signal processing than audio signal processing. For all the experiments of different languages and algorithms, ISR has infinite value. The mir_eval package is limited when it comes to auditory signification. For instance, SIR can hardly distinguish between two different variables. In addition, SDR does not compute the total perceived distortion in accurate manner.

5.3. Execution time

Figure 21 shows that Demucs separated (Music & Male) faster than (Conversation & Music), (Child & Music) and (Female & Music) of training and testing datasets respectively. Conv-TasNet separated (Female & Music) faster than (Music & Male), (Child & Music) and (Conversation & Music) in training dataset respectively. A slightly different in ms is between (Child & Music) and (Conversation & Music), but (Conversation & Music) was faster than (Child & Music) in testing dataset. Also, figure 21, shows the average execution time in nanoseconds per track of Conv-TasNet that was seven times faster than Demucs, for testing and training datasets in all languages' experiments. Conv-TasNet could be applied in real-time better than Demucs. Demucs needs further development, comparing with Conv-TasNet delay time. In [2], the author computed the speed of training of Demucs as 1.6 seconds and of Conv-TasNet as 0.7 seconds. However, this comparison is not accurate enough for consideration. The reason is that the Conv-TasNet was applied to only two seconds of track duration; also, for Demucs using a different sample size significantly affected the actual processing time. Noticeable, that there was a slightly different in execution time between different natural languages around 0.1 which is not necessary to be mentioned, because the separation process of algorithms does not affect by NL. The results from each algorithm were based on some equations to perform a separation process. The equations of each algorithm and their notations have mentioned before in research methodology section.

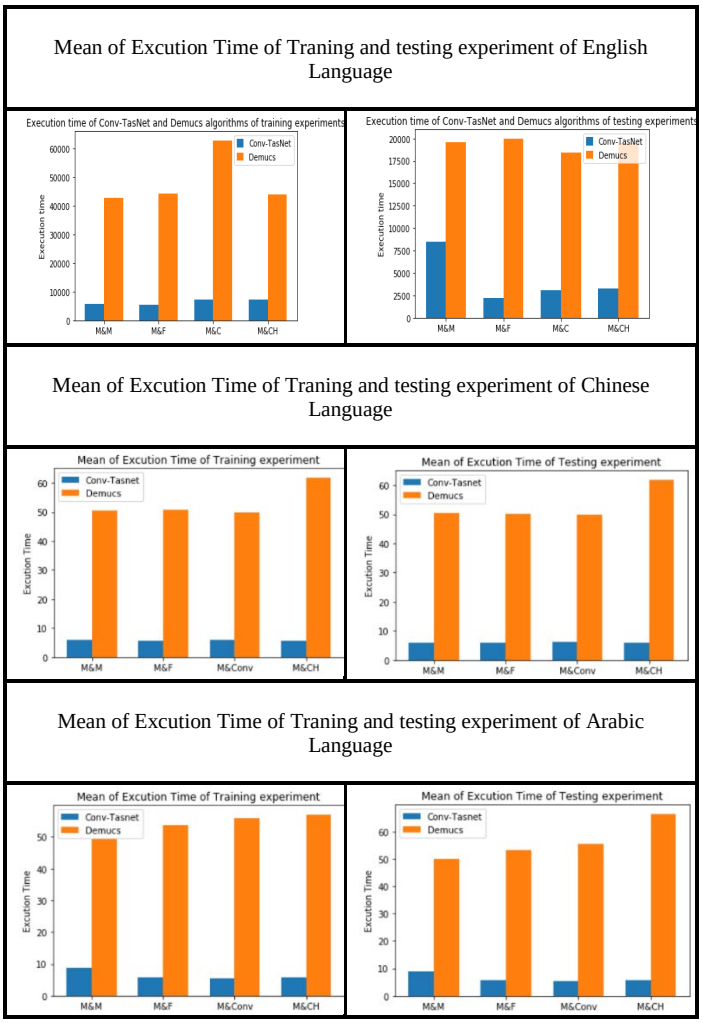


Figure 21: The mean of execution time of English, Arabic and Chinese languages.

In general, after calculating the sum of execution time in nanoseconds (ns) of each separation process of languages. training and testing Demucs experiment, English is 202.3 ns and 203.2 ns the fastest at separation process, respectively. However, Arabic is 216.1 ns and 225.1 ns the slowest at separation process, respectively. Training and testing Conv-TasNet experiment, English is 29.1 ns and 29.4 ns the slowest at separation process, respectively. Nevertheless, Chinese is 23 ns faster than English and Arabic at Training experiment. While Arabic is 25.9 ns faster than English and Chinese at Testing experiment, See Appendix Table 20-25.

6. Conclusion

This research study compared the Conv-TasNet and Demucs algorithms. Both algorithms use BSS approach during the separation process. Conv-TasNet and Demucs estimate the number of blind sources to be separated into four sources: drums, bass, vocals and other. However, this research reduces the number of sources to be two, which are vocal and music. These algorithms are based on two approaches: real-time and blind-source separation. A random 10GB audio set was taken from Youtube.com. These data separated into five categories: Male, female, Music, Conversation and Child. Using three Natural languages, which is English, Arabic and Chinese [25]. The performance with different and specific sources has been tested. The results show Conv-TasNet has excellent performance in separating music and speech sources. The highest SDR score of music is 9.21 in the female experiment, while the highest SDR score of speech at child experiment is 8.14. In addition, the average execution time of Conv-TasNet algorithm is seven times faster than Demucs algorithm. With the NL extend, in general, Chinese language has high performance at source separation process during Demucs and Conv-TasNet algorithms in training and testing experiments. However, there are some significant values of separation performance at (Music & Conversation) experiment. The values are (15.9 & 6.10) indicate high MAE error. Both algorithms have high error during source separation process of Chinese (Music & Conversation) experiments. The mixture signal contains Music and Male & Female Conversation.

Supplementarily, in general, Arabic language has high performance during Demucs separation process at (Music & Male). Normally, English language shows minor results comparing with Arabic and Chinese languages during both separation algorithms as it mentioned from table 6 to table 12.

There are some limitations in this research study. Noise reduction was attempted with multiple algorithms such as (spectral gating of noise reduction [38], Deep Convolutional Neural Networks for Musical Source Separation [39], aasp-noise-red [40], Deep Audio Prior - Pytorch Implementation [41], Joint Audio Correction Kit (J.A.C.K.) [42] and more). Nonetheless, those attempts were unsuccessful because the environment needs to be python 2. Old libraries are no longer supported for instance: J.A.C.K algorithm for noise reduction implemented on Python 3 via Pycharm. There was a lot of errors fixed by the researcher but after loading the input files there was no results because there is a runtime errors, which is one of the difficult errors to be tracked and fixed. The Demucs algorithm could not support large amount of data. The dataset was 10GB, but the Demucs algorithm started to lose values without separating after training 8GB. The Authors have environment limitation to run Demucs GPU version, because it needs to be run on a wider range of hardware.

The future work of this research will focus on training more audio sets or different audio sets, using different algorithms like deep learning approach to separate the data and increasing the number of data categories. In addition, noise reduction can be a function inside the BSS algorithms. The Conv-TasNet algorithm requires evolution in terms of actual time separation, computational cost and development of the separation process to achieve 100% interference separation. In addition, use the NL process functions in both algorithms could lead to a different and significant results.

7. References

[1] A. Alghamdi, G. Healy, H. Abdelhafez “Real Time Blind Audio Source Separation Based on Machine Learning Algorithms”, IEEE 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), 35-40, Oct, 2020, doi:10.1109/NILES50944.2020.9257891

[2] M. Pal, R. Roy, J. Basu, M. S. Bepari, “Blind Source Separation: A Review and Analysis,” 2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation, 1-5, 2013, doi:10.1109/ICSDA.2013.6709849.

[3] K. Ball, N. Bigdely-Shamlo, T. Mullen, K. Robbins, “PWC-ICA: A Method for Stationary Ordered Blind Source Separation with Application

to EEG," Computational intelligence and neuroscience **2016**(73), 1-20, 2016, doi:10.1155/2016/9754813.
[4] E. M. Grais, H. Wierstorf , D. Wa, R. Mason and M. D. Plumbley, "Referenceless Performance Evaluation of Audio Source Separation using Deep Neural Networks.," 27th European Signal Processing Conference (EUSIPCO). IEEE, 1-5, 2019.
[5] N. Makishima, S. Mogami, N. Takamune, D. Kitamura, H. Sumino, S. Takamichi , H. Saruwatari and . N. Ono, "Independent Deeply Learned Matrix Analysis for Determined Audio Source Separation," EEE/ACM Transactions on Audio, Speech, and Language Processing, **27**(10), 1601-1615, 2019.
[6] V. Leplat, N. Gillis, M. Shun Ang, "Blind Audio Source Separation with Minimum-Volume Beta-Divergence NMF," IEEE Transactions on Signal Processing, **68**, 3400 – 3410, 2020, doi: 10.1109/TSP.2020.2991801.
[7] Y. Luo and N. Mesgarani, "TASNET: TIME-DOMAIN AUDIO SEPARATION NETWORK FOR REALTIME, SINGLE-CHANNEL SPEECH SEPARATION," IEEE/ACM Transactions on Audio, Speech, and Language Processing, **27**(8), 1256-1266, 2019, doi: 10.1109/TASLP.2019.2915167.
[8] G. Wichern, M. Flynn, J. Antognini, E. McQuinn, "WHAM!: Extending Speech Separation to Noisy Environments," Sep. 2019. [Online]. Available: http://wham.whisper.ai/.
[9] Lab Neural Acoustic Processing, "NEURAL ACOUSTIC PROCESSING LAB," 2019. [Online]. Available: http://naplab.ee.columbia.edu/tasnet.html.
[10] D. Fourer G. Peeters, "Fast and adaptive blind audio source separation using recursive Levenberg-Marquardt synchrosqueezing," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 766-770, 2018, doi: 10.1109/ICASSP.2018.8461406.
[11] B. Pardo, "Interactive Audio Lab," 2018. [Online]. Available: https://interactiveaudiolab.github.io/.
[12] B. Loesch, B. Yang, "Blind source separation based on time-frequency sparseness in the presence of spatial aliasing," International Conference on Latent Variable Analysis and Signal Separation. Springer, Berlin, Heidelberg, 1-8, 2010.
[13] Informatics, UCD School of Computer Science and, "CHAINS characterizing Individual Speakers," 2006. [Online]. Available: https://chains.ucd.ie/.
[14] Y. Mori, T. Takatani, H. Saruwatari, K. Shikano, T. Hiekata, T. Morita, "High-presence hearing-aid system using DSP-based real-time blind source separation module.," IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, IV-609, 2007.
[15] S. Rickard, R. Balan and J. Rosca, "Real-time time-frequency based blind source separation," AJE, pp. 1-2, 2001.
[16] H. Abouzid and O. Chakkor, "Blind source separation approach for audio signals based on support vector machine classification.," Proceedings of the 2nd international conference on computing and wireless communication systems, pp. 1-6, 2017, doi: 10.1109/ICASSP.2007.366986.
[17] A. Ferreira, D. Alarcão, "Real-time blind source separation system with applications to distant speech recognition.," Applied Acoustics, 113, 170-184, 2016, doi: 10.1016/j.apacoust.2016.06.024.
[18] D. J. Watts. Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton Univ. Press, Princeton, USA, 1999.
[19] D. Stoller, S. Ewert, S. Dixon, "Wave-U-Net: A Multi-Scale Neural Network for End-to-End Audio Source Separation," 19th International Society for Music Information Retrieval Conference (ISMIR 2018), 334-340, 2018.
[20] A. Défossez, U. Nicolas, B. Léon and B. Francis, "Music Source Separation in the Waveform Domain," 1-16, 2019, arXiv preprint arXiv:1907.02404.
[21] Y. Luo, N. Mesgarani, "Conv-TasNet: Surpassing Ideal Time-Frequency Magnitude Masking for Speech Separation," IEEE/ACM transactions on audio, speech, and language processing **27**(8), 1256-1266, 2019, doi: 10.1109/TASLP.2019.2915167.
[22] C. Biemann "Chinese Whispers - an Efficient Graph Clustering Algorithm and its Application to Natural Language Processing Problems", Workshop on TextGraphs New York City, 73–80, June 2006.
[23] M. M. Awais; WaqasAhmad; S. Masud; S. Shamail, "Continuous Arabic Speech Segmentation using FFT Spectrogram", Innovations in Information Technology conference, 2006, doi: 10.1109/INNOVATIONS.2006.301939.
[24] D. Pani, A. Pani, L. Raffo, "Real-time blind audio source separation: performance assessment on an advanced digital signal processor," The Journal of Supercomputing **70** (3), 1555-1576, 2014, doi:10.1007/s11227-014-1252-4.
[25] Arwa-Data-Analytics, Machine-Learning-Algorithms-for-Real-Time-Blind-Audio-Source-Separation-with-Natural-Language-Detect, Retrieved from GitHub: https://github.com/ArwaDataAnalytics/Machine-Learning-Algorithms-for-Real-Time-Blind-Audio-Source-Separation-with-Natural-Language-Detect, July 23, 2021.
[26] "scipy.org," ENTHOUGHT, 26 Jul 2019. [Online]. Available: https://docs.scipy.org/doc/numpy/reference/generated/numpy.std.html. [Accessed 12 April 2019].
[27] W. Gragido, J. Pirc, D. Molina, Blackhatonomics, Science Direct .com, 2013.
[28] Vishal3096, "geeksforgeeks," geeksforgeeks, 2019. [Online]. Available: https://www.geeksforgeeks.org/scipy-stats-signaltonoise-function-python/. [Accessed 12 April 2019].
[29] A. HAYES, "investopedia," investopedia, 18 Mar 2020. [Online]. Available:https://www.investopedia.com/terms/r/r-squared.asp. [Accessed 14 April 2020].
[30] M. B. Editor, "Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of Fit?", 30 May 2013. [Online]. Available: https://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-iinterpret-r-squared-and-assess-the-goodness-of-fit. [Accessed 14 April 2020].
[31] T. Bock, "8 Tips for Interpreting R-Squared," display, 10 April 2020. [Online]. Available: https://www.displayr.com/8-tips-for-interpreting-r-squared/. [Accessed 14 April 2020].
[32] Stephanie, "What is Root Mean Square Error (RMSE)?," 25 October 2016. [Online]. Available: https://www.statisticshowto.com/rmse/. [Accessed 14 April 2020].
[33] K. GRACE-MARTIN, "Assessing the Fit of Regression Models," theanalysisfactor, [Online]. Available: https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/. [Accessed 14 April 2020].
[34] Craffel, "mir_eval github," Aug 2016. [Online]. Available: https://github.com/craffel/mir_eval/blob/master/mir_eval/transcription.py. [Accessed 11 March 2020].
[35] L. Roux, . J. S. Wisdom, H. Erdogan, "SDR – HALF-BAKED OR WELL DONE?" IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 626-630, 2019, doi: 10.1109/ICASSP.2019.8683855.
[36] J. M. Kates, "Cross-correlation procedures for measuring noise and distortion in AGC hearing aids," The Journal of the Acoustical Society of America, 107(6), 3407-3414, 2000, doi:10.1121/1.429411.
[37] S. Datasets, "#MUSDB18," Github, 9 10 2019. [Online]. Available: https://sigsep.github.io/datasets/musdb.html. [Accessed 27 Feb. 2020].
[38] K. amrr, "Noise reduction in python using spectral gating," Feb. 2020. [Online]. Available: https://github.com/timsainb/noisereduce. [Accessed March 2020].
[39] Nkundiushuti, "DeepConvSep," Github, 2015. [Online]. Available: https://github.com/MTG/DeepConvSep. [Accessed March 2020].
[40] lcolbois, "aasp-noise-red," 2019. [Online]. Available: https://github.com/lcolbois/aasp-noisered. [Accessed March 2020].
[41] Y. Tian, C. Xu and D. Li, "Deep Audio Prior - Pytorch Implementation," Jan. 2020. [Online]. Available: https://github.com/adobe/Deep-Audio-Prior. [Accessed March 2020].
[42] C. Barth, "Joint Audio Correction Kit (J.A.C.K.)," Github, 2019. [Online]. Available: https://github.com/cooperbarth/Joint-Audio-Correction-Kit. [Accessed March 2020].

Appendix

Training experiments:

Table 20: Execution time of Training Dataset of English languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	49.84469477	6.823944947
Female & Music (M&F)	51.39934172	6.003973062
Music & Conversation (M&C)	50.43340774	8.179383581
Music & Child (M&CH)	50.70981113	8.14998204

Table 21: Execution time of Training Dataset of Arabic languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	49.421875	8.90625
Female & Music (M&F)	53.703125	5.8046875
Music & Conversation (M&C)	55.8671875	5.4921875
Music & Child (M&CH)	57.109375	5.84375

Table 22: Execution time of Training Dataset of Chinese languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	50.421875	5.75
Female & Music (M&F)	50.890625	5.7265625
Music & Conversation (M&C)	49.8046875	6
Music & Child (M&CH)	61.8515625	5.5390625

Testing experiments

Table 23: Execution time of Testing Dataset English languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	50.24387019	6.735885814
Female & Music (M&F)	51.80440415	6.00008378
Music & Conversation (M&C)	50.3566125	8.143201463
Music & Child (M&CH)	50.78504406	8.641241776

Table 24: Execution time of Testing Dataset Arabic languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	50.0859375	8.7734375
Female & Music (M&F)	53.1015625	5.7578125
Music & Conversation (M&C)	55.4921875	5.53125
Music & Child (M&CH)	66.40625	5.828125

Table 25: Execution time of Testing Dataset Chinese languages

Experiments	Execution time of Demucs (ns)	Execution time of Conv-TasNet (ns)
Male & Music (M&M)	50.53125	5.9296875
Female & Music (M&F)	50.0546875	5.921875
Music & Conversation (M&C)	49.8046875	6.234375
Music & Child (M&CH)	61.8828125	8.669216296

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol.6, No.4, 138-146 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

The Gamification Design for Affordances Pedagogy

Wilawan Inchamnan[1,*], Jiraporn Chomsuan[2]

[1]*College of Creative Design & Entertainment Technology, Dhurakij Pundit University, Bangkok, Thailand*

[2]*College of Innovation Business and Accountancy, Dhurakij Pundit University, Bangkok, Thailand*

ARTICLE INFO

Article history:
Received: 12 May, 2021
Accepted: 22 June, 2021
Online: 10 July, 2021

Keywords:
Gamification
Affordance
Growth Mindset
Design method
Affordances Pedagogy
Motivation and Engagement

ABSTRACT

This study aims to design a gamification affordances pedagogy. Affordances are the ways in which we perceive environments to support the needs of learners in the educational system. The main questions are how gamification elements can influence student engagement to improve their affordances. Affordance behavior is a human behavior that refers to a mindset; an attitude or opinion, especially a habitual one. Motivational activities can change a learner's behavior. A skill-based mindset can be created through the use of affordance motivation. Affordance refers to the points, badges, and leaderboards in gamification elements. This research aims to improve the affordance mindset design of interactive systems with gamification. The affordance design will improve the pedagogy related to engagement. The research focuses on the mindset factors and the relationship between the factors that promote the desired learning outcomes. The findings may help in designing the gamification affordance design method for affordance pedagogy. The expected model could improve learners' affordances and instructional activities.

1. Introduction

In the 21st century, there is extensive research on growth mindset and intrinsic motivation in learning. The constructs of mindset and motivation are important for educators to determine the impact on student learning and outcomes [1]. Understanding the two constructs, mindset and motivation, and the relationship between them is necessary as it provides insight into student motivation and drive. Gamification is a tool that can increase and promote user motivation, especially in education. The educational concept requires that teaching and learning activities are more fun and interesting [2]. Today, learner engagement is still a challenge in the education system. Design is employed in education to increase the student desire to focus on the educational task [3], as an affordance mindset. Educational games and various forms of edutainment have gained more attention in the discipline of learning and teaching strategies. Educators believe learning can be enhanced through play and fun [4]. The increasing motivation to learn may affect the learner's affordances. An affordance concerns the possible actions that an item offers while learning. The term affordance is a somewhat ambiguous term [5]. and affordance could be improved in terms of its ability to influence learning outcomes.

Learner can conduct their learning lives using advanced technology that engender an affordance behavior mindset. The mindset is a crucial factor in leaner motivation. Affordances are a core opportunity for action [6]. A mindset can change the attitude to learning in the education system, which represents a step in the right direction.

This paper reviews several recent gamification studies that focus on growth mindset and motivation. The theoretical frameworks of Affordance Mindset and Motivation reflect how they are applied in educational gamification. The research design is divided into two parts. The first focuses on the engagement elements related to the gamified classroom activities. The gamification strategies are then designed by using the mindset factors and learner characteristics. Then the gamification affordances design method is applied through points, badges,

*Corresponding Author: Wilawan Inchamnan, E-mail: wilawan.inn@dpu.ac.th

leaderboards and ranks in gamification activities for the pedagogy strategies.

2. Literature Review

2.1. Mindset

The mindset is a set of both conscious and unconscious human beliefs, which relates to how humans view what they consider to be their personality. Mindset can be divided into two types, Fixed Mindset and Growth Mindset. These mindsets refer to the way people think about the nature of intelligence and learning. People with a growth mindset value effort, tend to set learning goals (e.g., mastery) rather than performance goals (e.g., grades), and attribute failure to lack of effort rather than lack of ability [7]. Learnersì mindsets can be influenced by school-based activities to help improve academic outcomes [7] through motivation and engagement.

2.2. Motivation and Engagement

Motivation and engagement indicate passion and emotional involvement in learning activities [8]. Engagement permits meaningful learning, which includes the quality of student effort, student interaction and their immersive experiences [8]. Some research divided engagement into three dimensions: behavioral, emotional, and cognitive engagement [3].

2.3. Motivation

Motivation is an abstract construct used to explain people behavior. The behavior represents the basis for peopleìs actions, desires, and needs. Motivation can be named as oneìs behavioral direction, or what justification a person to want to repeat a behavior [9]. Motivation can be allocated into two different types known as intrinsic (internal) motivation and extrinsic (external) motivation [10]. Intrinsic motivation is the desire to seek new things and new challenges, to analyze one's abilities, to observe and to gain knowledge [11]. It is driven by interest or enjoyment in the task itself and exists within the individual rather than relying on external factors or the desire for reward. Intrinsic motivation arises from within the individual, just as the idea of an affordance draws attention to a possible action [12]. Extrinsic motivation mentions to the performance of an activity to attain a desired consequence and is the opposite of intrinsic motivation [11]. Extrinsic motivation is the type of motivation that comes from outside the individual and often involves rewards such as trophies, money, social recognition, or praise. In education, motivation is a major cause of differences in student learning outcomes, considered a possible predictor of a student's academic performance; students with high academic motivation are more likely to succeed academically [13]. Students who are intrinsically motivated are more likely to be curious and inquire about the process, focusing on the task itself rather than just the outcome. In contrast, students who are extrinsically motivated are more concerned with the outcome (e.g., grades, prizes) than with the process of completing the task itself [14]. Game activities could encourage player experiences, which is called immersive engagement. Engagement also influences peopleìs adoption. Behavior usage is how frequently or for what purpose the behavior is used while behavior adoption is the degree to which the behavior is utilized. Psychological outcome is a measure of effort. The level of effort is influenced by emotional engagement (pleasure, excitement, and persistence), behavioral engagement (effort), cognitive engagement (attention, reflection), and learning performance (perceived competence, perceived improvement) [3]. The student can repeat desired behaviors by reinforcement providing. Reinforcement incentives people into two forms of motivation. The first one is intrinsic motivation that refers to engage in behaviors for enjoyment, challenge, pleasure, or interest [5]. Then, the extrinsic motivation can engage in an activity to earn an external reward when learner is motivated to perform their behavior [1].

2.4. Engagement

Engagement in the extent to which a learner connects with the gaming environments and indicates a positive psychological state of mind when so doing [15]. Games' activities provide immediate feedback, which is more effective and efficient than traditional learning strategies. Gamification strategies can encourage the learner to be acquire experience during play. Game experiences drive personal change and transformation by generating an attitude of acceptance about the challenge, motivation to achieve, and constant innovation by simulation. Simulation encourages the participant to immerse themselves in learning [16]. Adaptation is the process by which strategies are moderated by engagement. Adaptation is a consequence of activities and events that are enhanced, developed, and implemented. Success in learning depends on the learner's desire to learn, which is known as behavioral intention [17]. Learners can change their behavior as a result of motivation. Engagement may encourage behavior. Usage can be influenced by behavioral assumption. The focus of usage behavior is on specific activities performed using specific sources of information. These include general knowledge acquisition, learning, and the pursuit of purposes [18].

2.5. Gamification Affordances

As a concept, affordance provides a useful bridge to explain the interplay between the artefact and the human user [5]. Gamification elements could provide a thematic evaluation of subsequent formulations of the term affordance. This is the method designed to provide a preliminary overview of how the notion of affordance can be interpreted. The figure 1 show the direct perception of affordances for the user along with a consideration of the usersì skills. The researcher refers to the "flow channel" as a linear function on a plane with skills and challenges as axes. An increase in the learner's skills is due to learning, and an increase in the challenges of performing a task is due to novelty [19]. The pedagogy design may keep the two in balance between challenges and skills. Subjects experienced Flow when they first encountered a task with a high balance between skills and challenges [20]. Skills is due to learning and an increase in the challenges of performing tasks [21], which relate to an affordance. The design of the

pedagogy allows for a practical mapping flow onto the gameplay. The approach enhances the player's interaction with the game elements and provides useful insights into the learner's skills [21]. The next section presents gamification affordance strategies including points, badges, and leaderboards.

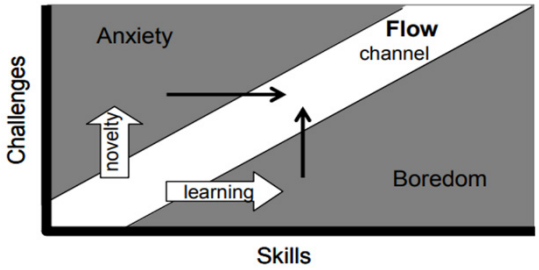


Figure 1:The Flow Channel

2.5.1. Points

Point is the activity outcome. Game point can trigger and reinforce situational awareness competencies. Points can provide information about the playerìs progressive such as response times, correct answers, and the procedures followed during play. The awarding of points is connected to behaviors that require crucial competencies [22].The competencies can develop as a training tool for learning goals [23].

This paper aims to design affordance mindset processes through the player data.The affordance for learning is measured by questionnaire questions regarding learnersì sense of pride and community acceptance. The satisfaction of basic psychological needs (competence, autonomy, and relatedness [24]), is a fundamental requirement for being autonomously motivated. Competence refers to the experience of success by fulfilling challenging tasks and gaining mastery within an environment [25] and is reflected by the number of points scored.

2.5.2. Badges

Badges reflect performance results. Digital badges indicate the achievements or skills acquired while playing the game. These badges are collected and displayed to the other players [25]. Gamification elements are used to encourage performance and skill acquisition, which are the desired learning outcomes.

2.5.3. Leaderboards

The most-used game element is leaderboard that refers to a ranking board of the players in a competitive event. The leaderboard is to illustrate player where they are ranked in a gamified system. Leaderboardsì mechanic can be employed in various ways to offer goals and to increase motivation [26]. Leaderboard ranking can motivate players to compete, which increases participation [27] and facilitates comparison and competition.Leaderboard is the basic elements that make up games that combined to deliver a system of mastery to end users [28].

2.6. Cognition

Cognition refers to the mental processes involved in the acquisition of knowledge and understanding. These cognitive processes include thinking, knowing, remembering, judging, and problem solving [29]. Cognitive processes affect every aspect of life, from school to work, to relationships. Some specific uses for these cognitive processes include the following.

- Learning New Things that require being able to take in new information and form new memories. The learner makes connections with other things that they already know.

- The formation of memories is a major topic in the field of cognitive psychology. Memories refer to how people remember, what they remember, and what they forget, and reveal much about how cognitive processes work.

Making decisions means making judgments about things you have experienced and processed. This may involve comparing new information with previous knowledge or integrating new information with new knowledge before making a decision [30]. Behavioral action refers to use, i.e. the fact that it is used, or habit. Relationships have been found between adoption, post-adoption variables, and usage behavior in the post-adoption process. Continued use based on experience and satisfaction in the post-adoption process represents high quality use [31]. Post-adoption, each individual is engaged in change behaviors to varying degrees. In this context, several factors may influence the relationships between adoption and post-adoption variables [31].

2.6.1. Positive Feedback

Positive feedback is a game mechanism. This mechanism is designed to accelerate or enhance ongoing output [32]. Gamification can be applied to stimuli of increasing intensities through intrinsic rewards and feedback. The user-centered design activities require an interactive feedback. The learning activity feedback is determined by peopleìs motivation and cognitive mindset.

2.6.2. Growth Mindset

The term mindset refers to implicit beliefs that have been shown to influence the thoughts and actions of individuals [33]. A learner's mindset has been shown to influence his motivation and academic performance [34], [35]. Growth mindset type means that a learner can improve his talents and abilities through effort. Growth mindset belief is a type of intelligence that can be improved through hard work and the use of strategies [36]. Growth Mindset can be generated through motivation and achievement [37], [38], could promote learner's engagement.

A growth mindset helps learners improve their skills and knowledge over time, and mindset research studies the power of such beliefs in influencing human behavior [39]. Mindset is a soft skill of great importance [40]. For example, athletes are driven by success and can realize their potential through effort, practice, and instruction. In the education system, some research has shown that students with a growth mindset can greatly improve their success and achievement [41].

2.7. The Growth Mindset Stimulus Model

According to the previous research, the figure 2 illustrates the gamification model, which encourages learner to conduct their learning by using gamification elements. The learning behavior refers to the taking steps in the right direction guideline. Gamification activities conceptual model can be influenced by gamified activities [42]. The growth mindset is fostered by the positive feedback through cognition and motivation. The motivation can be stimuli by competency and leaderboard. The cognition could be influence by level [42].

The figure 2 shows the gamification workflow for growth mindset processes that influence the positive feedback and growth mindset by gamified activities. The learnerìs growth mindset is measured through feedback, level, motivation, their position on the leaderboard, and their competencies. The model shows the gamification activities can increase the likelihood of achieving the learning objectives by motivating students to learn. This illustrates the relationship between the stimulus activities and growth mindset processes [43]. Game activities can encourage the learner to practice their skills and their position on the leaderboard can enhance their pride and social acceptance. Motivation and cognition can enhance their performance through positive feedback and higher scores (GPA).

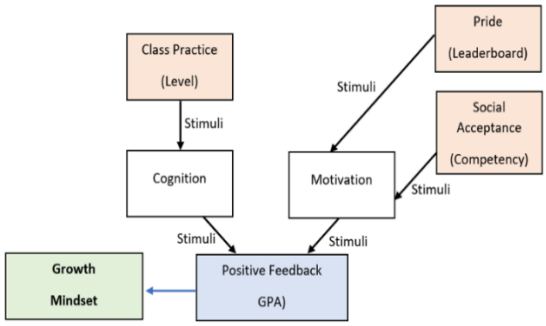


Figure 2: The Growth Mindset Stimulus Model [42]

3. Methodology

Having reviewed a scholarly source on a topic concerning mindset then mindset questionnaire is developed to classify participants into a different mindset characteristic. This research aims to link the mindset factors and gamification elements. The experiment design aims to cluster the mindset criteria and determine the relationships between the factors and the learning outcomes. The data collection uses a mindset questionnaire to facilitate comparisons with learner subjects. The pedagogy design gathers the university lecturer for the strategiesì class activity. The advisor experiences help to map the teaching or pedagogy style with gamification elements. Figure 3 illustrates the processes of data evaluation.

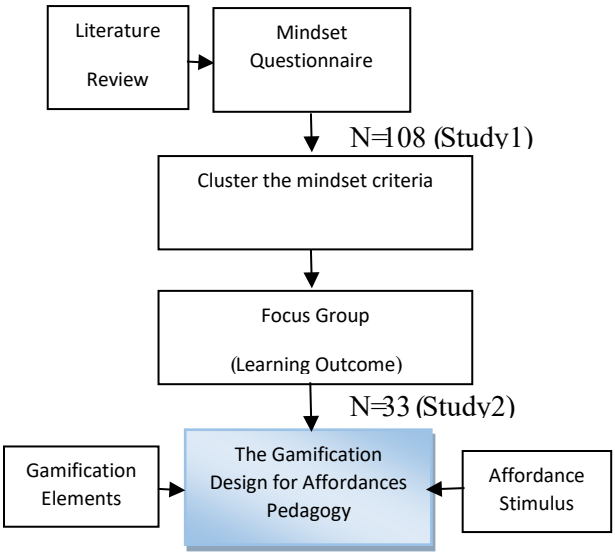


Figure 3: The Methodology of the gamification design for affordances pedagogy

3.1. Participants

Study1: According to factors finding, the participants in this study were 108 university Dhurakij Pundit University students. Participants included 72%from the Business faculty, 24%from the Creative Design faculty and 4%from the Journalism faculty all aged between 18 and 23. All participants in this study were volunteers. The average GPA was 3.55.

Study2: Then, the pedagogy design participants in this study were 33 university lecturers, randomly. The questionnaire design used the factors from the first findings. The average of experiences was 13 years. Normally, their teaching style were lecture and practical (50:50). They preferred to use gamification in the classroom and showed the score or progress to their student.

3.2. Clustering the mindset criteria: Factor Analysis (Study1)

This study uses the factor analysis. The mindset measures by questionnaires survey. The students may respond to questions about their opinions, which are all associated with the latent fix and growth variable mindset as below.

- Fix1: It is difficult for my intelligence /level of intelligence to change.
- Fix2: There are certain activities /subjects that I cannot be good at.
- Fix3: I think people do not have to try to do what they are capable of.
- Fix4: I can always learn new things, but I do not think that learning can increase my intelligence.
- Fix5: I am frustrated to see that people can do better than me.
- Fix6: If I have try something new and fail, I am not good at it.

- G1: I believe I can develop in all areas, albeit a little.
- G2: When I do something wrong, I feel that I would learn more from it.
- G3: I feel great when people see that I am good /talented, that means I am successful.
- G4: I am inspired by successful people.
- G5: I feel that I can help others to be successful.
- G6: If I try something new and fail, I want to repeat it until I can.

3.3. The Pedagogy Design (Study2)

The design process aims to link the pedagogical design and gamification elements and this paper focuses only on the factors affecting the mindset. The conceptual model can be employed in future work. The model can apply to the pedagogical context design includes the pedagogy plan.

The second findings from the university lectures show the relationship between factors.

- G1: The learnerìs skills can develop
- G2: Learnerìs practicing in the classroom
- G3: The level of learnerìs activities during learning
- G4-1: lectureriìs Comment for lesson activities individual
- G4-2: lectureriìs Comment for lesson activities to all of students
- Gamification: The class activities during learning
- Learn for smart:- Learn new things can increase the intelligence.

4. Results

4.1. Clustering the mindset criteria

The relationship of each variable to the underlying factor is expressed by the factor loading. The result of factor analysis, which deals with indicators of mindset, with twelve variables (opinions) and four resulting factors is shown below. Factor loadings can be interpreted like standardized regression coefficients, so the opinion in response 1 (Fix1) has a correlation of 0.72 with Factor 1. Five others, responses 2,3,4,5, and 6 (Fix 2,3,4,5,6), all representing fixed-attitude characteristics, are also associated with Factor 1. Based on the loading of the variables, the finding points high on factor 1, it could be considered as "Fixed Mindset".

However, opinions on questions 9 and 10 (Growth3 and 4), have high factor loadings on the other factor, Factor 2. They seem to indicate an external stimulus; that is, students are driven by other people, so Factor 2 would be classified as "Exogenous growth mindset." Similarly, Factor 4, is constituted by an answer of Growth 5 and 6 showing that students have positive feelings whenever they help others, which is an external stimulus, hence Factor 4 would also be regarded as "Exogenous growth mindset".

Table 1: Component Analysis

Rotated Component Matrix[a]

	Component			
	1	2	3	4
Fix1	.721	-.353	.071	.170
Fix6	.708	.174	-.121	-.193
Fix5	.690	.235	-.206	-.221
Fix2	.675	-.091	-.105	-.077
Fix3	.673	-.246	.091	.106
Fix4	.611	.131	.060	-.243
G3	.072	.784	.178	.011
G4	-.077	.638	.328	.221
G2	-.231	.164	.827	.081
G1	.106	.185	.796	.070
G6	-.107	.021	.176	.853
G5	-.167	.496	-.040	.647

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 11 iterations.

Opinions on questions 7 and 8 (Growth 1 and 2) have strong factor loadings to Factor 3. Students agreed they are likely to do anything as a result of intrinsic motivation, thus Factor 3 is regarded as îEndogenous growth mindsetï.

Table 2: KMO test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.730
Bartlett's Test of Sphericity	Approx. Chi-Square	292.755
	df	66
	Sig.	.000

The Kaiser-Meyer-Olkin test (KMO- Table 2) is a measure of how appropriate the data are for factor analysis. The KMO value in this study was .730, which indicates that the sampling is appropriate. Bartlett's test is performed before applying factor analysis to cheque whether the data reduction technique can reasonably compress the data. In this study, the test statistic Chi-Square was 292.755 and the corresponding p-value was 0.000, which is less than the significance level (0.05). Thus, the data are suitable for factor analysis.

In column labelled (table3) " Extraction Sums of Squared Loadings. The" Total" column shows the Eigenvalues for each factor extracted which higher than 1. The second column "%of Variance" indicates the variance is explained by each factor (or component). The "Cumulative %"column shows the percentages of the total variance explained by the factors (62.418%).

Table 3: Total Variance Explained

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.276	27.298	27.298	3.276	27.298	27.298	2.893	24.105	24.105
2	2.059	17.158	44.456	2.059	17.158	44.456	1.626	13.546	37.651
3	1.134	9.451	53.906	1.134	9.451	53.906	1.575	13.123	50.774
4	1.021	8.511	62.418	1.021	8.511	62.418	1.397	11.644	62.418
5	.857	7.140	69.558						
6	.694	5.782	75.340						
7	.643	5.358	80.697						
8	.599	4.991	85.688						
9	.538	4.486	90.175						
10	.457	3.809	93.983						
11	.391	3.255	97.238						
12	.331	2.762	100.000						
Extraction Method: Principal Component Analysis.									

4.2. Learning Outcome: Mindset Assessment Processes

The results (Table 4) show that learning outcomes have a positive relationship with a fixed mindset ((Sig, .008) Fix3: Try to do what they are capable of doing. Fix4: Learn new things, but I don't think learning can increase my intelligence).

Fix3 refers to the affordances of learning that learners believe they can develop in all areas, even if only a little (Growth Mindset: G1). The results show a significant relationship between Fix3 and G1 (.016). Affordance of Learning is also positively related to Growth Mindset (sig. .001) that the learner believes they will learn more if they do something wrong (G2). Having a Fixed Mindset has a positive influence on G4, which refers to the learner being inspired by successful people (Sig. .000).

Table4 refers to what learners think about intelligence. They love to learn new things, but intelligence is fixed and cannot be increased by learning. The results show significant relationships between Fix4 and G1, G2, G3, and G4 (sig. .025,.013,.000,.002).

Table 4: Chi-Square tests

Variable	Value	df	Asymptotic Significance (2-sided)
GPA and Fix3	32.692a	16	.008
GPA and Fix4	26.695a	16	.045
G1 and Fix3	30.486a	16	.016
G1 and Fix4	28.788a	16	.025
G2 and Fix3	31.970a	12	.001
G2 and Fix4	25.368a	12	.013
G3 and Fix4	38.382a	12	.000
G4 and Fix3	50.059a	16	.000
G4 and Fix4	37.716a	16	.002

Having a Growth Mindset means believing in skill development, having a desire to learn more, seeing others, and being inspired by successful people. The mindset can promote the fixed mindset, which refers to how people think about the nature of intelligence and learning.

The findings (table5) show the positive relationship between G1 and G2. The lectures believe that the students can develop and practice in the classroom (Sig. .015). The pedagogy design such as practical tasks (G1) will encourage with comment or feedback (G4) during the activities (Sig. .033, .026).

The level design for learner's activities is positive impact on the feedback (G4) (Sig. .028) and Gamification strategy (Sig. .045). The class activities during learning such as gamification can encourage learners in terms of the new things can increase the intelligence (Sig. .039).

Table 5: ANOVA Tests (The lecture participants)

Variable	Sum of Squares	Df Between Groups	F	Sig.
G1 and G2	.742	1	6.576	.015
G1 and G4-1	.970	1	5.010	.033
G1 and G4-2	1.227	1	5.471	.026
G3 and G4-1	1.478	2	4.036	.028
Gamification and G3	3.976	2	3.436	.045
Gamification and learn for smart	3.480	2	3.625	.039

The results show in the table 6 that include the activities in the classroom. The participants illustrate the teaching experience to catch up the students' attention. The advisor experiences help to map the teaching or pedagogy style with gamification elements.

Table 6: Means comparison

Variable	Mean	Std. deviation
Gamification	4.36/5	.603
The feedback of progress	4.33/5	.816
The percentage level of progress(100)	37.33	29.49
The Practical activities	4.85/5	.364
The percentage of practical activities(100)	54.24	20.620
Inborn Intelligence	2.33/5	.924
Learning the New things can improve the intelligence	4.06/5	.747
Time in the Activities	4.15/5	.667
Time for Activity/each (minute)	41.66	18.819

Quest in the Activity	3.96/5	.728
Time for Quest/each activity (minute)	44.54	17.781
Social media feedback	3.60/5	1
Reward for the activity	4.15/5	.618
Percentage for reward/each activity (100)	15.64	12.36
Comment for individual	4.69/5	.466
Comment for all of students	4.54/5	.505

5. The Gamification Design for Affordances Pedagogy

From the previous study that map to the results in this study, the finding shows the relationship between factors. The significance of fixed mindset measures was positive impact through the growth mindset (Fig 4.). According to the research background and findings, the gamification mechanics can helps develop deeper insights into the capacity for pedagogy design. The findings support the ideas to develop the optimal psychology or flow theory [19] to suggests a remarkable activity in lesson plan. The gamification elements including points, badges, and leaderboards can derive her/his optimal experience [21]. For instance, the level and point can motivate people for engagement. An individualìs capacity to concentrate will impact their ability to experience flow [21].

5.1. The Gamification Elements

The design process aims to link the pedagogical design and gamification elements and this paper focuses only on the factors affecting the mindset. The conceptual model can be employed in future work. The model can apply to the pedagogical context design includes the promotion of tasks in each faculty, increasing motivation and encouraging desirable learning behavior. Data are collected through classroom observations and stimulated recall interviews. The core characteristics of growth mindset pedagogy include focus on process, mastery orientation, persistence, and individual student support [44].

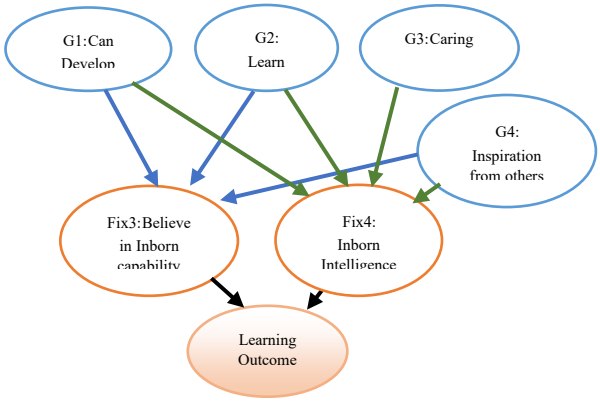


Figure 4: The relationships between factors

According to figure2, figure 5 shows the stimulus model for a growth mindset through motivation activities. The growth mindset factors are the stimuli, which can influence the learnerìs affordances.

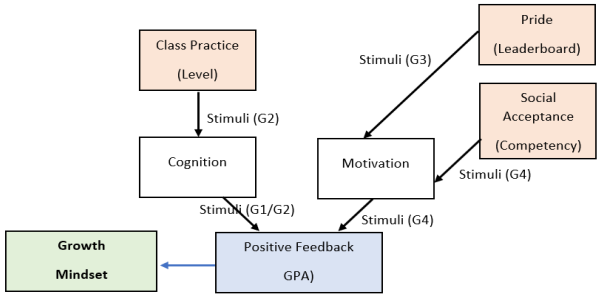


Figure 5: The Conceptual Gamification Affordances Design Model

5.2. The Affordances Pedagogy

The table 7 shows the findings that include the flow and pedagogy elements. These methods refer to the lecturersì experiences (Study 2) and studentsì feedback (Study 1). The gamification design for affordances pedagogy examines the literature review in terms of affordance, motivation and gamification. The research background aims to design the pedagogy elements below:

- Class assignment
- Quiz
- Group discussion
- Discussion
- Case study
- Play game

The pedagogy elements focus on the engagement that enable a practical mapping flow onto gamification. The approach enhances the studentìs interaction with the gamification elements such as points, badges, and leaderboard.

Table 7: The pedagogy elements of flow and gamification design (Adopted from Jones [45]).

	Element of Flow	Manifestation in Pedagogy Elements
1.	Task that we can complete	Class Assignment: The feedback of progress during class **Game element: Point**
2.	Ability to concentrate on task	Quiz (test): The feedback of progress or show the score **Game element: Point**
3.	Task has clear goals	Group Discussion: Quest in the Activity with friends and lecturer **Game element: Badges**
4.	Task provides immediate feedback	Class Assignment: Comment for individual/Comment for all of students

		Game element: Leaderboard
5.	Deep but effortless involvement	Discussion: social media/browsing the internet **Game element: Leaderboard**
6.	Exercising a sense of control over their actions	Case Study: The Practical activities/The percentage of practical activities around 50% per class **Game element: Badges**
7.	Concern for self disappears during flow, but sense of self is stronger after flow activity	Play game: The activity provides for a class environment as a simulation of life that refers to gamification such as reward, point and leader board in the class activity **Game element: Badges, point and leader**
8.	Sense of duration of time is altered	Class Assignment: Time in the Activities/Time for Activity and Time for case study for each activity around 40-45 minute **Game element: Leaderboard**

6. Discussion/Conclusion

Gamification design for affordances pedagogy focuses on engagement that provides a practical mapping flow to gamification activities during instruction. This approach could enhance student interaction through points, badges, and a leaderboard. A learning mindset could be created using motivation and affordance pedagogy. Affordance refers to points, badges, leaderboards and ranks in gamification elements. This research aimed to determine the affordance mindset factors of gamification interactive systems. The results show the relationships between factors that lead to desired learning outcomes. In accordance with Sailer, M., Homner, L. examined research topic "The Gamification of Learning: a Meta-analysis", the meta-analysis supports the claim that gamification of learning works because the results showed significant, positive effects of gamification on cognitive, motivational, and behavioral learning outcomes [46]. Similarly, there is a large body of work that clearly shows that incorporating gamification into the instructional process can lead to better student learning outcomes and helps to increase student achievement [47]-[50]. The findings derived from the current research could be used in a gamification affordances design method. The use of the model could help to improve the learner's affordances. The pedagogical guide could encourage the learner in the classroom. Engagement enables meaningful learning, which includes the quality of student effort, student interaction, and their immersive experiences during activities. The framework idea is a theoretical construct used to shape pedagogy and learner behavior. It represents the reasons for learners' actions, desires, and needs during instruction. Future research could test and revise the gamification design for the pedagogy of affordances.

References

[1] B. Ng, îThe neuroscience of growth mindset and intrinsic motivation,ï Brain Sciences, **8**(2), 20, 2018, doi: 10.3390/brainsci8020020.

[2] G.P. Kusuma, E.K. Wigati, Y. Utomo, L.K.P. Suryapranata, îAnalysis of gamification models in education using MDA framework,ï Procedia Computer Science, **135**, 385ó392, 2018, DOI: 10.1016/j.procs.2018.08.187.

[3] C. Silpasuwanchai, X. Ma, H. Shigemasu, X. Ren, îDeveloping a comprehensive engagement framework of gamification for reflective learning,ï in Proceedings of the 2016 ACM Conference on Designing Interactive Systems, 459ó472, 2016, doi: 10.1145/2901790.2901836.

[4] K. Rapeepisarn, K.W. Wong, C.C. Fung, M.S. Khine, îThe relationship between game genres, learning techniques and learning styles in educational computer games,ï in International conference on technologies for E-learning and digital entertainment, 497ó508, 2008, doi: 10.1007/978-3-540-69736-7_53.

[5] S. Harwood, N. Hafezieh, îíAffordanceí-what does this mean?,ï in 22nd UKAIS Annual Conference, St Catherineís College Oxford, UK, 4th-5th April 2017, 2017.

[6] T. McClelland, îThe mental affordance hypothesis,ï Mind, **129**(514), 401ó427, 2020, doi: 10.1093/mind/fzz036.

[7] S.K. Patrick, E. Joshi, îSet in Stoneï or íWilling to Growì? Teacher sensemaking during a growth mindset initiative,ï Teaching and Teacher Education, 83, 156ó167, 2019, doi: 10.1016/j.tate.2019.04.009.

[8] R.S. Alsawaier, îThe effect of gamification on motivation and engagement,ï The International Journal of Information and Learning Technology, 2018, doi: 10.1108/IJILT-02-2017-0009.

[9] A.J. Elliot, îApproach and avoidance motivation and achievement goals,ï Educational Psychologist, **34**(3), 169ó189, 1999, doi: 10.1023/A:1009009018235.

[10] M.R. Lepper, J.H. Corpus, S.S. Iyengar, îIntrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates.,ï Journal of Educational Psychology, **97**(2), 184, 2005, doi: 10.1037/0022-0663.97.2.184.

[11] R.M. Ryan, E.L. Deci, îIntrinsic and extrinsic motivations: Classic definitions and new directions,ï Contemporary Educational Psychology, **25**(1), 54ó67, 2000, doi: 10.1006/ceps.1999.1020.

[12] J. Tranquillo, M. Stecker, îUsing intrinsic and extrinsic motivation in continuing professional education,ï Surgical Neurology International, **7**(Suppl 7), S197, 2016, DOI: 10.4103/2152-7806.179231.

[13] F.A. Hodis, L.H. Meyer, J. McClure, K.F. Weir, F.H. Walkey, îA longitudinal investigation of motivation and secondary school achievement using growth mixture modeling.,ï Journal of Educational Psychology, **103**(2), 312, 2011, doi: 10.1037/a0022547.

[14] P.R. Clinkenbeard, îMotivation and gifted students: Implications of theory and research,ï Psychology in the Schools, **49**(7), 622ó630, 2012, doi: 10.1007/978-981-13-3021-6_15-1.

[15] D. Sharek, E. Wiebe, îMeasuring video game engagement through the cognitive and affective dimensions,ï Simulation \& Gaming, **45**(4ó5), 569ó592, 2014, doi: 10.1177/1046878114554176.

[16] A. Ahmed, M.J.D. Sutton, îGamification, serious games, simulations, and immersive learning environments in knowledge management initiatives,ï World Journal of Science, Technology and Sustainable Development, 2017, doi: 10.1108/WJSTSD-02-2017-0005.

[17] F.D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q., 319ò340, 1989, doi: 10.2307/249008.

[18] What is Usage Behavior. (Access March, 31 2021) https://www.igi-global.com/dictionary/usage-behavior/71901, 2021.

[19] M. Csikszentmihalyi, Play and intrinsic rewards, Springer: 135ó153, 2014, doi: 10.1007/978-94-017-9088-8_10.

[20] M. Csikszentmihalyi, M. Csikzentmihaly, Flow: The psychology of optimal experience, Harper \& Row New York, 1990, doi: 10.1080/00222216.1992.11969876.

[21] B. Cowley, D. Charles, M. Black, R. Hickey, îToward an understanding of flow in video games,ï Computers in Entertainment (CIE), **6**(2), 1ó27, 2008, doi: 10.1145/1371216.1371223.

[22] E.Kuindersma, J.van der Pal, J.van den Herik, A.Plaat, îBuilding a game to build competencies,ï in International Conference on Games and Learning Alliance, 14ó24, 2017, DOI:10.1007/978-3-319-71940-5_2.

[23] P.Wouters, E.D.der Spek, H.Van Oostendorp, îCurrent practices in serious game research: A review from a learning outcomes perspective,ï Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices, 232ó250, 2009, DOI: 10.4018/978-1-60566-360-9.ch014.

[24] R.M. Ryan, E.L. Deci, îSelf-determination theory and the facilitation of intrinsic motivation, social development, and well-being.,ï American Psychologist, **55**(1), 68, 2000, doi: 10.1037/0003-066X.55.1.68.

[25] O.Perski, A.Blandford, R.West, S.Michie, îConceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis,ï Translational Behavioral Medicine, **7**(2), 254ó267, 2017, doi:10.1007/s13142-016-0453-1.

[26] M.K. Pedersen, N.R. Rasmussen, J.F. Sherson, R.V. Basaiawmoit, îLeaderboard effects on player performance in a citizen science game,ï ArXiv Preprint ArXiv:1707.03704, 2017.

[27] R.Farzan, J.M.DiMicco, D.R.Millen, C.Dugan, W.Geyer, E.A.Brownholtz, îResults from deploying a participation incentive mechanism within the enterprise,ï in Proceedings of the SIGCHI conference on Human factors in computing systems, 563ó572, 2008, doi: 10.1145/1357054.1357145.

[28] G.Zichermann, J.Linder, îGamification revolution,ï 2013.

[29] A.P.A. STYLE, îGUIDE TO THE 6TH EDITION PUBLICATION MANUAL OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION.ï

[30] R.R.Irwin, Cognition, Springer:65ó78, 2002.

[31] K.Park, îInnovative product usage behavior in the post-adoption process,ï ACR Asia-Pacific Advances, 1998.

[32] K.A. Abdel-Sater, îPhysiological positive feedback mechanisms,ï Am J Biomed Sci, **3**(2), 145ó155, 2011, ; doi:10.5099/aj110200145.

[33] R.Ronkainen, E.Kuusisto, K.Tirri, îGrowth mindset in teaching: A case study of a Finnish elementary school teacher,ï 2019, DOI: 10.26803/ijlter.18.8.9.

[34] P.Bouvier, E.Lavoué, K.Sehaba, îDefining engagement and characterizing engaged-behaviors in digital gaming,ï Simulation \& Gaming, **45**(4ó5), 491ó507, 2014, DOI:10.1177/1046878114553571.

[35] K. Cherry, îPositive reinforcement and operant conditioning,ï VeryWell Mind, 2018.

[36] G.Norman, J.Norcini, G.Bordage, Competency-based education:milestones or millstones?, 2014, doi:10.4300/JGME-D-13-00445.1.

[37] C.M.Mueller, C.S.Dweck, îPraise for intelligence can undermine childrenis motivation and performance.,ï Journal of Personality and Social Psychology, 75(1), 33, 1998, doi: 10.1037/0022-3514.75.1.33.

[38] E.OiRourke, E.Peach, C.S.Dweck, Z.Popovic, îBrain points:A deeper look at a growth mindset incentive structure for an educational game,ï in Proceedings of the third (2016)acm conference on learning@ scale, 41ó50, 2016, doi: 10.1145/2876034.2876040.

[39] R.P.French II, îThe fuzziness of mindsets:Divergent conceptualizations and characterizations of mindset theory and praxis,ï International Journal of Organizational Analysis, 2016, doi: 10.1108/IJOA-09-2014-0797.

[40] A.Derler, Growth Mindset Culture, Harvard Business Review, July 23.2018.

[41] C.S.Dweck, D.S.Yeager, îMindsets:A view from two eras,ï Perspectives on Psychological Science, **14**(3), 481ó496, 2019, doi: 10.1177/1745691618804166.

[42] W.Inchamnan, J.Chomsuan, îGamification Workflow for Growth Mindset Processes,ï in 2020 18th International Conference on ICT and Knowledge Engineering (ICT\&KE), 1ó6, 2020, DOI: 10.1109/ICTKE50349.2020.9289879.

[43] A.Hochanadel, D.Finamore, others, îFixed and growth mindset in education and how grit helps students persist in the face of adversity,ï Journal of International Education Research (JIER), **11**(1), 47ó50, 2015, DOI:10.19030/JIER.V11I1.9099.

[44] I. Rissanen, E. Kuusisto, M. Tuominen, K. Tirri, îIn search of a growth mindset pedagogy: A case study of one teacherìs classroom practices in a Finnish elementary school,ï Teaching and Teacher Education, **77**, 204ó213, 2019, doi: 10.1016/j.tate.2018.10.002.

[45] M.G.Jones, îCreating Electronic Learning Environments:Games, Flow, and the User Interface.,ï 1998.

[46] M.Sailer, L.Homner, The gamification of learning:A meta-analysis, 2020, DOI:10.1007/s10648-019-09498-w.

[47] J.C. Burguillo, îUsing game theory and competition-based learning to stimulate student motivation and performance,ï Computers \& Education, 55(2), 566ó575, 2010, DOI:10.1016/j.compedu.2010.02.018.

[48] C.-H. Chen, C.-H. Chiu, îEmploying intergroup competition in multitouch design-based learning to foster student engagement, learning achievement, and creativity,ï Computers \& Education, **103**, 99ó113, 2016, doi: 10.1016/j.compedu.2016.09.007.

[49] A.Dominguez, J.Saenz-de-Navarrete, L.De-Marcos, L.Fernández-Sanz, C.Pagés, J.-J. Martinez-Herráiz, îGamifying learning experiences: Practical implications and outcomes,ï Computers \& Education, **63**, 380ó392, 2013, doi: 10.1016/j.compedu.2012.12.020.

[50] W. Frkacz, îAn empirical study inspecting the benefits of gamification applied to university classes,ï in 2015 7th Computer Science and Electronic Engineering Conference (CEEC), 135ó139, 2015, DOI: 10.1109/CEEC.2015.7332713.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 147-153 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Vibration and Airflow Tactile Perception as Applied to Large Scale Limb Movements for Children

Hung-Chi Chu[1], Fang-Lin Chao[*,2], Liza Lee[3]

[1]*Department of Information and Communication Engineering, Chaoyang University of Technology, 436, Taiwan R.O.C.*

[2]*Department of Industrial Design, 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: 26 April, 2021
Accepted: 21 June, 2021
Online: 20 July, 2021

Keywords:
Vibrators
Zigbee module
Visual impaired
Airflow
Large movement

ABSTRACT

This study aimed to develop an airflow-vibrator motivated facility and assess exercise behaviors. The combination design involved computer-controlled airflow/ vibrators, a user interface program, and an adjustable structure presenting interaction options. The teacher and the participants can choose specific music with adjustable speed. The researcher did interviews during the initial test and field study. During the intervention, all participants succeeded in following the impinged flow with a positive emotional display. A wireless module and gas flow clue lifted the distance limitation of the vibration connection and enabled prompts in a larger area covered by radio waves. The flexible structure fit individuals ergonomic and the affordance consideration. After practicing, the students knew exactly how to pass and asked for the ball from the classmate. Wireless switch and signals give students more confidence in pitching — the participants successfully swap the body to follow the airflow.

1. Introduction

Rhythmic activities involve a combination of music, rhythm, and movement, which is intended to relax the mind and body. Today's teenagers show their creativity through hip-hop, and visually impaired individuals learn physical activities through touch and spoken instructions. For the visually impaired, spoken instructions are often unclear, and physical contact may cause discomfort to the other party. This study aimed to develop an airflow-vibrator-motivated facility and assess exercise behaviors. Although these technologies have gradually matured, there is no appropriate combination of teaching materials to form a reasonable price auxiliary tool. Thus, there is a gap connecting the technology elements and assistive teaching need of motivated facility and assessing exercise behaviors in school. The research work's contribution extends assistive technology of tactile perception to vibration and airflow, the combination of IoT devices, and the haptic feedback with a better user experience.

The combination of technology and design through tactile perception has been used to reduce the rhythmic learning difficulties of visually impaired children. Vibro-tactile feedback enhancement for orientation and obstacle avoidance can be obtained through the use of discreet actuators and obstacle detector sensors [1]. It provides frequent indications of useful dynamic information, such as level of proximity or distance. People can also perceive rhythm through tactile senses. In [2], the author described rhythm combined with frequency and amplitude to systematically produce 84 distinguishable tactile stimuli icons to help user perceptual tactile rhythm. In [3], the author presented unimodal and cross-modal rhythm perception with auditory, tactile, and visual modalities. Based on these findings, auditory and tactile modalities are suitable for presenting rhythmic information. Many assistive technologies have been successfully applied in education. Games that require physical activity are commonly used in rehabilitation to help restore physical function and balance [4]. Rehabilitation at home can reduce medical expenses and treatment time.

Visually impaired individuals often exhibit limited ability to participate in physical activities. Studies have found that visually impaired individuals take 50% longer to complete tasks with equivalent accuracy compared to sighted participants [5]. In [6], the author explained that movement is often limited by requiring outside guidance, fear of injury, and ridicule of others. However, when visually impaired individuals participated in bowling and

[*]Corresponding Author: Fang-Lin Chao, E-mail: flin@cyut.edu.tw

tennis video games with vibration and sound prompts, they exhibited improved mood and satisfaction following the activities.

Further study by [7] demonstrate the feasibility of real-time sensory substitution as a cost-effective approach for making gesture-based video games. In [8], the author presented a method of real-time video analysis to detect the presence of a particular visual cue as a cost-effective approach for making gesture-based video games. Users can express meaning in terms of vibration intensity or frequency, and haptics is also used for control, such as feedback when tapping an electronic pet. Feedback can be used as a reminder of movements in different parts, and studies have found that tactile sensation improves user engagement [9]. In [10], the author described the evaluation of human-computer interaction with the decision-making model. The information flow is helpful to monitor the effectiveness of the collective's activities. In [11], the author presented ubiquitous touch interaction with haptic feedback and show the movable world object supplies an accurate detection through a user study.

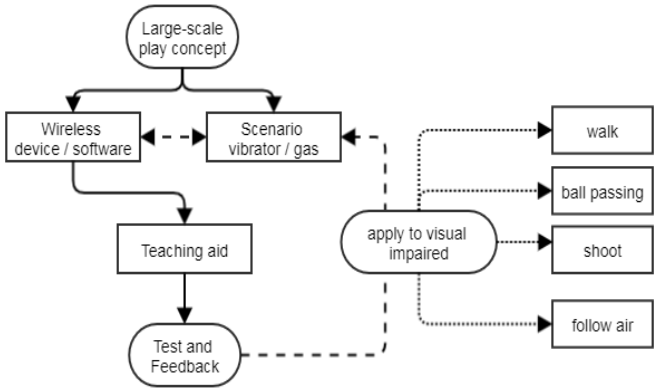
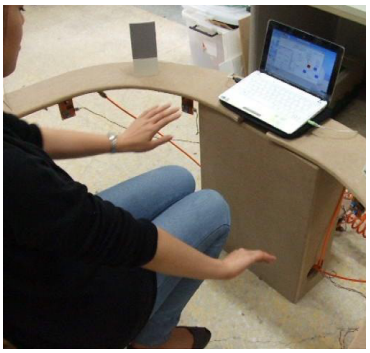


Figure 1: Design flow and applications for visually impaired

2. Detail Design

With the shrinking of the original size of the Internet of Things and the popularization of open platforms, researchers try to present the above ideas from the sensors and control components available on the market. This assistive tool involved a software development platform in performing human-machine interaction. Figure 1 indicates the design flow and possible applications for the visually impaired. First, start from the technical side, let those components be combined into the required system, and then look at the field of application and the idea of what kind of teaching materials teachers need. After the experiment in school, we extended it to the visually impaired (on the right) to apply it to different situations in their school life. The development platform and detailed parts in the design also show below.

As seen in Figure 2a, the vent is placed under the table for interacting with the user. The seated height of the participant was adjustable so that their legs remained at the same height as the air outlet so that their body could fully sense the airflow. As seen in Figure 2b, a large air compressor accommodates more compressed air. Figure 2b, top right shows the electromagnetic valve that controls the air outlet; bottom right shows the outlet of an exhaust pipe, which was an elongated slit that adjusted to the subject's position



(a)



(b)

Figure 2: Prototype of airflow assisted design: (a) use state; (b) compressor and spout.

A Phidget PC-USB interface board are utilized to provide an on/off sequence. As displayed in Figure 3, the interface control program has a graphical user interface using Microsoft Visual Studio with the current state of each switch. The button with a dark background indicates that it is currently activated. The text area on the right-hand side lists the sequence of the switching actions. A user can also manually control the switching by directly clicking on the buttons.

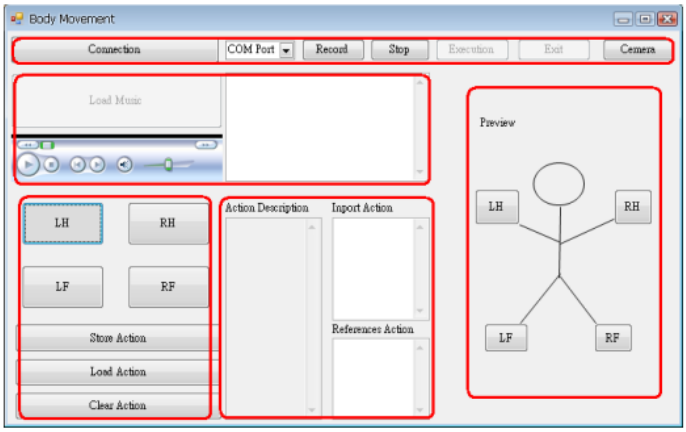


Figure 3. Software interface layout design with a set of the airflow outlet and timing control.

The total weighing of the vibrator is below 150 grams through circuit miniaturization to prevent additional loading during bodily movement. The vibrator is located in the skin contact surface with adjustable belts for proper contact. The control command could turn on/off the vibrator at a specific time. 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 extend the module's usage. The ZigBee interface is IP-Link5501, and the sensor node is IP-Link1223. 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.

3. Collocation and Research of Auxiliary methods

After observing a class of visually impaired children, the researchers proposed tactile perception to assist in directing physical activity. Engineer submitted proposals for the method of tactile prompting: vibration and airflow through software control. For example, during the design of the airflow, the compressor air was controlled by the researchers. In addition, musical elements ensemble games required the synchronization of music with a pair of wireless vibrators.

Research simulated visual impairment with blindfolds and made detailed adjustments to the study after testing the design setup. Participants were expected to don vibrators on their left and right hands and beat a tambourine with the help of cues from the vibrators (Figure 4).



Figure 4: Wireless vibrators on group's dual hand and wireless control software

Vibration and airflow cues complement each other: Vibration is more subtle and can prompt specific body movement. In contrast, airflow is more generally perceivable and can direct a wide range of group or individual actions. Experiments show that vibration stimulates activity, while the sound field of an airstream can indicate the direction of the movement. This study attempts to implement the combination of vibration and airflow into teaching: first by using vibration cues to instruct physical movement and airflow cues to guide group activity. Second, apply these findings to educate visually impaired individuals in physical activities.

3.1. Method

Although there are many ways to combine assistive technology and applications, this study selected vibration /airflow solutions for testing. We tested two separate groups:

(1) Children between the ages of 9 and 10 without visual impairments. These children had previously participated in rhythmic activities.

(2) Visually impaired children with general amblyopia or total blindness between the ages of 9 and 12 whose physical ability was close to that of children with normal vision. However, careful observation revealed that these students exhibited limited balance and restricted movement. Their teacher reported hoping to guide them to participate in more activities.

Eight independent wireless vibration modules were set up in a system where a single host controlled the signal. When configured between different people in the same group, they could be guided by the system to work together. The advantage of wireless communication was that subjects using it could move freely.

3.2. Group teaching

During group teaching, the participants were given different individual activities through the prompts of the vibrator after being linked. As a result, the following ideas were put forward to expand applications in teaching.

Activity 1: Ensemble: increase children's concentration and sensitivity to rhythm through touch. The teacher instructed children to strike a tambourine when they felt a vibration, forming an ensemble rhythm.

Activity 2: AB mode: Increase children's focus through two touching ways: A for a single short vibration and B for continuous vibration. Participator performs specific actions corresponding to the different vibrations when prompted at irregular intervals (Figure 5). The goal was to see if the child could concentrate and distinguish between different vibration patterns.

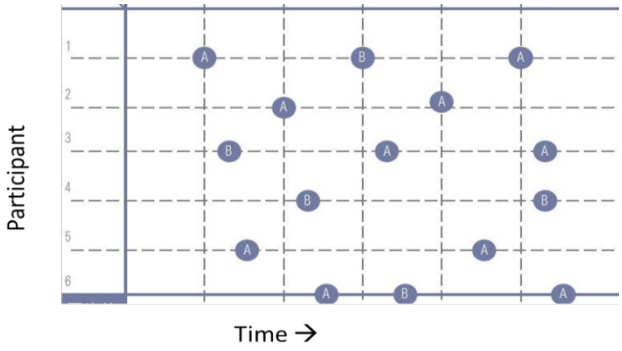


Figure 5: Simultaneous A/B vibration test

Activity 3: Consecutive A/B: Promote children's memory through tactile sensation and movement; promote children's creativity through music activities. One child was instructed to perform a motion after feeling the vibration signal. The next child was asked to imitate the action of the previous child in response to the vibration signal and then add a second motion to the sequence. A third child was then asked to perform the two last moves and add a third motion when they felt the vibration signal.

The jet sound is received through tactile and auditory cues. Air rhythmic action is the primary stimulus, supplemented by vibrations to prompt a change in the application, such as in the three activities above. When the child's back was facing the airflow, the study design proposed a scenario of moving a wall to instruct the child to push the invisible wall with his hand so that the child could express creativity during movement. When vibration is combined with airflow, large-scale airflow can guide actions in large areas to different locations, such as forward or backward ambulation, walking around a circle, or pointing in a specific direction.

3.3. Evaluation Criteria and Behavioral Assessment Tools

Researcher used the following pediatric behavioral assessment tools:

(1) Wenlan Adaptation Behavior Scale

The Wenlan Adaptation Behavior Scale (Motor Skills Section), which is used to measure motor skills development in children,

was completed at baseline. The test was conducted to track the changes in the children's physical ability.

(2) Observation and Evaluation Record Form for children's activities

The observation records of children were also used as a research tool for participatory observation research. Under controlled circumstances, systematic observations of phenomena or individual behaviors were based on established research goals. Observers recorded the performance of the assessed scores in writing. Finally, observer made an objective explanation of the phenomenon or individual behavior. The observation variables are defined as following: week-activity-participant-vibration mode; so that notation (1-2-M-B) means teaching week-1, playing activity-2, with participant-M, using continuous vibration mode-B.

The content of the observation table was divided into structured assessments, supplemented by available text supplements, and recorded by observing the child's reaction and concentration during movement. The teacher adopted a Likert five-point scoring method for evaluation. A score of 5 indicated frequent behavior observed during the activity. Three observers rated the frequency of each behavior according to the situation described in the items and finally calculated the average result of the observation score. The evaluation items included:

(1) Movement:

Students respond to vibrator and waves limbs. Uses instrument correctly according to instructions and beats instrument in time to the music. Responds to and moves limbs according to airflow position.

(2) Attention

Senses change in vibration mode. Uses hearing and physical perception to pay attention and engage to the airflow.

(3) Emotion

The student expresses positive emotion through spoken words, movement, and expression. Maintains positive feeling throughout the airflow process.

(4) Creativity

Moves in varied ways in the designated space according to the vibration prompts. Swings body and waves limbs were corresponding to airflow.

4. Results and Discussion

During the first week, observers found that the children were unsure how to react to the vibrators, which caused them to respond incorrectly. In addition, they did not make prominent movements in response to a prompt for free creative actions because of personal shyness. In the second week, the teacher added explanations, introduced movement simulation, and carried out group activities using airflow stimulation. We found that children exhibited significantly different motion feedback and even actively added new motions. The research found that children were integrated into the activity as a whole and actively with new movements in third week. Further details are as follows:

4.1. Week 1 record

Activity 1: Ensemble

1. The children were unfamiliar with the vibrator and could not accurately sense it. Therefore, the performance is not apparent (1-1-M-A) (Figure 6).

2. The children immediately shook their arms when both hands were shaken and showed their arms after a pause (1-1-M-B).

3. After putting on the vibrator, the children kept looking at it. They started by making fists and frowning. (1-1-A-B).

4. The children smiled before the start of the activity. They laughed and jumped on the tambourines. (1-1-E-B)

5. The children beat the tambourines correctly from beginning to end, maintaining a single movement pattern without change. (1-1-E-B)



(a)



(b) (c)

Figure 6: Student behaviors of three events: (a) ensemble, (b) A/B mode, and (c) A/B two-point.

Activity 2: A/B mode

1. The children could not clearly distinguish between A and B vibration patterns and demonstrated a high rate of action repetition. (1-2-M-A)

2. A small number of children could make the correct movements but could not clearly distinguish patterns when making inappropriate movements (1-2-M-B).

3. When the children wore the vibrator, they frowned and looked at it. They appeared quiet during the activity. (1-2-A-B)

4. Most of the children smiled during the activity, indicating excitement to participate. (1-2-E-B)

5. The child exhibited no change in movement and remained in place. (1-2-C-B)

Activity 3: Consecutive A/B

The instructions were not followed precisely and the children required cueing from the instructor. (1-3-M-A)

Some children responded quickly to the signal, but some needed a reminder from their peers. (1-3-M-B) (Figure 6b)

The children devoted themselves to the activity and paid attention to their vibrator. One child said, "Yeah, I'm here," immediately after the vibration. (1-3-A-B)

The children exhibited happiness through expression and body language engaged in the activity. (1-3-E-B)

The movements of the children were almost all performed while running (one was jumping). (1-3-C-B)

In response to wearing a wireless vibrator, the children's overall body language was stiff, and their movement rarely changed. While some exhibited creativity and came up with their ideas, most of them repeated the body movements of others. Some children could focus on the vibrator on their wrists and listen to the instructions of the tester. The children looked nervous but exhibited no negative emotions.

4.2. Week 2 record

Activity 1: Ensemble

Under the instructor's guidance, the children's ability to tap to the music improved, but the tapping method did not change. (2-1-C-A)

The children correctly responded to changes in vibration and beat the tambourines at the same time. When only one vibrator signaled them, most of them only moved one hand. (2-1-M-B)

The children were fully involved in the activity and could primarily respond correctly. (2-1-A-B)

The children laughed frequently and waved their arms, exhibiting positive emotions. (2-1-E-B)

The children's movements were primarily similar, and the rhythm was maintained when the facilitator participated. (2-1-C-B) (Figure 7a)

Activity 2: AB Mode

The children were still unable to create new movements because of nervousness. (2-2-C-A)

The children were able to perform AB mode correctly and respond to different prompts. (2-2-M-B)

The children demonstrated a high willingness to participate. Children were smiling and showing positive emotions during activities. (2-2-E-B)

They maintained the same movement with no other limb movements demonstrated. (2-2-C-B)



(a) (b)

Figure 7. Week 2 study: (a) ensemble (b) A/B two-point

Activity 3: Consecutive A/B

The children correctly judged the location of the airflow. (2-3-M-A) (Figure 7b)

The children made significant progress in focus and emotions. (2-3-A, E-A)

There were innovative movements. (2-3-C-A)

Most children could feel the vibration correctly, but one child happily ran around with the airflow; others only had physical movements. (2-3-M-B)

All of the children could focus on activities and follow the action. (2-3-A-B)

The children demonstrated a high willingness to participate and showed positive emotions throughout. (2-3-E-B)

The children exhibited 3 distinct movements; in addition to the original running and jumping, they were also sidestepping. (2-3-C-B)

After the participating children donned wireless vibrators and performed activities with airflow, the children occasionally need to be reminded by the testers. The overall magnitude of limb response became more significant, and the children were willing to cooperate actively to maintain rhythm (Figure 8). After growing familiar with the rhythm, they could focus on vibrational changes and react instantly. When students improved their sense of familiarity, the children had a high willingness to participate, showing positive emotions and smiling.

4.3. Week 3 record

The child feels the way of shaking (3-1-M-A). They demonstrated a high rate of repetitive movements, so they did not perform well in creativity. (3-1-C-A)

Most children could correctly sense the vibration pattern, but a few children exhibited slightly longer reaction times. (3-1-M-B)

The children engaged in the activity. (3-1-A-B)

The children demonstrated a high willingness to participate, showing positive emotions. (3-1-E-B)

Two children performed somersaults, swinging the tambourine from top to bottom. (3-1-C-B)

The children's energetic performance improved significantly, and they could respond to tactile vibrators, and air prompts. They were also more engaged in activities. During the exercise, they cooperated with a high degree of concentration. The children talked naturally with the tester and performed activities with enthusiasm; at the end, they actively asked when they could participate again (Figure 8). During this period, children engaged in simple competition with their peers, so more innovative actions appeared, such as somersaults and trunk rotations. Overall, the children's performance in terms of limb movement, concentration, emotions, and creativity improved significantly.



(a) (b)

Figure 8: Week 3 study: (a) airflow + vibrator, (b) ensemble and creative move

4.4. Overall Performance

Haptic perception successfully assisted motor skill development in children. In terms of overall performance, children have noticeably changed four goals: movement, attention, emotion, and creativity. The children's performance may have been limited due to unfamiliarity with the tester in the early stages. In the middle period, movements and expressions changed gradually as the children grew more comfort. The sensation of airflow was highly attractive to the children. When airflow appears in activities, the children seemed excited and did not want to leave so they could continue playing.

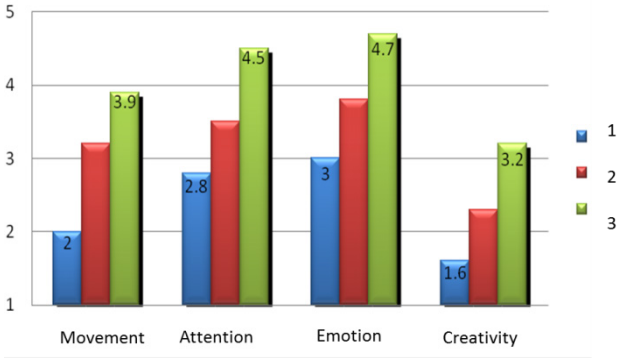


Figure 9: Children's overall performance during three weeks

The quantitative data (using the SPSS) of three class activities are analyzed and discussed below. Figure 9 shows the evaluation results using the Likert five-point scoring method. The children's overall performance in the later period improved, not only in more physical movements but also in the position of the vibrator and airflow. New actions appeared, such as somersaults. After the activity, the children asked if they could participate again, indicating that the activity was attractive. With appropriate syllabus design, children can demonstrate positive physical and emotional merges.

5. Applications for Visually Impaired Children

Visually impaired children with general amblyopia or total blindness, aged between 9-12 years, possess physical ability close to children with regular sight. Before the activity, however, observation revealed that the visually impaired students exhibited limited balance ability. In addition, they usually exhibit restricted movement. Because of safety concerns, the teacher gave them many constraints; for example, they had to go in pairs one by one when they went outside the classroom. When the teacher asked them to stand on one foot, they appeared unstable and eager to reach for the objects next to them for support. Their movements were less varied, and they presented with stiff posture. The teacher hoped to guide them to engage in more activities using the assistive devices. The wireless module lifted the distance limitation of the vibration connection and enabled prompts in a larger area covered by radio waves signal.

The scenario items included: ensemble, orientation, and passing ball, shoot the ball, and follow the change of airflow. Firstly, a small-scale practice using the previous ensemble and A/B game. The students were given different instruments instructed to tap the instrument through music and vibrator cues. The vibrator cues provided a rhythmic tacit understanding of ensemble participation. Visually impaired students have a good sense of sound and rhythm, students successfully played with musical instruments and A/B game.

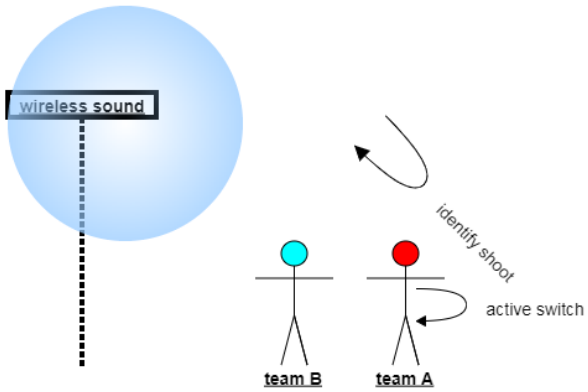


Figure 10: Schematic drawing of the large-scale activities: vibrator cues provided walking instructions, students improved pitching in a different location.

Secondly, six pupils with amblyopia in special schools participated in four kinds of large-scale movement play for 60 minutes (Figure 10).

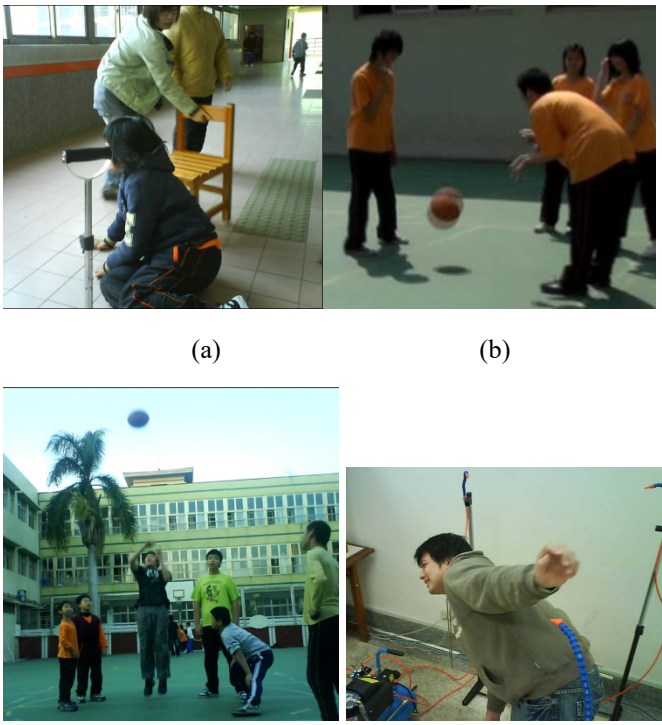
(1) Orientation: Vibrator cues provided directions and walking instructions. Testers used vibrators to indicate right and left to help children reach their destination.

(2) Passing practice: Vibrator cues are used to instruct children to throw a ball back and forth and improve communication efficiency. The vibrator provided the teacher's instructions to pass and receive the ball.

(3) Shooting: We set up a wireless sounder on the basketball frame. When students receive the ball and want to shoot, they start the sound source with the wireless switch and judge the direction base on the sound from target area. We found that wireless signals give students more confidence in pitching.

(4) Quickly follow the change of airflow: promptly change the air outlet point, the participants rotate the body to monitor the airflow to form a continuous action creatively.

Sometimes students were out of sync during orientation due to the speed and circuit delay. After practicing, the students knew exactly how to pass and asked for the ball from the team classmate (Figure 11). Wireless switches give students more confidence in pitching it. The participants swap the posture to follow the airflow successfully. Through the interview, the vibrator and gas clues enabled more interaction in large-scale exercises. Wireless technology and moving gas help remote instructions to assist users in interacting with other people, guidance from teacher, and the signal in the surrounding environment.



(a) (b)

(c) (d)

Figure 11. Large scale activities: (a) vibrator cues provided walking instructions in orientation practice (children walk along the path), (b) students pass the ball according to vibration cues, (c) students improved pitching in a different location, and (d) dynamic gas clues for posture following.

6. Conclusion

The aim of the research is developing tactile perception assistive technology to activate the physical activity of children. The design method utilized vibration and airflow actuators and a control system to provide proper feedback and acknowledgment. The study also developed the corresponding instructions and introducing context to promote children's participation. Three weeks evaluation process conducted through field practice in elementary school with guiding activities. Both explaining hints and free creative activities were encouraged in classroom group cooperation activities. They were also able to recognize different vibration patterns that respond to the corresponding limb movements. The field testing results indicated children's positive responses and spontaneous creative actions. Their overall performance in exercise, attention, emotion, and creativity improved significantly during the evaluation instructions. The research extends assistive technology of tactile perception to vibration and airflow. The combination of new IoT devices of wireless sensor networks and remote control of haptic feedback are helpful to visually impaired individuals, which support better instruction in the future classroom.

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] G. Ghiani, B. Leporini, F. Paternò, "Vibrotactile feedback to aid blind users of mobile guides," Journal of Visual Languages & Computing, **20**(5), 305-317, 2009, doi: 10.1145/1409240.1409306.

[2] D. Ternes, K. E. Maclean, "Designing large sets of haptic icons with rhythm," in 2008 International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, 199-208, Heidelberg, 2008, doi: 10.1007/978-3-540-69057-3_24.

[3] M. Jokiniemi, R. Raisamo,J. Lylykangas, "Crossmodal rhythm perception," in International Workshop on Haptic and Audio Interaction Design, 111-119, Heidelberg, 2008, doi: 10.1007/978-3-540-87883-4_12.

[4] B. Lange, C.Y. Chang, E. Suma, B. Newman, "Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor," in 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1831-1834, 2011, doi: 10.1109/IEMBS.2011.6090521.

[5] S.A. Douglas, S. Willson, "Haptic comparison of size (relative magnitude) in blind and sighted people," in 2007 Proceedings of the 9th international ACM SIGACCESS Conference on Computers and Accessibility, 83-90, 2007, doi: 10.1145/1296843.1296859.

[6] D. Morelli, E. Folmer, "Real-time sensory substitution to enable players who are blind to play video games using whole body gestures," Entertainment Computing, **5**(1), 83-90, 2014, doi: 10.1016/j.entcom.2013.08.003.

[7] D. Morelli, J. Foley, E. Folmer, "VI-Tennis: a vibrotactile/audio exergame for players who are visually impaired," in 2010 Proceedings of the Fifth International Conference on the Foundations of Digital Games, 147-154, 2010, doi: 10.1145/1822348.1822368.

[8] L. Zoccolillo, D. Morelli, M. Iosa, "Video-game based therapy performed by children with cerebral palsy: a cross-over randomized controlled trial and a cross-sectional quantitative measure of physical activity," Eur J. Phys Rehabil Med, **51**(6), 669-76, 2015, doi: 10.14288/hfjc.v10i1.225.

[9] J. Yim, N. Graham, "Using games to increase exercise motivation," in 2007 Proceedings of Conference on Future Play, 166-173, 2007, doi: 10.1145/1328202.1328232.

[10] A.H. Hoppe, F. Marek, F.v.d Camp, R. Stiefelhagen, "Extending movable surfaces with touch interaction using the virtualtablet: an extended view," Advances in Science, Technology and Engineering Systems Journal., **5**(2), 328-337, 2020, doi: 10.25046/aj050243.

[11] N. Bakanova, A. Bakanov, T. Atanasova, "Modelling human-computer interactions based on cognitive styles within collective decision-making," Advances in Science, Technology and Engineering Systems Journal, **6**(1), 631-635, 2021, doi: 10.25046/aj060169.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 154-160 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Novel De-rating Practice for Distributed Photovoltaic Power (DPVP) Generation Transformers

Bonginkosi Allen Thango*, Jacobus Andries Jordaan, Agha Francis Nnachi

Dept. of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 08 March, 2021
Accepted: 28 April, 2021
Online: 20 July, 2021

Keywords:
Transformer
Losses
Hotspot temperature rise
Finite Element Method
De-rating

ABSTRACT

Transformers are habitually designed and manufactured for operation at a fundamental frequency of 50Hz and sinusoidal load current. Transformers are susceptible to non-linear loads. The inception of switching action characterises Non-linear loads and consequently nonsinusoidal load current which brings about higher transformer service losses, hotspot temperature rise, and degradation of cellulosic and liquid insulation, and consequently untimely failure of transformers during service. This phenomenon yield current with different components that are multiples of the fundamental frequency of the distributed photovoltaic power (DPVP) generation system. In order to obviate these challenges, the continuous power rating of the transformer, which is intended to facilitate non-linear loads must be minimised using procedure ascribed by the standards as de-rating. This work, an extension of previous work, proposes a novel procedure by means of Finite Element Method (FEM) for the de-rating of DPVP transformers serving non-linear loads during their service life. The proposed procedure considers parameters such as skin effect, proximity effect, and the magnetic flux leakage on the windings that were not included in the IEEE recommended de-rating procedure. The theoretical examination is substantiated on a 500kVA, three-phase, oil-filled transformer.

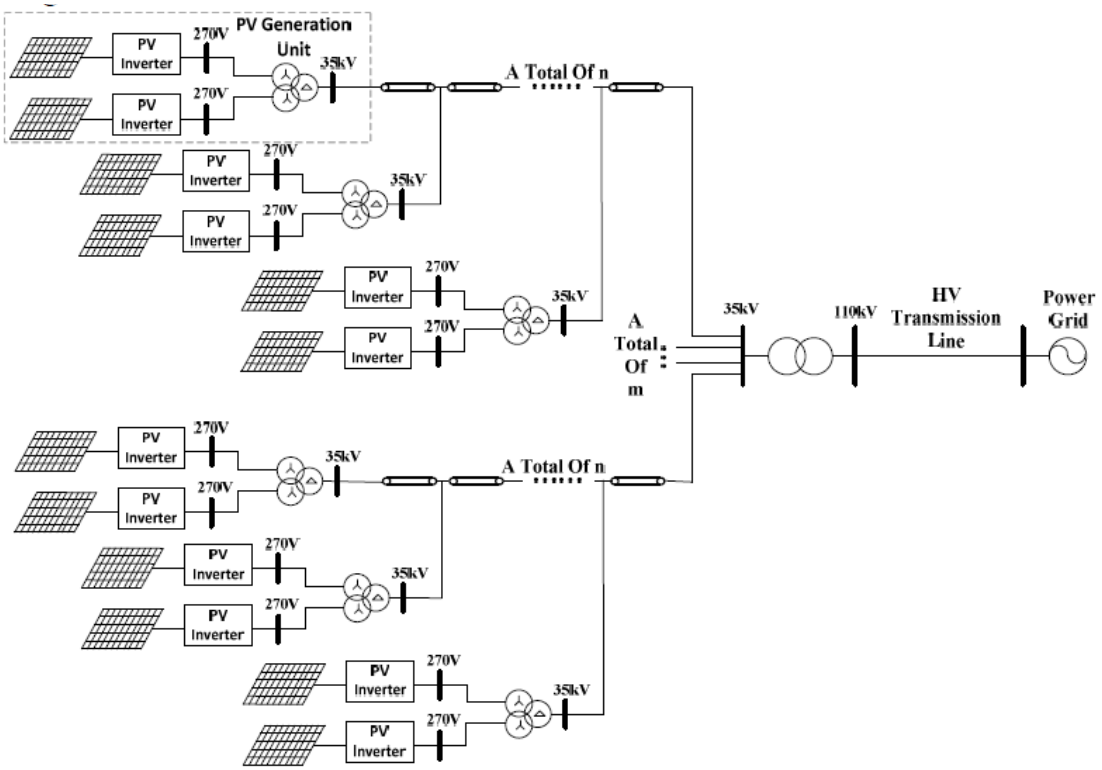
1. Introduction

In this day and age, distributed photovoltaic power (DPVP) generation producers are apprehensive concerning allotting of power ratings for transformers that are projected to operate in harmonically contaminated environment. The use of regular distribution transformers has so far proven to be impotent in handling the operational requirements of DPVP generation during service. To pledge future reliability of DPVP generation, transformers facilitating this application must embed these requirements, in which includes sporadic loading cycle, harmonic and distortion and de-rating in the design philosophy. Relying on conventional regular distribution transformer design philosophies can result in conditions where surged voltages and current could abbreviate a DPVP transformer's service life. Owing to the economic status of competitive rates for DPVP projects and since transformers are ordinarily not operated at rated loading, there is a tendency to de-rate the transformer. There may even be a propensity to oversize the transformers to increase the winding Eddy losses due to harmonic currents seen by the transformer during service. Considering that DPVP transformers are not rated at regular distribution transformer ratings, effort to adjust it to the standard kVA rating is laborious. Standard DPVP generating ratings generally lie between standard distributing transformer ratings, so there appears to be an inclination to select the nearest in spite of the power being de-rated.

Studies that consider the de-rating of transformer ratings when supplying harmonic currents are prevalent in the publications [1] – [7]. These publications are based on case studies of measured results and analytical formulations. The approach on this study have a similar downside of not accounting for significant parameters suchlike skin effect, proximity effect, and the magnetic flux leakage on the windings at fundamental and under harmonic conditions.

A configuration of a DPVP plant is demonstrated in figure 1 [8] and comprises of a cluster of PV generation inverter schemes, in which the generated power is collected by the power retrieval system into a 35kV bus. In the PV generation inverter schemes, two PV arrays are connected with PV inverters and into a 500kVA transformer through a LCL filter and then accessible to the station PV energy retrieval system. The generated energy is fed into the national grid through a high voltage (HV) transmission line.

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com



In view of the intermittent nature of the DPVP plant and the switching action of inverters, in case the harmonic current output of an individual PV generation inverter scheme is minimal, the output harmonic current of the cluster of PV generation inverter schemes has the potential to exceed the limitations recommended by the standard [9].

In [10], transformer losses that occur when facilitating solar PV farm environment were investigated based on seasons. The study in particular draw attention to the injection of harmonics and distortion using the IEEE Std. C57.110-2018 recommended practice. This work, an extension of the previous work in [10], examines the influence of harmonics and distortion on transformers during their service lifetime. The work further review the de-rating procedure recommended by IEEE for transformers under harmonic conditions. The work then proposes a novel procedure for de-rating transformers using the computational power of FEM. The proposed procedure takes into account of parameters suchlike skin effect, proximity effect, and the magnetic flux leakage that the analytical method (AM) recommended by the IEEE fails to take into consideration.

2. Effect of harmonic currents

During service the total transformer losses (P_{TOT}) are comprised of the no-load loss (P_{NL})and the load losses (P_{LOAD}) as expressed in eq. (1) below.

$$P_{TOT} = P_{NL} + P_{LOAD} \quad (1)$$

The service no-load losses are as a result of the core excitation when the transformer is connected to the supply voltage and is independent of the loading profile. Under harmonic conditions, the harmonic current passing through resistance and leakage reactance of the transformer may deform the output voltage. Practical experience has demonstrated that the rise in temperature in the core is not a factor limiting the evaluation of the permissible current under harmonic conditions. Evidently, the C57.110-2018 [11] recommended practice for establishing transformer capability under harmonic conditions does not take the account the no-load losses under such conditions.

The load losses (P_{LOAD}) are comprised of the copper losses (I^2R) and the winding stray losses in which are constituted by the winding Eddy losses (P_{WEC}) and stray loss in structural parts (P_{OSL}) as expressed in eq. (2) below [11].

$$P_{LOAD} = I^2R + P_{WEC} + P_{OSL} \quad (2)$$

The copper losses signifies the heat dissipated by the load current in the transformer winding conductors. Under harmonic conditions if the load current increases, then the copper losses will also experience an increase. During the factory acceptance testing of a transformer there is no methodology available for the testing of the winding stray losses. Although, the impedance test can be employed to ascertain the total transformer losses. Then the winding stray losses can be acquired by deducting the copper losses from total transformer losses. In the event that the rated winding Eddy loss of a transformer is known, then this loss under harmonic conditions can be evaluated as expressed in in eq. (3) [11].

$$P_{WEC} = P_{WEC_R} \times \sum_{h=1}^{h_{max}} \left[\frac{I_h}{IR}\right]^2 \times h^2 \quad (3)$$

The stray loss in structural parts can be evaluated for harmonic conditions by applying a similar procedure. In the 4th edition of

the UL 1561 standard published in 2011[12], the Factor-K for de-rating a transformer is specified as expressed in eq. (4) below.

$$K_{factor} = \left[\sum_{h=1}^{hmax} {I_h}^2\right] \times F_{HL} \quad (4)$$

The Factor-K indicates the effect of harmonic conditions upon the increase in the winding Eddy losses of the transformer. In the C57.110-2018 standard [11], the harmonic loss factor to account for the increase in the winding Eddy losses is expressed as follows in eq. (5).

$$F_{HL} = \frac{P_{WEC}}{P_{WEC_R}} = \frac{\sum_{h=1}^{hmax} {I_h}^2 \times h^2}{\sum_{h=1}^{hmax} {I_h}^2} \quad (5)$$

On the empirical studies conducted in [13], for evaluating the winding Eddy losses, eq. (6) and eq. (7) are established.

$$P_{WEC_R} = \frac{0.8 \times P_{WEC_R}}{{I_2}^2 I_L} \quad (6)$$

$$P_{WEC_R} = \frac{2.8 \times P_{WEC_R}}{3 \times {I_2}^2 I_L} \quad (7)$$

The maximum allowable harmonic load current formula recommended by the C57.110-2018 standard [11] of a transformer is the current whereupon the maximum winding Eddy loss ratio as expressed in eq. (8) below.

$$I_{max} = \sqrt{\frac{1+P_{WEC_R}}{1+F_{HL} \times P_{WEC_R}}} \quad (8)$$

3. Case scenario

In this case scenario, the technical characteristics of the studied 500kVA, three-phase, oil-immersed, DPVP transformer are presented in Table 1. The objective herein is to investigate the operational performance of the studied transformer that facilitate a harmonically distorted load current during service.

Table 1: Technical specification

Item	Value
kVA rating	500kVA
HV Voltage/HV Current	11kV/1.44A
LV Voltage/LV Current	400V/79.2A
Load Losses (@75% loading)	1375W
HV winding resistance	121.5 ohms
LC winding resistance	0.03 ohms

The corresponding harmonic load current considered in the analysis is presented in Figure 2. Further, the de-rating information of the studied transformer according to the analytical method and the proposed FEM procedure will be computed based upon the supplied harmonic load current.

In preparation to shed light on the concept of skin effect on the winding conductors at the fundamental frequency and under harmonic conditions, the two simulations with a round cross-section were investigated. These simulations substantiate the behaviour of the load current due to the applied frequency. At fundamental frequency, the effect of the skin effect upon the winding conductors will be as presented in Figure 3.

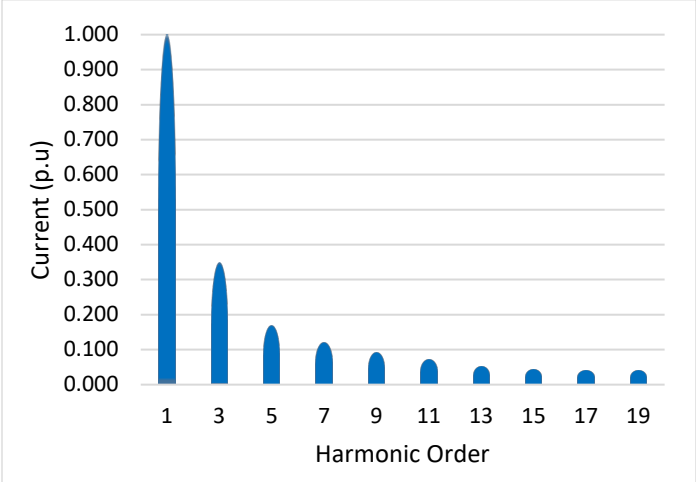
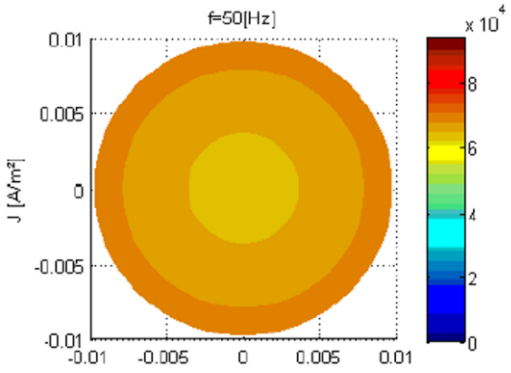
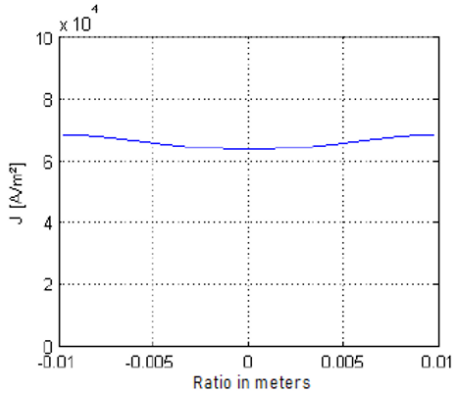


Figure 2: Harmonic spectrum.



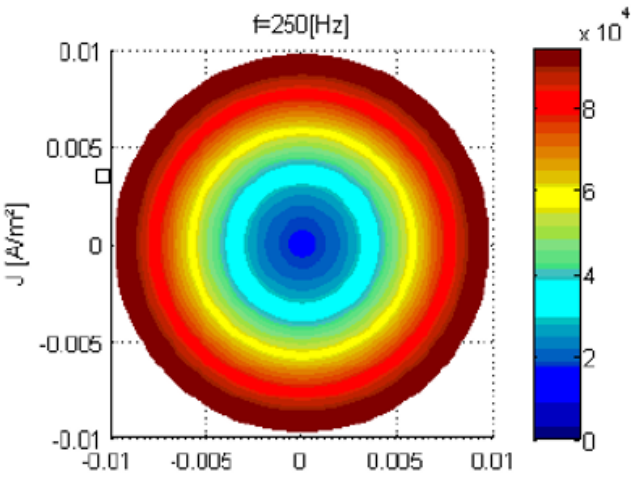
(a)



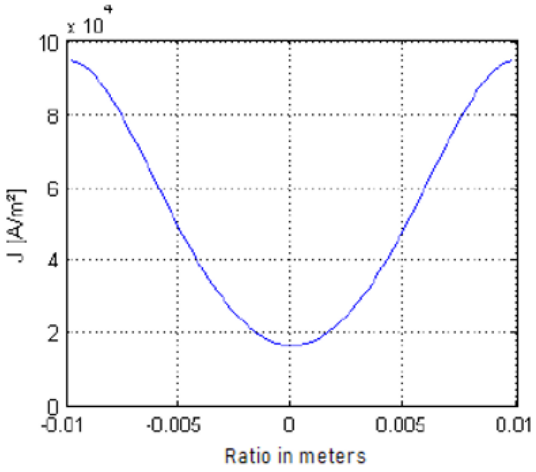
(b)

Figure 3: Winding conductor's skin effect at a fundamental frequency

As anticipated, Figure 3 (a) illustrate a uniform dispersion through the winding conductor. The graphical representation of this effect is shown in Figure 3 (b). The 5th order of the supplied harmonic profile was investigated and the corresponding results in Fig. 4 demonstrate a concentration of the current upon the surface of the winding conductor as demonstrated by Figure 4(a). The graphical representation of this effect is also presented in Figure 4 (b).



(a)



(b)

Figure 4: Winding conductor's skin effect at 5th harmonic order

4. Results

In this section, the results for de-rating a transformer to facilitate the supplied harmonic load current using the analytical method and the proposed 3D FEM procedure.

4.1. Classical Approach: Analytical method

The winding Eddy losses at fundamental and under harmonic conditions are acquired by the difference between the total transformer losses and copper losses as:

$$P_{WEC_R} = 1375 - 3 \times [1.44^2(121.5) + 79.2^2(0.03)] = 54{,}635W$$

The peak ratio of the winding Eddy losses is evaluated in p.u using eq. (6) given that the studied unit is within the category up to 650kVA.

$$P_{WEC_R} = \frac{0.8 \times 54.635}{79.2^2 \times 0.03} = 0.232\ p.u$$

In order to attain the Factor- K, the supplied harmonic load current is applied, as shown in Table 2.

Table 2: K-factor estimation

h	I_h(A)	$I^2{}_h$(A)	h^2	$I^2{}_h \times h^2$(A)
1	1,000	1,000	1,000	1,000
3	0,350	0,123	9,000	1,103
5	0,170	0,029	25,000	0,723
7	0,120	0,014	49,000	0,706
9	0,092	0,008	81,000	0,686
11	0,071	0,005	121,000	0,610
13	0,051	0,003	169,000	0,440
15	0,043	0,002	225,000	0,416
17	0,040	0,002	289,000	0,462
19	0,039	0,002	324,000	0,493
Σ		**1,187**		**6,637**

The maximum allowable current of the studied transformer based in eq. (8) is evaluated as follows:

$$I_{max} = \sqrt{\frac{1 + 0.232}{1 + 6.637 \times 0.232}} = 0.696$$

The maximum allowable load current in Ampere (A) is evaluated as:

$$I_{max} = 0.696 \times 79.2 = 55A$$

The continuous kVA rating of the transformer during service is evaluated as:

$$Equivalent\ kVA = 500 \times 0.696 = 347.263kVA$$

During service, this equivalent kVA is the valuation of the continuous power rating the transformer will be operating on under the supplied harmonic spectrum. In the event the utility owner changes the harmonic loading seen by the transformer, the kVA must be re-evaluated

4.2. Finite Element Method: Proposed Procedure

In the development of the proposed procedure, Ansys Maxwell, an electromagnetic field simulation software, is employed. A 3D cross-section of the three-phase, oil-immersed, DPVP transformer is developed for the finite element software. The short circuit test of the transformer is evaluated using the circuit model at fundamental and under harmonic conditions. The superposition property of the winding Eddy losses is employed in the event when the winding conductor dimensions are not more than the skin and proximity effect as well as when the flux degree is lesser than the saturation level. At large, the superposition property is employed to evaluate the transformer load losses under harmonic conditions. This loss is thereby the arithmetic sum of the losses due to different harmonic load current and harmonic orders. For the proposed FEM procedure, the windings' leakage flux and resistance are used to calculate the different loss components.

In Figure 5, the 3D FEM cross-section of the studied transformer is presented. The model takes into account the actual geometries, material properties and the supplied harmonic spectrum. Undoubtedly, the finer meshing of the geometries takes an extended processing time but leads to the most optimised solution. Consequently, in the assessment of the various harmonic

orders, a concession has been formed between the optimised solution and extended processing time.

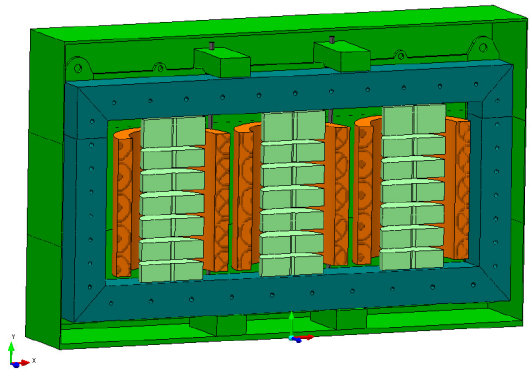


Figure 5: Proposed 2D FEM model

For the purpose of obtaining the results of the proposed 3D FEM procedure, the following steps are carried out on the transformer model:

- Excitation of the windings by the harmonic load current of each harmonic order,
- The corresponding flux density distribution in the active components are then calculated,
- The triggered currents generated by the distributed flux along with the material core, are determined, and
- The resistances of the windings are then presented to the geometries and the winding Eddy losses are then computed premised on the copper.

In the event the transformer under study is operating at a fundamental frequency, the load losses by employing the proposed FEM procedure are tabulated as shown in Table 3.

Table 3: Technical specification

Loss component	Value
HV load loss (Watts)	879,16
LV Load loss (Watts)	455,45
Total Load loss (Watts)	1334,59

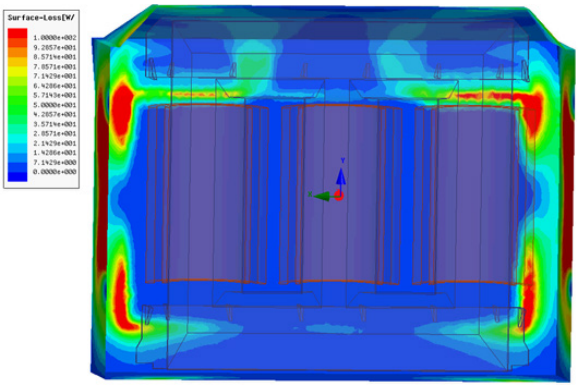


Figure 6: Magnetic flux leakage distribution.

In the magnetic circuit of the studied transformer, the distribution of the magnetic flux density is illustrated in Figure 6. At loading condition during service, the short circuit phenomena of the transformer may be presumed and the amplitude of the magnetisation current and leakage flux are minimal. As a result of the short circuit of the winding phases, significant density seeps through the windings. The latter presents a crucial factor in this condition. Under the supplied harmonic spectrum supplied in Figure 2, the total transformer losses are calculated as illustrated in Table 4 below.

The services load losses for each harmonic order is attained and tabulated as shown in Table 4 as 1410.53W.

Table 4: K-factor estimation

h	I_h(A)	P_H(W)	P_L(W)	P_{TOT}(W)
1	1,000	668,09	346,10	1014,19
3	0,350	167,02	86,53	253,55
5	0,170	58,28	29,68	87,95
7	0,120	20,05	11,49	31,54
9	0,092	6,68	2,87	9,56
11	0,071	5,51	2,83	8,34
13	0,051	1,82	0,74	2,56
15	0,043	0,61	0,49	1,10
17	0,040	0,81	0,39	1,20
19	0,039	0,37	0,18	0,55
Σ		**929,24**	**481,29**	**1410,53**

The winding Eddy losses at fundamental and under harmonic conditions are acquired by the difference between the total transformer losses and copper losses. The overall load current is at both conditions is presumed to be 1 p.u. The winding copper losses at a fundamental frequency and current as evaluated as:

$$3 \times [1.44^2(121.5) + 72^2(0.03)] = 1320.36W$$

The winding Eddy losses at fundamental frequency are evaluated as:

$$1410.53 - 1320.36 = 14.23W$$

The winding Eddy losses under harmonic conditions is evaluated as:

$$1410.53 - 1320.36 = 90.18W$$

By employing eq. (5), the harmonic loss factor is the ratio of the winding Eddy losses under harmonic conditions and at fundamental frequency.

$$F_{HL} = \frac{90.17}{14.23} = 6.337$$

Given that harmonic load current is equivalent to the rated current of the studied transformer (i.e. 1 p.u), the harmonic factor for the proposed FEM procedure is equal to the K-Factor as 6.337.

4.3. Method Comparison

The most significant outcomes of de-rating the transformer under study are presented in this sub-section. To critically evaluate the performance of the analytical method and the proposed FEM procedure, the harmonic loss factors under the supplied harmonic spectrum are compared.

Table 5: Harmonic loss factor comparison

Method	F_{HL}
AM	6,693
PFEM	6,337

Comparatively, the proposed 3D FEM procedure envisage a lower harmonic loss factor than the analytical method as evidenced in Table 5. The assumption by the analytical method on the direct proportion between the winding Eddy losses and the harmonic load current is glitch, which then yield hidebound results. The maximum allowable current of the studied transformer based in eq. (8) is evaluated as follows:

$$\mathrm{I}_{max} = \sqrt{\frac{1 + 0.232}{1 + 6.337 \times 0.232}} = 0{,}71p.u$$

The maximum allowable load current in Ampere (A) is evaluated as:

$$\mathrm{I}_{max} = 0.71 \times 79.2 = 55{,}93A$$

The equivalent power rating (in kVA) of the transformer using the proposed FEM procedure is:

$$Equivalent\ kVA = 500 \times 0.71 = 353{,}10kVA$$

The performance of the analytical procedure recommended by the C57.110-2018 standard and the proposed 3D FEM procedure are compare and tabulated as shown in Table 6. Comparatively, the results indicate that the two methods are close, regardless of the fact that the classical analytical procedure is hidebound.

Table 6: Transformer rating comparison

Method	kVA	I_{max}(A)
AM	347,26	55,01
PFEM	353,10	55,93

A further look into the results indicates that since the analytical method cannot consider parameters such as skin effect, proximity effect, and the magnetic flux leakage on the windings, it can underestimate the equivalent kVA and maximum allowable current. Based on practical perspective during service, this unit may experience issues such as stray gassing. The authors presented the work related to this phenomenon in the publications [14] and [15]. The embedding of FEM into the design philosophy proves to have enhanced results. The authors herein have also presented additional work on this application in [16] and [17].

5. Conclusion

In this work, an extension of previous work [10], the impact of harmonics and distortion upon a DPVP transformer predicated on the classical method has been investigated with the ambition to examine its de-rating capability. A 3D FEM procedure has been proposed to evaluate the equivalent power rating and maximum allowable load current of a 500kVA, three-phase, oil-immersed, DPVP transformer. The following conclusions can be drawn from the case scenario and the practical experience of the transformer under study:

- The most crucial impact of the harmonics load current upon transformers planned for the DPVP application is the service winding Eddy losses and the load losses.
- A surge in the transformer service losses under harmonic conditions carries off premature insulation materials and designed transformer service lifetime. The power rating (kVA) of a DPVP transformer must consequently be de-rated under harmonic conditions.
- The conservative assumption of the C57.110-2018 standard of the direct proportion between the winding Eddy losses and the harmonic load current is a glitch as it does not take into account parameters suchlike skin effect, proximity effect, and the magnetic flux leakage on the windings

The proposed 3D FEM procedure as a highly accurate procedure for evaluating the transformer service losses under harmonic conditions as it takes into account parameters that cannot be captured by the procedure proposed by the C57.110-2018 standard may be employed in the final design stage for de-rating.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] E. Cazacu, L. Petrescu and V. Ionita, "Derating of power distribution transformers serving nonlinear industrial loads," in 2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) & 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP), Brasov, 90-95, 2017, doi: 10.1109/OPTIM.2017.7974953.

[2] D. Yildirim and E. F. Fuchs, "Measured transformer derating and comparison with harmonic loss factor (F/sub HL/) approach," in IEEE Transactions on Power Delivery, **15** (1), 186-191, 2000, doi: 10.1109/61.847249.

[3] E. Cherian and G. R. Bindu, "Minimizing Harmonics and Transformer Derating in Low Voltage Distribution Networks by DC Distribution," in 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR), Ernakulam, 2018, 1-6, doi: 10.1109/ICETIETR.2018.8529063.

[4] B. P. Das and Z. Radakovic, "Is Transformer kVA Derating Always Required Under Harmonics? A Manufacturer's Perspective," in IEEE Transactions on Power Delivery, **33** (6), 2693-2699, 2018, doi: 10.1109/TPWRD.2018.2815901.

[5] D. W. Egolf and A. J. Flechsig, "Harmonics-transformer derating," in Proceedings of Industrial and Commercial Power Systems Conference, Irvine, CA, USA, 79-84, 1994, doi: 10.1109/ICPS.1994.303557.

[6] E. Cazacu and L. Petrescu, "On-site derating of in-service power distribution transformers supplying nolinear loads", Revue Roumaine des Sciences Techniques - Serie Électrotechnique et Énergétique **59**(3), 259-268, 2014, doi: 18845326 .

[7] A.Y Shklyarskiy and A I Bardanov , "Test of the method for calculation of derating of workshop transformers on engineering plants", IOP Conf. Series: Materials Science and Engineering, **177**, 012054, 2017, doi:10.1088/1757-899X/177/1/012054.

[8] N. Xie, A. Luo, Fujun Ma, et al., "Harmonic interaction between large-scale photovoltaic power stations and grid", Proceedings of the Chinese Society of Electrical Engineering, **33**(34), 9-16, 2013, doi: 288752772.

[9] IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems," in IEEE Std 519-1992 ,1-112, 1993, doi: 10.1109/IEEESTD.1993.114370.

[10] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Contemplation of Harmonic Currents Loading on Large-Scale Photovoltaic Transformers," in 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 2020, 479-483, doi: 10.1109/ENERGYCon48941.2020.9236514.

[11] IEEE Recommended Practice for Establishing Liquid-Immersed and Dry-Type Power and Distribution Transformer Capability When Supplying Nonsinusoidal Load Currents," in IEEE Std C57.110™-2018 (Revision of

IEEE Std C57.110-2008), 1-68, 2018, doi: 10.1109/IEEESTD.2018.8511103.
[12] UL 1561, UL Standard for Safety Dry-Type General Purpose and Power Transformer, 4th Edition, 2011, doi: 00096942.
[13] J. Faiz, M.B.B Sharifian, S.A. Fakheri , "Research report on effect of non-linear loads upon distribution transformers and correction factor estimation for optimal operation of transformer-Part I", (in Persian), Azarbaijan Regional Electricity Company, Tabriz, Iran, 2001.
[14] D.B. Nyandeni, M. Phoshoko (Pr. Eng.), R. Murray, B.A. Thango, "Transformer Oil Degradation on PV Plants – A Case Study", 8th South Africa regional conference, 14-17, 2017, doi: 10.1109/MEI.2014.6749569.
[15] B. A. Thango, Jacobus A. Jordaan and Agha F. Nnachi, "Stray Gassing of Transformer Oil in Distributed Solar Photovoltaic (DSPV) Systems" 6th IEEE R8 International ENERGY Conference, 13-16, 2020, doi: 10.1109/ENERGYCon48941.2020.9236522.
[16] B.A Thango, J.A Jordaan, A.F Nnachi, D.B Nyandeni "Solar Power Plant Transformer Loss Calculation under Harmonic Currents using Field Element Method", 9th CIGRE Southern Africa Regional Conference, 1st – 4th October 2019, doi: 10.1109/MEI.2014.6749569.
[17] B. A. Thango, J.A. Jordaan, A. F. Nnachi, "Step-Up Transformers for PV Plants: Load Loss Estimation under Harmonic Conditions", 19th International Conference on Harmonics and Quality of Power (ICHQP), 2020, doi: 10.1109/ICHQP46026.2020.9177938.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 158-166 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Numeric Simulation on the Waves from Artificial Anti-gravity upon General Theory of Relativity

Yoshio Matsuki[1,*], Petro Ivanovich Bidyuk[2]

[1]*National University of Ukraine "Kyiv-Mohyla Academy", Kyiv 04070, Ukraine*

[2]*Educational and Scientific Complex "Institute for Applied System Analysis" of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv 03056, Ukraine*

ARTICLE INFO

Article history:
Received: 20 June, 2021
Accepted: 25 September, 2021
Online: 27 September, 2021

Keywords:
Distorted Time and Space
General Theory of Relativity
Curvature Tensor

ABSTRACT

This paper reports the algorithm, the input data and the result of the numeric simulation on the flows of the waves emitted from a rotating object that forms the artificial anti-gravity. First, an object with a heavy mass is placed in the 4-dimensional time and space, which is described by a fundamental tensor. Then the first-order derivative of the tensor describes the gravity, and the second-order derivative describes the waves. If the gravity created by the heavy mass is strong enough, time and space become dependent on each other. The input data for the simulation are discrete numbers that surrogate the infinity of the 4-dimensional time-space. The object is assumed to rotate and the tensor equations are solved. Then the coefficients are calculated, which present physical properties of the waves. The result of the simulation shows that the rotating object emits the waves with positive and negative energy intensities, and they have the spin angular momentum that changes its spinning direction upon the selection of the rotation-speed of the object.

1. Introduction

This research is an extended version of the work presented at the 2020 IEEE 2nd International Conference on System Analysis_Intelligent Computing (SAIC) held in Kyiv, Ukraine, in October 2020 [1] as well as the consequently published paper [2] from the Advances in Science, Technology and Engineering Systems Journal in May 2021. These papers reported that a rotating object would produce anti-gravity, if it held a heavy mass that would distort time and space. The last paper [2] informed of the waves emitted from the rotating object; however, the report of the analysis was forwarded to this new paper because the numeric simulation required a separate research.

An explicit image of the anti-gravity of the rotating object is a flying craft having a disc-shaped body for interstellar travel. Such an object should hold a compact nuclear fusion reactor inside, because the fusion reaction forms the heavier nuclei by fusing the lighter ones such as hydrogen, deuterium and tritium, and this process should distort time and space. However, this is the theory of a black hole [3], and the simulation of the time-space distortion in the nuclear fusion reactor is still the topic of the next research after this paper.

*Corresponding Author: Yoshio Matsuki, matsuki@wdc.org.ua

The waves emitted from the source of the anti-gravity are byproducts, but they can indicate the presence of the source if they have any physical properties such as the energy intensity and the spin angular momentum. The goal of this research is to calculate the energy intensity and the angular momentum of the waves.

In order to describe this phenomenon, the 4-dimensional curvilinear coordinates of the curved space are selected because the heavy object should warp the rectilinear coordinates of the flat space. To start the simulation, the curvature tensor, $R_{\mu\nu}$, is given to describe time and space where the object with a heavy mass is placed [3].

$$R_{\mu\nu} = 0 \tag{1}$$

$$R_{\mu\nu}=\Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}+\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} = 0 \tag{2}$$

$$\Gamma_{\mu\nu\sigma} = \frac{1}{2}\left(g_{\mu\nu,\sigma}+g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu}\right) \tag{3}$$

This curvature tensor includes the first-order differentials of a fundamental tensor ($g_{\mu\nu}$), $g_{\mu\nu,\sigma} = \partial g_{\mu\nu}/\partial x^{\sigma}$, where x^{σ} is the σ-th vector (variable) in the given coordinates. Before describing the phenomenon in the curvilinear coordinates, the flat space provides

a significant introduction to the theory of the waves. Therefore, at first a series of the notations derives the tensor of the waves in an approximate flat space in this introductory section, and then the tensor of the waves is derived in the curved space later in the next section 2.1.

If the gravitational field is weak, the curvature is also weak and $g_{\mu\nu}$ is approximately constant; therefore,

$$-\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}+\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}=0, \tag{4}$$

and

$$R_{\mu\nu}=\Gamma^{\alpha}_{\mu\alpha,\nu}-\Gamma^{\alpha}_{\mu\nu,\alpha}=0. \tag{5}$$

On the other hand,

$$R_{\mu\nu\rho\sigma}=\frac{1}{2}\left(g_{\mu\sigma,\nu\rho}-g_{\nu\sigma,\mu\rho}-g_{\mu\rho,\nu\sigma}+g_{\nu\rho,\mu\sigma}\right)$$
$$+\Gamma_{\beta\mu\sigma}\Gamma^{\beta}_{\nu\rho}-\Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma}. \tag{6}$$

By interchanging ρ and μ, and neglecting $\Gamma_{\beta\mu\sigma}\Gamma^{\beta}_{\nu\rho}$ - $\Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma}$, we get

$$R_{\mu\nu}=g^{\rho\sigma}R_{\mu\nu\rho\sigma}$$
$$=g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu}-g_{\nu\sigma,\mu\rho}-g_{\mu\rho,\nu\sigma}+g_{\mu\nu,\rho\sigma}\right)=0. \tag{7}$$

Then

$$g^{\mu\nu}\left(g_{\mu\nu,\rho\sigma}-g_{\mu\rho,\nu\sigma}-g_{\mu\sigma,\nu\rho}+g_{\rho\sigma,\mu\nu}\right)=0. \tag{8}$$

In general, the moving particle in a scalar field of the potential energy V follows d'Alembert equation,

$$\nabla V=g^{\mu\nu}\left(V_{,\mu\nu}-\Gamma^{\alpha}_{\mu\nu}V_{,\alpha}\right)=0. \tag{9}$$

In order to describe a particle moving in the 4-dimensional time and space, V is replaced by vectors x^{λ} of the 4-dimensional coordinates in which $x^{\lambda}{}_{,\alpha}=g^{\lambda}{}_{\alpha,}$, then d'Alembert equation becomes

$$g^{\mu\nu}g^{\lambda}_{\alpha,\nu}-g^{\mu\nu}g^{\lambda}_{\alpha}\Gamma^{\alpha}_{\mu\nu}=0. \tag{10}$$

Then

$$g^{\mu\nu}\Gamma^{\lambda}_{\mu\nu}=0. \tag{11}$$

Meanwhile,

$$\Gamma^{\lambda}_{\mu\nu}=g^{\lambda\rho}\Gamma_{\rho\mu\nu}=\frac{1}{2}g^{\lambda\rho}\left(g_{\rho\mu,\nu}+g_{\rho\nu,\mu}-g_{\mu\nu,\rho}\right). \tag{12}$$

Therefore,

$$g^{\mu\nu}\Gamma^{\lambda}_{\mu\nu}=\frac{1}{2}g^{\lambda\rho}g^{\mu\nu}\left(g_{\rho\mu,\nu}+g_{\rho\nu,\mu}-g_{\mu\nu,\rho}\right)$$
$$=g^{\lambda\rho}g^{\mu\nu}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)=0, \tag{13}$$

where

$$g_{\rho\nu,\mu}=y_{n,\rho\mu}y^{n}_{,\nu}+y_{n,\nu\mu}y^{n}_{,\rho}=y_{n,\rho\nu}y^{n}_{,\mu}+y_{n,\mu\nu}y^{n}_{,\rho}=g_{\rho\mu,\nu}, \tag{14}$$

where μ and ν are in symmetrical relation in the equation, therefore they are exchangeable; and,

$$y^{n}_{,\mu}=\frac{\partial y^{n}(x^{\mu})}{\partial x^{\mu}}, \tag{15}$$

where $\mu = 0, 1, 2, 3$, while x^{μ} are located in N-dimensional physical space of y^{n}, where $n = 1, 2, \ldots..,$ N.

Therefore,

$$g^{\mu\nu}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)=0. \tag{16}$$

Then, in order to describe the waves moving in the gravitational field, it is differentiated by x^{σ}.

$$\frac{d}{dx^{\sigma}}g^{\mu\nu}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)$$
$$=g^{\mu\nu}_{,\sigma}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)+g^{\mu\nu}\left(g_{\mu\rho,\nu\sigma}-\frac{1}{2}g_{\mu\nu,\rho\sigma}\right)$$
$$=g^{\mu\nu}\left(g_{\mu\rho,\nu\sigma}-\frac{1}{2}g_{\mu\nu,\rho\sigma}\right)=0 \tag{17}$$

In the equation (17),

$$g^{\mu\nu}_{,\sigma}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)=0 \tag{18}$$

because $g_{\rho\mu}$ is approximately constant in the weak gravitational field. Therefore,

$$g_{\rho\mu,\nu}=0 \tag{19}$$

By interchanging ρ and σ of (17),

$$g^{\mu\nu}\left(g_{\mu\sigma,\nu\rho}-\frac{1}{2}g_{\mu\nu,\sigma\rho}\right)=g^{\mu\nu}\left(g_{\mu\sigma,\nu\rho}-\frac{1}{2}g_{\mu\nu,\rho\sigma}\right)=0. \tag{20}$$

By adding (8), (17) and (20), we get

$$g^{\mu\nu}g_{\rho\sigma,\mu\nu}=0. \tag{21}$$

It satisfies d'Alembert equation, therefore it describes the waves that move in an empty space where there is only the weak gravitational field. (Note: This equation includes the second-order differentials, $g_{\mu\nu,\rho\sigma}=\partial^{2}g_{\mu\nu}/\partial x^{\rho}\partial x^{\sigma}$.)

The equation (21) describes the waves moving in the weak gravitational field that is approximately a flat 4-dimensional time and space; however, the general theory of relativity predicts the flow of the waves in curvilinear coordinates of the stronger gravitational field. In order to simulate the flow of the waves in the stronger gravitational field, this research includes the

following terms that have been neglected in the classic reference [3]:

$$-\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}+\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \neq 0 \tag{22}$$

instead of the equation (4),

$$\Gamma_{\beta\mu\sigma}\Gamma^{\beta}_{\nu\rho} - \Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma} \neq 0 \tag{23}$$

for the equation (6), and

$$g^{\mu\nu}_{,\sigma}\left(g_{\rho\mu,\nu} - \frac{1}{2}g_{\mu\nu,\rho}\right) \neq 0 \tag{24}$$

instead of the equation (18).

In this research, the energy intensity and the spin angular momentum of the flow of the waves are simulated as the calculated coefficients of the tensor equation in the spherical polar coordinate system, which is one of the curvilinear coordinate systems. The methodology and the result of the simulation are reported in the following sections.

2. Methodology

2.1. Algorithm

For the research of the strong gravitational field, $g_{\mu\nu}$, is not constant, but given by the following matrix, [3]:

$$g_{\mu\nu} = \begin{bmatrix} g_{00} & 0 & 0 & 0 \\ 0 & g_{11} & 0 & 0 \\ 0 & 0 & g_{22} & 0 \\ 0 & 0 & 0 & g_{33} \end{bmatrix} \tag{25}$$

$$g_{00} = 1 \tag{26}$$

$$g_{11} = -\frac{2m}{\mu(\rho-\tau)^{\frac{2}{3}}} \tag{27}$$

$$g_{22} = -\mu^2(\rho-\tau)^{\frac{4}{3}} \tag{28}$$

$$g_{33} = -\mu^2(\rho-\tau)^{\frac{4}{3}}sin^2\theta \tag{29}$$

And the inverse of $g_{\mu\nu}$ is $g^{\mu\nu}$. For example,

$$g^{11} = -\frac{\mu(\rho-\tau)^{\frac{2}{3}}}{2m} \tag{30}$$

Then t is time, m is the mass of the object, r is the distance from the center of the object, θ is the angle between r and the rotation axis, z, and ϕ is the angle of the rotation around the rotation axis in the spherical polar coordinates shown in Figure 1.

τ is the distorted time and ρ *is* the distorted distance as defined by the following equations [3].

$$\tau=t+f(r) \tag{31}$$

$$\rho=t+g(r) \tag{32}$$

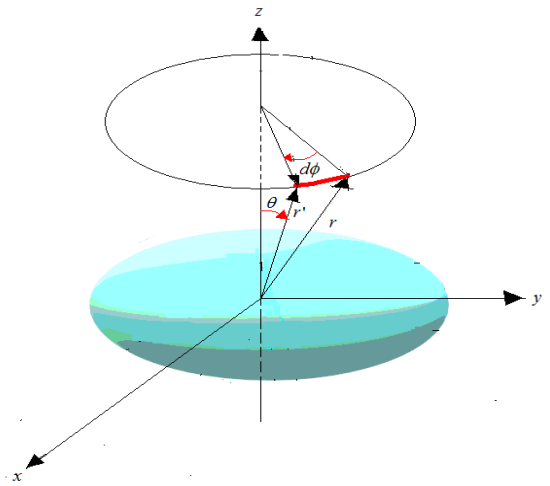


Figure 1: Spherical polar coordinates for the simulation

f(r) and g(r) are the functions later defined in the section 2.2. μ is given by the following equation with the mass, *m*, of the object ([3] p. 35):

$$\mu=\left(\frac{3}{2}\sqrt{2m}\right)^{\frac{2}{3}} \tag{33}$$

The equations below from (34) to (38) show how to derive the equation of the waves in the curved space. It follows the same logic from (7) to (21) in the introductory section above, which derives the equation of the waves in the approximate flat space. But, the equations below include the neglected terms in order to derive the equation in the curved space.

By adding the neglected terms of the equation (23) to the equation (7), we get:

$$\begin{aligned} R_{\mu\nu}=&g^{\rho\sigma}R_{\mu\nu\rho\sigma}=g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma}+g_{\mu\nu,\rho\sigma}\right) \\ &+g^{\rho\sigma}\Gamma_{\beta\rho\sigma}\Gamma^{\beta}_{\mu\nu} - g^{\rho\sigma}\Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma} \\ =&g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma}+g_{\mu\nu,\rho\sigma}\right) \\ &+g^{\rho\sigma}g_{\beta\beta}\Gamma_{\beta\rho\sigma}\Gamma_{\beta\mu\nu} - g^{\rho\sigma}g_{\beta\beta}\Gamma_{\beta\mu\rho}\Gamma_{\beta\nu\sigma} \\ =&g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma}+g_{\mu\nu,\rho\sigma}\right) \\ &+g^{\rho\sigma}g_{\beta\beta}\frac{1}{2}\left(g_{\beta\rho,\sigma}+g_{\beta\sigma,\rho} - g_{\rho\sigma,\beta}\right)\left(g_{\beta\mu,\nu}+g_{\beta\nu,\mu} - g_{\mu\nu,\beta}\right) \\ &-g^{\rho\sigma}g_{\beta\beta}\frac{1}{2}\left(g_{\beta\mu,\rho}+g_{\beta\,\rho\,,\,\mu} - g_{\mu\rho,\beta}\right)\left(g_{\beta\nu,\sigma}+g_{\beta\sigma,\nu} - g_{\nu\sigma,\beta}\right) \\ =&g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma}+g_{\mu\nu,\rho\sigma}\right) \\ &+\left(\frac{1}{2}g^{\rho\sigma}g_{\beta\beta}g_{\rho\sigma,\beta}g_{\nu\sigma,\beta}\right) = 0. \end{aligned} \tag{34}$$

Interchanging ρ, μ, σ and ν, we get

$$\begin{aligned} &g^{\mu\nu}\left(g_{\mu\nu,\rho\sigma} - g_{\mu\rho,\nu\sigma} - g_{\mu\sigma,\nu\rho}+g_{\rho\sigma,\mu\nu}\right) \\ &+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta}g_{\nu\mu,\beta}\right) = 0. \end{aligned} \tag{35}$$

By adding the neglected terms to the equation (17), we get

$$\frac{d}{dx^{\sigma}}\left\{g^{\mu\nu}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)+\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta}g_{\mu\nu,\beta}\right\}$$
$$=g^{\mu\nu}_{,\sigma}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)+g^{\mu\nu}\left(g_{\mu\rho,\nu\sigma}-\frac{1}{2}g_{\mu\nu,\rho\sigma}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta}g_{\nu\mu,\beta\sigma}\right)+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta\sigma}g_{\nu\mu,\beta}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}_{,\sigma}g_{\beta\beta}g_{\rho\mu,\beta}g_{\nu\mu,\beta}\right)$$
$$+\left(\tfrac{1}{2}g^{\rho\mu}g_{\beta\beta,\sigma}g_{\rho\mu,\beta}g_{\nu\mu,\beta}\right)=0. \quad (36)$$

Interchanging ρ and σ in (36), we get

$$g^{\mu\nu}_{,\rho}\left(g_{\sigma\mu,\nu}-\frac{1}{2}g_{\mu\nu,\sigma}\right)+g^{\mu\nu}\left(g_{\mu\sigma,\nu\rho}-\frac{1}{2}g_{\mu\nu,\sigma\rho}\right)$$
$$+\left(\frac{1}{2}g^{\sigma\mu}g_{\beta\beta}g_{\sigma\mu,\beta}g_{\nu\mu,\beta\rho}\right)+\left(\frac{1}{2}g^{\sigma\mu}g_{\beta\beta}g_{\sigma\mu,\beta\rho}g_{\nu\mu,\beta}\right)$$
$$+\left(\tfrac{1}{2}g^{\sigma\mu}_{,\rho}g_{\beta\beta}g_{\sigma\mu,\beta}g_{\nu\mu,\beta}\right)+\left(\tfrac{1}{2}g^{\sigma\mu}g_{\beta\beta,\rho}g_{\sigma\mu,\beta}g_{\nu\mu,\beta}\right)=0. \quad (37)$$

Adding (35), (36) and (37), we get

$$g^{\mu\nu\rho\sigma,\mu\nu}+g^{\mu\nu}_{,\sigma}\left(g_{\rho\mu,\nu}-\frac{1}{2}g_{\mu\nu,\rho}\right)+g^{\mu\nu}_{,\rho}\left(g_{\sigma\mu,\nu}-\frac{1}{2}g_{\mu\nu,\sigma}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta}g_{\nu\mu,\beta\sigma}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta}g_{\rho\mu,\beta\sigma}g_{\nu\mu,\beta}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}_{,\sigma}g_{\beta\beta}g_{\rho\mu,\beta}g_{\nu\mu,\beta}\right)$$
$$+\left(\frac{1}{2}g^{\rho\mu}g_{\beta\beta,\sigma}g_{\rho\mu,\beta}g_{\nu\mu,\beta}\right)$$
$$+\left(\frac{1}{2}g^{\sigma\mu}g_{\beta\beta}g_{\sigma\mu,\beta}g_{\nu\mu,\beta\rho}\right)$$
$$+\left(\frac{1}{2}g^{\sigma\mu}g_{\beta\beta}g_{\sigma\mu,\beta\rho}g_{\nu\mu,\beta}\right)$$
$$+\left(\tfrac{1}{2}g^{\sigma\mu}_{,\rho}g_{\beta\beta}g_{\sigma\mu,\beta}g_{\nu\mu,\beta}\right)+\left(\tfrac{1}{2}g^{\sigma\mu}g_{\beta\beta,\rho}g_{\sigma\mu,\beta}g_{\nu\mu,\beta}\right)=0. \quad (38)$$

The equation (38) describes the waves that move in the strong gravitational field. The numbers, 0, 1, 2, and 3, are assigned to each of μ, ν, ρ, σ and β of the equation (38); then by the rule of tensor, all terms are summed up. Then the mathematical forms of the equation (38) are derived by the functions shown in the equations from (26) to (29). Then according to the equation (25), non-diagonal components such as g_{12} and g_{23} vanish; then only diagonal components remain as shown in the equations from (39) to (41). Hereinafter this research focuses on the spatial distribution of the waves but doesn't consider the time. Time is taken into account only by the distorted distance, ρ, of the equation (32). Therefore, only three diagonal components of a 3×3 matrix that represent the spatial distribution of the waves are considered.

$$W_{11}=\frac{16}{9(\rho-\tau)^2}+\frac{8}{3(\rho-\tau)^2}+\frac{64m}{27\mu(\rho-\tau)^{\frac{11}{3}}}$$
$$-\frac{144m}{27\mu(\rho-\tau)^{\frac{13}{3}}}+\frac{8\mu^3}{9m}sin^2\theta+\frac{16\mu^3}{18m}. \quad (39)$$

$$W_{22}=8cos^2\theta-8sin^2\theta. \quad (40)$$

$$W_{33}=\frac{128}{9(\rho-\tau)^2}+16cot^2\theta. \quad (41)$$

And,

$$W_{\mu\nu}=W_{11}+W_{22}+W_{33}$$
$$=c_1X_1+c_2X_2+\ldots+c_lX_l. \quad (42)$$

$16\mu^3/18m$ of the equation (39) is a constant term, and it is not included in the equation (42) for the simulation.

The algorithm to simulate the relative energy intensity of the waves is as follows:

It is assumed that the stress energy tensor, T, reflects the energy intensities of the waves as shown in the equation (43), where k is a constant.

$$W_{\mu\nu}=kT \quad (43)$$

The idea of the stress energy tensor is taken from the classical mechanics [4] (p. 327).

Each of the coefficients, c_1, $c_2 \cdots c_l$, relates to a component of the waves such as $1/(\rho-\tau)^2$, $1/(\rho-\tau)^{3/11}$, and so on. With the equation (44) we calculate the coefficients, c_1, $c_2 \cdots c_l$.

$$H=kT-W_{\mu\nu}=kT-(c_1X_1+c_2X_2+\ldots+c_lX_l)$$
$$=kT-\left(c_1\frac{1}{(\rho-\tau)^2}+c_2\frac{1}{(\rho-\tau)^{\frac{11}{3}}}-c_3\frac{1}{(\rho-\tau)^{\frac{13}{3}}}\right.$$
$$\left.+c_4sin^2\theta+c_5(cos^2\theta-sin^2\theta)+c_6cot^2\theta\right) \quad (44)$$

These coefficients make a column vector, c, which holds l rows. For calculating c, X (a n×l matrix) is set as shown in (45), where each of X_1, X_2, $\cdots$ X_l is a vector that holds n rows, and n is also the number of rows of each column of X. In this simulation the value of n is 23 as later explained in the section of the input data, and l is 6 for the equation (44).

The vectors of the matrix X should hold the projected image of the energy intensity on the surface of the sphere in the spherical polar coordinates. (Note: The surface of the sphere refers to Figure 14 of [2].

$$X=[X_1 \quad X_2 \quad \ldots \quad X_l] \quad (45)$$

Then H is set for matrix algebra as shown below.

$$H=kT-Xc \quad (46)$$

To solve this equation, a constraint is given as shown below, where X' is the transposed matrix of X.

$$X'H=X'(kT-Xc)=0 \quad (47)$$

The matrix algebra continues as shown below to calculate the values of c.

$$X'Xc=X'kT \quad (48)$$

$$c=(X'X)^{-1}X'kT \quad (49)$$

here $(X'X)^{-1}$ is an inverse matrix of $X'X$.

If the object rotates as shown in Figure 1, its coordinate system is transformed by the transformation matrix D of the Euler's angles [4] as shown below.

$$D=\begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (50)$$

To simulate the rotation of the object around the axis of z, the matrix, $W_{\mu\nu}$, is multiplied by D; then it is transformed as shown below.

$$D \cdot W_{\mu\nu}$$

$$=\begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} W_{11} & 0 & 0 \\ 0 & W_{22} & 0 \\ 0 & 0 & W_{33} \end{pmatrix}$$

$$=\begin{pmatrix} cos\phi \cdot W_{11} & sin\phi \cdot W_{22} & 0 \\ -sin\phi \cdot W_{11} & cos\phi \cdot W_{22} & 0 \\ 0 & 0 & W_{33} \end{pmatrix} \quad (51)$$

To calculate the energy intensity of the waves, the diagonal components are taken from the above transformed matrix, $D \cdot W_{\mu\nu}$, then a matrix (52) is set.

$$\begin{bmatrix} dW_{11} & 0 & 0 \\ 0 & dW_{22} & 0 \\ 0 & 0 & W_{33} \end{bmatrix}$$

$$=\begin{pmatrix} cos\phi \cdot W_{11} & 0 & 0 \\ 0 & cos\phi \cdot W_{22} & 0 \\ 0 & 0 & W_{33} \end{pmatrix} \quad (52)$$

Then an equation of H is set as shown below to calculate the coefficients of the diagonal components.

$$H=kT-\{c_1\frac{cos\,\phi}{(\rho-\tau)^2}+c_2\frac{cos\phi}{(\rho-\tau)^{\frac{11}{3}}}-c_3\frac{cos\phi}{(\rho-\tau)^{\frac{13}{3}}}$$
$$+c_4 cos\phi sin^2\theta + c_5 cos\phi(cos^2\theta - sin^2\theta)$$
$$+c_6\frac{1}{(\rho-\tau)^2}+c_7 cot^2\theta\} \quad (53)$$

Hereinafter the same procedure follows as explained from the equation (44) to (49) to calculate the coefficients.

To calculate the spin angular momentum, a matrix is set as shown in (54), by taking out the anti-symmetrical components of $D \cdot W_{\mu\nu}$ from (51).

$$\begin{bmatrix} 0 & sin\phi \cdot W_{22} & 0 \\ -sin\phi \cdot W_{11} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (54)$$

Then a column vector is set as shown below.

$$\begin{bmatrix} sin\,\phi\ W_{22} \\ -sin\,\phi\ W_{11} \\ 0 \end{bmatrix} = \begin{bmatrix} W_{22} \cdot d\Omega_3 \\ -W_{11} \cdot d\Omega_3 \\ 0 \end{bmatrix} \quad (55)$$

With the above column vector, an equation of H is set as shown below for the calculation of the coefficients.

$$H=kT-\{-c_1\frac{-sin\phi}{(\rho-\tau)^2}-c_2\frac{-sin\phi}{(\rho-\tau)^{\frac{11}{3}}}+c_3\frac{-sin\phi}{(\rho-\tau)^{\frac{13}{3}}}$$
$$-c_4(-sin\phi)sin^2\theta + c_5 sin\phi(cos^2\theta - sin^2\theta) \quad (56)$$

The spin angular momentum is to be projected on the flat surface, which is perpendicular to the curved spherical surface in the spherical polar coordinate system. (Note: The flat surface refers to Figure 15 of [2].

2.2. Input data

Figure 2 and Figure 3 show the input data that simulate the distortion of time and space, which we used also for the previous research [1, 2]. In these figures, r is the distance from the center of the strong gravity; t is the time to travel for the distance; τ is the distorted time by the equation (31), which expands and shrinks depending on r and t; and, ρ is the distorted distance by the equation (32), which expands and shrinks depending on t and r.

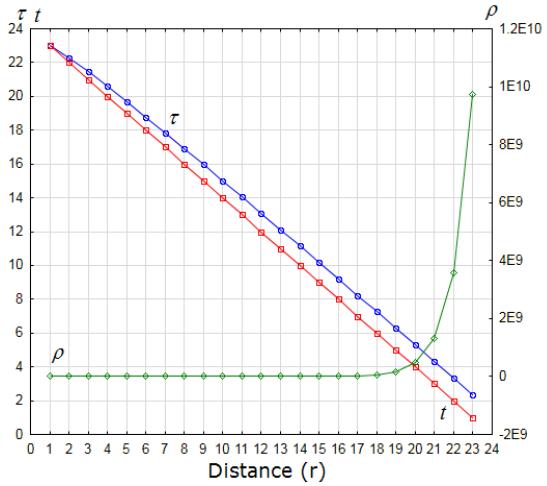


Figure 2: Time and distance from the center of the gravity, Case-1 (non-linear model): f(r)=log r and g(r)= e^r

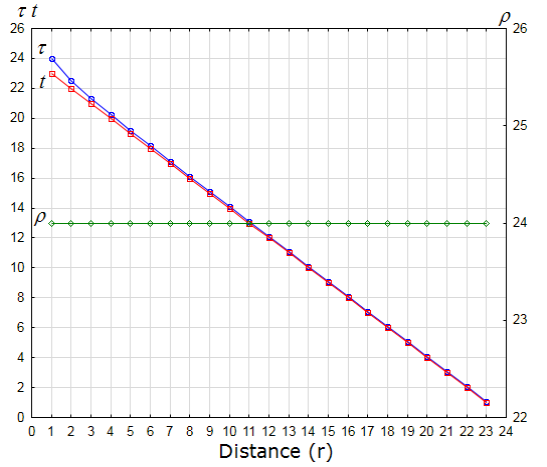


Figure 3: Time and distance from the center of the gravity, Case-2 (linear model): f(r) = 1/r and g(r) = r

For the distorted time and space, two models are set: Case-1 (non-linear model) and Case-2 (linear model) as shown in Table

1, which assign the functions of f(r) of the equation (31) and g(r) of the equation (32).

Table 1: Two models for simulating the distorted time-space

	Case-1 (Non-linear model)	Case-2 (Linear model)
f(r)	*log·r*	*1/r*
g(r)	e^r	*r*

Then the distributions of r and t are set to simulate the description of the reference [3], "Any signal, even a light signal, would take an infinite time to cross the boundary of a black hole". In our research the term, a black hole, is replaced by the strong gravity that distorts time and space. However it is not possible to set the infinite value for the numeric simulation of this research that uses a personal computer. Therefore we used 24 discrete finite values as the surrogate that simulates "It takes more time to travel closer to the center of the strong gravity".

The time, t, and the distance, r, are set as shown in Figure 2 for Case-1 and Figure 3 for Case-2. For simulating the waves, we assumed that θ would become smaller in far distance from the object as shown in Figure 4, because the spin of the waves should stabilize the flow as the waves move forward from the center of the gravity; then the angle θ should be focused and smaller in the far distance.

For simulating the rotation of the object, two modes of the frequencies are set: ϕ_1 (rotation 1) and ϕ_2 (rotation 2) also as shown in Figure 4. The frequency ϕ_2 is higher than ϕ_1; therefore, rotation 2 represents the faster rotation of the object than rotation 1. With these settings, sin θ, cos θ, cot θ, sin ϕ_1, sin ϕ_2, cos ϕ_1 and cos ϕ_2 change their values as shown in Figure 5.

The value of the stress energy tensor is set as 1; therefore, this simulation doesn't calculate absolute values of the intensities, but relative intensities. The calculated coefficients have dimensions. For example, the coefficient of the first term of the equation (39), $16/\{9(\rho-\tau)^2\}$, has the dimension of the inverse of $(\rho-\tau)^2$ that is the squared dimension of the difference between the distorted distance, ρ, and the distorted time, τ.

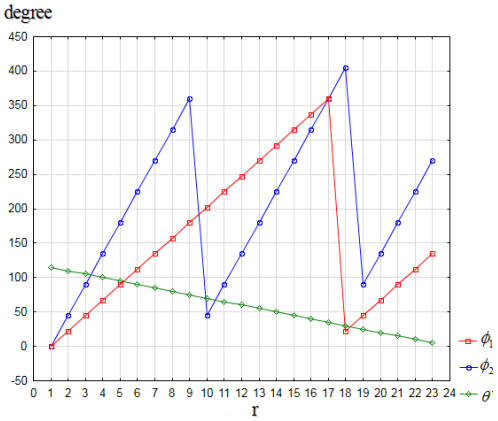


Figure 4: Angles, θ, ϕ_1 and ϕ_2 for the simulation

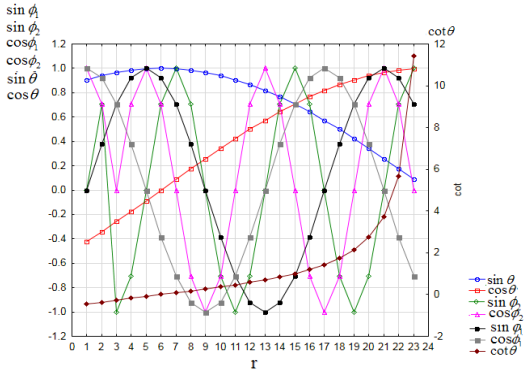


Figure 5: Trigonometric functions, sin *θ*, *cos θ*, *cot θ*, sin ϕ_1, sin ϕ_2, cos ϕ_1 and cos ϕ_2 for the simulation

3. Result

3.1. Energy intensity of the waves

The calculated energy intensities of the waves are shown in Table 2 for Case-1 (non-linear model) and Table 3 for Case-2 (linear model). The first and the third columns of these tables are the mathematical forms of the components of the waves, which are taken from the equations from (39) to (41), but each of the constant factors is simplified to unity. For example, $16/\{9(\rho-\tau)^2\}$ is simplified to $1/(\rho-\tau)^2$ because the factor, 16/9, is not necessary as it is to be included in the coefficient when it is calculated by the equation (49).

The calculated coefficients have either positive or negative values. For example, in Case-1 (non-linear model) the calculated coefficient of $1/(\rho-\tau)^2$ is $4.081\cdot10^{-6}$ with no rotation; and, it changes to $-1.465\cdot10^{4}$ with rotation 1, and then $4.178\cdot10^{3}$ with rotation 2. It suggests that the energy intensity of this component of the waves changes its sign from negative to positive when the object rotates faster from rotation 1 to rotation 2.

In order to interpret the calculated results, hereinafter we define that the positive flow holds positive (+) coefficient, and the negative flow holds negative (–) coefficient. For example in the above case of $1/(\rho-\tau)^2$, the wave is a negative flow with rotation 1; but, it changes to a positive flow with rotation 2.

Table 2: Results of the simulation of energy intensity of the waves, Case-1

Components of Waves (No rotation)	Coefficients (No rotation)	Components of Waves with the rotation	Coefficients (Rotation 1)	Coefficients (Rotation 2)
$1/(\rho-\tau)^2$ *	$4.081\cdot10^{-6}$	$\cos\phi/(\rho-\tau)^2$	$-1.465\cdot10^{4}$	$4.178\cdot10^{3}$
$1/(\rho-\tau)^{11/3}$	$-1.892\cdot10^{-4}$	$\cos\phi/(\rho-\tau)^{11/3}$	$1.682\cdot10^{5}$	$-2.311\cdot10^{5}$
$-1/(\rho-\tau)^{13/3}$	$-3.270\cdot10^{-4}$	$-\cos\phi/(\rho-\tau)^{13/3}$	$2.836\cdot10^{5}$	$-4.028\cdot10^{5}$
$\sin^2\theta$	2.000	$\cos\phi\sin^2\theta$	$8.397\cdot10^{-2}$	$3.875\cdot10^{-2}$
$\cos^2\theta-\sin^2\theta$	1.000	$\cos\phi(\cos^2\theta-\sin^2\theta)$	1.436	$-6.610\cdot10^{-2}$
–	–	$1/(\rho-\tau)^2$ ⁂	$1.041\cdot10^{4}$	$3.825\cdot10^{2}$
$\cot^2\theta$	$-4.358\cdot10^{-10}$	$\cot^2\theta$ ⁂	$7.534\cdot10^{-2}$	$8.819\cdot10^{-2}$

Notes for Table 2 and 3.

*: This component belongs to both W_{11}, $(16/9(\rho-\tau)^2 + 8/3(\rho-\tau)^2$, of the equation (39) and W_{33}, $(128/9(\rho-\tau)^2$, of the equation (41).

⁑: These components are on the axis of the rotation and belong to W_{33}; therefore, cosϕ is not multiplied.

Table 3: Results of the simulation of energy intensity of the waves, Case-2

Components of Waves (No rotation)	Coefficients (No rotation)	Components of Waves with the rotation	Coefficients (Rotation 1)	Coefficients (Rotation 2)
$1/(\rho-\tau)^2$ *	$-1.563\cdot10^{-7}$	$\cos\phi/(\rho-\tau)^2$	–17.34	–63.97
$1/(\rho-\tau)^{11/3}$	$2.206\cdot10^{-6}$	$\cos\phi/(\rho-\tau)^{11/3}$	$-1.313\cdot10^{2}$	$1.156\cdot10^{3}$
$-1/(\rho-\tau)^{13/3}$	$2.493\cdot10^{-6}$	$-\cos\phi/(\rho-\tau)^{13/3}$	$-1.460\cdot10^{2}$	$1.390\cdot10^{3}$
$\sin^2\theta$	2.000	$\cos\phi\sin^2\theta$	0.1066	$6.768\cdot10^{-2}$
$\cos^2\theta-\sin^2\theta$	1.000	$\cos\phi(\cos^2\theta-\sin^2\theta)$	0.4353	–0.5905
–	–	$1/(\rho-\tau)^2$ ⁑	28.03	12.58
$\cot^2\theta$	$-4.306\cdot10^{-10}$	$\cot^2\theta$ ⁑	$8.474\cdot10^{-2}$	$8.236\cdot10^{-3}$

The summations of the values of each of the positive coefficients and the negative coefficients are shown in Table 4 for both Case-1 and Case-2, and Figure 6 for Case-1 and Figure 7 for Case-2. In both cases, positive and negative flows of the waves are not visible in the figures when the object doesn't rotate; but, they become visible when the object rotates.

Table 4: Energy intensities of positive and negative flows

	Positive flow	Negative flow
No rotation (Case-1)	3.000	$-4.358\cdot10^{-10}$
Rotation 1 (Case-1)	$4.622\cdot10^{5}$	$-1.465\cdot10^{4}$
Rotation 2 (Case-1)	$4.560\cdot10^{3}$	$-6.339\cdot10^{5}$
No rotation (Case-2)	3.000	$-1.567\cdot10^{-7}$
Rotation 1 (Case-2)	$2.865\cdot10^{1}$	$-2.947\cdot10^{2}$
Rotation 2 (Case-2)	$2.558\cdot10^{3}$	$-6.456\cdot10^{1}$

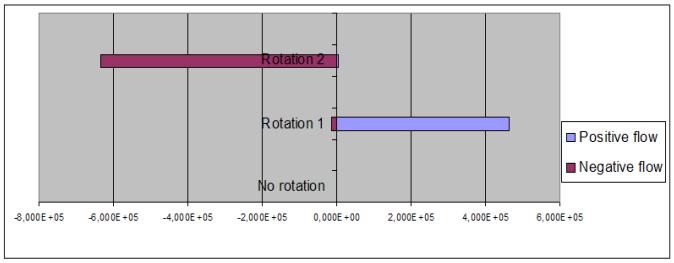


Figure 6: Energy intensity of positive and negative flows: (Case-1)

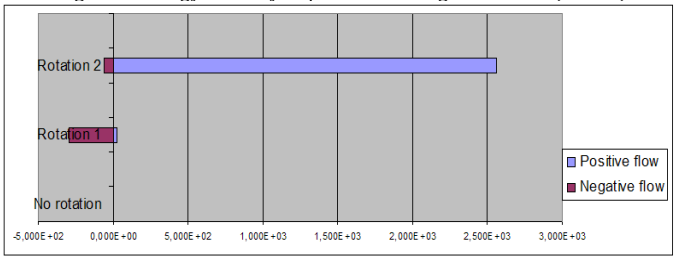


Figure 7: Energy intensity of positive and negative flows: (Case-2)

3.2. Spin angular momentum

Table 5 and Table 6 show the calculated relative intensity of the spin angular momentum of the waves, which are projected in the directions of $\sin\phi\cdot W_{22}$ and $-\sin\phi\cdot W_{11}$. (Note: $\sin\phi\cdot W_{22}$ and $-\sin\phi\cdot W_{11}$ are perpendicular to the rotation axis.)

The calculated coefficients of the spin angular momentum hold either positive or negative signs. For example, in Case-1 (non-linear model) the coefficient of $-\sin\phi/(\rho-\tau)^2$ is $-6.003\cdot10^{3}$ with the rotation 1, but it changes to $7.215\cdot10^{3}$ with the rotation 2. $-\sin\phi/(\rho-\tau)^2$ is one of the projected momenta of $1/(\rho-\tau)^2$ in the direction of $-\sin\phi\cdot W_{11}$, where $1/(\rho-\tau)^2$ represents the first and the second terms of the equation (39), $16/\{9(\rho-\tau)^2\}$ and $8/\{3(\rho-\tau)^2\}$.

Then we compare the changes of the spin angular momentum with the changes of the energy intensity. The energy intensity of $1/(\rho-\tau)^2$ in Table 2 for Case-1 (non-linear model) changes its sign from negative to positive when the rotation of the object becomes faster from rotation 1 to rotation 2. It suggests that this component is a negative flow of the waves when the object rotates with the frequency of ϕ_1, but when the object rotates faster with the frequency of ϕ_2, the waves become a positive flow; and, the spin angular momentum also changes its spinning direction from negative to positive as shown in Table 5.

Table 5: Relative intensity of spin angular momentum, Case-1

Components of the waves	Rotation 1	Rotation 2
$-\sin\phi/(\rho-\tau)^2$	$-6.003\cdot10^{3}$	$7.215\cdot10^{3}$
$-\sin\phi/(\rho-\tau)^{11/3}$	$1.400\cdot10^{6}$	$-1.734\cdot10^{6}$
$-\sin\phi/(\rho-\tau)^{13/3}$	$4.592\cdot10^{6}$	$-5.688\cdot10^{6}$
$-\sin\phi\sin^2\theta$	$-1.163\cdot10^{-2}$	$-1.234\cdot10^{-2}$
$\sin\phi(\cos^2\theta-\sin^2\theta)$	0.2898	–0.5123

Table 6: Relative intensity of spin angular momentum, Case-2

Components of the waves	Rotation 1	Rotation 2
$-\sin\phi/(\rho-\tau)^2$	–71.08	7.644
$-\sin\phi/(\rho-\tau)^{11/3}$	$5.946\cdot10^{2}$	$1.031\cdot10^{2}$
$-\sin\phi/(\rho-\tau)^{13/3}$	$6.159\cdot10^{2}$	$1.605\cdot10^{2}$
$-\sin\phi\sin^2\theta$	$-2.250\cdot10^{-4}$	$-1.270\cdot10^{-2}$
$\sin\phi(\cos^2\theta-\sin^2\theta)$	1.129	–0.5360

The summations of the values for positive and negative coefficients of the spin angular momenta are shown in Table 7 for Case-1 and Case-2, and Figure 8 for Case-1 and Figure 9 for Case-2. In both cases, the values change upon the selection of the rotation-mode of the object, rotation 1 or rotation 2. As shown in

the figures, the spin angular momentum changes its sign between rotation 1 and rotation 2 in Case-1 (non-linear model), while such a rule of changing sign is not so clear in Case-2 (linear model).

Table 7: Relative intensity of spin angular momentum

	Positive flow	Negative flow
Case-1 Rotation 1	$5.992 \cdot 10^{6}$	$-6.003 \cdot 10^{3}$
Case-1 Rotation 2	$7.215 \cdot 10^{3}$	$-7.421 \cdot 10^{6}$
Case-2 Rotation 1	$1.212 \cdot 10^{3}$	$-7.108 \cdot 10^{1}$
Case-2 Rotation 2	$2.713 \cdot 10^{2}$	$-5.486 \cdot 10^{2}$

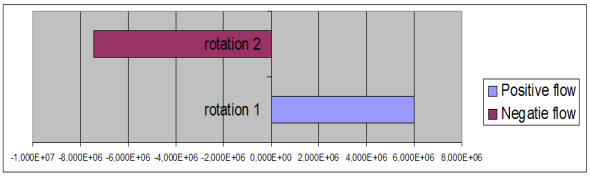


Figure 8: Relative intensity of spin angular momentum Case-1

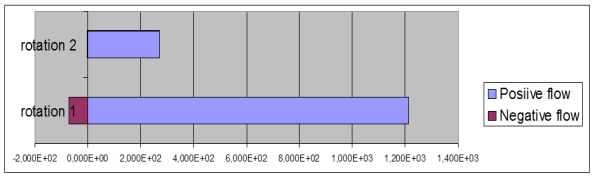


Figure 9: Relative intensity of spin angular momentum Case-2

3.3. Discussion

The results show that positive and negative flows of the waves appear if the object rotates, and some of the components that indicate positive and negative flows have opposite spinning directions. The results also show that the rotation-mode (rotation 1 or rotation 2) of the object relates to the energy intensity of the emitted waves, as well as their spinning directions.

Table 8 and Table 9 for Case-1 (non-linear model) and Table 10 and Table 11 for Case-2 (linear model) summarize the changes of the signs of the energy intensity and the spin angular momentum.

In Case-1 (non-linear model), except $sin^2\theta$, these components change their energy intensities either from positive to negative or from negative to positive, and simultaneously change their signs of the spin angular momenta either from positive to negative or from negative to positive. For example, the component of $1/(\rho-\tau)^2$ is a negative flow with rotation 1, but it is a positive flow with rotation 2; while its spinning direction changes from negative to positive when the rotation-mode changes from rotation 1 to rotation 2.

The coefficients of $1/(\rho-\tau)^{11/3}$, $-1/(\rho-\tau)^{13/3}$ and $cos^2\theta-sin^2\theta$ show the same rule of changing sign, but in the opposite directions of the flow and the spin of $1/(\rho-\tau)^2$.

This finding in Case-1 (non-linear model) is consistent with Figure 10 of our previous research [5] upon the analysis in rectilinear coordinates, in which the negative flow has the clockwise spin, while the positive flow has the anti-clockwise spin.

The exception, $sin^2\theta$, relates to $cos\phi sin^2\theta$ of Table 2 for the energy intensity and to $sin\phi sin^2\theta$ of Table 5 for the spin angular momentum. The absolute values of these calculated coefficients are smaller than those of the other components; and, the flow and the spin don't change their signs. Therefore this component is not considered significant.

In Case-2 (linear model) shown in Table 10 and Table 11, the rule of changing sign as observed in Case-1 applies only to $cos^2\theta-sin^2\theta$.

Table 8: The sign of the coefficient, energy intensity, Case-1

	No rotation	Rotation 1	Rotation 2
$1/(\rho-\tau)^2$	+	−	+
$1/(\rho-\tau)^{11/3}$	−	+	−
$-1/(\rho-\tau)^{13/3}$	−	+	−
$sin^2\theta$	+	+	+
$cos^2\theta-sin^2\theta$	+	+	−

Note for Tables 8, 9, 10 and 11: The component of $cot^2\theta$ doesn't appear in these tables because it is on the rotation axis, and not considered. $1/(\rho-\tau)^2$ is a simplified form of $16/9(\rho-\tau)^2$ and $8/3(\rho-\tau)^2$; $-1/(\rho-\tau)^{11/3}$ is a simplified form of $-64m/27\mu(\rho-\tau)^{11/3}$; $sin^2\theta$ is a simplified form of $8\mu^3 sin^2\theta/18m$ of the equation (39); and $cos^2\theta-sin^2\theta$ is a simplified form of $8cos^2\theta-8sin^2\theta$ of the equation (40). All these components are multiplied by $cos\phi$ for calculating the energy intensity and by $-sin\phi$ and $sin\phi$ for calculating the spin angular momentum.

Table 9: The sign of the coefficient, Spin angular momentum, Case-1

	Rotation 1	Rotation 2
$-sin\phi/(\rho-\tau)^2$	−	+
$-sin\phi/(\rho-\tau)^{11/3}$	+	−
$-sin\phi/(\rho-\tau)^{13/3}$	+	−
$-sin\phi sin^2\theta$	−	−
$sin\phi(cos^2\theta-sin^2\theta)$	+	−

Table 10: The sign of the coefficient, energy intensity, Case-2

	No rotation	Rotation 1	Rotation 2
$1/(\rho-\tau)^2$	−	−	−

$1/(\rho-\tau)^{11/3}$	+	−	+
$-1/(\rho-\tau)^{13/3}$	+	−	+
$sin^2\theta$	+	+	+
$cos^2\theta-sin^2\theta$	+	+	−

Table 11: The sign of the coefficient, Spin angular momentum, Case-2

	Rotation 1	Rotation 2
$-sin\phi/(\rho-\tau)^2$	−	+
$-sin\phi/(\rho-\tau)^{11/3}$	+	+
$-sin\phi/(\rho-\tau)^{13/3}$	+	+
$-sin\phi sin^2\theta$	−	−
$sin\phi(cos^2\theta-sin^2\theta)$	+	−

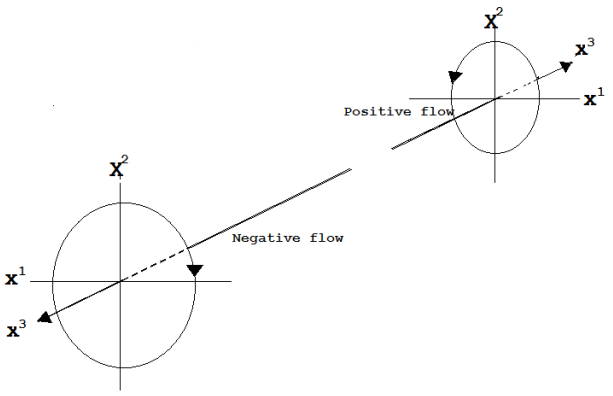


Figure 10: Geometrical relation between positive and negative flows in flat space (from [5])

4. Conclusion and recommendation

The result indicates that the rotation of the object can produce both positive and negative energy intensities of the waves, and they are interpreted as positive and negative flows. Both positive and negative flows appear when the object rotates, although the selection of the input data, the non-linear model or the linear model, leads to differences of the energy intensities.

With the non-linear model, the calculated spin angular momentum of the waves changes its sign, + or –, when most of the components change the signs of their energy intensities either from positive to negative or from negative to positive, depending on the speed of the object's rotation. This finding means that most of the components of the waves change their spinning directions when the energy intensity of the flow of the waves changes its sign + or –, also depending on the speed of the object's rotation. This particular finding on the spin angular momentum is consistent with the previous research [5]. However, with the linear model, most of the components of the waves don't follow the same rule of changing sign.

A parallel image of the source of the anti-gravity is a rotating black hole, and the same rule may apply. Then the positive flow of the waves may be gravitational waves and the negative flow may be the gravitational waves that flow in opposite directions. If those gravitational waves are observed from a rotating black hole and if they indicate the same rule of the energy intensity and the angular momentum as shown in this paper, it will certify the theory.

On the other hand, the theory of the time-space distortion should be verified with a compact nuclear fusion reactor for developing the flying craft. At the next step of the research, it is necessary to calculate the probability of proton-capture or proton-charge-exchange with the time-space distortion model that is used for the simulation reported in this paper.

References

[1] Y. Matsuki, P.I. Bidyuk, "Theory and Simulation of Artificial Antigravity," 2020 IEEE 2nd International Conference on System Analysis Intelligent Computing, 2020, doi: 10.1109/SAIC51296.2020.9239195.

[2] Y. Matsuki, P.I. Bidyuk, "Numeric Simulation of Artificial Antigravity upon General Theory of Relativity," Advances in Science, Technology and Engineering Systems Journal, **6**(3), 45-53, 2021, doi: 10.25046/aj060307.

[3] P.A.M. Dirac, General Theory of Relativity, New York: Florida University, A Wiley Inter-science Publication, John Wiley & Sons, 1975.

[4] H. Goldstein, C.P. Poole, J.L. Safko, Classical Mechanics, 3rd Edition published by Pearson Education. Inc., 2002.

[5] Y. Matsuki, P.I. Bidyuk, "Analysis of Negative Flow of Gravitational Waves," System Research & Information Technology, N4, 7-18, 2019, doi: 10.20535/SRIT.2308.8893.2019.4.01.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 164-171 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Hiding Information in DNA Sequence Data using Open Reading Frame Guided Splicing

Amal Khalifa*

Computer Science Department, Purdue University Fort Wayne, Fort Wayne, IN- 46814, USA

ARTICLE INFO

Article history:
Received: 02 February, 2021
Accepted: 04 May, 2021
Online: 20 May, 2021

Keywords:
Steganography
DNA sequences
Playfair cipher
Blind extraction
ORF
Sequence splicing

ABSTRACT

Encouraged by the huge publicly available genomic databases, research in the field of steganography was recently extended to utilize DNA sequence data to conceal secret information. As an extension of the work presented earlier by the author, this paper proposes an approach for a secure data communication channel between two parties. At one side of the communication, the sender starts the hiding process by encrypting the secret message using a bio-inspired 8x8 play-fair ciphering algorithm. Next, the secret sequence is randomly spliced and merged into the cover sequence replacing its non-coding regions. Using the secret key shared in advance, the receiver, on the other side of communication, can extract and concatenate the segments of the encrypted message and reveals the original message after deciphering. The method was proven to be robust to brute-force attacks while providing a hiding capacity up to two bit-per-nucleotide. A comparison with some existing techniques showed that the proposed method outperforms most of them not only in terms of the hiding capacity but also for the feature of blind extraction.

1. Introduction

Despite the proven efficiency of cryptography in protecting the security of information during communication, this protection is not guaranteed after the sent information is received and decrypted. On the other hand, Steganography introduced innovative ways to hide the existence of the secret information into innocent looking cover media in such a way that the resultant stego-media is hardly distinguishable from the original one. This makes it almost impossible to discover the concealed information or even suspect its existence. The first use of steganography dates back to the fifth century BC as documented in "Histories" of Herodotus [1]. The story happened during the conflict between Persia and Greece when Demaratus sent a secret message to warn the Spartans about the surprise attack planned by Xerxes, the tyrant king of Persia. The message was carved on a wooden tablet and covered with wax. Being apparently blank, the tablet was not intercepted by the Persian guards along the road. When the tablet reached its destination, the message was revealed after scrapping the wax off. With the element of surprise lost, the Persian fleet was lured in the harbor and damaged in less than a day. A modern model for a steganographic system was formulated by Simmons in terms of a prisoner's problem [2]. It assumes that Alice and Bob are trying to prepare an escape plan and need to exchange messages without drawing the attention of the warden. So, they embed their secret messages some cover-object using a secret key. Only using the same key, the message can be extracted from the transmitted stego-object.

In the digital era, modern steganographic techniques use a variety of digital media for the purpose of data hiding. Examples of such media include: images [3], audio tracks [4], video files [5], 3D Objects [6], and even file systems [7]. Recently, researchers attempted to use Deoxyribonucleic Acid (DNA) sequence data as a cover medium. In 1999, a paper published in Nature [8] presented one of the earliest methods for using biological DNA as a cover for the purpose of information hiding. The authors used a synthesized a DNA strand to encode the secret data that was then copied and camouflaged within an enormous number of similarly sized fragments of human DNA. Later, a small amount of the resultant DNA-containing solution was printed as a dot on a period in a typed letter. When mailed, the recipient of the letter was able to successfully recover the secret message after laboratory analysis. Another interesting example is live data storage, where digital data can be stored in the genome of a living organism; preserved for thousands of years and protected even from nuclear explosions [9]. The authors of [10] proposed a watermarking technique for RNA sequences in such a way the functionality of the organism remains intact. The main application for such a technique is protecting genetic discoveries such as gene therapy, transgenic crops, and tissue cloning. A similar approach was

*Corresponding Author: Amal Khalifa, Email: khalifaa@pfw.edu

proposed in [11] where the message is encrypted before hiding and block-sum checking is used to detect errors in mutations.

Another category of DNA-based steganographic methods was motivated by the huge publicly available genomic databases that store DNA sequence data on digital files following a simple format. In [12], the authors proposed that both the sender, and the receiver should agree or share a reference sequence before the communication takes place. The secret message can then be embedded into the reference sequence using one of three different methods. With the help of the reference sequence, the reverse of the hiding process can be performed by the rreceiver to extract the hidden message. However, communicating a sequence twice can draw suspicion and may affect the security of the steganographic channel itself. Secondly, the reference sequence is randomly modified without any consideration to the biological functionality of the DNA sequence. Addressing this later point, the authors of [13] introduced a hiding method that exploits the codon redundancy feature of DNA to hide data into sequence data without affecting the type nor the structure of protein it encodes for. Reversible hiding techniques, on the other hand, works in a way such that the cover sequence can be recovered from the stego-DNA sequence. an example of such technique was proposed in [14] where the cover sequence is coded into symbols that have integer values and the hiding process is then implemented using multilevel histogram shifting. Other methods added the power of encryption to provide more security to the hidden messages [15]. Some of these ciphers are actually bio-inspired and can be used to represent the message into a DNA sequence such as [16], [17] and [18]. A more recent research investigated using DNA steganography and PCR technology for the purpose of quantum key distribution (QKD) [19]. The authors of [20] provide a more comprehensive review on recent DNA-based steganographic methods.

In this paper, we present a method for securely hiding information into DNA sequences. The method is an extension of the work published by the author in [21]. The original research hides a secret message into a reference DNA sequence where both the message and the reference sequences are divided into random-length splices that are eventually merged to form the setgo-sequence. Before embedding, the binary message is encrypted and encoded into a DNA sequence using an 8x8 playfair cipher. In order to perform the extraction process correctly, it was proposed to add a header section before the embedded message to store the length of the hidden message. In this research, however, we propose a novel splicing technique based on the genes detected in the cover sequence. Therefore, the need for the message-size header information is eliminated. Finally, since the message splices are replacing the non-coding regions of the cover, the resultant stego-sequence is shorter. Although the aforementioned modifications don't increase the hiding capacity but it enhances both applicability and security of the method. The rest of the paper is organized as follows: section 2 provides a brief overview on some useful characteristics of DNA sequence data. Section 3 describes the details of the hiding and the extraction modules of the proposed method. In section 4, the performance of the proposed approach is measured, analyzed, and compared with some other techniques. Finally, section 5 concludes the paper.

2. DNA Sequence Data

The genetic information of all living organisms, as well as viruses, is stored in DeoxyriboNucleic Acid (DNA) molecules. A DNA molecule consists of two polynucleotide strands coiled around each other in the form of double helix structure. Each individual strand of DNA is made up with 4 different types of nucleotides. Nucleotides can contain either a purine base : adenine (A) and guanine (G) or pyrimidine base: thymine (T) and cytosine (C). In nature, A pairs only with T and G pairs only with C [22]. As a data medium, DNA can be represented as string of characters over the alphabet {A, C, G, T}. Using some coding rule, a DNA string can be converted into a string of binary digits where each base is mapped into two bits. Figure 1 shows an example of such coding rule along with a sample sequence encoded using this rule.

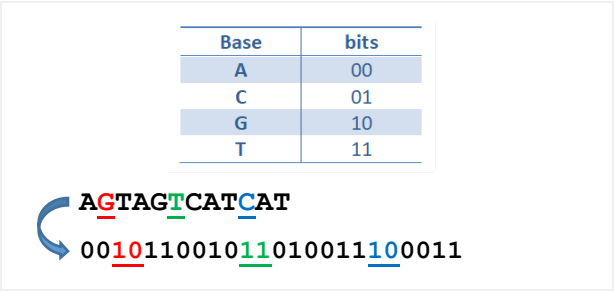


Figure 1: An example digital coding of DNA sequence data

The information stored in the DNA molecule plays a vital role in controlling all aspects of cell functionality. Through the complicated process of Central Dogma, the DNA sequence is read, copied, and eventually translated into a chain of amino acids that forms a protein [23]. Although the reading and the copying process is made one base at a time, the translation process reads the sequence into units of three adjacent nucleotides; called codons. With only 4 possible bases; there are 64 (4^3) distinctive ways to form 3-base long codons. As shown in figure 2, a three-letter abbreviation designates the type of amino acid molecule as in "Phe" and "Leu". Notice that some codons code for more than one amino acid. This is a feature called codon degeneracy [22]. Furthermore, three of the codons; indicated as STOP, identifies the end of the protein chain and doesn't actually code for any amino acid.

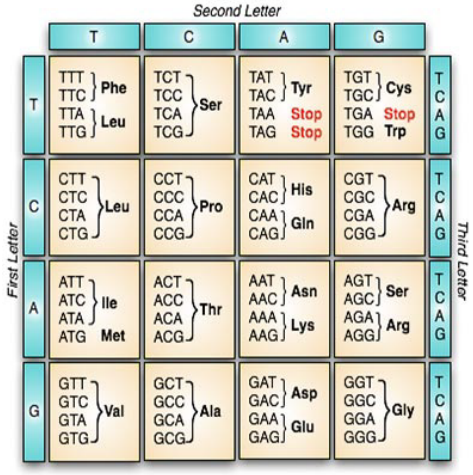


Figure 2: The genetic code table

In fact, not all parts of the DNA sequence code for proteins. That segment of the DNA that code for a protein is called a gene and hence other segments are referred to as non-coding regions. Genes can be identified by searching the DNA sequence for open reading frames (ORF). An ORF is a long stretch of codons that is preceded by a start codon (ATG) and will be uninterrupted by stop codons (TGA, TAG or TAA) [24]. This searching process proceeds by dividing the DNA sequence into a set of consecutive, non-overlapping triplets. This can start at the first, the second, or the third base in the sequence resulting in three different reading frames in that direction. The reading frame that has the potential to be translated into protein is identified as an ORF. Since DNA is double stranded and either strand could include a gene, there is a total of six reading frames: three in forward direction and three in the reverse complement direction. Figure3 shows an example of a sample a sample sequence and the three reading frames in its forward direction. The potential ORFs found in each frame are highlighted well. Notice that this sequence has three potential genes, one in each reading frame. The gene identified in frame 2 for instance consists of exactly three codons in between the start and the stop codons.

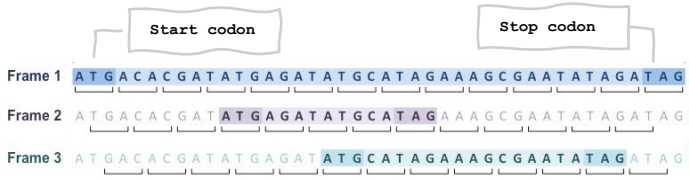


Figure 3: An example of Open Reading Frames (ORF) for gene identification

3. The Steganographic method

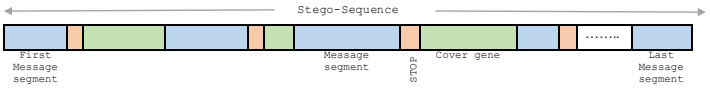
A steganographic communication channel consists of two main processes: hiding and extraction. Whereas the hiding process is normally carried out by the sender, the extraction process is performed by the receiver to reveal the secret message. As shown in figure 6, the proposed hiding process starts with ciphering the message and encoding it into a DNA sequence. Next, the cover sequence is spliced to identify the ORFs. The embedding process then moves forward by infusing the message sequence into the cover sequence splices. On the other side of the communication, the receiver extracts the secret message from the stego-sequence simply by reversing the hiding steps. Here we assume that both of the sender and the receiver share a secret key in advance. This key is used in different steps of the hiding process to make sure that only the intended recipient will be able to retrieve the hidden message.

3.1. The Hiding Process

The detailed steps of the hiding process are listed in Algorithm 1. It starts with an encryption step that utilizes a DNA-inspired implementation of Playfair cipher [17]. The classical Playfair cipher uses a 5 by 5 table to substitute a pair of letters (digraph) and with another one following a few simple rules. This implementation was originally proposed to work plaintext consisting only of alphabets without punctuation, or even numerical values. The cipher also requires a preprocessing step in order to remove spaces and handle double-letter digraphs. However, in [17], the author proposed using a randomly constructed 8x8 matrix representing the 64 different DNA codons. This technique is capable of ciphering any type of digital data not only text. In addition, the preprocessing step is no longer required since the digraphs are represented over the DNA not English alphabet. The final output of this cipher can be coded back into characters or left in its DNA data representation. In this research, we decided to use the cipher proposed in [17] to encrypt the secret message and use its intermediate DNA representation of the cipher-text. The encrypted message sequence is then infused into the cover sequence following a structured splicing methodology.

The splicing process is performed on both the cover and the message sequences. However, the splicing is carried out differently on each one of them. First, the cover sequence (C) is divided into coding and non-coding regions through an ORF analysis. Guided by the ORF result, the coding genes (C_g) are identified, and the cover sequence is spliced out at those specific locations. On the other hand, the message sequence (M) is randomly spliced into random-length segments (m_i) similar to the insertion method proposed in [12]. Finally, the message splices and the cover genes are merged to form the stego-sequence (S). The merging process is done in a very special way. As shown in figure 4, the merging process starts with a message segment followed by the first cover gene and then another message segment that is followed by the second cover gene and so on.

In fact, the base composition of message sequence can vary depending on several factors such as the message content, the secret key, and binary coding rule. Therefore, and like any random sequence, a message sequence will comprise of coding and non-coding regions as well. This may introduce a challenge in the extraction process on how to distinguish between a cover gene and a message gene. Therefore, we suggest adding a separator codon at the end of each message segment to uniquely identify the start of a cover gene.

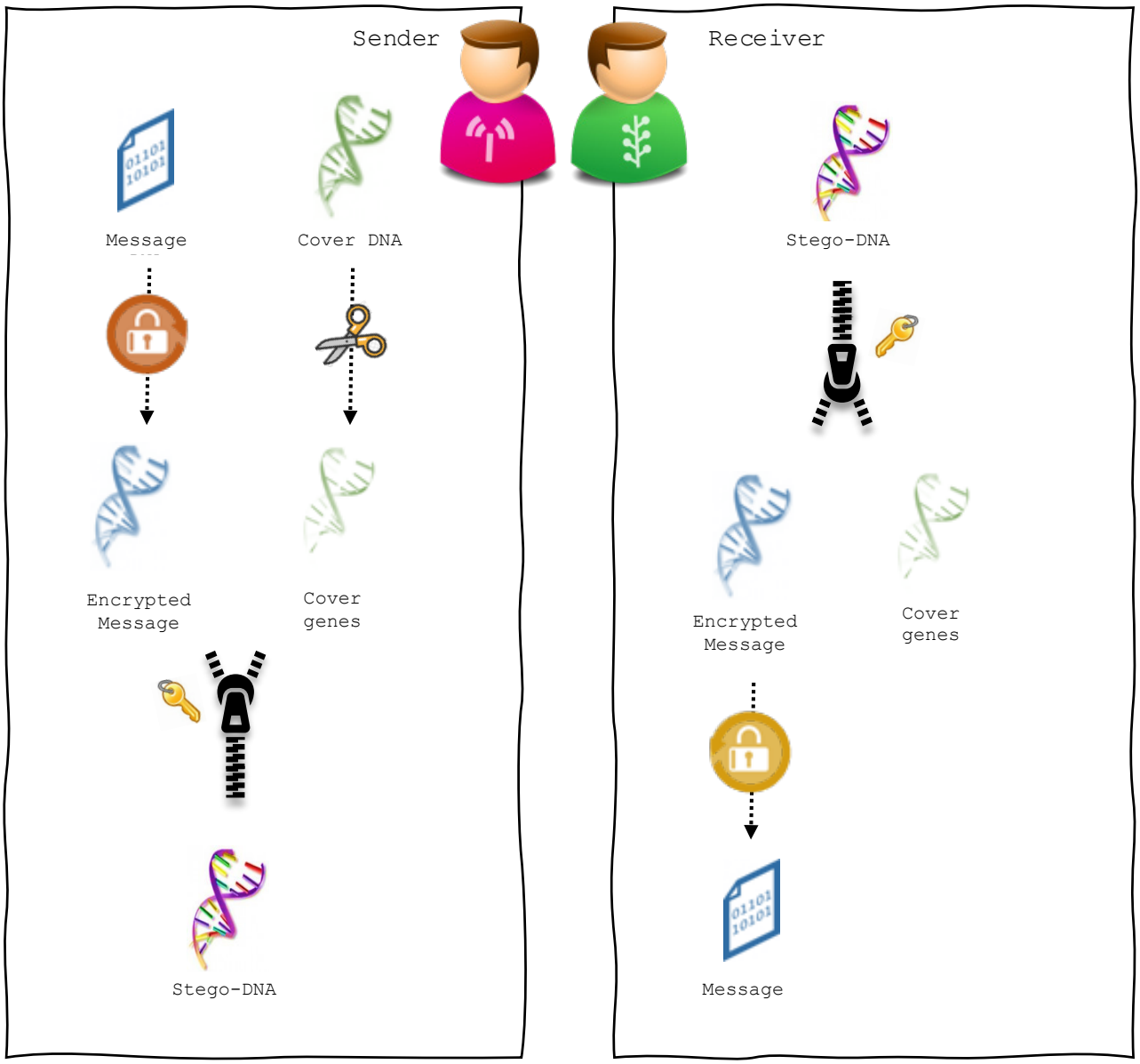


: Merging message splices with cover genes.

Theoretically, this codon can be any of the 63 codons (except the start codon). However, we suggest using one of the three stop codons (STOP) to avoid forming new genes in the message sequence that may overlap with the cover genes causing issues during the extraction process. Figure 6 shows an example of such a situation. The figure shows four ORFs (highlighted in blue). The first two of them are part of the message sequence, while the other two are cover genes since they are proceeded with the stop codon (TAA in this case).

```
000001  CCTCTTACATTCGGCTTGCCTTATGAAAGTTGGGTCAGCTGTAACGCACCTGCATTTCAGTCCT
000065  GATCGATCAGTGCTCCCCGTTTGTATGTTATGTTAGGAAGATTTGTCTCATCTTATGATTCCAC
000129  CGACAGTACCGTCCGCGCTGAAGGGCTGCGTGCAGACAACGGTCTTAAATTTGTTGTCCGATTC
000193  GGAGTGTTCCGCGCTCTTACATTCGGCTTGCCTTATGAAAGTTGGGTCAGCTGTAACGCACCTG
000257  CATTTCAGTCCTGATCGATCAGTGCTCCCCGTTTGTATGTTATGTTAGGAAGATTTGTCTCATC
000321  TTATGATTCCACCGACAGTACCGTCCGCGCTGAGAGGTAAATGCATAATTTTCTAAGAGAAAGC
000385  AGATTTCAAAAATAGCTGCGTGCAGACAACGGTCTTAAATTTGTTGTCCGATTCGGAGTGTAAA
000449  TGGAAGCTGCTGTGTGTTGTCATATCCAAGGGCGATTTCTCAGCAAGGGGAGCTGGTTAGTGCT
000513  CAGGAATACAGACAGATGATTCCGCGCTCTTACATTCGGCTTGCCTTATGAAAGTTGGGTCAGC
```

Figure 5: An example of a ORFs formed in message splices.



3.2. The Extraction Process

The extraction is obviously the reverse process of hiding. As listed in Algorithm 2, the message retrieval starts with sequence splitting based on the specified reading frame. The locations of the start and the stop codons of the genes can then be identified. It is expected to find more genes in the stego-sequence than those used from the cover during the hiding process. As mentioned above, depending on the message contents and the randomness of the encryption algorithm, the message sequence can form new genes. Therefore, the proposed algorithm checks, for every detected gene, the codon that proceeds its start codon. If it is the specified STOP codon utilized during the hiding process, the gene will be skipped and the message segment that follows will be extracted. Otherwise, this detected gene actually belongs to the message sequence.

Figure 7 uses the same stego-sequence example shown in figure 6. Here only the highlighted segments belong to the message sequence. Notice how the first message segment spans the range of bases from 1 to 357 but when extracted, it will be concatenated from four different segments because of the two ORFs formed in its content. Furthermore, since the third and the fourth detected ORFs are cover genes, the segment following each one of them is extracted and appended to the message sequence extracted so far. Once all ORFs are processed, all segments of the message will be extracted, concatenated, and deciphered. Thus, there is no need to embed any header information to store the length of the hidden message as required in [21]. It is important to mention that, the extraction process here is done blindly without the need to reference the reference cover sequence.

Algorithm 1: Message Hiding

Input: *C* : A reference DNA sequence, used as a cover
Msg : a secret message
Key : a secret key-word
Result: *S* : A Stego-DNA sequence

1. Message Ciphering
 1.1 Encode *Msg* into Msg_{DNA} using a DNA binary coding rule
 1.2 Build an 8x8 codon matrix
 1.3 Shuffle the matrix based on the *Key*
 1.4 **for each** pair of codons in Msg_{DNA} **do**
  if the pair appears on the same row of the matrix, replace each codon with its immediate right with wrap around
  else if the pair appears on the same column of the matrix, replace each codon with its immediate below with wrap around.
  else replace the pair with the pair at the corners of the rectangle defined by the original pair.
  end
 end
 1.5 Build a column-wise square matrix filled with bases in Msg_{DNA} in reverse order
 1.6 **for each** column and row in sequence matrix **do**
  1.6.1 Rotate circular upward on the column
  1.6.2 Rotate circular left on the row
 end
 1.7 Rearrange encrypted message into a linear sequence Msg_{cph}

2. Cover Splicing
 2.1 Locate genes in *C* considering one of the ORFs.
 2.2 Let (C_1 , C_2 , C_3 , … … C_{t_c}) be the detected cover genes.

3. Message Splicing:
 3.1 Let *i* be a value derived from *Key*
 3.2 Generate a sequence of random numbers (i_1 , i_2 , i_3 …) using *i* as the seed
 3.3 Find the smallest integer t_m such that :
  $\sum_{k=1}^{tm} i_k > |Msg_{cph}|$
 3.4 **if** $t_m < t_c$
  Divide Msg_{cph} into t_m *-1* segments (M_1 , M_2 , M_3 , … … M_{t_m-1}) with lengths (i_1 , i_2 , i_3 , … … i_{t_m-1}) respectively and keep the residual part of Msg_{cph} in M_{t_m}
 else
  return.
 end

4. Segment merging:
 4.1 Initialize *S* as an empty sequence
 4.2 **for each** $k = 1$ to t_m *-1* **do**
  Append M_k to *S*
  Append *STOP* to *S*
  Append C_k to *S*
 4.3 Append M_{t_m} to *S*
5. return *S*

end

Algorithm 2: Message retrieval

Input: *S* : A Stego-DNA sequence
Key : a secret key-word
Result: *Msg* : the embedded secret message

1. ORF Splicing
 1.1 find location of genes in *S* for the specified ORF
 1.2 Let (S_1 , S_2 , S_3 , … … S_t) be the detected cover genes
 1.3 Let $start(S_i)$ return the location of the start codon for the gene S_i
 1.4 Let $stop(S_i)$ return the location of the stop codon for the gene S_i

2. Message extraction:
 2.1 Initialize Msg_{cph} as an empty sequence
 2.2 let k = 1
 2.3 **for** $k = 1$ to t *-1*
  if $start(S_k)$ is proceeded by STOP
   extend Msg_{cph} by adding the bases from $stop(S_k)$ to $start(S_{k+1})$
  else
   extend Msg_{cph} by adding the bases from $start(S_k)$ to $start(S_{k+1})$
  end

 end
 2.4 Append the bases following $stop(S_t)$ to Msg_{cph}

3. Message Deciphering:
 3.1 Build an 8x8 codon matrix
 3.2 Shuffle the matrix based on the *Key*
 3.3 **for each** pair of codons in Msg_{DNA} **do**
  if the pair appears on the same row of the matrix, replace each codon with its immediate right with wrap around
  else if the pair appears on the same column of the matrix, replace each codon with its immediate below with wrap around.
  else replace the pair with the other pair at corners of the rectangle defined by the original pair.
  end

 end
 3.4 Build a column-wise square matrix filled with bases in Msg_{DNA} in reverse order
 3.5 **for each** column and row in sequence matrix
  Rotate circular upward on the column
  Rotate circular left on the row
 end
 3.6 Rearrange deciphered message into a linear sequence Msg_{DNA}
 3.7 Convert Msg_{DNA} into *Msg* using the specified coding rule

4. return *Msg*

end



Figure 7: An example of detecting and extracting message splices.

4. Performance Analysis

4.1. Hiding Capacity

Usually, the hiding capacity of a steganographic technique is measured by the maximum number of bits that can be hidden into the cover media. For DNA data, this capacity is measured in bit-per-nucleotide (*bpn*) which reflects how many bits can be hidden for each nucleotide in the cover sequence. Since the proposed algorithm requires only that the number of message segments should be less than or equal to the number of genes in the cover sequence, this doesn't impose any restrictions on the length of message. Theoretically, fine tuning the embedding parameters would allow us to hide a message sequence that is as long as the cover sequence itself. Furthermore, since each encoded nucleotide in a message sequence represents 2 bits, the hiding capacity of the proposed method can be expressed as follows:

$$Capacity = \frac{size\ of\ message\ in\ bits}{size\ of\ cover\ in\ bases}$$

$$= \frac{2*|C|}{|C|} = 2\ bpn \qquad (1)$$

where C represents the cover sequence and |C| refers to the length of C in base pairs (bp).

4.2. Security

Generally speaking, the harder for an attacker to crack the implementation of a steganographic method, the more secure this approach is. Therefore, the proposed method was designed in way that the details of the hiding process are based on several parameters the attacker needs to guess correctly in order to extract the hidden message. Beside the value of the secret key, the attacker needs to know: the binary rule used to encode the DNA nucleotides, the random number generator used during the message splicing, the ORF selected for cover splicing, and the stop codon used to separate between message and cover segments.

First, given the fact that there are only 4 nucleotides, there are 4! = 24 possible binary coding rules. So, the probability of a successful guess on this parameter is 1/24 . Similarly, there are 6 different ORFs and 3 different stop codons. This makes the probability of predicting each one of them to be 1/6 and 1/3 respectively. More importantly, to find the sequence of numbers generated to randomly slice the message sequence, an attacker may need to make a number of guesses [12] up to the value computed in (2). Where *n* represents the length of the message sequence and $\binom{n}{k}$ is the set of all *k*-combinations of *n*.

$$\binom{n}{n-1} + \binom{n}{n-2} + \binom{n}{n-3} + \ldots + \binom{n}{0}$$

$$= \sum_{k=0}^{n-1} \binom{n}{n-1-k} = 2^n - 1 \qquad (2)$$

Since the summations of the *k*'s can be as long as the cover sequence, the probability of a successful guess for this parameter can reach $\frac{1}{(2^{|C|}-1)}$, where C represents the cover sequence and |C| refers to the length of C in base pairs (*bp*). In conclusion, the probability of cracking the hiding process of the proposed method can be estimated as follows:

$$P_{bf}\ \frac{1}{(2^{|C|}-1)}\ \text{x}\ \frac{1}{24}\ \text{x}\ \frac{1}{6}\ \text{x}\ \frac{1}{3} \qquad (3)$$

Considering the fact that |C| may reach hundreds of thousands bases, it is almost impossible for an attacker to retrieve the hidden message based on a successful guess. In addition, the message itself is protected by a powerful cipher that adds another layer of security to the hidden message in case the steganographic method fails.

5. Results and Discussions

In this section, the proposed method is tested using 10 different cover sequences. The discussion is then followed by a comparison with some existing methods based on different performance measures.

5.1. Experimental results

The sequence data was drawn from the Genbank database on the National Center for Biotechnology Information (NCBI) website using their unique accession numbers. The downloaded files in the FASTA format which is a text-based format that starts with a single-line description followed by lines of sequence data using the {A, C, T, G} alphabet. Furthermore, in this set of experiments, we choose to randomly generate 30KB of textual data and use it as a sample secret message. However, this should not limit the applicability of the proposed algorithm to hide any type of digital data as long as it is encoded using a rule similar to the one given in Fig. 1.

Table 1 summarizes the experimental results for 10 sequences using the secret key "12345". It lists the lengths of the cover- and stego- sequences as well as the number of ORFs detected in the first reading frame of the sequences in each case. The payload is computed as the ratio of message to cover nucleotides in the stego-sequence. Notice that the results show an increase in the number of genes detected in the stego-sequence compared with those of the cover sequence. That was expected because the encoded message sequence will probably form some genes that will add to the detected ORFs in the resultant stego-sequence. For example, in the case of the AAEX03000038 sequence, the message was sliced into 343 segments. this means that only 342 genes out of the 364 cover genes are needed for hiding. However, during the extraction process, 657 genes were detected in the stego-sequence representing not only cover genes but also genes formed by the message content.

Furthermore, the embedding process was not successful for sequences AAEX03000001, AAEX03000010, and AAEX03000999 because none of them has enough ORFs to hide the message segments. For example, *AAEX03000999* is a sequence that is 22,099 bases long with only 48 ORFs. Since the message is sliced into 832 segments, we need at least 831 cover genes to infuse those segments into. It is also worth to notice that the number of message segments changes in each case. In fact, in our implementation, the randomness of the splicing process does not only depend on the secret key, but also on the number of bases in the longest ORF in the identified the cover sequence.

5.2. Comparisons

In this set of experiments, the performance of the proposed technique was compared with some existing methods in terms of capacity, blind extraction, and security against brute force attacks. With blind extraction we mean that the original cover sequence is not required during the extraction process to retrieve the embedded message. With non-blind methods, the cover sequence needs to be communicated in advance which can draw suspicion and affect the security of the channel.

The results listed in table 2, show that the insertion method proposed in [12] may be more robust to brute-forth attacks than the proposed method, but it is not blind. On the other hand, the proposed method achieves a higher hiding capacity than all the methods in [12] while maintaining the advantage of blind extraction. The same is true for the GCBC method, despite it is also a blind technique, its hiding capacity is less than that offered by the proposed method. The methods proposed in [12] were not evaluated in terms of robustness, so there is no way to compare them to the proposed method in that regard. However, it is obvious though that they did not succeed to offer a better hiding capacity. Finally, although the technique introduced in [21] offers the same hiding capacity, the details of the proposed method added more parameters to the hiding process that enhanced its security.

Table 1: Experimental results for hiding 30KB of text

Cover Sequence			**Number of message segments**	*Stego-Sequence*		
Accession number	**Length (bp)**	**Number of detected ORFs**		**Length (bp)**	**Number of detected ORFs**	**Payload (*bpn*)**
AL645637	207,629	588	222	147,239	525	1.69
AC168892	197,841	563	452	167,027	743	1.49
AAEX03000080	305,811	823	437	162,125	750	1.54
AAEX03000038	133,800	364	343	157,811	657	1.59
AC153526	200,117	521	281	151,259	592	1.65
AC168874	206,488	523	466	169,277	771	1.48
AL645625	226,754	553	490	178,616	790	1.40
AAEX03000001	24,025	59	526	--	--	--
AAEX03000010	44,186	98	403	--	--	--
AAEX03000999	22,099	48	832	--	--	--

Table 2: A comparison with some existing techniques

Author	**Method**	**Capacity (bpn)**	**Security (P_{bf})**	**Blind?**
[12], 2010	Insertion	0.58	$\frac{1}{1.63\ x\ 10^8}\ x\frac{1}{n-1}\ x\ \frac{1}{2^m-1}\ x\ \frac{1}{2^{s-1}} x\ \frac{1}{24}$	No
	Complementary	0.07	$\frac{1}{1.63\ x\ 10^8}\ x\ \frac{1}{24^2}$	No
	Substitution	0.82	$\frac{1}{(2^{\lvert S\rvert}-1)^2}\ x\ \frac{1}{6}$	No
[15], 2016	Generic Complementary Base Substitution (GCBS)	1.5	$\frac{1}{(2^{\lvert S\rvert}-1)^2}\ x\ \frac{1}{6}\ x\ \frac{1}{24}$	Yes
[16], 2018	Noncircular type (NHS)	1.243	NA	yes
	Circular type (CHS)	1.865	NA	yes
[21], 2020	Random Splicing	2	$\frac{1}{(2^{\lvert C\rvert}-1)^2}\ x\ \frac{1}{24}$	yes
Proposed	ORF-guided Splicing	2	$\frac{1}{(2^{\lvert C\rvert}-1)}\ x\ \frac{1}{24}\ x\ \frac{1}{6}\ x\ \frac{1}{3}$	yes

6. Conclusions

This paper describes a steganographic method that extends the work presented in [21] and uses DNA sequence data for data hiding purposes. The proposed technique consist of two processes: hiding and extraction. The hiding module starts with an encryption step that encodes the secret message as a DNA sequence. The cover sequence is then spliced into coding and non-coding segments in a given reading frame. The encrypted sequence is also spliced; but is a random fashion, and infused into the cover's gene-coding segments. The extraction module, on the other hand, reverses the hiding process starting from gene detection and ending with a decryption step.

The proposed method is blind, which means that parties of the communication don't need to exchange anything in advance other than the secret key. Experimental results showed a superior performance of the proposed technique in terms of capacity and security. The proposed method achieved a hiding capacity of 2 bit per nucleotide, which is the highest among all methods except [21]. However, the proposed technique succeeded to eliminate the need for embedding header information about the size of the hidden message in [21]. Furthermore, the proposed method showed strong robustness against attacks making it almost unbreakable. Not to mention the ciphering step that protects the contents of the secret message in case the steganographic shield was cracked.

This research can be extended in one of two directions. First, introducing randomness in the merging process of the message and cover segments. That is, message segments can be infused in a random order in between the cover ORFs. Secondly, encryption techniques other than the play-fair cipher can be explored as well.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Herodotus, Herodotus : The History, University of Chicago Press, 1987.

[2] G. Simmons, "The Prisoners' Problem and the Subliminal Channel," Advances in Cryptology, Springer, 51-67, 1984, doi:10.1007/978-1-4684-4730-9_5.

[3] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital Image Steganography: Survey and Analysis of Current Methods," Signal Processing, **90**, 727-752, 2010. doi: 10.1016/j.sigpro.2009.08.010.

[4] P. Balgurgi, S. Jagtap, "Audio Steganography used for Secure Data Transmission," in International Conference on Advances in Computing. Advances in Intelligent Systems and Computing, 2013, doi:10.1007/978-81-322-0740-5_83.

[5] M. Sadek, A. Khalifa, M. Gad-El Haaq, "Robust Video Steganography Algorithm using Adaptive Skin-tone Detection," Multimedia Tools and Applications, **76**(2), 2017, doi:10.1007/s11042-015-3170-8.

[6] Y. Tsai, I. Chi, C. Chan, "A Vertex-Based 3D Authentication Algorithm Based on Spatial Subdivision," Symmetry, **10**(10), 2018, doi:10.3390/sym10100422.

[7] J. MacLean, D. Dampier, J. Davis, "Methods of Information Hiding and Detection in File Systems," in Fifth International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE), 66-69, 2010, doi:10.1109/SADFE.2010.17.

[8] C. Clelland, R. Taylor, C. Bancroft, "Hiding Messages in DNA Microdots," Nature, **399**, 533-534, 1999, doi:10.1038/21092.

[9] S. Jiao, R. Goutte, "Hiding Data in DNA of Living Organisms," Natural Science, **1**(3), 181-184, 2009, doi:10.4236/ns.2009.13023.

[10] D. Heider, M. Pyka, A. Barnekow, "DNA Watermarks in Non-Coding Regulatory Sequences," BMC Research Notes, **2**(125), 2009, doi:10.1186/1756-0500-2-125

[11] D. Na, "DNA Steganography: Hiding Undetectable Secret Messages Within the Single Nucleotide Polymorphisms Of A Genome and Detecting Mutation-Induced Errors," Microb Cell Fact, **19**(128), 2020, doi:10.1186/s12934-020-01387-0

[12] H. Shiu, K. Ng, J. Fang, R. Lee, C. Huang, "Data Hiding Methods based upon DNA Sequences," Information Sciences., **180**(11), 2196-2208, 2010. doi: 10.1016/j.ins.2010.01.030.

[13] A. Khalifa, "LSBase: A Key Encapsulation Scheme to Improve Hybrid Crypto-Systems using DNA Steganography," in 8th International conference on Computer and Engineering, 105-110, 2013, doi:10.1109/ICCES.2013.6707182.

[14] S. Lee, "Reversible Data Hiding for DNA Sequence using Multilevel Histogram Shifting," Security and Communication Networks, 2018. doi:10.1155/2018/3530969.

[15] A. Khalifa, A. Elhadad, S. Hamad, "Secure Blind Data Hiding into Pseudo DNA Sequences Using Playfair Ciphering and Generic Complementary Substitution," Applied Mathematics & Information Sciences, **10**(4), 2016, doi:10.18576/amis/100427

[16] M. Sabry, M. Hashem, T. Nazmy, M. Khalifa, "A DNA and Amino Acids-Based Implementation of Playfair Cipher," (IJCSIS) International Journal of Computer Science and Information Security, **8**(3), 129-136, 2010.

[17] S. Hamad, "A Novel Implementation of an Extended 8x8 Playfair Cipher Using Interweaving on DNA-encoded Data," International Journal of Electrical and Computer Engineering, **4**(1), 93-100, 2014, doi:10.11591/ijece.v4i1.4969.

[18] M. Biswas, K. Alam, S. Tamura, Y. Morimoto, "A Technique for DNA Cryptography based on Dynamic Mechanisms," Journal of Information Security and Applications, **48**, 2019, doi: 10.1016/j.jisa.2019.102363.

[19] M. Cui, Y. Zhang, "Advancing DNA Steganography with Incorporation of Randomness," ChemBioChem, **21**(17), 2503-2511, 2020, doi:10.1002/cbic.202000149.

[20] D. Zebari, H. Haron, S. Zeebaree, "Security Issues in DNA Based on Data Hiding: A Review," International Journal of Applied Engineering Research, **12**(24), 15363-15377, 2017.

[21] A. Khalifa, "A Blind DNA-Steganography Approach using Ciphering and Random Sequence Splicing," in 10th International Conference on Information Science and Technology (ICIST), 86-90, 2020, doi:10.1109/ICIST49303.2020.9202036.

[22] A. Ghosh, M. Bansal, "A glossary of DNA structures from A to Z," ACTA Crystallography D, Biological Crystallography, **59**(4), 620-626, 2003, doi:10.1107/s0907444903003251.

[23] F. Crick, "Central Dogma of Molecular Biology," Nature, **227**, 561–563, 1970, doi:10.1038/227561a0.

[24] M. Adam, H. Ardinger, R. Pagon, et al, Gene Reviews, University of Washington, 1993-2021.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 167-176 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Acoustic Scene Classifier Based on Gaussian Mixture Model in the Concept Drift Situation

Ibnu Daqiqil Id*, Masanobu Abe, Sunao Hara

Graduate School of Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama, 700-0001, Japan

ARTICLEINFO

Article history:
Received: 15 July, 2021
Accepted: 22 September, 2021
Online: 30 September, 2021

Keywords:
Acoustic Scene Classifier
Concept Drift
Kernel Density Estimation
Gaussian Mixture Model

ABSTRACT

The data distribution used in model training is assumed to be similar with that when the model is applied. However, in some applications, data distributions may change over time. This situation is called the concept drift, which might decrease the model performance because the model is trained and evaluated in different distributions. To solve this problem for scene audio classification, this study proposes the kernel density drift detection (KD3) algorithm to detect the concept drift and the combine–merge Gaussian mixture model (CMGMM) algorithm to adapt to the concept drift. The strength of the CMGMM algorithm is its ability to perform adaptation and continuously learn from stream data with a local replacement strategy that enables it to preserve previously learned knowledge and avoid catastrophic forgetting. KD3 plays an essential role in detecting the concept drift and supplying adaptation data to the CMGMM. Their performance is evaluated for four types of concept drift with three systematically generated scenarios. The CMGMM is evaluated with and without the concept drift detector. In summary, the combination of the CMGMM and KD3 outperforms two of four other combination methods and shows its best performance at a recurring concept drift.

1. Introduction

Human–computer interaction through audition requires devices to recognize the environment using acoustic sound analysis. One of the primary research topics in this area is acoustic scene classification (ASC), which attempts to classify digital audio signals into mutually exclusive scene categories. ASC is an important area of study covering various applications, including smart homes, context-aware audio services, security surveillance, mobile robot navigation, and wildlife monitoring in natural habitats. Machine audition applications have a high potential to lead to more innovative context-aware services.

We intend to develop an ASC system for environmental or scene audio in specific locations (i.e., beach, shop, bus station, and airport) with different acoustic characteristics. The scene audio contains an ensemble of background and foreground sounds. One of the most important aspects of the audio scene in real life is the concept drift [1], whose data distribution might evolve or change in the future. For example, the foreground event sounds in a bus station, such as an ambulance siren, wind noise, and rustling sounds, might change because of the physical environment, human activities, or nature [2]. These changes cause the acoustic data distributions to change, potentially causing a lower performance in the trained model [3].

The simplest solutions for handling the abovementioned problems to maintain the model performance are periodic retraining and redeployment of the model. Nevertheless, these solutions can be time consuming and costly. Moreover, the decision for the frequency of retraining and redeployment is a difficult task. Another promising approach is to use an evolving or incremental learning method [4], [5], where the model is updated when a new subset of data arrives [6]. Each iteration is considered as an incremental step toward revisiting the current model.

In this study, we propose a combine–merge Gaussian mixture model (CMGMM) and kernel density drift detection (KD3) to solve the concept drift problem [7]. The CMGMM is an algorithm based on the Gaussian mixture model (GMM) that adapts to the concept drift by adding or modifying its components to accommodate the emerging concept drift. The algorithm's advantages are adaptation and continuous learning from stream data with a local replacement strategy to preserve previously learned knowledge and avoid catastrophic forgetting.

In [7], we compared the CMGMM to the incremental GMM (IGMM) [8] and KD3 to adaptive windowing (ADWIN) [9] and HDMM [10] in two approaches, namely the active and passive

*Corresponding Author: Ibnu Daqiqil Id, Okayama, Japan, daqiqil@s.okayama-u.ac.jp

approaches. In the active approach, the concept drift is detected using a specific algorithm, then adapts the model. In the passive approach, the model is continuously adapted at a specific fixed interval. The result is that the combination of CMGMM and KD3 outperforms other combinations in two of three evaluation scenarios.

The work described herein extends and improves that of previous publications [7,11] in the following respects:

- The algorithm has been modified to use component pruning to overcome the overfitting problem and support the Scikit-Multiflow Framework [12].
- The KD3 hyperparameter is optimized, and the algorithm is evaluated using prequential evaluation for better results and online performance monitoring in several concept drift types and scenarios.

The rest of this paper is organized as follows: Section 2 presents the related work of this research and the fundamental equations used in our proposed solutions; Section 3 describes the proposed CMGMM and KD3; Section 4 outlines the experimental setup; Section 5 discusses the experimental results; and finally, Section 6 provides our conclusions.

2. Related Work

This section provides a brief overview of the related work on the concept drift and the fundamental theory used in the proposed method.

2.1. Concept Drift in Audio Scenes

A concept is defined as a set of object instances [13]. Probabilistically, a concept is defined using prior class probabilities $p(y)$ and class conditional probabilities $p(X|y)$ [4]. $p(y)$ and $p(X|y)$ determine the joint distribution $p(X,y)$ [3]; hence, a concept is defined as the joint probability distribution of a set of input features X and the corresponding label y in dataset $\mathfrak{D}$. In this paper, X is an acoustic scene sound defined as a mixture of specific event sounds ($\hat{x}$) perceived and defined by humans [14], and y is the label of X. X has numerous types of acoustic event sounds and background noises $\hat{x}$ that often overlap with each other. In other words, the relationship between of X and $\hat{x}$ determines $p(X,y)$, $X \in (\hat{x}_1, \hat{x}_2, \hat{x}_3, .\hat{x}_i,)$, where i denotes the number of $\hat{x}$ in X.

In the future, the relationship of $\hat{x}$ in X might change, which then changes the relationship of $p(X,y)$. For example, another event sound might appear, or some existing event sounds may disappear. This situation is called the concept drift, which is expressed as follows:

$$\exists X: p_{w_0}(X,y) \neq p_{w_n}(X,y). \quad (1)$$

Eq. (1) and Figure 1 describe concept drift as the change in the joint probability distribution between two-time windows, w_0 and w_n. Models built on previous data at w_0 might not be suitable for predicting new incoming data at w_n. This change may be caused by a change not only in the number of $\hat{x}$, but also in the underlying data distribution of $\hat{x}$. These changes require model adaptation because the model's error may no longer be acceptable with the new data distribution [15].

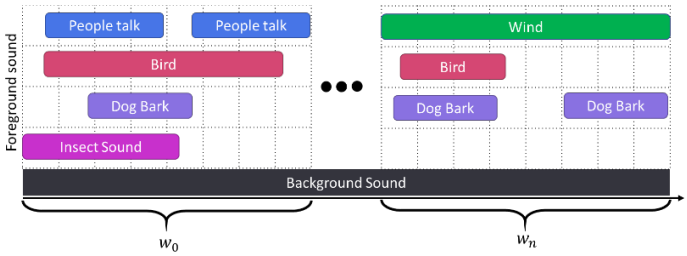


Figure 1: Illustration of the concept drifts in an acoustic scene audio at a park

The change in the incoming data at w_n depends on a variety of different internal or external influences (e.g., event sounds that exist in a park depending on the season). The initial data recorded in the winter may only consist of people talking, bird calls, and dogs barking. However, the event sounds change in the summer, and new event sounds, such as insect and wind sounds, emerge.

2.2. Concept Drift Detection Methods

Several methods have been proposed to detect concept drifts from a data stream. This study focuses on window-based methods that use fixed windows as a reference for summarizing previous information. This approach has more accurate results than other more straightforward methods, such as cumulative sum [16]. However, the computational time and space used are higher [17]. This approach usually utilizes statistical tests or mathematical inequalities to compute the change in data. Some of the state-of-the-art methods used in this paper is presented below:

- ADWIN is a sliding window-based concept drift detection algorithm. The size of ADWIN windows *SW* might change depending on the instance in the distribution. Its size increases when the instance in the stream continues in the same distribution. *SW* shrinks when distribution changes occur [9]. ADWIN detects concept drifts when the averages between these windows are higher than a given threshold.
- The HDDM is a concept drift detection algorithm based on fixed windows and probability inequalities [10]. The author proposes two types of HDDM, namely HDDMA and HDDMW. The HDDMA uses moving averages, whereas the HDDMW uses weighted moving averages to detect the concept drift. The HDDMA is suitable for detecting the abrupt concept drift, whereas the HDDMW is suitable for detecting the gradual concept drift.
- KSWIN [18] is a window-based concept drift detection method that utilizes the Kolmogorov–Smirnov statistic test (KS-Test) to compare the distances of two distributions. This test is a non-parametric test that does not require any assumptions about the underlying data distribution.

Each method has its optimal hyperparameters, which differ based on the datasets used and the type of drift in those datasets.

2.3. Merging Gaussian Mixture Component

A component of a Gaussian distribution is represented by (w, μ, P) and $\{(w_1, \mu_1, P_1), (w_2, \mu_2, P_2), \ldots (w_n, \mu_n, P_n)\}$ to denote a mixture of n Gaussian components, where w, μ, and P are the weight or prior probability, distribution means, and covariance matrix, respectively. This mixture must satisfy $w_1 + w_2 + \cdots + w_n = 1$ and has the pdf defined in Eq. (2).

$$f(X) = \sum_{i=1}^{n} w_i N(x; \mu_i, P_i), \quad (2)$$

where,

$$N(x; \mu_i, P_i) = \frac{1}{\sqrt{(2\pi)^d \det P_i}} e^{\left[-\frac{1}{2}(x-\mu_i)^T P_i^{-1}(x-\mu_i)\right]}. \quad (3)$$

Suppose we wish to merge two Gaussian components $\{(w_i, \mu_i, P_i), (w_j, \mu_j, P_j)\}$, where $w_i + w_j \leq 1$, and approximate the result as a single Gaussian. The new Gaussian candidate $(w_{ij}, \mu_{ij}, P_{ij})$ must preserve the zeroth-, first-, and second-order moments of the original Gaussian. The moment-preserving merge is shown in Eqs. (4)–(6).

$$w_{ij} = w_i + w_j \quad (4)$$

$$\mu_{ij} = \frac{w_i}{w_i + w_j}\mu_i + \frac{w_j}{w_i + w_j}\mu_j \quad (5)$$

$$P_{ij} = \frac{w_i}{w_i + w_j}P_i + \frac{w_j}{w_i + w_j}P_j + \frac{w_i w_j}{(w_i + w_j)^2}(\mu_i - \mu_j)(\mu_i - \mu_j)^T \quad (6)$$

2.4. Kullback–Leibler (KL) Dissimilarity

Kullback–Leibler (KL) discrimination, known as KL divergence or relative entropy, is a tool for measuring the discrepancy between two probability distributions. The KL discrimination between $f(x)$, a probability distribution for random variables X and $g(x)$, another probability distribution is the expected value of the log-likelihood ratio. The KL divergence of $f(x)$ and $g(x)$ is defined in Eq. (7), where $\Re^d$ is the sample space of the random variable X.

$$d_{KL}(f, g) = \int_{\Re^d} f(x) log \frac{f(x)}{g(x)} dx. \quad (7)$$

Based on Eq. (7), KL is not a perfect distance metric because it is asymmetric $d_{KL}(f, g) \neq d_{KL}(g, f)$ and does not satisfy the triangle inequality, $d_{KL}(f, g) + d_{KL}(g, h) \geq d_{KL}(f, h)$. However, we can use it as metric distance because the KL discrimination of two probability distributions is larger than zero, $d_{KL}(f, g) \geq 0$, and the discrimination of the identical two probability distributions is zero, $d_{KL}(f, f) = 0$.

Accordingly, we apply the KL dissimilarity by computing the Kullback–Leibler discrimination upper bound of the post-merge mixture with respect to the pre-merge mixture. In the case of the Gaussian mixture, where $f(x) = N(w_1, \mu_1, P_1), g(x) = N(w_2, \mu_2, P_2)$ and $w_1 + w_2 < 1$, the KL dissimilarity between $f(x)$ and $g(x)$ is shown in Eq. (8).

$$d_{KL}(\{(w_1, \mu_1, P_1)\}, \{(w_2, \mu_2, P_2)\}) = \frac{1}{2}\left[tr(P_2^{-1}[P_1 - P_2 + (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T]) + log\frac{\det(P_2)}{\det(P_1)}\right] \quad (8)$$

Please refer to [19] for more details about the KL dissimilarity of the Gaussian distributions.

3. Proposed Method

This study extends the combine–merge Gaussian mixture model (CMGMM) [7] to classify audio scenes in concept drift situations. The CMGMM was developed based on the GMM algorithm that can incrementally adapt to the new identified component. This algorithm can add new components as new concepts and update existing components as a response to the change of the current existing concept in the current data. The CMGMM implementation is available in our public repository[1], with the algorithm pipeline shown in Figure 2.

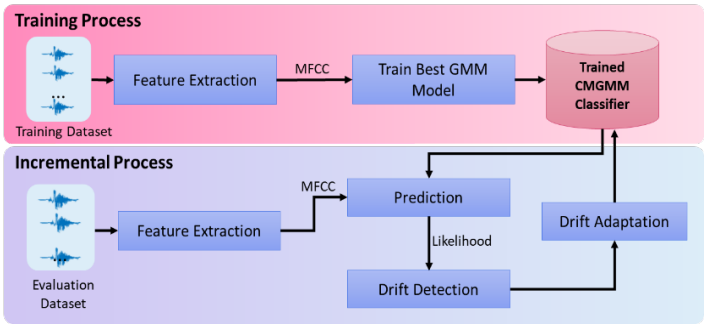


Figure 2: Combine–merge Gaussian mixture model (CMGMM) general workflow

In the training process, we extract the feature of the scene audio from the training dataset D_0 and train an optimal model $M_{optimal}$. We use the Expectation maximization (EM) [20] algorithm to train the model and the Bayesian information criterion (BIC) [21] to select the best model.

In the incremental process, the $M_{optimal}$ performance is observed through the prediction likelihood. When KD3 detects a significant likelihood change, the model activates the concept drift adaptation process. The concept drift adaptation process then begins by creating a local model $M_{drifted}$ from the new coming data. $M_{drifted}$ represents the new concepts or concept updates in the incoming data. Finally, we combine the $M_{optimal}$ and $M_{drifted}$ components to include any new concepts from $M_{drifted}$ that may not exist in the $M_{optimal}$ at the initial training and merge similar components to update the existing component in $M_{optimal}$.

The CMGMM pipeline process is detailed in the subsections that follow.

3.1. Feature Extraction

Feature extraction is the first step of both the training and incremental processes. In this research, we use normalized Mel-frequency cepstral coefficients (MFCCs) that represent the short-term power spectrum of audio in the frequency domain of the Mel scale. MFCCs are commonly used as features in audio processing and speech recognition. The first step is pre-emphasis for enhancing the quantity of energy in high frequencies. The next step is windowing the signal and computing the fast Fourier transformation to transform the sample from the time domain to the frequency domain. Subsequently, the frequencies are wrapped

[1]https://github.com/ibnudaqiqil/CMGMM

on a Mel scale, and the inverse DCT is applied [22]. Finally, each of the MFCCs is normalized using mean and variance normalization based on Eq. (9):

$$MFCC_{norm} = \frac{(MFCC-\mu)}{S}, \quad (9)$$

where, μ and S denote the mean and the standard deviation of the training samples, respectively.

3.2. Model Training

The training process is intended to build a set of models from the training dataset D_0 containing training data $x = \{x_1, x_2, \dots, x_n\}$, where x_n denotes the MFCC vector. The models are trained Q times using the EM algorithm. For each training cycle, a different number of components K ranging from K_{min} to K_{max} is used, where $Q = K_{max} - K_{min}$. Consequently, a set of models $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, .., \mathcal{M}_Q\}$ is obtained based on the different numbers of components.

The next step is model selection using the BIC. In [23], the BIC value of a model $\mathcal{M}_K$ trained over the dataset X with K components, BIC(X, M_K), is defined as follows:

$$BIC(X, M_k) \equiv -2 \log L(X, M_K) + v \log N, \quad (10)$$

where, L denotes the model likelihood; v denotes the degree of freedom of the model parameters; and N denotes the number of training data points. The model with the lowest BIC value is selected because it maximizes the log-likelihood [6]. Algorithm 1 presents the steps of the learning process.

Algorithm 1: Training the Optimal Model

Input: Initial Dataset D_{init}, Minimum Component Number K_{min}, and Maximum Component Number Kmax
Result: Best GMM Model
$BIC_{best} = \infty$
for K_{min} to K_{max} **do**
  $M_{candidate}$= EMTrain(D_{init}, k)
  $BIC_{candidate}$ = ComputeBIC (D_{init}, $M_{candidate}$)
  if $BIC_{candidate} < BIC_{best}$ **then**
    $M_{best} = M_{candidate}$
  end
end
return M_{best}

3.3. Concept Drift Detection

We propose Kernel Density Drift Detection (KD3) to detect the concept drift. KD3 is a window-based algorithm for concept drift detection. It works based on estimating the window density using the Kernel Density Estimation (KDE) or the Parzen's window [24]. The KDE is a non-parametric probability density estimator that automatically estimates the shape of the data density without assuming the underlying distribution. The concept drift can be detected by comparing the probability functions between these windows. The greater the variation between the windows, the more evidence obtained for the concept drifts. Aside from detecting concept drifts, KD3 also collects data for adaptation (D_{drift}) by identifying a warning zone when data begin to show indications of concept drift.

KD3 requires three hyperparameters, namely α, β, and h, which denote the margins for detecting the concept drift and accumulating the density distance and the window length, respectively. α is used to determine the threshold of the density variation in the concept drift, while β is employed to determine the threshold of the density variation in the warning zone. Therefore, α must be greater than β. KD3 accepts a set of likelihood windows z_c as input. z_c is the current likelihood window that contains a sequence of log-likelihood ℓ from the model prediction, $z_c = \{\ell_1, \ell_2, \ell_3, \dots, \ell_h\}$.

First, this algorithm aims to estimate the density (f_{kde}) of the current z_c and previous z_{c-1} windows. Let ℓ_n be the latest generated ℓ. Let z_c contain the h-latest ℓ from ℓ_n, $z_c \in [\ell_{n-h}, \ell_n]$, and let z_{c-1} contain the h-latest ℓ from ℓ_{n-h}, $z_{c-1} \in [\ell_{n-2h}, \ell_{n-h}]$. To detect a concept drift, the distance $\dot{d}_t$ between f_{kde} of z_c and z_{c-1} is computed using Eq. (11) within the bounds of $b1$ and $b2$. The bounds are computed based on the maximum and minimum values of the joined ℓ of z_c and z_{c-1}.

$$\dot{d}_t = \frac{1}{2} \int_{b1}^{b2} |f_{kde}(z_c) - f_{kde}(z_{c-1})| \, dz, \text{where} \quad (11)$$

$$z_c \in [\ell_{n-h}, \ell_n],\ z_{c-1} \in [\ell_{n-2h}, \ell_{n-h}],$$

$$b1 = \min(\ell_{n-2h}, \ell_n), b2 = \max(\ell_{n-2h}, \ell_n).$$

Finally, the algorithm compares $\dot{d}_t$ to α and β.hen $\dot{d}_t > \beta$. Suppose that the accumulative distance is equal to or greater than α. In that case, the algorithm sends the collected data to the model for adaptation. Figure 3 and Algorithm 2 illustrate the detailed KD3 process.

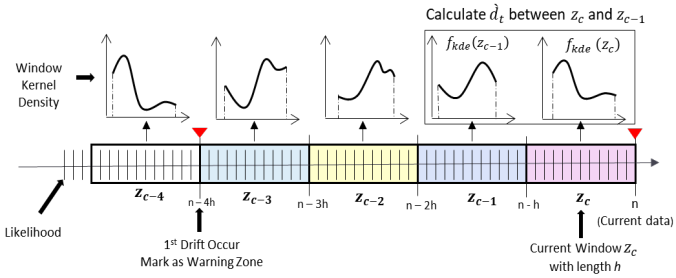


Figure 3: Illustration of the Kernel density drift detection (KD3) concept

Algorithm 2: Detecting the Concept Drift

Input: Set of t likelihood ℓ, drift margin α, warning margin β (α > β), window length h,
Result: Drift Concept Signal, Drifted Dataset (D_{drift})
$Window_1 = \ell[c-h:c]$;
$Window_2 = \ell[c-2h:c-h]$;
B_{min}, B_{max} = CalculateWindowBound($\ell[c-2h:c]$);
KDE_1 = EstimateKDE($Window_1$)
KDE_2 = EstimateKDE($Window_2$)
diff = distance(KDE_1, KDE_2, B_{min}, B_{max})
if (diff ≥ α) **then**
  resetWarningZoneData()
  return true, $[c-h:c]$
end
if (diff ≥ β) **then**
  accumulativeWarning += diff
  if (accumulativeWarning ≥ α) **then**
    resetWarningZoneData();

return true, $[c - h: c]$
end
return false, $[c - h: c]$
end
return false, null

3.4. Model Adaptation

Model adaptation aims to revise the current model upon newly incoming data that might contain new concepts or concept changes. The result of this adaptation is an adapted weighted mixture component that respects the original mixture.

The model adaptation method starts by training a new model M_{drift} from data drifts D_{drift} using Algorithm 1, and then combining the existing model $\mathcal{M}$. Consequently, the newly adapted model $\mathcal{M}$ accommodates the new concept represented by the components in $\mathcal{M}_{drift}$. The next step is to calculate the pairwise distance between the components in $\mathcal{M}$ using KL discrimination. The KL discrimination formula (Eq. (8)) enables us to set an upper bound on the discrimination of the mixture before and after the merging process. According to this formula, components with low weights means close to their variances, and similar covariance matrices are selected for merging. When two components are merged, the moment-preserving merging method [25] is used to preserve the mean and the covariance of the overall mixture (Eqs. (4)–(6)). Figure 4 illustrates the CMGM adaptation process.

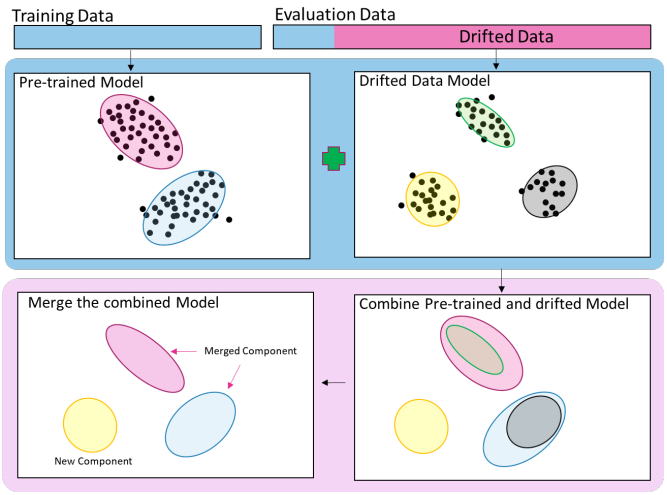


Figure 4: Illustration of the combine–merge Gaussian mixture model (CMGMM) adaptation process

As a result, the reduction process generates a set of merged models $\mathcal{M}_{merge.}$ To select the best $\mathcal{M}_{merge}$ model, the accumulative BIC is computed by combining sampling data from $\mathcal{M}_{curr}$. D_{drift} then computes the BIC value using Eq. (10). The smaller the value of the accumulative BIC, the better the newly adapted model.

Based on [7] and [11], the CMGMM tends to increase the number of components because it combines and merges them. This mechanism leads to an overfitting problem because the adaptation frequency increases due to the sensitive KD3 hyperparameter.

To maintain the compactness of the CMGMM and avoid overfitting, we design a strategy to merge statistically equivalent components into one component, then prune the inactive components. The inactive components are identified by the proximity of the ratio of w^2 and P^2 of the merged component to zero. In practice, components with w that are very close to zero are ignored by the model, whereas those with a large covariance tend to overlap with other components. Algorithm 3 presents in detail the steps of the proposed CMGMM-based method.

Algorithm 3: Model adaptation

Input: Current Model M, Drifted Dataset D_{drift}
Result: Adapted Model
M_{drift} = findBestGMM(D_{drift})
$M_{combine}$ = CombineGMMComponent($M_{drift,}$ M)
distanceMatrix = KLDissimalarity($M_{combine}$)
ds = D_{drift} + M.generateData()
$nComp_{min}$ = M.number_component
$nComp_{max}$ = $M_{combine}$.number_component
$BIC_{best} = \infty$
for targetComponent = $nComp_{min}$ **to** $nComp_{max}$
  M_{merge} = mergeComponent(target, distanceMatrix)
  if useComponentPrune **then**
    ComponentPrune(M_{merge})
  End
  $BIC_{candidate}$ = ComputeBIC (M_{merge}, ds)
  if $BIC_{candidate} < BIC_{best}$ **then**
    $M_{best} = M_{merge}$
  End
end
return M_{best}

4. Experiment

This section provides information about the datasets and experimental setup used in this study to train, optimize, and evaluate the proposed method.

4.1. Datasets

We used three types of datasets in this experiment, that is, training, optimization, and evaluation. The training dataset consisted of audio signals extracted with a 10-seconds window from 15 scenes in the TUT Acoustic Scenes 2017 [26] and TAU Urban Acoustic Scenes 2019 datasets [26]. The scenes were home, airport, beach, office, cafe, grocery store, bus, tram, metro, city center, residential area, street pedestrian, and shopping mall.

To simulate the concept drift in the datasets, the optimization and evaluation datasets were generated by overlay new additional event sounds from and UrbanSound8K datasets [27] and the BBC Sound Class Library [28]. When the sounds were added, the numbers of additions (1–10), positions in the time axis (0–9000 ms), and loudness (−20,0) of the sounds were randomly changed at random. In total, the dataset had 46 new event classes and 371 added event sounds.

We generated four concept drift types with three scenarios. The four types of concept drift are as follows:

- abrupt concept drift (AB), where ongoing concepts are replaced with new concepts at a particular time;
- gradual concept drift (GR), where new concepts are added to an ongoing concept at a particular time;

- recurring concept drift type 1 (R1), where an ongoing concept is replaced at a particular time with concepts that previously appeared; and
- recurring concept drift type 2 (R2), where concepts that previously appeared are added to an ongoing concept at a particular time.

Figure 5 illustrates the scenario generation. The three scenarios were designed to have different data distribution complexities. The scenario details are described below:

- Scenario 1 (Sc1): A unique event sounds from a specific event sounds is repeatedly introduced with a random number of times, gain, and timing. For example, in the airport scene, unique sounds representing the airplane sound, crowd background, and construction site are overlaid with a random number of times (1–10), position (0–9000 ms), and loudness (−20,0).
- Scenario 2 (Sc2): Several event sounds are randomly selected from a set of the same sound labels in Sc1 and added using the same rule as Sc1.
- Scenario3 (Sc3): This scenario differs from Sc1 and Sc2 in that event sounds coexist among scenes. For example, a set of rain sounds exists in other scenes (e.g., beach, city center, and forest paths). The methods of selection and addition are the same as those in Sc2.

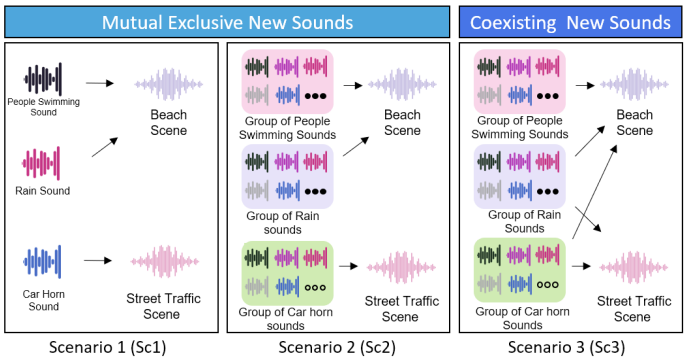


Figure 5: Concept drift scenario

Table 1 shows a list of event sounds that appear at scene types in every concept drift scenario. The mutually exclusive event sounds appear in all scenarios, but coexisting sounds only appear for Sc3.

Table 1: Setting of the novel sounds in scene audio for Sc1, Sc2, and Sc3

Scene	Mutually Exclusive Sounds in Sc1, Sc2, and Sc3	Additional Coexisting Sounds in Sc3
Airport	Helicopter, crowd, and construction site	Airplane, footsteps, and children playing
Beach	People swimming, footsteps on the sand, and rain	Teenage crowd, dog, and birds
Bus	Car horn, engine, and city car	Kitchenware, phone ringing, children playing, and teenage crowd
Café /restaurant	Washing machine, food mixer, and kitchenware	Phone ringing, children playing, and teenage crowd
City center	Sound of bird, ambulance, and wind	Footsteps, phone ringing, and children playing
Grocery store	Footsteps, children playing, and shopping cart	Vacuum cleaner, phone ringing, and footstep
Home	Frying, door, and vacuum cleaner	Clock and phone ringing
Metro station	Siren, road car, and thunder	Footsteps, crowd, and wind
Office	Typing, phone ringing, and sneeze	Broom, camera, and footsteps
Public square	People running, music, and airplane	Birds, rain, and teenagers talking
Residential area	Wind, camera, and cat	Birds, sneeze, and clock
Shopping mall	Clock, camera, and teenage crowd	Children playing, phone ringing, dog
Street pedestrian	Dog, bicycle, and bird	Footsteps and children playing
Street traffic	Motorcycle, horn, and train	Siren, airplane, and bell
Tram	Coughing, bell, and footsteps on the pavement	Teenage crowd and children playing

Finally, we have one training dataset, four optimization datasets, and 12 evaluation datasets in this experiment. Each training dataset contained 3,000 scene audio, while the optimization and evaluation dataset contained 15,000 scene audio. The datasets are available in our repository[2].

4.2. Experimental Setup

The CMGMM accuracy was evaluated under four concept drift types in three scenarios (i.e., Sc1, Sc2, and Sc3). The evaluations are performed using the two following approaches:

- Active CMGMM adaptation: In this approach, the CMGMM actively detects the concept drift using a certain method and only adapts the model when the concept drift is detected. In this study, we compared KD3 to ADWIN [9], HDDMA, HDDMW [10], and KSWIN [18].
- Passive CMGMM adaptation: In this approach, the CMGMM adapts as soon as a particular datum is received without requiring the explicit prior detection of the concept drift. Several adaptation cycle sizes were tested, that is, 25, 50, 100, 150, and 200.

5. Experiment Result

The experimental result of the proposed method are presented herein.

5.1. Hyperparameter Optimization

The first step in this experiment is the systematic optimization of the KD3 hyperparameter. We used the grid-search method using a combination of hyperparameters α from 0.1 to 0.001, β from 0.0001 to 0.000001, and $\hat{n}$ from 45 to 300. We prepared a particular

[2]https://bit.ly/CMGMM_Dataset

dataset for the KD3 hyperparameter optimization in four types of concept drift.

In [11], we reported that hyperparameters β and ɦ did not have a significant effect on accuracy. Therefore, during the initial step, we observed the performance change according to β and ɦ. Figure 6 shows the average accuracy in all concept drift types according to hyperparameters β and ɦ. In this experiment, the best β and ɦ were set at 0.0001 and 45, respectively.

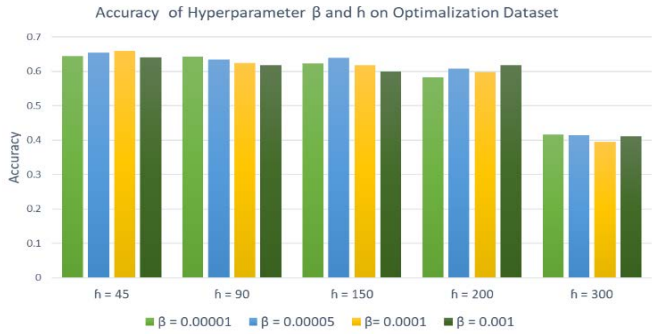


Figure 6: Result of hyperparameters β and ɦ in four types of concept drift

Table 2 lists the experimental results of α in the optimization dataset in four types of concept drift. Based on this experiment, every concept drift type has its respective hyperparameter α according to the concept drift characteristics. AB and GR have similar characteristics. There are no repeating concepts in the future; hence, a more sensitive concept drift detector than R1 and R2 is required.

Table 2: KD3 hyperparameter optimization result

Concept Drift Types	Hyperparameter α (β = 0.001, ɦ = 45)				
	α = 0.1	**α = 0.05**	**α = 0.01**	**α = 0.005**	**α = 0.001**
AB	0.6568	0.7113	0.7069	**0.7137**	0.7066
GR	0.6007	0.6173	**0.7050**	0.6922	0.7002
R1	**0.7332**	0.7232	0.7158	0.7054	0.6927
R2	**0.7268**	0.7133	0.7228	0.7090	0.6823
Overall	0.6793	0.6912	**0.7126**	0.7050	0.6950

In AB and GR, a sensitive hyperparameter α accelerated the update frequency. In the experimental result for these types of concept drifts, a high adaptation frequency reduced the loss received. However, a less-sensitive hyperparameter showed a better result in recurring concept drifts, where an old concept reappears in the future. A less-sensitive hyperparameter α provided the model with longer data compared to the sensitive hyperparameter.

We also selected the overall hyperparameter setting based on this experiment. The overall hyperparameter was selected from the best average performance of the hyperparameter optimization (α = 0.01, β = 0.001, and ɦ = 45). We used this hyperparameter for further CMGMM and KD3 evaluations.

5.2. Active Combine–Merge Gaussian Mixture Model (CMGMM) Adaptation Result

Table 3 presents the experimental results of the active CMGMM adaptation. In general, the model performance without a concept drift detector is low in all concept drift types and scenarios. The adaptations of the CMGMM on R1 and R2 showed better accuracy than those on AB and GR. On average, AB exhibited the lowest accuracy, while R2 showed the highest accuracy. This high accuracy on recurring was caused by the CMGMM being designed to preserve the old concept, even though the new concept is adapted in the model. Thus, the model can recognize the previously learned concept if it is repeated in the future.

The CMGMM experiment result depicted that KD3 outperformed other combinations in two of the four concept drift types in GR and R2. Meanwhile, ADWIN showed the best accuracy in AB. KSWIN demonstrated the best accuracy in R1, whereas HDDM was unsuitable for this case. Despite getting the highest overall accuracy score, the combination of the CMGMM and KD3 needed more frequent adaptations than ADWIN and KSWIN. In contrast, both HDDM-based methods showed worse performances compared to all others. HDDMA overdetected the concept drift in all concept drift types for more than 3000 times in GR.

The abovementioned results illustrated that the concept drift detector plays a vital role in the concept drift adaptation. The model performance decreased over time if the drift detector failed to detect or delay detecting or over detecting the concept drift.

Performance of the combine–merge Gaussian mixture model (CMGMM) and kernel density drift detection (KD3)

KD3 showed the best average accuracy of 0.6983 compared to ADWIN, KSWIN, and HDMM with 209 adaptations. This combination also showed its best results on R2 with 0.7321 accuracy, followed by R1 with 0.7373 accuracy, GR with 0.6999 accuracy, and AB with accuracy 0.6469. Furthermore, this combination was the most stable in all scenarios. The maximum performance decrements in AB, GR, R1, and R2 were 1.38%, 1.2%, 1.34%, and 0.94%, respectively.

Despite achieving a good performance in all concept drift types, the number of concept drifts detected in this combination was higher than ADWIN and KSWIN. The most significant number of adaptations occurred in AB. The disadvantage of a high number of adaptations is the higher computation time required to finish the task and possible overfitting. In this case, the higher numbers of adaptations in AB and GR are obtained because the concept constantly changes over time, and the learned concept becomes obsolete in the future; hence, the higher the adaptation, the better the performance.

Performance of the combine–merge Gaussian mixture model (CMGMM) and ADWIN

In general, the combination of CMGMM and ADWIN showed a good performance in every concept drift type, especially on the abrupt datasets, where this combination showed its best performance. The overall accuracy was 0.6371 with 83 times of adaptation. The overall accuracies of this combination in AB, GR, R1, and R2 were 0.6369, 0.6475, 0.6912, and 0.7169, respectively. Furthermore, this combination had the advantage of a small number of adaptations in all concept drift types. Hence, ADWIN showed an effective performance in using resources and had a reasonably good performance. This combination performed very well on Sc1, but showed a performance drop in Sc2 and Sc3. For

Table 3: Experiment result of the CMGMM with the concept drift detector

Concept Drift Detector		Accuracy				F1				Number of Concept Drift Detection			
		Sc1	Sc2	Sc3	Overall	Sc1	Sc2	Sc3	Overall	Sc1	Sc2	Sc3	Overall
AB	**ADWIN**	**0.6989**	**0.6525**	**0.6214**	**0.6369**	**0.713**	**0.6495**	**0.6353**	**0.6599**	83	89	85	85
	HDDM_A	0.4157	0.4476	0.4345	0.4326	0.4586	0.4804	0.477	0.472	3287	3350	3359	3332
	HDDM_W	0.4041	0.4455	0.4546	0.4347	0.4485	0.4902	0.5004	0.4797	489	413	409	437
	KD3*	0.6469	0.6359	0.6331	0.6386	0.6476	0.6467	0.6395	0.6446	207	236	220	221
	KSWIN	0.6611	0.6241	0.6345	0.6317	0.6722	0.6369	0.6411	0.6501	132	121	123	125
	Without Detector	0.4121	0.4054	0.4095	0.4090	0.4235	0.4125	0.4095	0.4151				
GR	ADWIN	0.6723	0.6341	0.6363	0.6475	0.6845	0.6506	0.6478	0.6609	92	81	83	85
	HDDM_A	0.4134	0.4463	0.4311	0.4302	0.4580	0.4946	0.4701	0.4742	3306	3308	3265	3293
	HDDM_W	0.4131	0.4497	0.4544	0.439	0.4524	0.487	0.4816	0.4736	0	409	401	270
	KD3*	**0.6999**	**0.6942**	**0.6879**	**0.694**	**0.7044**	**0.7004**	**0.6867**	**0.6971**	190	218	221	209
	KSWIN	0.6532	0.6241	0.618	0.6322	0.6631	0.6326	0.6204	0.6387	129	125	107	120
	Without Detector	0.3554	0.3524	0.3489	0.3522	0.3571	0.3542	0.3501	0.3538				
R1	ADWIN	0.7222	0.6948	0.6568	0.6912	0.7393	0.6981	0.6333	0.6902	78	83	87	82
	HDDM_A	0.4721	0.5204	0.4896	0.494	0.5206	0.5566	0.5392	0.5388	3154	3258	3201	3204
	HDDM_W	0.4847	0.4815	0.5068	0.491	0.5305	0.5215	0.5357	0.5292	0	662	647	436
	KD3*	0.7373	0.7334	0.7239	0.7315	0.7389	0.7341	0.7247	0.7325	208	201	204	204
	KSWIN	**0.7818**	**0.7351**	**0.7357**	**0.7508**	**0.7876**	**0.7404**	**0.7416**	**0.7565**	142	135	126	134
	Without Detector	0.4512	0.4458	0.4257	0.4409	0.4512	0.4458	0.4257	0.4409				
R2	ADWIN	0.7353	0.7066	0.7089	0.7169	0.7398	0.7131	0.7163	0.723	75	86	82	81
	HDDM_A	0.5919	0.6034	0.6259	0.607	0.6369	0.6398	0.6658	0.6475	3052	2957	3017	3008
	HDDM_W	0.5789	0.5996	0.5875	0.5886	0.6369	0.6398	0.6658	0.6475	759	604	543	635
	KD3*	**0.7321**	**0.7325**	**0.7227**	**0.7291**	**0.7352**	**0.7315**	**0.7221**	**0.7296**	195	207	205	202
	KSWIN	0.6914	0.6964	0.677	0.6882	0.6977	0.6976	0.6811	0.6921	156	143	141	146
	Without Detector	0.4254	0.4205	0.3985	0.4148	0.4279	0.4198	0.3968	0.4148				
Overall	ADWIN	**0.7072**	0.6720	0.6559	0.6731	**0.7191**	0.6778	0.6581	0.6835	82	84	84	83
	HDDM_A	0.4733	0.5044	0.4953	0.4910	0.5185	0.5428	0.5380	0.5331	3199	3218	3210	3209
	HDDM_W	0.4702	0.4941	0.5008	0.4883	0.5170	0.5346	0.5458	0.5325	312	522	500	444
	KD3*	0.7041	**0.6990**	**0.6919**	**0.6983**	0.7065	**0.7031**	**0.6932**	**0.7009**	200	215	212	209
	KSWIN	0.6969	0.6699	0.6663	0.6757	0.7051	0.6768	0.6710	0.6843	139	131	124	131
	Without Detector	0.4110	0.4060	0.3956	0.4042	0.4149	0.4080	0.3955	0.4061				

(*) Proposed method

example, the accuracies of AB, GR, R1, and R2 decreased in Sc3 by 4.64%, 3.82%, 2.74%, and 287%, respectively.

Performance of the combine–merge Gaussian mixture model (CMGMM) and Hoeffding's bounds-based method (HDDM)

In this experiment, both Hoeffding's inequality-based algorithms showed underperformance results for all concept drift types. Both algorithms were less effective in detecting the concept drift in this case. The overall accuracies of HDDMA in AB, GR, R1, and R2 were 0.4326, 0.4302, 0.494, and 0.607, respectively. The number of HDDMA adaptations exceeded 3000 times of adaptation. This high adaptation process was ineffective because the amount of trained data for each adaptation was too small. This condition led to an overfitting and decreased the model performance.

HDDM W also experienced the same problem. In some cases, HDDMA failed to detect the drift concepts, such as GR, R1, and R2 in Sc1. The overall accuracies of this HDDMW in AB, GR, R1, and R2 were 0.4347, 0.439, 0.491, and 0.5886, respectively.

Performance of the combine–merge Gaussian mixture model (CMGMM) and KSWIN

The combination of CMGMM and KSWIN showed the best performance in R1, with an overall accuracy of 0.750 with 134 adaptations. The accuracies of this combination in AB, GR, R1, and R2 were 0.6317, 0.6322, 0.7508, and 0.6882, respectively. On average, KSWIN required eight to nine adaptations per scene in all dataset types. This algorithm seems able to detect occurring changes in data and supports the concept drift handling process with good indicators at a given time.

5.3. Passive Combine–Merge Gaussian Mixture Model (CMGMM) Adaptation Result

Table 4 lists the experimental results of the passive CMGMM adaptation. The best performance in AB, GR, and R1 was obtained with a cycle size of 50, and that in R2 was obtained with a cycle size of 100. The best accuracies of AB, GR, R1, and R2 were 0.7152, 0.7139, 0.7323, and 0.7155, respectively.

Table 4: The experiment result of CMGMM without concept drift detector

Concept Drift Types	Cycle Size	Accuracy			
		Sc1	Sc2	Sc3	Average
AB	25	0.6290	0.6236	0.6256	0.6260
	50	**0.7122**	**0.7159**	**0.7177**	**0.7152**
	100	0.6285	0.5925	0.5955	0.6055
	150	0.6361	0.5776	0.5851	0.5996
	200	0.5580	0.5059	0.5119	0.5252
GR	25	0.6294	0.6232	0.6004	0.6176
	50	**0.7133**	**0.7169**	**0.7115**	**0.7139**
	100	0.6173	0.5887	0.6089	0.6049
	150	0.6371	0.5818	0.5797	0.5995
	200	0.5334	0.5094	0.5076	0.5168
R1	25	0.6186	0.5799	0.5608	0.5864
	50	**0.7235**	**0.7320**	**0.7416**	**0.7323**
	100	0.7211	0.7105	0.6988	0.7101
	150	0.7332	0.7086	0.7120	0.7179
	200	0.6639	0.7018	0.7146	0.6934
R2	25	0.5133	0.5444	0.5669	0.5415
	50	0.7396	0.6991	0.7011	0.7132
	100	**0.7431**	**0.7012**	**0.7023**	**0.7155**
	150	0.6865	0.6639	0.6803	0.6769
	200	0.6502	0.6011	0.6089	0.6200

Similar to active adaptation, R1 and R2 showed good performances compared to AB and GR, but better performances in passive adaptation. Although R1 exhibited the best adaptation at cycle size 50, it also showed good result at cycle sizes 100 and 150. If you consider the time and the computing resources used, then cycle sizes 100 and 150 are recommended.

In passive adaptation, the cycle size is vital in achieving a good performance. This cycle size determines the adequacy of the data for adaptation. If the cycle size is too short, the number of data adapted is small, leading to overfitting problems.

5.4. Suggestions and Limitations

The experiment results showed that the combination of CMGMM and KD3 has a higher number of adaptations compared to that of ADWIN and KSWIN due to the selection of KD3 hyperparameters that are sensitive to accommodating GR and AB. When the number of adaptations is increased, then in certain cases, such as GR and AB, the accuracy is improved, albeit with a higher computing cost. The advantage of KD3 compared to the other methods is that it could be applied in multi-dimensional probability distribution; hence, it is more flexible to apply in other models and cases.

In cases where the time or location of the concept drift can be predicted, the use of a passive adaptation strategy is more beneficial and has a lower computational cost than the active strategy. However, if the adaptation cycle is too far from the concept drift, then the model performance will decrease over time.

6. Conclusion

This paper presented KD3 hyperparameter optimization and the CMGMM performance evaluation in four types of concept drift. All concept drift types had optimum hyperparameter configurations. AB and GR showed similar patterns that required a smaller α value than the recurring concept drift. In this type of concept drift, concepts continuously appear; hence, a sensitive drift detector is needed to update the model early or have a high-frequency model adaptation. However, in recurring concepts, the new concept may be repeated in the future; thus, a lower-frequency adaptation shows a better performance.

The evaluation results demonstrated that the proposed algorithms work well in detecting and adapting to four types of concept drift and three scenarios. Overall, the CMGMM works better in R1 and R2 than in AB and GR because it is designed to preserve old concepts to preserve previously learned knowledge. Component pruning is only performed on components with a minimal impact on the prediction.

In the active adaptation strategy, the proposed combined method of CMGMM and KD3 outperformed two of four other combination methods in GR and R1. These methods showed the most stable performance among Sc1, Sc2, and Sc3 in all concept drift types. Furthermore, ADWIN showed the best results on AB. This algorithm is efficient in computing resources, but less stable in more challenging scenarios, such as Sc2 and Sc3. KSWIN showed the best results in recurring drift and good performance for all concept drift types. HDDM overdetected or failed to detect the concept drift and was not suitable in this case.

In the passive adaptation strategy, a short or long adaptation cycle could reduce the performance, and a short cycle size makes the adaptation less effective. A short cycle could disrupt the model component because the model is trained with insufficient data, leading to an underfitting problem.

In AB and GR, the model requires a high-frequency adaptation to maintain the performance; therefore, sensitive hyperparameters or a short cycle size is required. Based on the experimental results, the passive method is more suitable in these concept drift types. However, the computational costs are higher than those of the active method. Furthermore, in recurring drift scenarios, a less-sensitive hyperparameter, or a moderate cycle size is needed. The active method is recommended for recurring drift types considering the computation cost and better accuracy in experiments.

As part of our future work, we plan to improve the proposed KD3 concept drift detection algorithm to achieve adaptive weight, reduce the kernel density computation, and optimize the number of data points considered when detecting the concept drift and decreasing the computation time.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP20K12079.

References

[1] R. Elwell, R. Polikar, “Incremental Learning of Concept Drift in Nonstationary Environments,” IEEE Transactions on Neural Networks, **22**(10), 1517–1531, 2011, doi:10.1109/TNN.2011.2160459.

[2] S. Ntalampiras, “Automatic analysis of audiostreams in the concept drift environment,” in 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), IEEE, 2016, doi:10.1109/MLSP.2016.7738905.

[3] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, A. Bouchachia, "A Survey on Concept Drift Adaptation," ACM Computing Surveys, **46**(4), 1–37, 2014, doi:10.1145/2523813.
[4] T.R. Hoens, R. Polikar, N. V. Chawla, "Learning from Streaming Data with Concept Drift and Imbalance: An Overview," Progress in Artificial Intelligence, **1**(1), 89–101, 2012, doi:10.1007/s13748-011-0008-0.
[5] I. Žliobaitė, "Learning under Concept Drift: an Overview," 1–36, 2010, doi:10.1002/sam.
[6] C. Chen, C. Wang, J. Hou, M. Qi, J. Dai, Y. Zhang, P. Zhang, "Improving Accuracy of Evolving GMM under GPGPU-Friendly Block-Evolutionary Pattern," International Journal of Pattern Recognition and Artificial Intelligence, **34**(3), 1–34, 2020, doi:10.1142/S0218001420500068.
[7] I.D. Id, M. Abe, S. Hara, "Concept Drift Adaptation for Acoustic Scene Classifier based on Gaussian Mixture Model," in IEEE Region 10 Annual International Conference, Proceedings/TENCON, IEEE, Osaka: 450–455, 2020, doi:10.1109/TENCON50793.2020.9293766.
[8] J.M. Acevedo-Valle, K. Trejo, C. Angulo, "Multivariate regression with incremental learning of Gaussian mixture models," Frontiers in Artificial Intelligence and Applications, **300**, 196–205, 2017, doi:10.3233/978-1-61499-806-8-196.
[9] A. Bifet, R. Gavaldà, "Learning from Time-changing Data with Adaptive Windowing," Proceedings of the 7th SIAM International Conference on Data Mining, (April), 443–448, 2007, doi:10.1137/1.9781611972771.42.
[10] I. Frías-Blanco, J. Del Campo-Ávila, G. Ramos-Jiménez, R. Morales-Bueno, A. Ortiz-Díaz, Y. Caballero-Mota, "Online and Non-Parametric Drift Detection Methods Based on Hoeffding's Bounds," IEEE Transactions on Knowledge and Data Engineering, **27**(3), 810–823, 2015, doi:10.1109/TKDE.2014.2345382.
[11] I.D. Id, M. Abe, S. Hara, "Evaluation of Concept Drift Adaptation for Acoustic Scene Classifier Based on Kernel Density Drift Detection and Combine Merge Gaussian Mixture Model," in 2021 Spring meeting of the Acoustical Society of Japan, 2021.
[12] J. Montiel, J. Read, A. Bifet, T. Abdessalem, "Scikit-multiflow: A Multi-output Streaming Framework," Journal of Machine Learning Research, **19**, 1–5, 2018, doi:10.5555/3291125.3309634.
[13] G.I. Webb, R. Hyde, H. Cao, H.L. Nguyen, F. Petitjean, "Characterizing Concept Drift," Data Mining and Knowledge Discovery, **30**(4), 964–994, 2016, doi:10.1007/s10618-015-0448-4.
[14] T. Zhang, J. Liang, B. Ding, "Acoustic scene classification using deep CNN with fine-resolution feature," Expert Systems with Applications, **143**(1), 113067, 2020, doi:10.1016/j.eswa.2019.113067.
[15] A. Tsymbal, "The Problem of Concept Drift: Definitions and Related Work," 2004.
[16] E.S. Page, "Continuous Inspection Schemes," Biometrika, **41**(1/2), 100, 1954, doi:10.2307/2333009.
[17] Y. Yuan, Z. Wang, W. Wang, "Unsupervised concept drift detection based on multi-scale slide windows," Ad Hoc Networks, **111**, 102325, 2021, doi:10.1016/j.adhoc.2020.102325.
[18] C. Raab, M. Heusinger, F.M. Schleif, "Reactive Soft Prototype Computing for Concept Drift Streams," Neurocomputing, 2020, doi:10.1016/j.neucom.2019.11.111.
[19] A.R. Runnalls, "Kullback-Leibler Approach to Gaussian Mixture Reduction," IEEE Transactions on Aerospace and Electronic Systems, **43**(3), 989–999, 2007, doi:10.1109/TAES.2007.4383588.
[20] A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society. Series B (Methodological), **39**(1), 1–38, 1977.
[21] A.D.R. McQuarrie, C.-L. Tsai, Regression and Time Series Model Selection, WORLD SCIENTIFIC, 1998, doi:10.1142/3573.
[22] P. Pal Singh, "An Approach to Extract Feature using MFCC," IOSR Journal of Engineering, **4**(8), 21–25, 2014, doi:10.9790/3021-04812125.
[23] C. Fraley, "How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis," The Computer Journal, **41**(8), 578–588, 1998, doi:10.1093/comjnl/41.8.578.
[24] E. Parzen, "On Estimation of a Probability Density Function and Mode," The Annals of Mathematical Statistics, **33**(3), 1065–1076, 1962, doi:10.1214/aoms/1177704472.
[25] J.L. Williams, P.S. Maybeck, "Cost-function-based Gaussian mixture reduction for target tracking," Proceedings of the 6th International Conference on Information Fusion, **2**, 1047–1054, 2003, doi:10.1109/ICIF.2003.177354.
[26] A. Mesaros, T. Heittola, T. Virtanen, TUT Acoustic scenes 2017, Development dataset, 2017, doi:10.5281/ZENODO.400515.
[27] J. Salamon, C. Jacoby, J.P. Bello, "A Dataset and Taxonomy for Urban Sound Research," MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia, (November), 1041–1044, 2014, doi:10.1145/2647868.2655045.
[28] BBC, BBC Sound Effects, https://sound-effects.bbcrewind.co.uk/search, Mar. 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 175-184 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

A Hybrid NMF-AttLSTM Method for Short-term Traffic Flow Prediction

Dejun Chen[1], Congcong Xiong[*,1], Li Guo[2], Ming Zhong[3]

[1]*School of Information Engineering, WuHan University of Technology, Wuhan, 430079, China*

[2]*Department of Information Engineering, Wuhan Electric Power Technical College, Wuhan, 430079, China*

[3]*Intelligent Transport Systems Research Center, WuHan University of Technology, Wuhan, 430079, China*

ARTICLE INFO

Article history:
Received: 13 January, 2021
Accepted: 14 February, 2021
Online: 10 March, 2021

Keywords:
ITS
Traffic Flow Prediction
Non-negative Matrix Factorization
Attention Mechanism
LSTM

ABSTRACT

In view of the current short-term traffic flow prediction methods that fail to fully consider the spatial correlation of traffic flow, and fail to make full use of historical data features, resulting in low prediction accuracy and poor robustness. Therefore, in paper, combining Non-negative Matrix Factorization (NMF) and LSTM model Based on Attention Mechanism (AttLSTM), the NMF-AttLSTM traffic flow prediction algorithm is proposed. The NMF algorithm is used to extract the spatial characteristics of traffic flow and reduce the data dimension. The attention mechanism can extract more valuable features from a long sequence of historical data. First, select high-correlation upstream and downstream roads, use NMF algorithm to perform dimensionality reduction and to extract historical data features of these roads, then combine with the historical data of this road as input. Finally, use the AttLSTM model to predict. Experiments with the PeMS public data set and Wuhan core roads data show that the method has higher prediction accuracy than other prediction models and is an effective traffic flow prediction method.

1 Introduction

Short-term traffic flow forecasting in Intelligent Transportation Systems (ITSs) has always been an important component.[1] Based on current and past traffic flow data, it can predict traffic flow ranging from a few minutes to a few hours in the future, providing a basis for decision-making for traffic dispatch and planning. Traditional traffic flow prediction models can be divided into two types, parametric and non-parametric. Parametric models include some time series models, such as Exponential Smoothing (ES), Autoregressive Integrated Moving Average (ARIMA)[2] model and Kalman Filtering model, etc. Non-parametric models include K-Nearest Neighbor (K-NN) method, Artificial Neural Network (ANN), and SupportVector Machine (SVM) [3]–[5] etc. However, because traffic flow data is affected by various environmental factors and has the characteristics of non-linearity and suddenness, the above models are affected by random factors and become fragile, which makes it difficult for these models to obtain high prediction accuracy. In recent years, deep learning has developed rapidly, and its applications have penetrated into all walks of life. Since the deep neural network used in deep learning can simulate deep complex nonlinear relationships through hierarchical feature representation, it can extract hidden features in the data. Therefore, in line with the characteristics of traffic flow data, deep learning methods can be used to predict traffic flow.

Recurrent Neural Network (RNN) is a type of deep learning model that can effectively predict time series data. In recent years, many researchers have conducted in-depth research on its application and achieved many results. However, with the increase of the input time series length, the traditional RNN will have the problem of gradient explosion and gradient disappearance. It can only use the information on time steps close to itself. In [6], the author uses Long Short Term Memory Network (LSTM), which can effectively overcome the problems of gradient disappearance existing in RNN and capture the characteristics of time series in a longer time span. There are also many variants of LSTM networks. Gated Recurrent Unit (GRU) can be seen as a simplification of the LSTM network. It combines the forgetgate and input gate in the LSTM network into one gate unit, which simplifies the structure of the model. In [7], the author compared the performance of LSTM model and GRU model on short-term traffic flow. Although the above deep learning models have achieved good results on prediction tasks, they do not consider the spatial correlation between traffic flows, and the spatial

[*]Corresponding Author: Congcong Xiong, WuHan University of Technology, window3cc@qq.com.

[1]This paper is an extension of work originally presented in the 10th International Conference on Information Science and Technology (ICIST), 2020 [1].

topology of traffic sections has an important impact on traffic flow. Therefore, considering the temporal and spatial characteristics of traffic flow, [8] combines the Autoencoder and the LSTM model, and uses the Autoencoder to obtain the traffic flow characteristics of adjacent locations. The adjacent locations represent the upstream and downstream locations of the current location. The LSTM model is used to predict the traffic flow at the current location. However, using the autoencoder model to extract features requires pre-training of the data, which is time-consuming, and the LSTM model does not perform well when the historical input sequence is long.

In [9], the author uses the attention mechanism to perform translation and alignment at the same time on machine translation tasks. At present, it is believed that the paper applies the attention mechanism to the NLP for the first time. In [10], the author points out that attention mechanisms have been successfully combined with existing models in machine translation. In other areas, attention mechanisms have also been applied, for instance, [11] based on LSTM neural network combined with attention mechanism for visual analysis of human behavior. The paper [12] applied neural network based on attention mechanism to medical diagnosis. The attention mechanism can be understood as, for a long time series data, after calculating the correlation between each element and the current element, assigning different weights to these elements, used to measure their impact on the current element. We focus on those elements that have a greater impact on the current element.

Traffic flow data with typical time series characteristics can be predicted by the RNN model. In recent years, researchers have used this model to predict traffic flow. In view of the above-mentioned improvements of the RNN model in recent years, in order to further improve its predictive ability, based on the above research, a hybrid model traffic flow prediction method based on NMF-AttLSTM will be proposed. The method first uses the weekly average spatial correlation coefficient based on the Pearson correlation coefficient to filter the upstream and downstream roads, and uses the NMF algorithm to reduce the dimensionality of the historical data matrix composed of the filtered upstream and downstream data to obtain the reduced feature matrix. Compared with the Autoencoder, the NMF algorithm is fast, does not require a pre-training process, and is suitable for processing high-dimensional data. The advantage of using the LSTM model based on the attention mechanism is that for long historical input sequence data, it can measure the similarity between each historical data and the current observation, and determine the selection of features in the historical data according to the similarity. Compared with the LSTM model, its feature extraction method is more reasonable. The experimental results in this paper are based on the PeMS public data set [13] and the real traffic data of the core area of Wuhan. The effectiveness of the proposed method is verified by experiments on these two data sets.

The main work of this paper is as follows:

- Propose a method to extract the temporal and spatial characteristics of traffic flow, combine the upstream and downstream features of the current road segment, and use a non-negative matrix factorization algorithm to reduce the dimension of the data and extract the feature matrix.
- The attention mechanism is added to the LSTM model to enhance the predictive ability of the LSTM model.
- The NMF-AttLSTM model was verified using public data sets and real traffic data in Wuhan, and the experimental results proved the effectiveness of the method.

The rest of this paper is organized as follows. In the section II, the NMF algorithm is introduced first, and then the process of adding the attention mechanism to the LSTM model is given. Finally, the traffic flow prediction framework based on the NMF-AttLSTM model is proposed. In section III, Experiments were performed on two data sets and the experimental results were evaluated. The section IV summarizes the whole paper.

2 NMF-AttLSTM Traffic Flow Prediction Framework

2.1 Non-negative Matrix Factorization Algorithm

The traffic flow data is time series data, and the change patterns of the traffic flow of adjacent roads in the spatial position have high similarity, which implies the common characteristics of the traffic flow changes. In order to extract this specific feature, we can construct the historical traffic data on all relevant roads into a matrix form, As shown in the following formula, suppose there are related roads in total, and the time series length is t, and f_{ij} represents the traffic value of the road number i at time j.

$$F_{mt} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1t} \\ f_{21} & f_{22} & \cdots & f_{2t} \\ \cdots & \cdots & f_{ij} & \cdots \\ f_{m1} & f_{m2} & \cdots & f_{mt} \end{bmatrix}, i \in [1, m] \;\; j \in [1, t] \tag{1}$$

The NMF algorithm is used to analyze and decompose matrix data. Its definition is that for any given two-dimensional non-negative matrix, it can be decomposed into two non-negative submatrices to satisfy the multiplication of these two non-negative matrices to get the original matrix. As shown in (2).

$$F_{mt} = W_{mk} \times H_{kt} \tag{2}$$

Among them, W is called the basis matrix, and H is called the coefficient matrix or the characteristic matrix. In paper, the H matrix is used as the feature matrix of the original data to achieve the goal of dimensionality reduction and feature extraction. Since both W and H matrices are unknown, they need to be solved by algorithms. Although there are many ways to reduce the matrix dimension, we need to ensure that the values in W and H are non-negative, because all traffic flow values need to be non-negative.

Therefore, the NMF(F, k) algorithm is used to solve the characteristic matrix. At this time, the problem becomes that, given the original non-negative matrix and parameters, matrix factorization can be formulated as a non-negative factorization minimization problem[14]. As shown in the (3):

$$f(W, H) = \underset{\{W,H\}}{arg\,\min} \|F - W \cdot H\|^2 \tag{3}$$

The NMF(F, k) algorithm first needs to define a cost function to quantify the degree of approximation between F and WH. One way is to use the square of the Euclidean distance between F and WH, as follows:

$$\|F - W \cdot H\|^2 = \sum_{ij} \left(F_{ij} - (W \cdot H)_{ij}\right)^2 \tag{4}$$

Another way is to use Kullback-Leibler divergence (KL divergence) as the cost function, as shown in the (5):

$$D(F\|(W \cdot H)) = \sum_{ij} \left(F_{ij} \log \frac{F_{ij}}{(W \cdot H)_{ij}} - F_{ij} + (W \cdot H)_{ij}\right) \tag{5}$$

Since this paper needs to use the NMF algorithm to extract the flow characteristics of the relevant roads, it pays more attention to the distribution characteristics of the flow instead of the absolute difference in the flow value. And the KL divergence is often used to measure the similarity of the two distributions. Therefore, the KL divergence is selected as the cost function, and the problem of solving W and H is transformed into: Under the constraint condition $W, H \geq 0$, with W and H as parameters, minimize the cost function $D(F\|WH)$.

Use the multiplication update rule to iteratively update the parameters W and H. The update equation is as shown in the (6):

$$W_{ia} \leftarrow W_{ia} \frac{\frac{\sum_\mu H_{a\mu} V_{i\mu}}{(WH)_{i\mu}}}{\sum_v H_{av}} \quad H_{a\mu} \leftarrow H_{a\mu} \frac{\frac{\sum_i W_{ia} V_{i\mu}}{(WH)_{i\mu}}}{\sum_k W_{ka}} \tag{6}$$

In the above update rule, when W and H are at the stagnation point of the divergence formula, the divergence will no longer be updated. The proof of the convergence of the above update equation is given in [14]. Based on the NMF(F, k) algorithm, the basic matrix W and the corresponding feature matrix H are obtained. In the NMF-AttLSTM model, we mainly input H as a feature into the AttLSTM model to provide upstream and downstream spatial features for traffic flow prediction.

2.2 Construction of Attention-LSTM Model

LSTM is a recurrent neural network, as shown in Figure 1. RNN is usually used to deal with time series problems. RNN uses a series of historical data as input, extracts features through non-linear functions, and stores the features extracted from each layer to provide feature information for subsequent calculations. Circulation makes information continue to pass to the next layer.

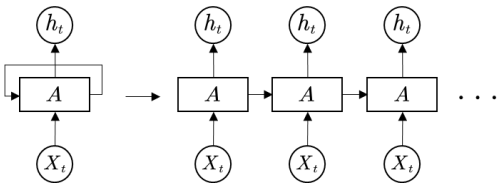


Figure 1: Recurrent neural network (RNN)

However, compared with the conventional RNN, the structure of this repeated module A of LSTM is more complicated, as shown in Figure 2.

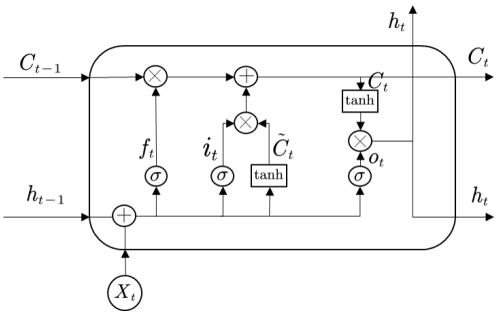


Figure 2: The structure of LSTM cell

This module consists of three parts, the forgotten gate, the input gate and the output gate. σ is the Sigmoid function, output a value between 0 and 1, describing how much of each part can pass.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{7}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{8}$$

$$\tilde{C}_t = \tan\mathrm{h}(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{9}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{10}$$

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t \cdot \tan\mathrm{h}(C_t) \tag{12}$$

Among them, f_t determines how much information we want to discard. i_t determines how much new information we should add. o_t determines how much information we want to output. x_t is the input at time t. h_{t-1} is the output of the previous gate, W_f, W_i, W_c and W_o is the weight, b_i, b_f, b_c and b_o is the bias, C_{t-1} is the cell state at the previous moment, C_t is the cell state at the current moment. Since the model is difficult to learn information at a time far from the current time, and it may be important for the current value. To overcome the weakness, we tried to add an attention layer to the LSTM network. Referring to the attention implementation steps of [10], we can apply it to the LSTM model. As shown in Figure 3, the attention layer is added to the LSTM model.

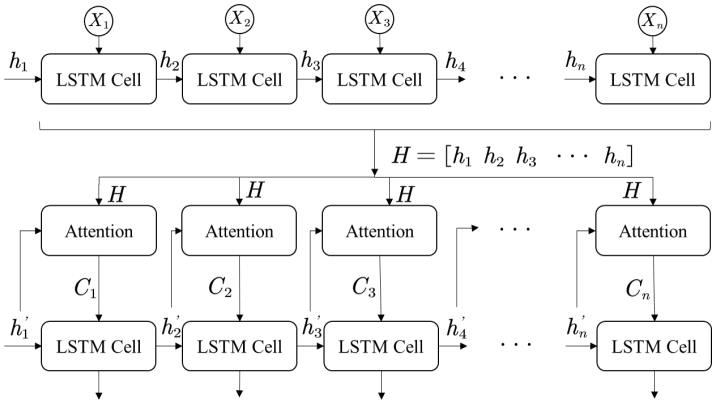


Figure 3: The process of adding an attention layer to the LSTM model

Among them, $X_i, i \in (1, n)$ is the input, h_i is the intermediate output result of each cell, all of h_i are input into each attention model as H, and the elements of the next layer h_i' are used as H_i'

to calculate the similarity and weight coefficient, and finally get the attention coefficient. The specific attention model is shown in Figure 4.

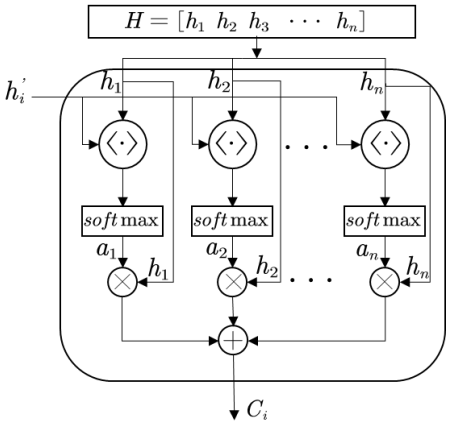


Figure 4: The internal structure of attention model

where $\langle\cdot\rangle$ represents the dot product operation, which is used to calculate the similarity between the current element and the intermediate output result in the previous layer, and then normalized by the softmax function to obtain the corresponding weight coefficient a_i. Finally, a weighted summation operation is performed to obtain the Attention value C_i. The equations used in the attention layer are as follows:

$$H = [h_1\ h_2\ \cdots\ h_n] \tag{13}$$

$$sim_i = h_i^{'} \cdot H^T \tag{14}$$

$$a_i = \frac{e^{sim_i}}{\sum_{j=1}^{L_h} e^{sim_j}} \tag{15}$$

$$C_i = \sum_{j=1}^{L_h} a_i \cdot h_j \tag{16}$$

In (14), it uses vector and to calculate similarity to obtain weights, (15) uses the softmax function to normalize the weight, (16) uses the normalized weight a_i and h_i weighted sum. The result of weighted summation is the attention weight value C_i. The implementation of the Attention layer is to retain the intermediate output results of the input sequence by the LSTM encoder, and then calculate the similarity between the intermediate output results of the previous layer and the current output to obtain the weight factor, and finally obtain the attention coefficient. Through the Attention mechanism, it is possible to find out the traffic flow in the past period that is most relevant to the forecast period in the time-based long traffic flow sequence, which improves the ability of the original model to predict a longer sequence traffic flow.

2.3 NMF-AttLSTM

In general, the traffic flow of a road is not only related to its own historical flow, but the flow of its upstream and downstream roads in space also has an important impact on the changes in its own flow. Full consideration of the flow changes in the upstream and downstream roads is of great significance to the current road flow prediction. The definition of the upstream and downstream roads in the spatial position of a road is as follows: based on the direction of vehicle movement, all roads connected to the entrance of the road are called upstream roads, and all roads connected to the exit of the road are called downstream roads, As shown in Figure 5.

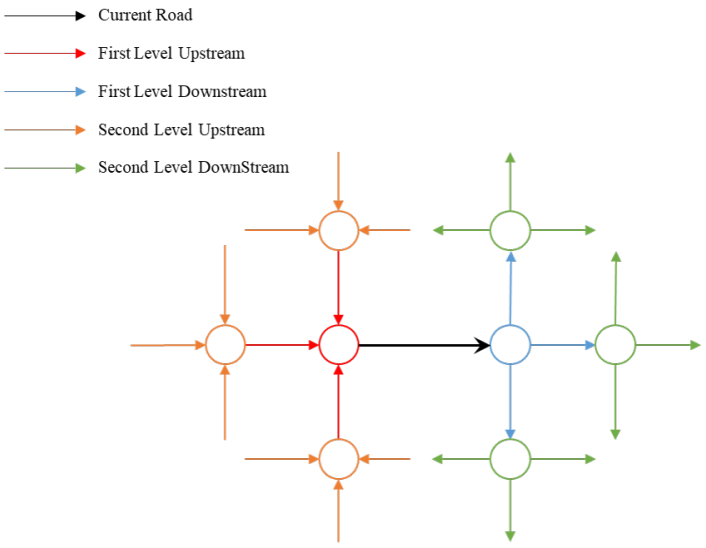


Figure 5: Schematic diagram of the first and second level upstream and downstream roads

As shown in Figure 5, we consider the first and second level upstream and downstream flow data of the current road. However, in the first and second level upstream and downstream roads, it cannot be guaranteed that all the road traffic has a high correlation with the current road flow. The similarity calculation method is used to calculate the similarity of the flow of all relevant roads with the current road, and the roads with weaker correlation are eliminated accordingly. The addition of the flow characteristics of these roads is not conducive to the improvement of the prediction accuracy. This paper proposes to use the weekly average Pearson correlation coefficient to measure the similarity of the flow changes between the current road and the related road. The Pearson correlation coefficient is generally used to analyze the similarity between two ordered vectors of equal length. The equation is as shown in (17):

$$R(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \tag{17}$$

Where E is the mathematical expectation. As the traffic flow changes have a certain periodicity, for example, the weekly traffic changes of a certain road section are similar, as shown in Figure 6:

It shows that a week's traffic flow change can represent the overall traffic flow change of the road section, and the average of the Pearson correlation coefficient of the flow data of a week can be used as the similarity value between the road section and other road sections.

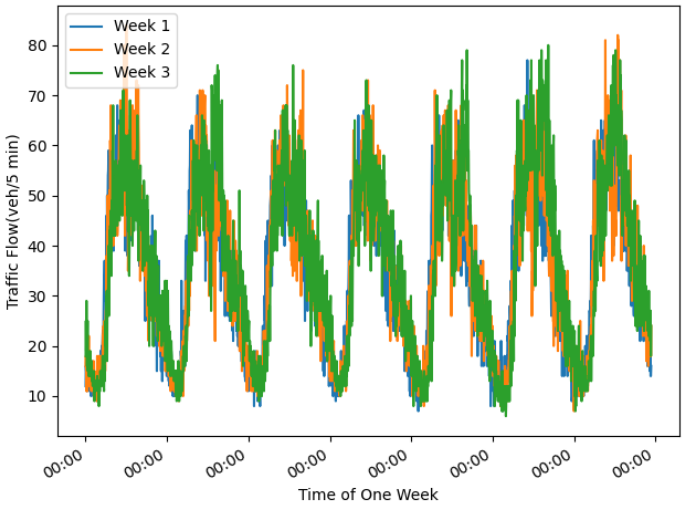


Figure 6: Traffic flows in link 10483 of Wuhan for three weeks

$$R_{week} = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{18}$$

Then the weekly average correlation coefficient can be expressed as the equation (18), n is the number of days, the value is set to 7, and R_i is the day correlation coefficient. Figure 7 is a flowchart of the overall architecture of NMF-AttLSTM.

3 Experimental and Analysis

3.1 Data Sources

(1) PeMS public data set

The Caltrans Performance Measurement System (PeMS) is a professional traffic flow data collection system. Approximately 15,000 detectors located in major cities in California collect traffic data every day. This paper uses a total of 32 relevant road traffic data from January to March 2017. The traffic data set is counted as 5 minutes, which means that there will be 12 traffic data points per hour. For some missing and abnormal data, the method of moving average is used to fill and correct the abnormal value. In order to obtain the similarity between each road section, a week of data was selected to calculate the weekly average Pearson correlation coefficient, the remaining data was divided into training set and test set, and the effectiveness of the model was evaluated on the test set.

(2) Wuhan City Traffic Data

To verify the prediction effect of the model on the actual road section, the road section 14394 in the core area of Wuhan was selected as the research section, and a total of 30 related road sections in the upstream and downstream were extracted. The relationship between the upstream and downstream is shown in Table 1.Including road traffic data from November to December 2020. The data collection granularity is 5 minutes, which means that 288 data will be collected for each road section every day. Same as the PeMS data set, the abnormal and missing data are also corrected accordingly. And use one week of data to calculate the weekly spatial correlation coefficient, to select high-correlation road sections to extract spatial features. The actual spatial relationship between section 14394 and its upstream and downstream sections is displayed in QGIS[15] by using OpenStreetMap, as shown in Figure 8.

Table 1: Link number 14394 corresponds to the link number of the upstream and downstream sections

Current Road	Upstream link	Downstream link
14394	14310,14318,14319, 14322,14388,14390, 14392,14393 ,14395, 14397,14398,14742, 14743,32433,32434, 14400,14405	14199,14197,14198, 14200,14313,14317, 14314,14315,14316, 14222,14223,14195, 32344

3.2 Experimental Setup

A sliding time window is used to construct a data set of historical data of this road section and its related road sections. The specific process is shown in Figure 9, where S_t represents the characteristic matrix of historical time series data. Assuming that the input sequence length is 10, the prediction lag time is 15 minutes. Since the data is at a granularity of 5 minutes, it moves backward two time units in turn to extract the current road flow as the label. Now the data set is $row_1 : (x_1 : s_1\ s_{10},\ f_1 : s_{13})$, $row_2 : (x_2 : s_2\ s_{11},\ f_2 : s_{14})$, ... By analogy, other data set with different input sequence lengths are constructed.

Due to the many parameters of the proposed prediction framework model, the prediction results are also affected. The important parameters are shown in Table 2. Through related experiments, the most suitable parameters are selected as the benchmark parameters for the experiment.

The model in Figure 9 is used for comparative analysis with the NMF-AttLSTM model.

- SVR[16]: Support Vector Regression (SVR) is a method that uses support vector machines (SVM) to solve regression problems, and is often used for time series forecasting. And use RBF as the kernel function.
- LSTM[6]: Which has good predictive ability for time series problems.
- AE-LSTM[8]: Use AutoEncoder to obtain the characteristics of the upstream and downstream traffic flow data of the internal relationship of the traffic flow, and then use the acquired feature data and historical data to predict through the LSTM network.
- AttLSTM[1]: Use the attention mechanism to extract features, add a layer of AttentionDecoder to the LSTM network, so that the model can extract more valuable information.

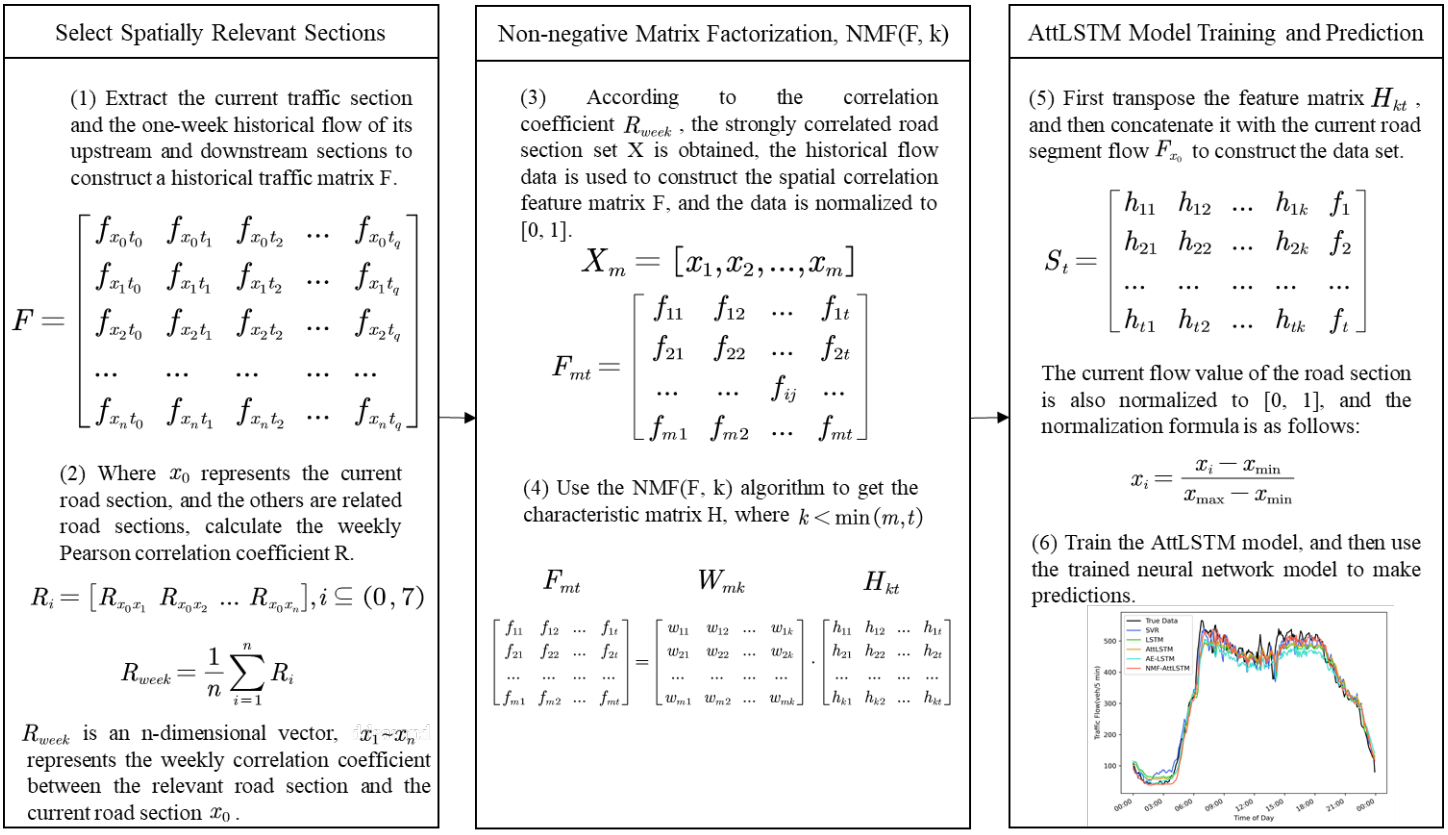


Figure 7: The flow chart of the traffic flow prediction architecture of the NMF-AttLSTM method. The entire prediction process is divided into selection of spatially relevant sections, Non-negative Matrix Factorization (NMF) and AttLSTM model training and prediction. The upstream and downstream sections selected through correlation analysis are used as the input of NMF(F, k), and the feature matrix after dimensionality reduction of the NMF algorithm is spliced with the current section flow as input to predict future traffic flow



Figure 8: Link 14394 and its upstream and downstream sections in Wuhan urban area

The size of the parameter k in the NMF algorithm is equal to the dimension of the feature matrix after dimensionality reduction. For the proposed model and the contrasted neural network model, the hyperparameters are set to the same value. The learning rate is 0.001, the optimization algorithm chooses the adam algorithm, and the mean square error (MSE) is used as the loss function.

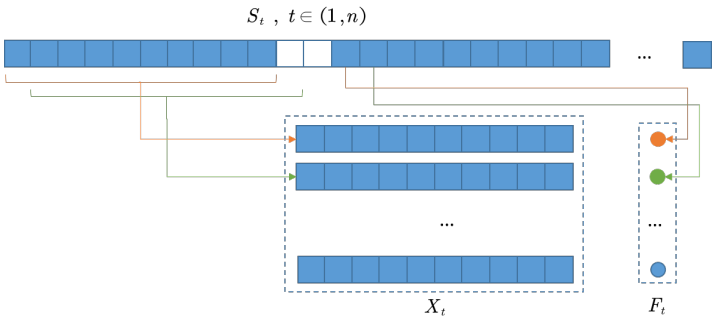


Figure 9: Constructing a data set through a sliding window

As shown in Table 2, we selected about 30 upstream and downstream road sections related to the current road section, and then based on related experiments, determined that the road section with a circumferential spatial correlation coefficient greater than 0.5 would be selected, and the data is reduced by the NMF algorithm.

3.3 Model Evaluation

In the experiment, the performance of the traffic flow prediction model is measured by three indicators: Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Mean

Absolute Error (MAE).

$$MAPE\,(\bar{x}, x) = \frac{1}{N}\sum_{n=1}^{N}\frac{|\bar{x}_n - x_n|}{x_n} \tag{19}$$

Table 2: Parameters set by default in the experiment

Parameter	Description	Value
n	Number of relevant sections of upstream and downstream	30
R_{week}	Weekly average Pearson correlation coefficient	>0.5
τ	Input time series data length	12,24,36
λ	Learning rate	0.001
	Batch Size	16
	Hidden layer neuron	64
	Prediction interval	15min

$$MAE\,(\bar{x}, x) = \frac{1}{N}\sum_{n=1}^{N}|\bar{x}_n - x_n| \tag{20}$$

$$RMS\,E\,(\bar{x}, x) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\bar{x}_n - x_n)^2} \tag{21}$$

$\bar{x}_n$ represents the predicted value, x_n represents the observed value, and N represents the number of data. MAPE considers the closeness of the true value to the predicted value and the ratio of the error to the true value, which is often used as a measure of the accuracy of the prediction in the prediction problem. RMSE is used to measure forecast stability. MAE is used to evaluate how close the predicted results are to the real data. The smaller the value, the better the fitting effect.

3.4 Experimental Results

First, use the PeMS public data set for analysis. For different NMF(F, k) algorithm parameters k, under the condition that the input sequence length is 12 and the prediction interval is 15 minutes, the distribution of MAPE is shown in Figure 10. When the parameter k is 5, it has the lowest MAPE error. Here, the parameter k of the NMF algorithm is also equivalent to the reduced dimensionality of the historical data matrix. In subsequent experiments, both the NMF-AttLSTM model and the AE-LSTM model use 5 as the reduced dimension.

We use one day's data to measure the performance of the NMF-AttLSTM model and other comparison models. At the same time, we set three different input sequence data lengths, 12, 24, and 36. They represent the use of 1-hour, 2-hour, and 3-hour historical data as the model input to predict the traffic flow after 15 minutes. Table 3 shows the performance of each model at different input sequence lengths. The results show that as the input sequence length increases, the MAPE, MAE, and RMSE of the AttLSTM model and NMF-AttLSTM model decrease, which proves that the model with the attention mechanism is more advantageous than other models when the length of the input sequence increases, because the model with the attention mechanism can extract more important features from the long sequence input. Under the same input sequence length, the NMF-AttLSTM model and the AE-LSTM model have better performance than other models. Due to the addition of upstream and downstream related road section features, the model provides the characteristics of traffic flow changes. In all the experimental results, the proposed NMF-AttLSTM model has the lowest MAPE, which proves that the model can not only use the spatial characteristics of traffic flow changes, but also obtain more important features from historical flow data, enhancing the prediction ability of neural network model for traffic flow.

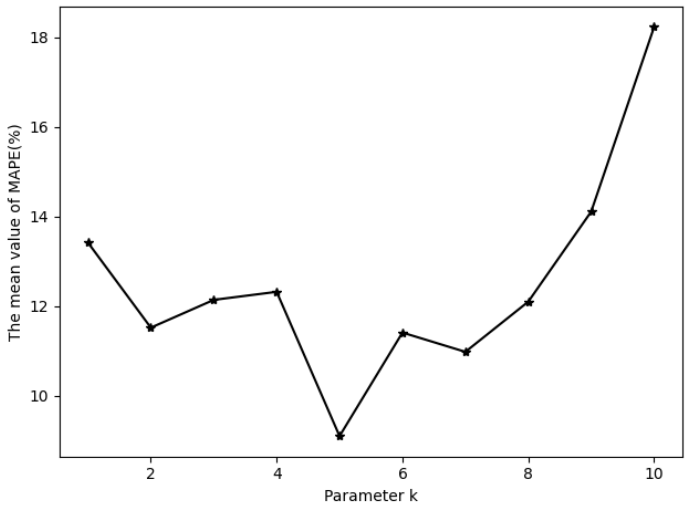


Figure 10: MPAE error of NMF-AttLSTM model under different dimensionality reduction parameter k (PeMS)

Table 3: Performance comparison of models under different input sequence lengths (PeMS)

Input Length	Model	Error Value		
		MAPE(%)	MAE	RMSE
12	SVR	12.43	24.00	**30.96**
	LSTM	12.52	27.68	35.65
	AttLSTM	10.96	26.02	34.92
	AE-LSTM	12.15	36.65	46.96
	NMF-AttLSTM	**9.10**	**22.72**	31.49
24	SVR	12.23	24.31	32.85
	LSTM	13.20	25.32	31.88
	AttLSTM	10.50	26.30	34.57
	AE-LSTM	8.60	22.23	30.41
	NMF-AttLSTM	**8.48**	**21.46**	**29.27**
36	SVR	12.91	29.28	37.59
	LSTM	11.53	25.84	33.45
	AttLSTM	9.07	24.61	32.54
	AE-LSTM	9.29	22.45	30.08
	NMF-AttLSTM	**7.89**	**15.21**	**19.92**

In order to verify the effectiveness of the proposed model on the actual road section, the road section 14394 in the core area of

Wuhan is selected for analysis. The specific location is shown in Figure 7. According to the method proposed in Section 2.3, firstly, the correlation between each upstream-downstream road section and the current road section is calculated by using the weekly spatial correlation coefficient, Finally, the correlation coefficients of the upstream and downstream sections with a correlation greater than 0.5 with 14394 are obtained, as shown in Table 4. Among the 30 upstream and downstream road sections, 14 relevant road sections are selected. The historical flow data of these 14 relevant road sections are used to construct the historical data matrix F, and then the NMF algorithm is used to obtain the traffic flow spatial characteristics H. For the NMF(F, k) algorithm, the effect of different k on the results was also tested. As shown in Figure 11, the results show that k is 7 with the lowest MAPE.

Table 4: Weekly spatial correlation coefficients of relevant sections in the upstream and downstream of road section 14394 of Wuhan

Upstream		Downstream	
Link ID	Correlation coefficient	Link ID	Correlation coefficient
14310	0.805	14198	0.639
14322	0.805	14200	0.942
14392	0.811	14317	0.941
14742	0.496	14315	0.941
14400	0.660	14316	0.936
14405	0.593	14222	0.638
		14223	0.926
		14195	0.634

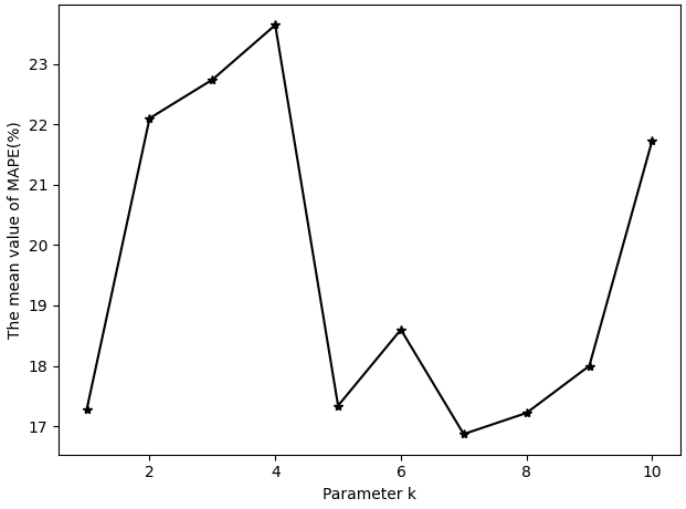


Figure 11: MPAE error of NMF-AttLSTM model under different dimensionality reduction parameter k (Wuhan)

For road section 14394, also use one day's data to predict the performance of the model. In the experiment, the input sequence length is 12 and the prediction interval is 15 minutes. Table 5 shows the prediction errors of different models. The results show that the proposed NMF-AttLSTM model has the lowest MAPE. However, because road section 14394 of Wuhan and its upstream and downstream sections are all branch sections, compared with PeMS, the flow value is generally low, so its MAPE error is higher, MAE and RMSE error are lower.

Table 5: Performance comparison of each model (Wuhan)

Model	Error Value		
	MAPE(%)	MAE	RMSE
SVR	22.92	3.27	4.39
LSTM	20.40	3.35	4.75
AttLSTM	18.35	3.29	4.75
AE-LSTM	18.75	3.16	4.22
NMF-AttLSTM	**16.54**	**3.01**	**4.13**

Table 6: Performance comparison of each model under two data sets

Data Set	Model	Error Value		
		MAPE(%)	MAE	RMSE
PeMS	SVR	12.52	25.86	33.80
	LSTM	12.41	26.28	33.66
	AttLSTM	10.17	25.64	34.01
	AE-LSTM	10.01	27.11	35.81
	NMF-AttLSTM	**8.49**	**19.79**	**26.89**
Wuhan	SVR	22.92	3.27	4.39
	LSTM	20.40	3.35	4.75
	AttLSTM	18.35	3.29	4.75
	AE-LSTM	18.75	3.16	4.22
	NMF-AttLSTM	**16.54**	**3.01**	**4.13**

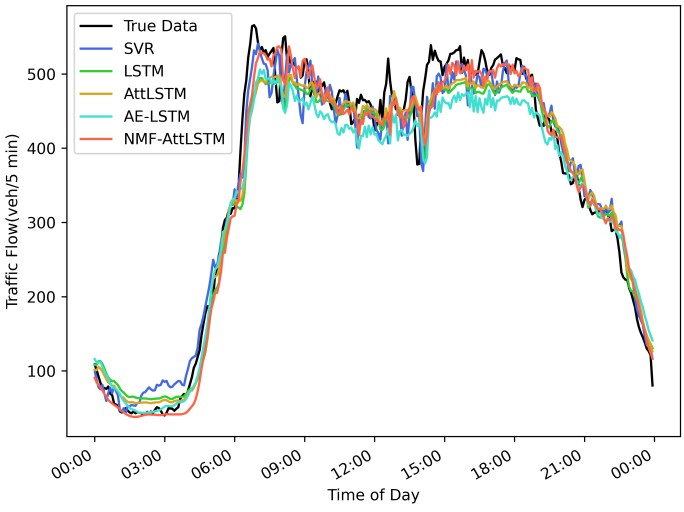


Figure 12: The prediction result of each model under the condition that the input sequence length is 12 (PeMS)

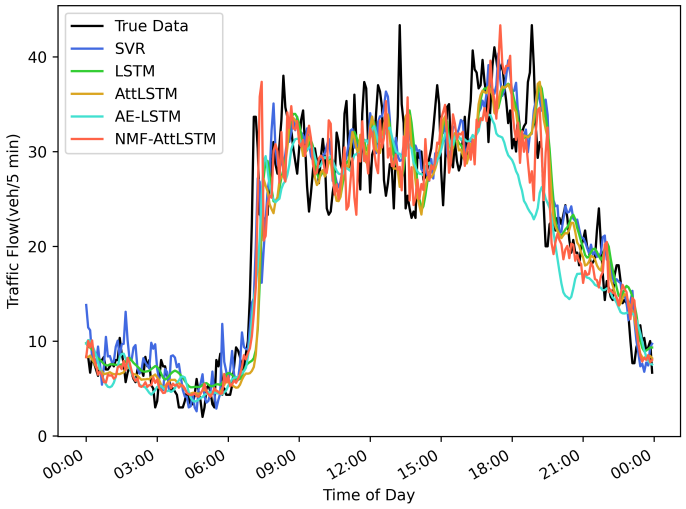


Figure 13: The prediction result of each model under the condition that the input sequence length is 12 (Wuhan)

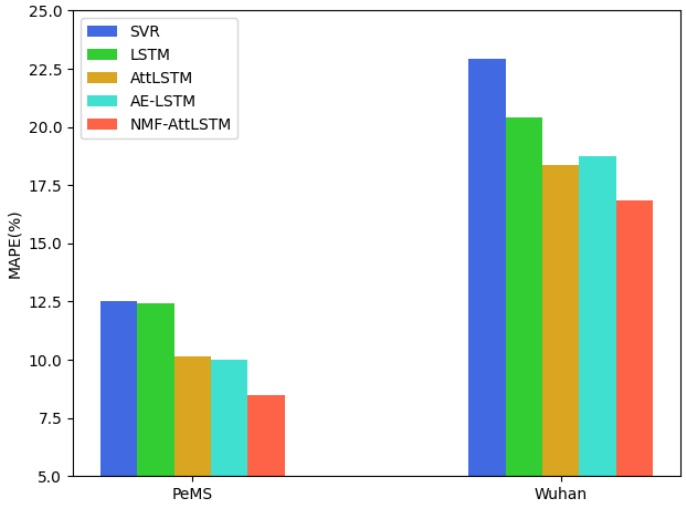
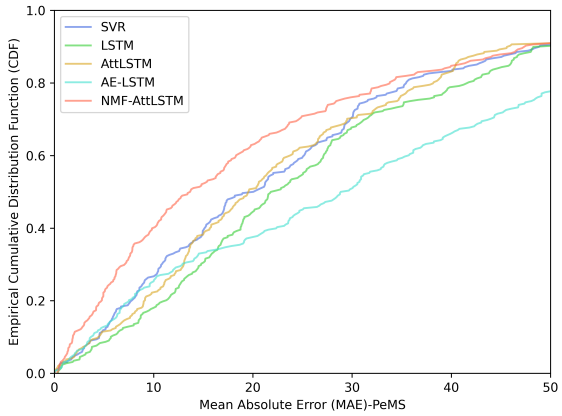


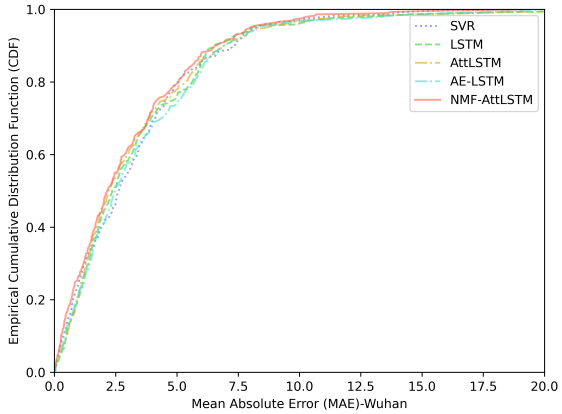
Figure 14: Comparison of MAPE errors of different models in PeMS and wuhan datasets

Table 6 shows the comparison results of the PeMS data set and the wuhan data set. Under the condition that the prediction interval is 15 minutes, the experimental results of the PeMS data set show that the MAPE of the NMF-AttLSTM model is reduced by 1.68% compared with the AttLSTM model and 1.52% compared with the AE-LSTM model. The experimental results of the wuhan data set show that the MAPE of the NMF-AttLSTM model is reduced by 1.81% compared with the AttLSTM model and 2.21% compared with the AE-LSTM model. As shown in Figure 14, NMF-AttLSTM has the lowest MAPE error on both data sets. The results on different experimental data sets prove the effectiveness of the proposed NMF-AttLSTM model.

Figure 12 and Figure 13 respectively describe the results of using one day's data prediction on the PeMS and Wuhan datasets. It can be seen that the results of the proposed NMF-AttLSTM model fit better with the true values. Figure 15 is the cumulative distribution function (CDF) diagram of the MAE error of the prediction results of each model. CDF can describe the probability distribution of the MAE error. The result shows that the MAE error of NMF-AttLSTM on PeMS is less than 20, accounting for more than 60%. The MAE error on the Wuhan dataset is less than 5, accounting for more than 80%, which is better than other models.



(a) PeMS



(b) Wuhan

Figure 15: The CDF of MAE error for different models of PeMS and Wuhan data sets

4 Conclusions

The paper studies the method of using the temporal and spatial characteristics of traffic flow to predict traffic flow, and proposes a short-term traffic flow prediction framework based on the NMF-AttLSTM model. First, analyze the correlation between each upstream and downstream road section and the current research road section by using the weekly spatial correlation coefficient, and eliminate the road sections with low correlation. Then use the NMF algorithm to extract the features of the selected road sections, and finally combine the extracted features with the historical traffic information of the current road section, and use the AttLSTM model for training and prediction. The algorithm considers the spatial relationship of traffic flow, effectively utilizes the flow information of upstream and downstream sections, and also reduces the dimension of data.

In addition, an attention mechanism is added to the LSTM model, so that the LSTM model can extract more valuable features from historical data. This research only considers the relevant flow information of the upstream and downstream sections. In the subsequent work, we can consider adding more complex road network topology information to improve the performance of the model.

5 Acknowledgment

This work is supported by "The wisdom decision-making construction of Wuzhi platform in Wuhan traffic management(ZSHJ-WHS-FW-2018-658)".

References

[1] D. Chen, C. Xiong, M. Zhong, "Improved LSTM Based on Attention Mechanism for Short-term Traffic Flow Prediction," in 2020 10th International Conference on Information Science and Technology (ICIST), 71–76, 2020, doi:10.1109/ICIST49303.2020.9202045.

[2] B. M. Williams, P. K. Durvasula, D. E. Brown, "Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models," Transportation Research Record, **1644**(1), 132–141, 1998, doi:10.3141/1644-14.

[3] Y. Zhang, Y. Xie, "Forecasting of Short-Term Freeway Volume with v-Support Vector Machines," Transportation Research Record, **2024**(1), 92–99, 2007, doi:10.3141/2024-11.

[4] M. Duan, "Short-Time Prediction of Traffic Flow Based on PSO Optimized SVM," in 2018 International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS), 41–45, 2018, doi:10.1109/ICITBS.2018.00018.

[5] M. Tan, S. C. Wong, J. Xu, Z. Guan, P. Zhang, "An Aggregation Approach to Short-Term Traffic Flow Prediction," IEEE Transactions on Intelligent Transportation Systems, **10**(1), 60–69, 2009, doi:10.1109/TITS.2008.2011693.

[6] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," Iet Intelligent Transport Systems, **11**(2), 68–75, 2017.

[7] R. Fu, Z. Zhang, L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 324–328, IEEE, 2016.

[8] W. Wei, H. Wu, H. Ma, "An autoencoder and LSTM-based traffic flow prediction method," Sensors, **19**(13), 2946, 2019.

[9] D. Bahdanau, K. Cho, Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.

[10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, "Attention is all you need," arXiv preprint arXiv:1706.03762, 2017.

[11] Z. Hao, M. Liu, Z. Wang, W. Zhan, "Human behavior analysis based on attention mechanism and LSTM neural network," in 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC), 346–349, IEEE, 2019.

[12] Z. Wang, J. Poon, S. Sun, S. Poon, "Attention-based multi-instance neural network for medical diagnosis from incomplete and low quality data," in 2019 International Joint Conference on Neural Networks (IJCNN), 1–8, IEEE, 2019.

[13] http://pems.dot.ca.gov/

[14] D. Lee, "Algorithms for non-negative matrix factorization," Advances in Neural Information Processing Systems, **13**, 2001.

[15] https://www.qgis.org/en/site/

[16] H. Drucker, C. Burges, L. Kaufman, A. Smola, V. Vapnik, "Support vector regression machines," Advances in neural information processing systems, **9**(7), 779–784, 2003.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 177-184 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Development of Miniaturized Monolithic Isolated Gate Driver

Hsuan-Yu Kuo, Jau-Jr Lin*

Department of Electrical Engineering, National Changhua University of Education, Changhua, 500 Taiwan, R.O.C.

ARTICLE INFO

Article history:
Received: 08 July, 2021
Accepted: 19 September, 2021
Online: 30 September, 2021

Keywords:
On-Chip Transformer
Gate Driver
Galvanic Isolation
Amplitude Shift Keying
Full-Wave Rectifier

ABSTRACT

Gate driver has been applied in many ways, exemplified by that, by using the DC-isolated and AC-pass characteristics of gate driver's primary and secondary sides, the problem of floating endpoint in semiconductor power switch can be solved. However, the conventional design of isolated gate driver provides circuit voltage blocking by optically coupled components. Due to the need for optoelectronic conversion, it requires III-VI semiconductor process and non-standard CMOS process, and the cost is always high. Therefore, in order to better solve the above mentioned problem, an electronic isolated gate driver is proposed. It employs an on-chip transformer to provide voltage isolation between the primary and secondary sides of the circuit, and converts the control signal in the circuit into a high-frequency modulated signal, which in the secondary side is then demodulated by the rectifier through the on-chip transformer to produce the original control signal. The miniaturized isolated gate driver proposed herein adopts TSMC T25HVG2 process and uses an on-chip transformer design in lieu of an optically coupled components. As the amplitude shift keying, on-chip transformer, full-wave quadruple rectifier and data buffer amplifier involved in this design are all integrated on the same chip, the integration can be improved. The size can be smaller than the generally separating electronic isolated gate driver, with the interference from external noise being reduced. In addition, the proposed gate driver generates large signals, in terms of chip layout, therefore, the circuit is put inside the on-chip transformer, which can further save area.

1. Introduction

Conventional galvanic isolated gate drivers use optically coupled components for a high degree of voltage isolation. Due to the need for optoelectronic conversion, it requires III-VI semiconductor process and non-standard CMOS process, and the cost is always high. In desirous of solving these problems, MOSFET electronic isolated gate driver has been studied as a new direction. It couples the energy to the secondary side of the transformer by the transformer's resonant coupling method in lieu of the optical coupling driver. Since the isolation is realized by optical and electrical energy conversion, many of the shortcomings in the traditional design can be improved, such as the larger size and higher process cost in the past.

This paper is an extension of work originally presented in the 2nd IEEE Eurasia Conference on IOT, Communication and Engineering 2020[1]. The proposed system concept in which is shown in Figure 1. The circuit is mainly composed of a VCO, a mixing circuit (switch), an on-chip transformer, a quadruple rectifier and a data buffer amplifier. We input a PWM-controlled signal to the chip, and then the ~1 GHz oscillation signal and the PWM signal are modulated by ASK (amplitude shift keying) at the switch, and then the modulated signal is brought to the secondary side by the on-chip transformer; the quadruple rectifier then further amplifies the small signal and rectify the DC voltage of the process to provide enough driving voltage for the data buffer amplifier to generate the PWM signal, and that's the expected result.

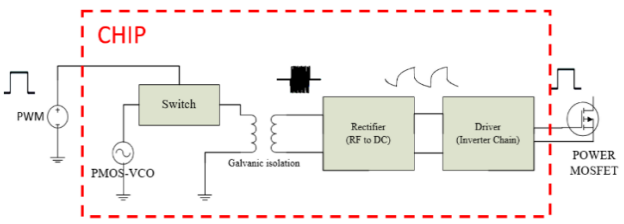


Figure 1: System concept of proposed isolated gate driver.

The circuit proposed is integrated on a single chip, so the size is much smaller than the separated design [2, 3]; besides, according

*Corresponding Author: Jau-Jr Lin, Email: jaujrlin@cc.ncue.edu.tw

https://dx.doi.org/10.25046/aj060520

to the simulation, if a PWM control signal of 100- kHz 50% duty circle is given at the primary side's input end, we hope that after the voltage isolation, the output can remain the same. However, on account of the transmission delay caused by the rise time and fall time, the output PWM control signal will be slightly degraded as the simulation shows that the duty circle of the output PWM control signal is 48.4%, a difference of 1.6%; the difference between the rise time and fall time is only 0.2 μs. In the next section, we introduce the architecture of the chip, and discuss the detailed circuit design and the model of the on-load transformer in detail in the implementation section. Finally, we present the results of different process variations under Post-Simulation in Full-chip Simulation, all of which are in line with the expected specifications.

2. Chip architecture

The system block diagram and schematic diagram proposed in this article are shown in Figure 2 and Figure 3, respectively. Its primary side is composed of a current mirror with a current controller and a PMOS VCO. The input PWM signal and the 1GHz signal provided by the oscillator, after modulation, are transmitted to the secondary side through the on-chip transformer. The full-wave quadruple rectifier can further amplify and demodulate the modulated signal into a DC voltage. However, the potential and ripple of the rectified DC voltage is not suitable for being directly used as a gate drive signal, and the data buffer amplifier can improve this problem by further pulling the voltage to 5 V and greatly improving the ripple, so as to increase its driving force.

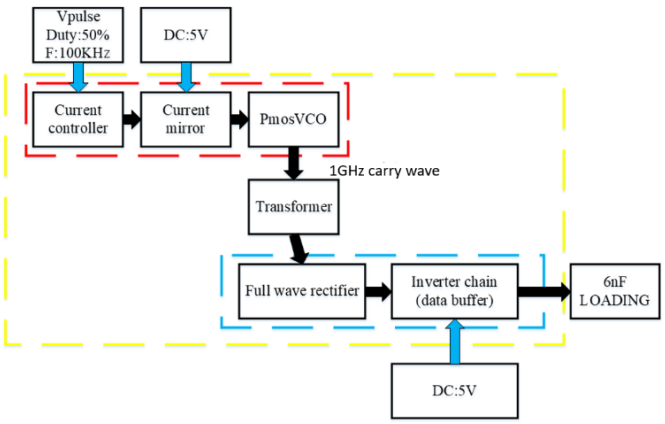


Figure 2: The system block diagram of the proposed design.

On and off of the switch can be controlled by putting a PWM control signal on the primary side. Once the switch is on, the current will increase. At this point, the large current causes the VCO to generate a larger amplitude oscillation signal. That is, as the current becomes larger or smaller, the PMOS VCO generates two signals with different amplitudes. The above phenomenon can be interpreted as that, the carrier wave generated by the PMOS VCO and the PWM signal are subjected to an ASK modulation action, and then the modulation signal is transmitted to the secondary side through the on-chip transformer. However, since this signal voltage level is not high enough to drive the data buffer amplifier, the full-wave rectifier (voltage multiplier) will increase the voltage level and rectify the RF signal to a DC voltage. Finally, the data buffer amplifier further enhances the driving capability. In addition, the data buffer amplifier converts the ASK signal into a PWM gate signal to drive the subsequent semiconductor power switch.

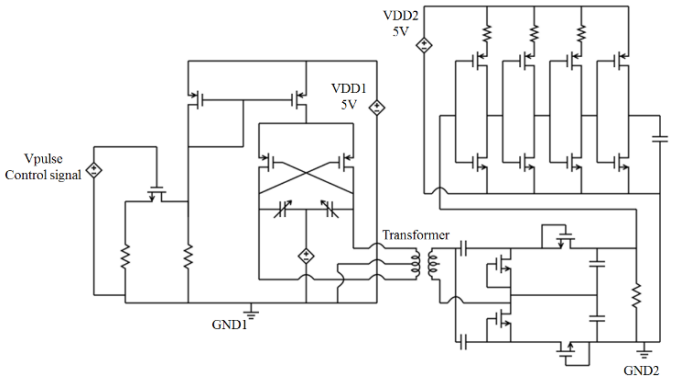


Figure 3: The schematic diagram of the proposed isolated gate driver.

There are two types of transformers for on-chip transformer design: stack transformer or tapped transformer [1, 4]. It provides galvanic isolation in the circuit proposed in this article. The stack transformer performs better in coupling, but its isolation is limited by process as the distance between the metal layers is fixed. In other words, although the coupling degree of the tapped transformer is less than that of the stack transformer, the distance between the two inductors is longer than that of the stack transformer. The transformer is more flexible. The transformer applied in this paper is a stack transformer which features higher coupling as mentioned above.

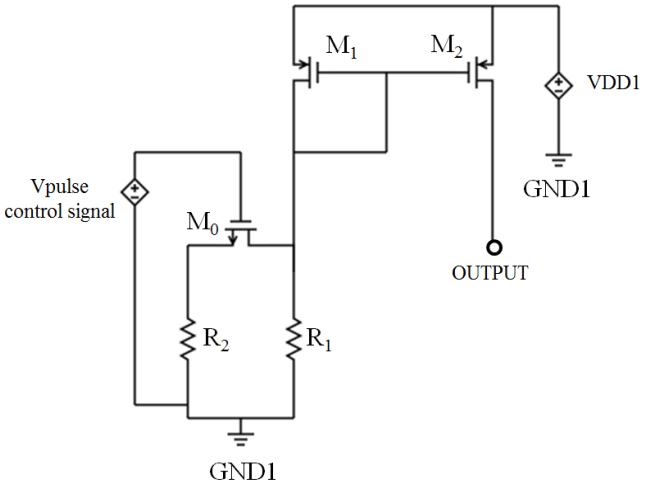


Figure 4: The schematic of current mirror and current controller

3. Circuit implementation

3.1. Current Mirror with Current Controller

The current controller, shown in Figure 4, consists of two resistors and a switch. The switch is controlled by a PWM control signal generated by a signal generator outside the chip. When the switch is on, the two resistors become parallel, which makes the impedance smaller and current higher (Figure 5). The lower part of Figure 5 is the input current of Figure 4 current mirror, and the upper part of Figure 5 is the output current after magnification of 3 times by the Figure 4 current mirror. This design serves the

purpose of performing ASK modulation of the PWM signal and the oscillation signal. Another reason for this design is that if the OOK modulation is achieved by simply turning off the oscillator instead of generating oscillation signals of different amplitudes, it takes more time for the system to enter the steady-state mode, which causes even longer transmission delay. Therefore, in order to achieve ASK modulation, the oscillation signal needs to be decreased so that it cannot drive the data buffer amplifier instead of directly closing the entire system.

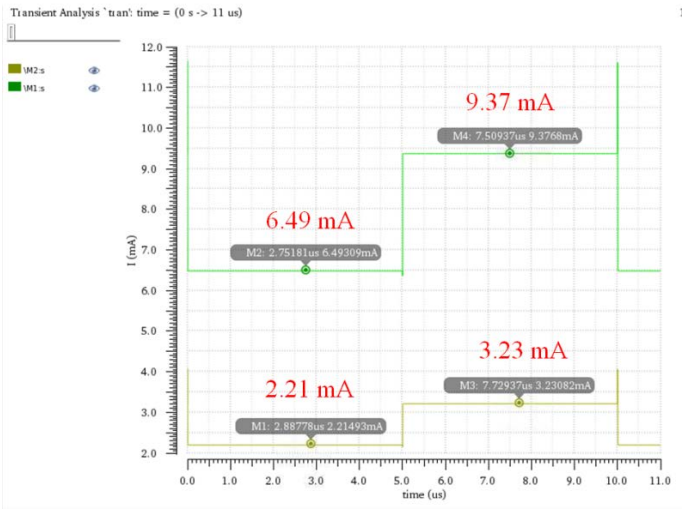


Figure 5: The output current of the current controller and the output current after going through the current mirror.

A current mirror is employed to increase the current, and the result is shown in Figure 5. The lower curve is the current value of current controller side and the upper curve is the current value that follows into VCO. When the voltage control signal is low, the current values of the current control and VCO are 2.21 mA and 6.49 mA, respectively. When the voltage control signal is high, the current values of the current control and VCO are 3.23 mA and 9.73 mA, respectively. The current mirror magnifies the current to a set magnification, such as (1). Since $L_1 = L_2$ and $W_2 = 3W_1$ in this design, the magnification should be 3 times the current in Figure 5, and the result is obtained. M is Magnification of current mirror.

$$M = (W_2 \cdot L_1) / (W_1 \cdot L_2) \quad (1)$$

3.2. PMOS Voltage Control Oscillator

The main function of the VCO, shown in Figure 6, is to change the capacitance value by adjusting V_{BIAS} [4-6], thereby generating different oscillation frequencies. As this LC resonant circuit is composed of an inductance with fixed inductance value and a variable capacitance, this study can therefore increase the exhaustion region in the diode by enhancing the reverse bias voltage of the VCD. Therefore, when the reverse bias voltage becomes larger, the distance between the two metals in the VCD increases, the capacitance value decreases, and the resonance frequency of the LC circuit reduces. In addition, the PMOS VCO design proposed in this paper must achieve the ASK modulation effect with a current mirror with a current controller, that is, the PWM signal and the oscillation signal generated by the PMOS VCO need to be modulated.

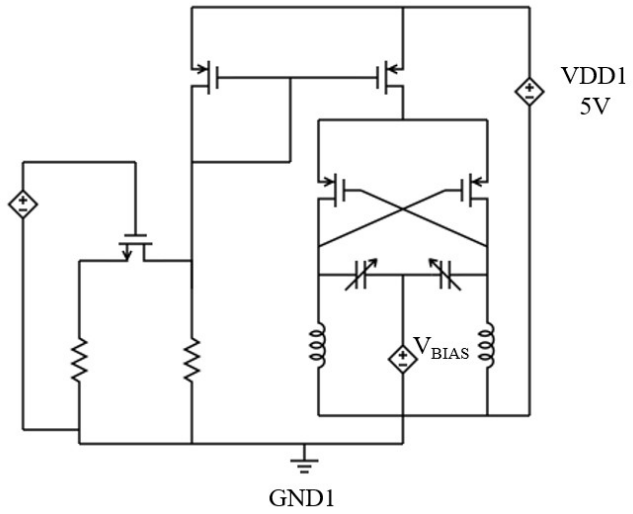


Figure 6: The circuit schematic of the PMOS VCO.

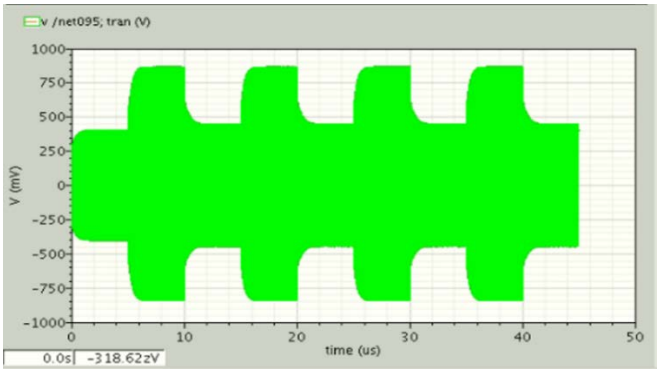


Figure 7: The output signal of the PMOS VCO with the current control voltage low and high.

As shown in Figure 7, the current controller and current mirror make the VCO generate low and high currents, which further makes the PMOS VCO generate two RF signals with different amplitudes. Generally, it takes longer to reach a stable state after the VCO oscillation stops. Therefore, it needs to make the small oscillation signal keep oscillating, and the signal must be small enough so as not to affect the components in the design proposed herein.

3.3. Stack transformer

The stack transformer [1, 4, 7], shown in Figure 8, is chosen because of its higher theoretical coupling, which means that, compared with the tapped transformer [8, 9], the stack transformer can transmit signal to the secondary side more effectively. However, the stack transformer still has shortcomings. The distance between the two inductors is limited by the process, that is to say, the distance between each layer is fixed, which means that the isolation is limited. A tapped transformer, however, uses a same layer (metal in the top), so the distance between the metals is more flexible despite the lower degree of coupling. The results are shown in Figure 9 and 10, indicating that the inductance value is about 3.57 nH, the quality factor is about 6.94, and the coupling coefficient is about 0.917 at 1 GHz.

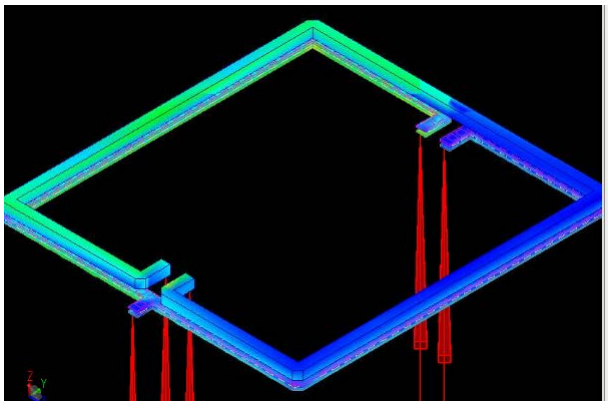


Figure 8: 3D graph of the stack transformer.

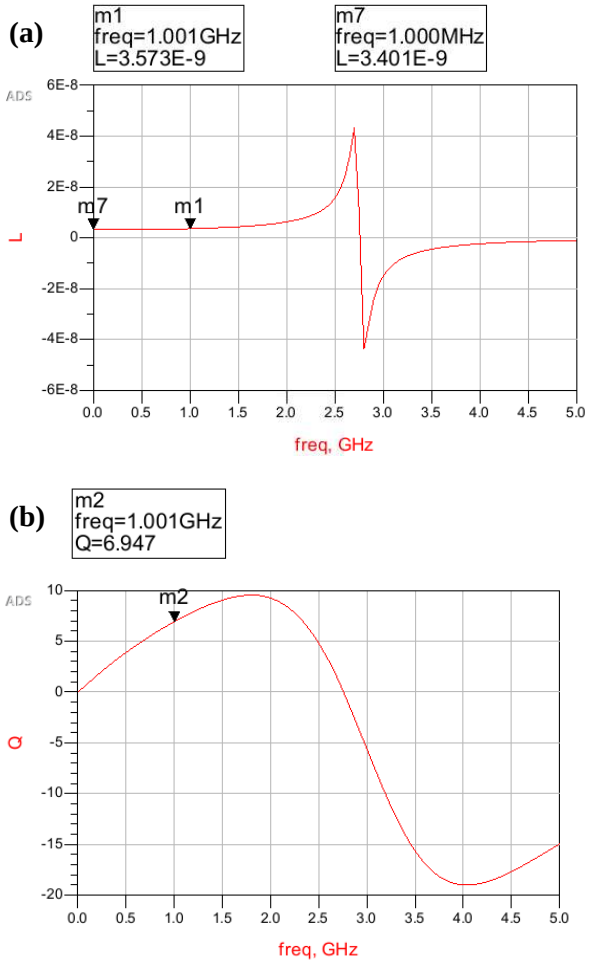


Figure 9: (a) The inductance of transformer and (b) Quality factor in the primary side.

Figure 11 shows the modeling of the on-chip transformer proposed in this paper [10]. R_{10} and R_{20} are the impedance of the transformer itself at low frequency. The values of L_{10} and L_{20} in Figure 11 are from the model of the ADS momentum transformer. L_{11}, L_{21} , R_{11}, R_{21} are the inductance and impedance generated by the proximity effect and skin effect [11,12] when the two on-chip inductors are at high frequencies. C_{ox} and R_{sub} are the parasitic resistance and capacitance between the metal layer and the substrate, and C_{ps} is the coupling capacitance between the two inductors.

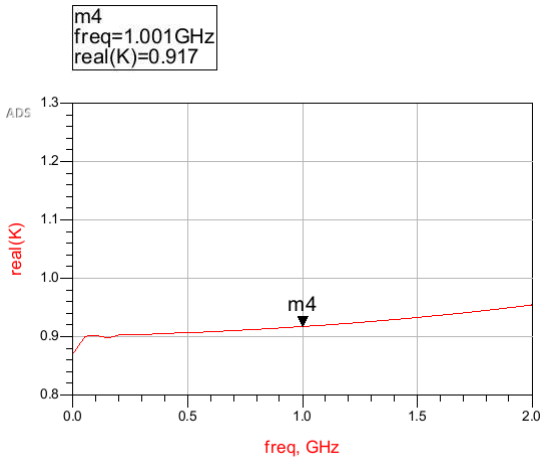


Figure 10: Coupling coefficient of the proposed transformer.

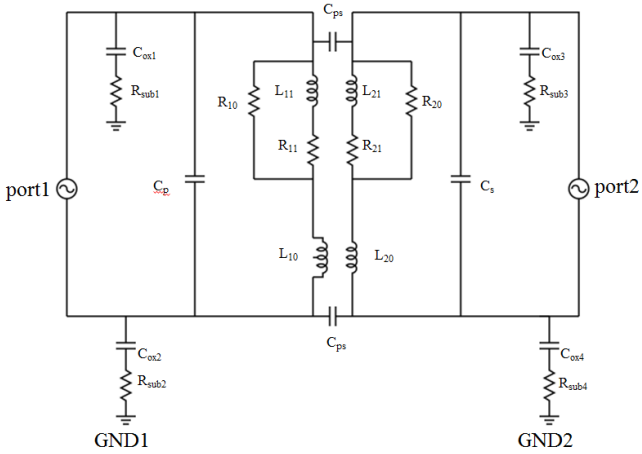


Figure 11: Modeling of the on-chip transformer.

In the circuit, L_{10} and L_{20} can be obtained by (2), and the result is shown in Figure 9. L_{10} and L_{20} are the inductance values in the circuit at low frequency, so we can get this data in Figure 9 (a). R_{11} and R_{21} are also the impedance of the circuit at low frequency, so its parameters can be obtained in smith and z_parameter by the same way. C_p and C_s are the parasitic capacitances in the inductance, which can be reversely obtained according to its natural vibration frequency, such as (3) (4). The natural vibration frequencies of the primary side and the secondary side are represented by $freq_p$ and $freq_s$ respectively [13].

$$L_{10} = imag(Z_{11}) / (2 \cdot pi \cdot freq) \quad (2)$$

$$freq_p = 1 / [2 \cdot pi \cdot sqrt(L_{10} \cdot C_p)] \quad (3)$$

$$freq_s = 1 / [2 \cdot pi \cdot sqrt(L_{10} \cdot C_p)] \quad (4)$$

As the chip itself is a 1:1 symmetrical architecture, $R_{sub1} = R_{sub2} = R_{sub3} = R_{sub4}$, and $C_{ox1} = C_{ox2} = C_{ox3} = C_{ox4}$. we can see the comparison between the simulation and the optimized modeling from Figure12. In Figure 12, curves of S(1,1), S(2,2) and S(1,2) are from ADS momentum simulation in Figure 8, and curves of modeling2..S(1,1), modeling2..S(2,2) and modeling2..S(1,2) are

from optimized modeling in Figure 11. Coupling coefficient K and quality factor Q can be obtained from (5) (6). The difference between the results of ADS momentum simulation and optimized modeling simulation is shown in Table 1.

$$K = sqrt(imag(Z_{12}) \cdot imag(Z_{21})) / sqrt(imag(Z_{11}) \cdot imag(Z_{22})) \quad (5)$$

$$Q = imag(Z_{11}) / real(Z_{11}) \quad (6)$$

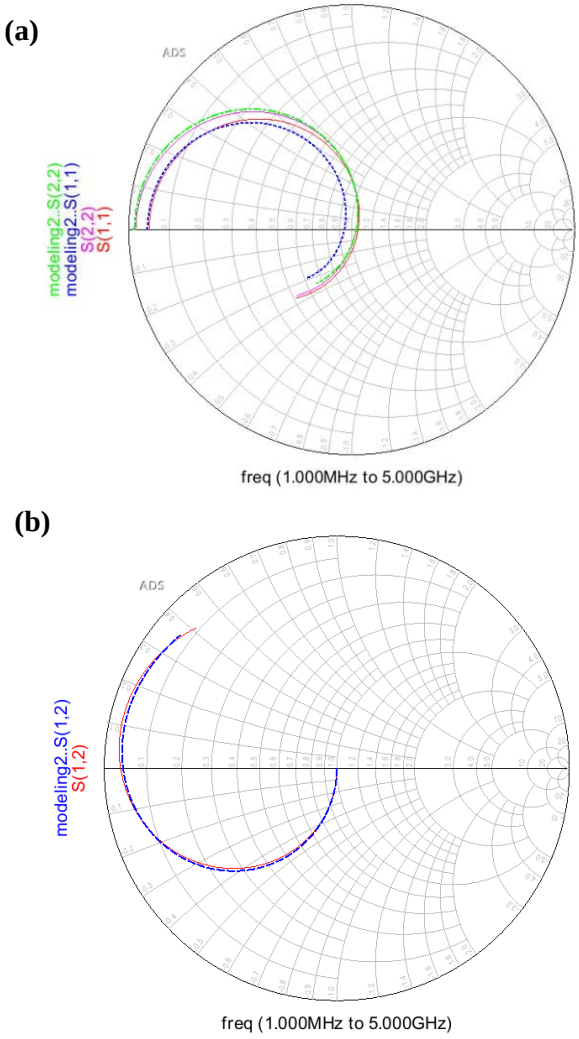


Figure 12: Comparison of the results between ADS momentum (Figure 8) and modeling (Figure 11) (a) S_{11} and S_{22} (b) S_{12} results

3.4. Full Wave Rectifier (Voltage Multiplier)

Since the RF signal received from the transformer output is low, the full-wave rectifier rectifies the RF signal to a DC voltage and multiplies the signal, making it has a greater ability to drive the data buffer amplifier. Figure 13 is the circuit schematic of the full-wave rectifier [14] proposed in this paper.

The traditional voltage multiplier is a diode voltage doubler in cascaded capacitor, where the voltage doubler uses voltage multipliers to connect N stages. If the threshold voltage is regarded as zero, it can theoretically reach N times of magnification. However, Zener diode threshold voltage and low threshold NMOS cannot be regarded as zero. In the formula (7), V_{IN} is the amplitude of the RF oscillation signal, V_T is the threshold voltage, and N is the series of voltage multiplier in cascaded capacitor.

$$V_{OUT} = 2N \cdot (V_{IN} - V_T) \quad (7)$$

However, conventional diode multipliers come with the disadvantages of high output impedance, large DC output ripple, and low conversion efficiency. Therefore, in this study, the proposed design replaces the Zener diode with a low threshold voltage NMOS (Figure 13).

Table 1: Comparison of s parameter simulation and modeling simulation

	ADS momentum	Optimized Modeling
R_{20}	0.78 ohm	0.65 ohm
L_{20}	3.16 nH	3.51 nH
R_{10}	2.36 ohm	2.17 ohm
L_{10}	3.40 nH	3.33 nH
$freq_p$	2.75 GHz	3.05 GHz
$freq_s$	2.75 GHz	3.05 GHz
K	0.91	0.93
Q	6.94	8.43

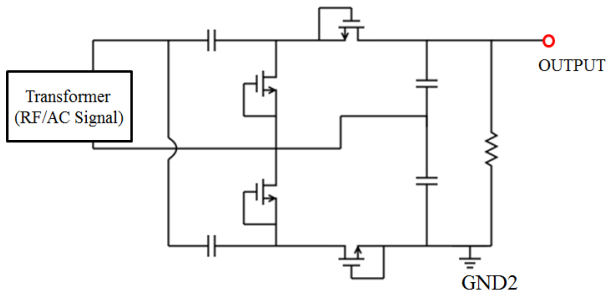


Figure 13: Proposed full-wave rectifier.

3.5. Data buffer

The data buffer amplifier consists of four inverters connected in series. The basis of the data buffer is to effectively enhance the driving capability of the gate signals [15]. In addition, it also reduces the transmission delay of the signal in inverters. Figure 14 shows the relationship between the delay time and the different stages of the inverter chain. N is the design time, f is the ratio of C_{in} to C_{out}, and tp is the transmission delay of signal through the data buffer amplifier. As N increases, tp can be reduced. This design serves the purpose of effectively reducing the ratio of electrical output capacitance to gate capacitance in each stage, and of gradually increasing the current size of each stage, so that it can shorten the rise and fall times by connecting a multi-stage inverter and thus reduce the transmission delay, which is better than the original small capacitor charging a large one. However, this stage connection has its limits as too many stages make the output of the signal generate excess delay, and the size is not easy to control. The advantage of the four-stage architecture adopted in this paper is that the phase of the output control signal is the same as that of

the primary side's PWM control signal, and the transmission delay can be minimized to avoid too many long stage connections.

Since the capacitance ratio of each stage in the four-stage architecture is three times larger than that of the previous stage, the gate capacitance (C_{in}) can be changed by application of CMOS technology, the width and length of MOSFET. In other words, since the inverter chain proposed in this paper consists of four inverters, the width of the next MOSFET stage needs to be modified so that it is three times larger than that of the previous MOSFET.

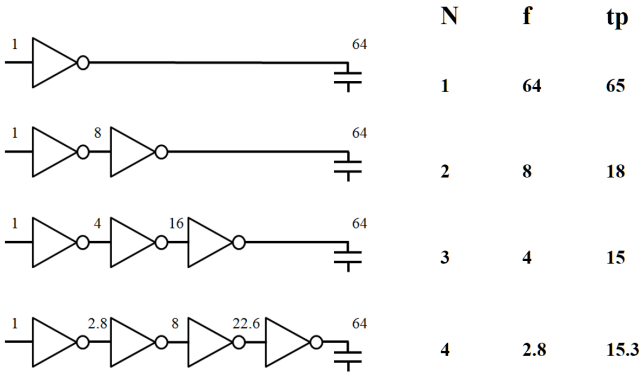


Figure 14: Signal delay time with different stage inverter.

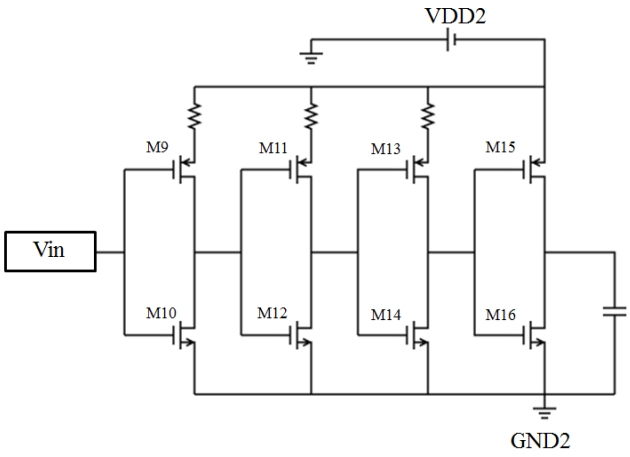


Figure 15: Proposed data buffer.

In the design, DC voltage cannot be used directly as a gate control signal for gate driver because of its large ripple and long rise and fall times. Therefore, the data buffer amplifier is an important element to improve the drawback by enhancing the drive capability. In the proposed design (Figure 15), the data buffer amplifier consists of four inverters and three resistors. Since the rectified DC voltage is not sufficient to drive the data buffer, the resistors are used to lower the threshold of the driving voltage. The four-stage data buffer can also amplify the input current by 64 magnitudes, which allows the circuit to generate enough current to drive the potential larger power MOSFET. In addition, the inverter chain eliminates ripples and also reduces fall and rise times.

4. All chip simulation

This monolithic isolated gate driver proposed in this paper is designed and manufactured with TSMC 250nm HV COMS technology and 5 V transistors. The on-chip transformer, with area 1.000 mm × 1.125 mm, provides galvanic isolation between the primary and secondary sides. The control signal sets a 100-kHz 50% duty cycle 5-V PWM signal for the current controller to switch the current value. The VCO generates 1-GHz oscillation signals of different amplitudes due to the different magnitudes of current from the current mirror and the current controller. The results are shown in Figure 16.

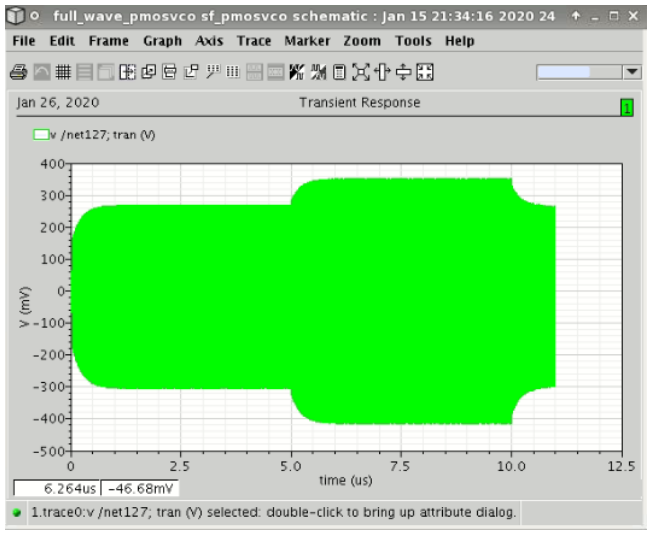


Figure 16: The oscillation signal of the PMOS VCO.

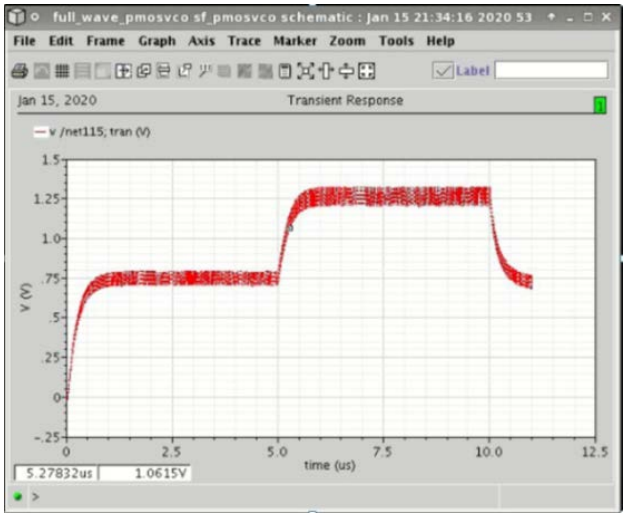


Figure 17: The output voltage of the full wave rectifier.

The signal generated by VCO is not high enough to drive the data buffer amplifier, so the full-wave rectifier amplifies the voltage and rectifies the RF signal to DC voltage, as shown in Figure 17. The driving threshold voltage of the data buffer amplifier is shown in Figure 18, and the threshold voltage of the four-stage chain is around 1V. When the driving voltage is higher than 1V, the rectified signal can trig the inverter chain. When the driving voltage is lower than 1V, it cannot trig. The results are shown in Figure 18(a). Figure 18(b) shows the result when

resistors are not added to the data buffer amplifier to reduce its threshold drive voltage, its original voltage is about 2.5 V. The threshold drive voltage that requires more voltage multipliers to be larger leads to extra energy loss and chip size increase due to the too many components.

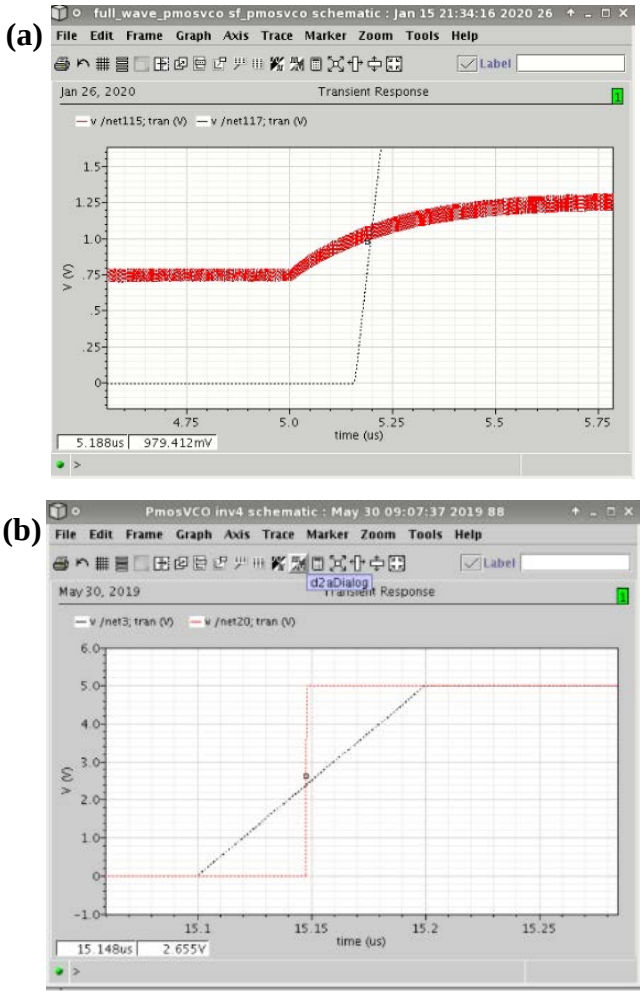


Figure 18: (a) The threshold of the data buffer is ~1V. (b) the threshold voltage in data buffer without resister is ~2.5 V

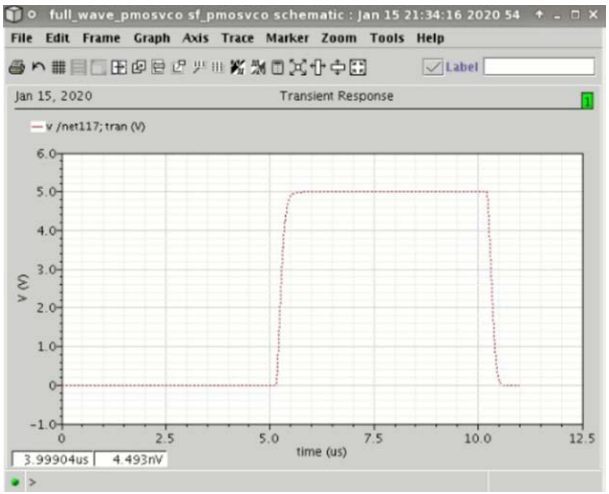


Figure 19: The output of the data buffer.

In the final design, after the signal passes through the whole system, the duty cycle of the PWM signal with 50% duty cycle is kept at 48.4% and the rise and fall times are kept below 10% of the cycle (rise and fall times are about 0.2 μs, and the cycle time at 100 kHz is 10 μs). The results are shown in Figure 19.

Table 2: In case of different corners, the expected result and the simulation result of changing the control DC bias of the variable vessel

	Specification	Post_sim (ss)	Post_sim (tt)	Psot_sim (ff)
VCO output signal (off)	0.2 ~ 0.8	0.24 ~ 0.73	0.35 ~ 0.74	0.43 ~ 0.89
VCO output signal (on)	> 0.9	0.69 ~ 1.19	0.72 ~ 1.21	0.77 ~ 1.26
Multiplier output (off)	< 1 V	0 ~ 0.65	0.05 ~ 0.85	0.12 ~ 1.03
Multiplier output (on)	> 1V	0.53 ~ 1.91	0.78 ~ 2.12	0.85 ~ 2.43
Threshold voltage for data buffer	1 V	0.95 V	1.01 V	1.09 V
Rise time	< 1 μs	0.242 μs	0.206 μs	0.166 μs
Fall time	< 1 μs	0.239 μs	0.227 μs	0.200 μs

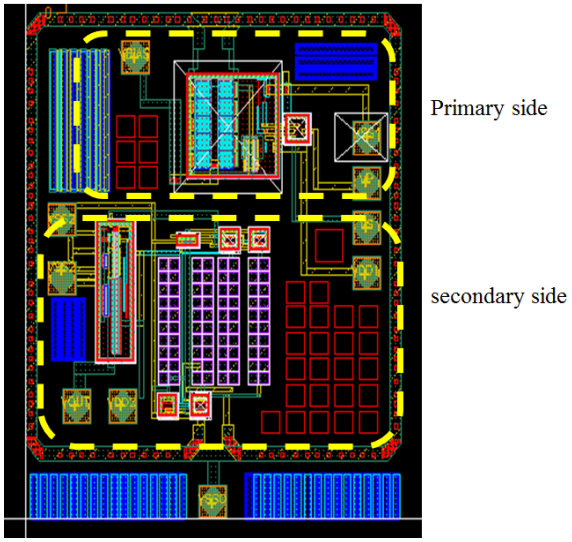


Figure 20: Chip layout

Table 2 shows the specifications and simulation results of the circuit proposed in this paper. The output voltages of the VCO and multiplier vary depending on the control voltage given to the varactor diode. It needs attention in the circuit, because in design, VCO must generate two different voltages, and after the two voltages are rectified by the voltage quadrupler, one of the rectified DC voltages needs to be able to drive the data buffer amplifier, while the other rectified DC voltage cannot drive the data buffer amplifier, so as to ensure that these two voltages can control the first-stage inverter of the data buffer amplifier, making the circuit generate PWM control signals. Assuming that this condition cannot be guaranteed to be established, it may cause the data buffer amplifier to become routinely open or closed, which leads to the failure to achieve our expected results. In addition, the varactor also makes it difficult for the VCO to start up as shown in Table 2. As to the output signal of VCO and the voltage

amplifier, the result also increases as the control DC bias of the varactor increases, which means that if insufficient or too much control DC bias is given, it may cause the first-stage structure of the inverter chain to be routinely open and closed, thus making the circuit lose its intended function. Figure 20 shows the layout of the chip to be delisted. This gate driver generates large signals, so the layout of this chip puts the circuit inside the on-chip transformer to save more area.

5. Conclusion

Traditional isolated gate drivers generally adopt optically coupled driver components to achieve high isolation, so they require III-VI semiconductor process and non-standard CMOS process, making the cost even higher. In order to solve this problem, electronic isolated gate drivers are gradually attracting more attention. On-chip transformers can reduce the energy loss caused by the optical coupling components, but a separate architecture is still used in many designs – devices with different functions are designed on different chips, which leads to other problems, such as larger size, and poor treatment of external noise. Therefore, the problem of high cost remains unsolved in such designs. The architecture proposed in this paper is to integrate the whole oscillator, transformer, full-wave multiplier and rectified DC signal processing system on one chip, so as to reach optimization in both integration and size.

On the primary side, the signal generator provides the chip with a 100 kHz PWM switching signal with a 50% duty cycle, which is then modulated by a ~1 GHz carrier wave generated by a VCO. After passing through a full-wave rectifier and a data buffer amplifier, it retains a 48.4% duty cycle and a 5-V output signal. All devices are integrated in a single chip. This will reduce the system dimension and enhance the system integration significantly.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is supported by the Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. MOST 107-2221-E-018-002- and MOST 109-2221-E-018-007-. We are thankful to the National Chip Implementation Center (CIC) for making its EDA tools and technology available, and the National Center for High-performance Computing (NCHC) for the use of its computer time and facilities.

References

[1] H.Y. Kuo, J.J.Lin, “Implementation of Miniaturized Monolithic Isolated Gate Driver,” in 2nd IEEE Eurasia Conference on IOT, Communication and Engineering 2020, ECICE 2020, 129–132, 2020, doi:10.1109/ECICE50847.2020.9301940.

[2] S. Nagai, Y.Yamada, N.Negoro, H.Handa, M.Hiraiwa, N.Otsuka, D.Ueda, “A 3-phase AC-AC matrix converter GaN chipset with drive-by-microwave technology,” IEEE Journal of the Electron Devices Society, **3**(1), 7–14, 2015, doi:10.1109/JEDS.2014.2364307.

[3] C. Liu, D.Mandal, Z.Yao, M.Sun, J.Todsen, B.Johnson, S.Kiaei, B.Bakkaloglu, “A 50-V Isolation, 100-MHz, 50-mW Single-Chip Junction Isolated DC-DC Converter with Self-Tuned Maximum Power Transfer Frequency,” IEEE Transactions on Circuits and Systems II: Express Briefs, **66**(6), 1003–1007, 2019, doi:10.1109/TCSII.2018.2869074.

[4] J. Lin, K.C.Lin, “A monolithic isolated gate driver with on-chip transformer and a Voltage Level Shifter,” in 2017 IEEE 3rd International Future Energy Electronics Conference and ECCE Asia, IFEEC - ECCE Asia 2017, 1256–1259, 2017, doi:10.1109/IFEEC.2017.7992223.

[5] M. DaTsai, Y.H.Cho, H.Wang, “A 5-GHz low phase noise differential colpitts CMOS VCO,” IEEE Microwave and Wireless Components Letters, **15**(5), 327–329, 2005, doi:10.1109/LMWC.2005.847696.

[6] X. Wang, B.Bakkaloglu, “Systematic design of supply regulated LC-tank voltage-controlled oscillators,” IEEE Transactions on Circuits and Systems I: Regular Papers, **55**(7), 1834–1844, 2008, doi:10.1109/TCSI.2008.918004.

[7] A. Ghadiri, K.Moez, “Bandwidth enhancement of on-chip transformers using negative capacitance,” IEEE Transactions on Circuits and Systems II: Express Briefs, **59**(10), 648–652, 2012, doi:10.1109/TCSII.2012.2213360.

[8] H.M. Hsu, C.W.Tseng, K.Y.Chan, “Characterization of on-chip transformer using microwave technique,” IEEE Transactions on Electron Devices, **55**(3), 833–837, 2008, doi:10.1109/TED.2007.914841.

[9] R. Wu, N.Liao, X.Fang, S.Member, “A Silicon-Embedded Transformer for High-Efficiency, High-Isolation, and Low-Frequency On-Chip Power Transfer,” **62**(1), 220–223, 2015. doi:10.1109/TC.2015.2517635.

[10] Q. El-Gharniti, E.Kerhervé, J.B.Bégueret, “Modeling and characterization of on-chip transformers for silicon RFIC,” IEEE Transactions on Microwave Theory and Techniques, **55**(4), 607–615, 2007, doi:10.1109/TMTT.2007.893647.

[11] J. Kang, L.Sun, J.Wen, M.Zhao, “An equivalent lumped-circuit model for on-chip symmetric intertwined transformer,” in ASICON 2009 - Proceedings 2009 8th IEEE International Conference on ASIC, IEEE: 674–677, 2009, doi:10.1109/ASICON.2009.5351336.

[12] J. Liu, L.Sun, Z.Yu, “An accurate compact model for on-chip vertically coiled transformers,” IEEE Electron Device Letters, **34**(4), 484–486, 2013, doi:10.1109/LED.2013.2247559.

[13] H. Liao, C.Wang, “A New Approach to Parameter Extraction for On-Chip Symmetric Transformers,” in 2009 IEEE International Symposium on Radio-Frequency Integration Technology, RFIT 2009, IEEE: 257–260, 2009.

[14] S. Iqbal, “A hybrid symmetrical voltage multiplier,” IEEE Transactions on Power Electronics, **29**(1), 6–12, 2014, doi:10.1109/TPEL.2013.2251474.

[15] X.J. Wang, T.Kwasniewski, “Propagation Delay-Based Expression of Power Supply-Induced Jitter Sensitivity for CMOS Buffer Chain,” IEEE Transactions on Electromagnetic Compatibility, **58**(2), 627–630, 2016, doi:10.1109/TEMC.2016.2517635.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 182-188 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Use of Unmanned Aerial Vehicles in Aircraft Inspection

Andrej Novák[*,1], Martin Bugaj[1], Alena Novák Sedláčková[1], Branislav Kandera[1], Anna Stelmach[2], Tomasz Lusiak[3]

[1]*University of Žilina, Air Transport Department, Žilina, 010 26, Slovakia*

[2]*Warsaw University of Technology, Faculty of Transport, Warsaw, 00-662, Poland*

[3]*Lublin University of Technology, Faculty of Mechanical Engineering, Lublin, 20-618, Poland*

ARTICLE INFO

Article history:
Received: 15 February, 2021
Accepted: 10 May, 2021
Online: 23 May, 2021

Keywords:
Smart Hangar
UAV
Pre-flight inspection

ABSTRACT

The article further extends the researched issue of the unmanned aircraft use in the pre-flight and post-flight visual check of aircraft. Procedures of pre-flight inspection are fulfilled by the aircraft maintenance certified staff or the crew member before flight. The process is similar for all categories of aircraft, but its implementation differs for individual specific types of aircraft. Therefore, the article will deal only with small training aircraft, which will be used to verify the use of UAV (Unmanned Aerial Vehicle) in normal operation. It identifies and defines the problem of using multiple UAV in swarms and their usage in standard activities in aircraft operation. The outcome should be a reduction of the number of possible failures cause by the human factor with impact on the safety in operations. Proportionately important fact is the desirable minimization in the time necessary to carry out a pre-flight inspection process, which will improve the final indicator of the efficiency in aeroplane operations.

1. Introduction

This article builds on the existing field of research by the authors on the use of UAV by maintenance and training organizations as well as commercial airlines. In this research field, the articles Use of Unmanned Aerial Vehicles in Aircraft Maintenance [1] and Unmanned Aerial Vehicles and Their Use for Aircraft Inspection [2] have been published. In aviation, defined legal regulations and standards are implemented and must be strictly applied. Procedures related to aircraft periodic inspections and checks are performed during aircraft operations by certified organizations, always supervised by the Continuing Airworthiness Management Organization (CAMO). According to Commission Regulation (EU) no. 965/2012 laying down technical requirements and administrative procedures related to air operations pursuant to Regulation of the European Parliament and Council Regulation (EC) No. 216/2008 are aircraft operators at the same time also holders of an Air Operator Certificate (AOC). Aviation companies are obliged to implement a management system in operation with working management of safety risks in whole process of operation (Safety Management System - SMS). On top of that airlines are prior to each flight expected to perform own routine pre-flight inspections. An urge to utilize new technologies and equipment that reflect changing nature of industry was a result of these rapid changes. Similarly, as industry puts an emphasis on the new technologies in accordance with Industry 4.0 principles, so must certified maintenance organizations (MROs) and continuing airworthiness management organizations (CAMOs) follow suit and utilize “Smart Technologies" [3] The collective response of maintenance and aviation industry organisations led to the creation of the concept "Smart Hangar", which enables CAMOs and MROs to conduct their business with a higher degree of objective accountability in detecting human error.

The UAV operation worldwide is growing in numbers, as are the areas of human activity in which UAV are operated professionally. They are part of all areas of human life and are already an integral part of the equipment of law enforcements and rescue units, photogrammetry, transport, construction, industry, research, logistics, they can furthermore be used for youngsters to play with. [4]. Technology evolves and changes with their increasing use. At the same time a relevant legislation for a safe utilisation should be in line with technological progress. This field must be protected by European legislation and national in a manner that will not restrict freedom [5].

[*]Corresponding Author: University of Žilina, Air Transport Department, Univerzitná 8215/1, 010 26 Žilina, Slovakia, Email: Novak30@uniza.sk

Recently Intel and Airbus have started a cooperation in a bid to further improve this process. Intel delivers UAV equipped with video recording devices that enable recording images or data which can later be used to generate whole three-dimensional models of Airbus aircraft. Airbus Aerial a subsidiary established by Airbus S.A.S provides equipment inspection services in different aviation industry sectors [6]. There are also other similar companies that operate in this market, such as French Ubisense S.A.S, MRO Drone Ltd from Great Britain and Canard Drones from Spain. These have lately developed particular procedures to provide inspection solutions for airports as well as aeroplanes [7], [8].

The goal of our project is to demonstrate the possible use of UAV system for aircraft inspection by a small aviation organization.

2. Definition of Smart Hangar concept

Smart Hangar can be defined as a set of actions including digitization of analytical tools, automation of its environment, smart manufacturing and logistics. If we should specify the concept of "Smart Hangar" in detail, it is necessary for the maintenance organization to expand key areas as they are listed in Figure 1.

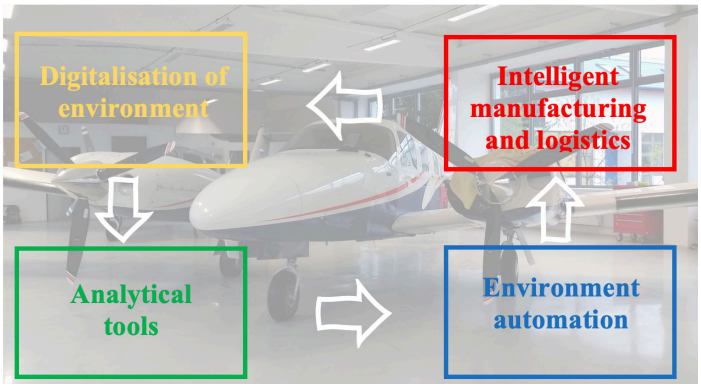


Figure 1: Concept of Smart Hangar for MRO. (Source: authors)

2.1. Digitization of environment and virtualization of tasks

In order to establish a functioning "Smart MRO" all work procedures and management systems that are effective at the moment need to be digitized to form a basic framework of the system. Maintenance organization, a continuing airworthiness management organization, hangar or even an airport that utilizes the electronic (paperless) system are able to achieve higher efficiency of operation. This can be done by online monitoring, assignment and approval of maintenance tasks. An MROs maintenance technician has access to the corporate network, therefore is able to work anytime from any location. Electronic handheld devices enable him to track, compare and record maintenance task [9].

2.2. Analytical tools of intelligent hangar

Digitization enables meaningful data analysis and improvement of internal processes. It also helps airlines optimize activities in cost planning, warehouse management and materials logistics. Employees can foresee errors in aircraft repair and maintenance and better plan their resources (both material and human). This is possible through the advanced analysis that results from data collection. The use of interactive prediction software bots not only enables better control of material logistics and supplies, it also helps to reduce costs in MROs and the amount of supplies [10].

2.3. Robotization and automation of an environment

Introduction of collaborative robots brought increased productivity, work performance efficiency and directly increased work procedures and processes. Utilization of digitized end-to-end workflow monitoring resulted in higher level of safety and quality. After establishment of such innovative work environment unmanned aerial vehicles will be able to perform aircraft visual inspections prior, during or after the maintenance tasks are carried out. Further implementation of Automated Guided Vehicles (AGV) as well as Intelligent cabinet will help in search for and point-to-point transport of tools and aircraft parts from the storage to the place of the maintenance (point-to-use) in real time [9].

2.4. Intelligent manufacturing and logistics

3D printing, free-form fabrication or rapid prototyping are terms used to describe process of additive manufacturing (AM). It describes a process of merging materials in order to create any object/product from 3D model data. Manufacturing aircraft components using additive manufacturing method of 3D Printing (3DP) is eco-friendly and presents an opportunity to optimize the design of components at fraction of costs. Adaptive manufacturing provides range of advantages compared to regular manufacturing. Waiting time when replacing specific low-volume components that need to be replaced as soon as possible can be significantly reduced. It also uses less resources compared to standard manufacturing process. On the other hand, significant engineering and certification experience is required to fully utilize such technology to meet exact requirements of national aviation authorities (or EASA) in order for new parts to be approved for installation.

3. Research on pre-flight inspection process using UAV at University of Žilina (UNIZA)

Organisation and performance of aircraft maintenance at the University of Žilina conforms to the legislation of European Union. UNIZA is an Approved Maintenance Organization (AMO) with the appropriate permits and licence that conform to the requirements of Commission Regulation (EU) No. 1321/2014 under number SK.MF06. It sets out requirements for aircraft (EASA aircraft) that meet the requirements set out in Article 2 (1) of Regulation (EU) No.182/2011 of the European Parliament and of the Council No.1139/2018. Maintenance organizations have to meet the following requirements:

- for complex engine-powered aircraft and aircraft used by air carriers licensed in accordance with Regulation (EC) No 1008/2008 and their components (Part 145).
- for aircraft other than complex engine-powered airplanes and their components (Part-M Subpart F or Part-CAO).

Aircraft at University of Žilina are maintained in line with the certification and according to the latest regulations. Only trained

and certified personnel with corresponding technician's licence is authorized to release aircraft into the service.

Pre-flight aircraft inspection is always performed in line with the approved aircraft manual as well as the organization's operations manual.

3.1. Pre-flight inspection

PA34-220T SENECA V is a modern twin-engine aircraft used by University of Žilina for pilot trainings as well as aerial work. We have selected this aircraft for the research purposes. Inspection instructions listed in the following paragraph are extracted from the flight manual of the aircraft. In addition, procedures that are unique for the aircraft type have been also considered. First, it is required to inspect the general state of aircraft. It has to be inspected from outside as well as from inside. In case of winter weather conditions any build-up of snow or even icing on the flight and control surfaces has to be removed. Functionality of all lights and flashlight has to be ensured prior to the planned night flight. Operational manual section of Airplane Flight Manual offers in-depth clarification of information regarding the pre-flight inspection procedures [12]. Inspection process of PA34-220T SENECA V aircraft must be done in order depicted in figure 2. Aircraft cabin is first on the pre-flight inspection checklist (Figure 2, Point 1). According to the flight manual, correct operation of individual lightning, power supply, navigation and communication elements has to be ensured. In addition, all mechanical and safety elements together with correct fixation of luggage and flight documentation have to be inspected as well.

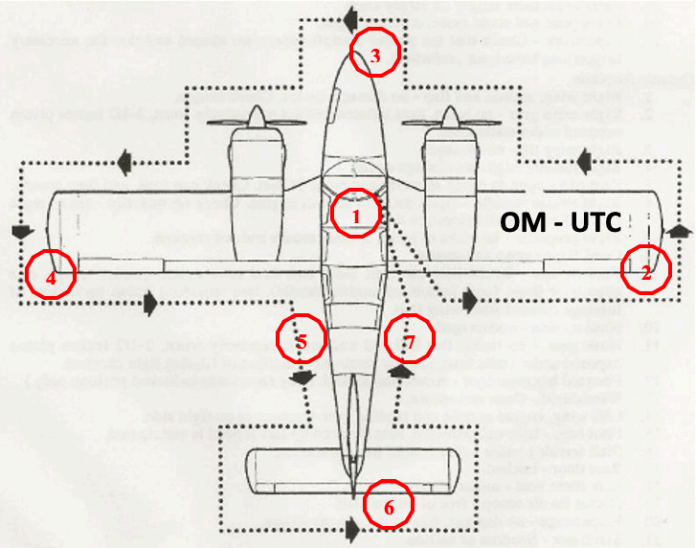


Figure 2: Inspection sequence of PA34-220T Seneca V aircraft. (Source: authors)

After this is completed an exterior inspection follows. Aircraft inspection starts with the check of the starboard wing following with a right engine in a counter clockwise direction. During this inspection every mechanical wing part such as flaps and ailerons are checked. Among other elements that are inspected are engine, propeller, amount of fuel, removal of fuel sludge as well as fuselage elements and chassis of the aircraft; at the front (Figure 2, point 3) we inspect windshield, tires, front landing gear leg, cargo compartment in the nose of aircraft together with landing reflector. While inspecting the top port wing (Figure 2 point 4), every mechanical wing parts such as flaps and ailerons and flaps are checked together with an engine, propeller, fuel quantity, fuel sludge removal, icing indication, pitot tube, chassis control and fuselage parts. Inspection then continues with the fuselage on the left side (Figure 2, point 5). This segment consists of checks of fuselage for the presence of snow and ice, antenna, cargo bay, pitot-static system and side door. Similarly, the rear part of the aircraft is inspected for the presence of any snow or ice residues. It is followed by an inspection of elevator, rudders, air intake system. Moving to the right (Figure 2, point 7) it is required to check the fuselage for the presence of snow and ice. As on the left side of aircraft antennas and pitot-static system must be checked for any debris or dirt that may hamper their operation.

4. Use of UAV in pre-flight inspection

When inspecting an aircraft using a UAV this has to be performed outside. Movement and equipment of the UAV has to meet certain requirements when operating in the Control Zone (CTR) of particular airport. These include propeller cover, outdoor movement sensors and navigation including GNSS navigation. The decisive factor for the use of UAV in an aircraft monitoring and control system is the software and subsequent applications for data evaluation.

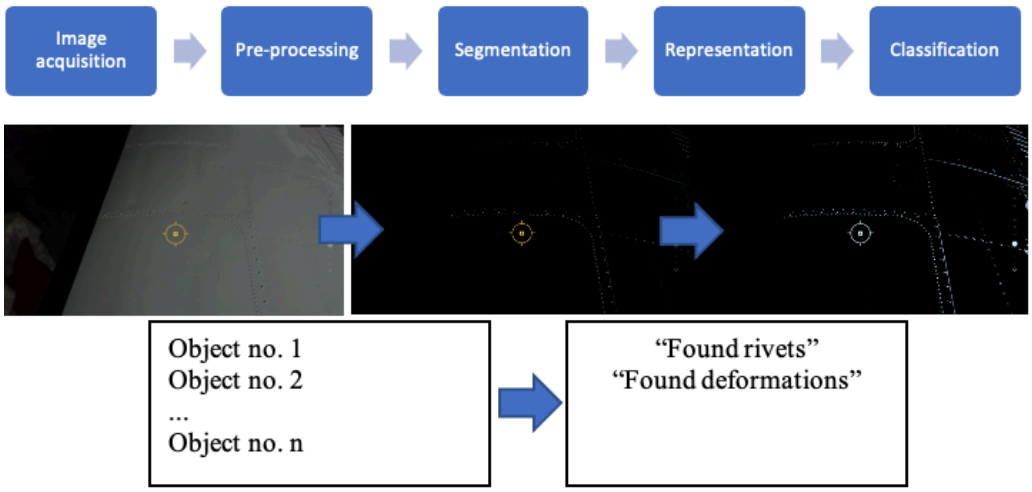
4.1. Methodology for aircraft object identification and changes

When performing pre-flight inspection, it is important to correctly recognize and identify the individual components and significant parts of the fuselage, engine and propeller blades to be inspected. Correct identification of individual components is a very important task. The recognition of any objects in general can be expressed by the following sequence of steps:

1. Image acquisition;
2. Pre-processing;
3. Segmentation;
4. Representation;
5. Classification.

After obtaining an image from a source (captured by a camera or from video), it is pre-processed, which means the image is prepared into such a form that it is possible to implement algorithms for further processing. This step can include converting a colour image into an image in shades of grey (see Figure 3), or using filters to remove, or suppress noise. Image processing algorithms are deployed in this step. The next step is segmentation, which extracts objects of interest, thus separating the foreground (what interests us) from the background (everything else). In the picture, the objects of interest are rivets on wings, but undesired objects (noise) also got there.

This is a general model, which means that not all blocks may always be followed. Their order may be changed, or the steps may be completely replaced by other procedures or combined into separate blocks. Formally, we can model the image as a continuous (image) function of two variables f(x,y), where f represents the value of brightness and (x,y) represent the coordinates on the surface. In the case of a dynamic image (video), we can also express the image as a function of three variables f(x,y,t), where *t* means time.



"Smart Hangar" solutions depend on partial systems and applications that safeguard the functioning of the whole system. Model - driven Software Development (MDSD) (see Figure 4) based on an automatic code generation of system requirements specification, it is typically used to providing specific properties to the whole system [13]. The module DeRoS safety specification DSL provides a simple and declarative syntax that makes the task of implementing the safety-related requirements accessible to robotics experts with less software engineering expertise. The risk of errors is reduced because the Declarative Robot Safety (DeRoS) declaration controls the automatic generation of all safety-related code. Our model approach directly enables implementation-independent reuse of the safety-relevant part of a robot controller between different versions, since the DeRoS declaration does not need to be changed when the underlying software changes (with the exception that names shared between DeRoS rules and component interfaces must be kept consistent) [14]. Such an approach can be utilized when enforcing safety regulations for unmanned systems or to state and improve system cross-sectional properties, as power unit energy balance, safety and timing or to supervise and organize the behaviour of number of UAV in a fleet system such as UAV swarm when inspecting more than one aircraft for example [15].

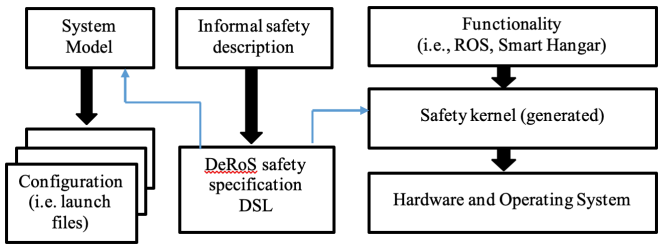


Figure 4: UAV Software platform for MDSD. (Source: authors)

PA34-220T Seneca V aircraft pre-flight inspection process has been detailed in chapter 3.1 of this article. If we are to carry out an aircraft pre-flight inspection, input variables and environmental properties, the core elements of the aircraft sensing model that identifies the state of the aircraft, need to be defined first. Next, the UAV flight and scanning trajectory, the time intervals of UAV movement relative to the defined trajectory, and the incremental termination of the scanning process must be defined and established. [16]. In the 3D-space of the proposed model, UAV will be characterised by a material point. We need to establish the start (and end) point of UAV take-off, then the altitude for individual reference points within specified trajectory, the total estimated period of its entire pre-flight inspection in the aircraft scanning phase [17, 18]. However, this describes the initial part of a complete aircraft pre-flight inspection. As soon as possible after the UAV inspection data are captured, they need to be processed and evaluated. Result of this operation should be a comprehensive information regarding the state of the aircraft which is then sent to the pilot. After inspecting the information, pilot decides whether an in-depth personal inspection of the aircraft or its parts that that deviate from the UAV sensing model is required. The pattern furthermore explains certain tasks included in the aircraft pre-flight examination, (see Figure 5) which must be carried out physically in accordance with the prescribed pre-flight inspections in the relevant aircraft documentation.

In order to create a functional model, it has to contain an information about the beginning point of UAV pre-flight inspection procedure. In the essence it surely is an exactly defined point at which the UAV take-off as well as the landing takes place. Such point is created during the planning of flight trajectory model by defining the default position. The flight route must respect the procedure of in-person pre-flight inspection carried out by a pilot as shown on the Figure 2. Model of pre-flight inspection can be seen on the Figure 6.

During the creation process of this functional model, based on the defined field of partial position points X1 to Xn; Y1 to Yn; Z1 to Zn. The initial take-off position of UAV is defined as:

$$X1 = 0; Y1 = 0; Z1 = 0 \quad (1)$$

The UAV initial position at its start should be identical to its final position after its flight performance. In order to model the entire trajectory (see Figure 6), a sequence of position points must be defined. To ensure the safety of the scanning process and to obtain relevant data, a requirement for the constant altitude of the

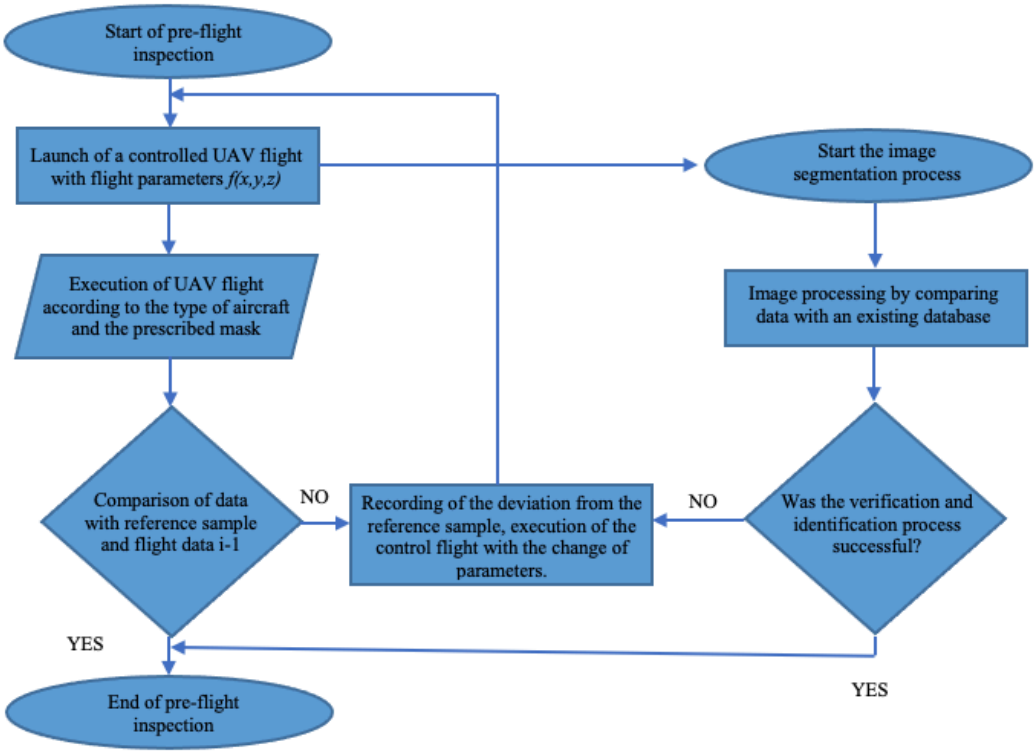


Figure 5: Aircraft inspection flowchart with image classification and evaluation (Source: authors)

UAV during the execution of the aircraft inspection was formulated. The determination of the constant altitude is based on several assumptions, namely:

- sufficient resolution value of the scanning device,
- security aspect,
- type of information obtained from the UAV.

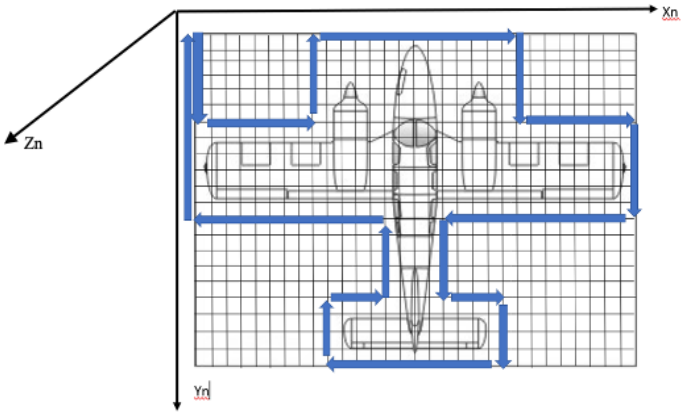


Figure 6: Model of UAV trajectory. (Source: authors)

In order to prevent potentially dangerous contacts with an aircraft it is of a paramount importance to design a so-called restricted no-go zone for UAV. Otherwise, the damage an aircraft may sustain can translate to considerable operational and financial consequences as a Figure 7. The minimum distance from any part of the aircraft has been set at:

$$\Delta l_{min}\,(Xn, Yn) = 0.5\text{m} \quad (2)$$

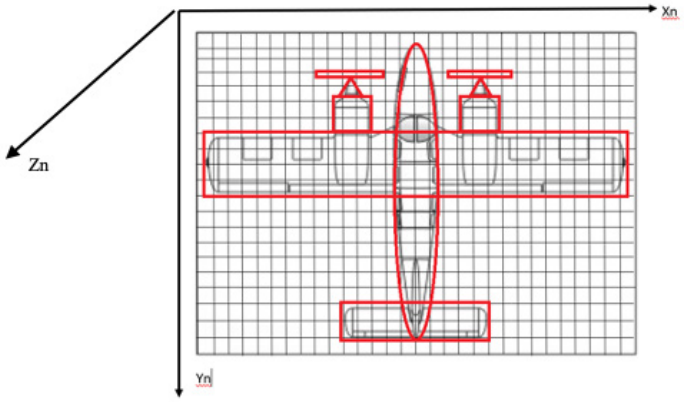


Figure 7: Restricted no-go zones for UAV. (Source: authors)

After the default starting position is defined, other reference points, meaning points where the flight trajectory has to be altered, are programmed. Each point is determined by its position:

$$B1\ (X_1,Y_1,Z_1);\ B2\ (X_2,Y_2,Z_2);\ \ldots\ldots\ Bn\ (Xn,Yn,Yn) \quad (3)$$

$$D = D_{x,y} = \sqrt{(Xi - Xj)^2 + (Yi - Yj)^2} \quad (4)$$

Where Xi and Yi determine the i-th point B, Xj and Yj coordinates present position of j-th point B of the set path of movement. The final trajectory is specified by reference coordinates and distances between determined UAV positions. The trajectory must be defined by data describing proposed restricted area around the aircraft. Suggested UAV flight trajectory model stands as:

$$D = D_{x,y} = \sqrt{(Xi - Xj)^2 + (Yi - Yj)^2}, (\forall B \notin RZ(P \in Pi, Pj)) \quad (7)$$

The B points are belonging to red zone (RZ) is described by points with coordinates "Pi" and "Pj" in such form. Model which defines the final UAV trajectory around aircraft is relevant and in such form also applicable in the phase of programming the UAV. Significance of this step lays in the mitigation of the damage that may be caused by a possible clash between the inspected aircraft and the UAV. Data gathered by the UAV during the pre-flight inspection must be processed. The processing takes place in the form of a comparison of the obtained data with an aircraft reference model. The reference model must meet all requirements placed on real flight aircraft condition before each flight inspection begins.

4.3. Discussion

The use of UAV during pre-flight inspection brings benefits in several areas. By eliminating the influence of the human factor, which can reduce the level of safety of the aircraft, it is possible to achieve an increase in safety. The use UAV pre-flight inspection tour around the aircraft following the manufacturer's checklist. It's necessary pay special attention to lose or "smoking" rivets (which have a residue around them), the safety of all bolts and nuts, and safety-wired devices. At the same time, it is possible to increase efficiency by introducing automation. UAV application could support both aviation and airport operations, it would decrease the maintenance costs and aircraft inspection procedures. [19]. Like other technologies, UAV bring a wide range of benefits. But besides the benefits, the sad truth about UAV is that there are certain burning flaws like sudden crashes, risks from hacking, and security and privacy issues that plague this part of the technology. [20]

Another aspect of the UAV pre-flight inspection system is the development of a UAV communication system with a hardware-software platform, which will be designed to automate the processes of processing and transfer of information between involved parties so that there are no security risks associated with the transfer and processing of acquired information. Although the use of UAV in any process in air transport is a very complex and lengthy process, its potential benefits are undoubtedly clear.

5. Conclusion

Significant technological developments in the field of UAV have recently caused conditions to be created for their use in specific areas of civil aviation. There are ever more use cases in this area for use of UAV, despite being perceived as a threat to civil aviation. The philosophy of Smart Airport consists of several advanced concepts, one of which is the use of UAV in the process of pre-flight inspections, this was also the subject of the presented article. Firstly, it was crucial to analyse and define technical and operational conditions for the UAV use in pre-flight inspection by detailed analysis of the environment, determining the nature of information that will form a set of input and output information and data and then creating a system to process and evaluate the data. The initial idea of creating this model of UAV application during a pre-flight inspection was to evaluate the overall aircraft condition. Potential threats were also described and a way of solving them was proposed. In the end, advantages of UAV application in the process of pre-flight inspections have been mentioned. Among the most important are the reduction of human factor influence in the performed pre-flight inspection, the time reduction allocated for this activity, which achieves a higher efficiency of the whole process. Further research actions would lead to a detailed definition of aircraft model reference samples for the purpose of obtained records comparison with the reference model. Furthermore, it will be required to develop a swarm model for UAV devices where several devices with diverse types of equipment (mechanical arm, laser scanner, digital camera, etc.) would cooperate with each other. Thanks to the use of modern digital and transmission technologies, it is possible to ensure real-time data comparison. However, it is always necessary to ensure the maximum level of safety of these processes when using UAV during the pre-flight inspection

Upcoming efforts will extend the experimental assessments to UAV simulations and actual flight tests to use probability numbers instead of frequencies which have been used in this case for demonstration purposes. In the author´s opinion and experience it is rather difficult to accomplish the certification and validation processes of safety functions as well as more efficient outcomes. For that reason, our strategy is to verify and validate already functioning safety system to guarantee that the safety system meets functional requirements and is suitable for risk mitigation. Besides, we will investigate how the UAV swarm can be implemented to monitor and ensure a safe flight operation of pre-flight inspection during this specific mission for example in aerodrome apron.

Acknowledgements

This publication was realized with support of Operational Program Integrated Infrastructure 2014 - 2020 of the project: "Intelligent operating and processing systems for UAV, code ITMS 313011V422", co-financed by the European Regional Development Fund".

References

[1] A. Novák, A. Novák Sedlackova, M. Bugaj, B. Kandera, T. Lusiak, "Use of Unmanned Aerial Vehicles in Aircraft Maintenance", Transportation Research Procedia, **51**, 160-170, 2020, DOI 10.1016/j.trpro.2020.11.018

[2] M. Bugaj, A. Novák, A. Stelmach, T. Lusiak, "Unmanned Aerial Vehicles and Their Use for Aircraft Inspection," 2020 New Trends in Civil Aviation (NTCA), Prague, 45-50, 2020. doi: 10.23919/NTCA50409.2020.9290929

[3] J. Holl, "Hangar of the future Excelling in MRO", 6. December 2016, Airbus S.A.S Source: https://www.airbus.com/newsroom/news/en/2016/12/Hangar-of-the-future.html

[4] P. Kurdel, A.N. Sedláčková, J, Labun, "UAV flight safety close to the mountain massif",Transportation Research Procedia, **43**, 319-327 2019. doi:10.1016/j.trpro.2019.12.047

[5] A.N. Sedláčková, P. Kurdel, B. Mrekaj, "Synthesis criterion of ergatic base complex with focus on its reliability", INFORMATICS 2017 - Proceedings, 318-321, 2018. doi:10.1109/INFORMATICS.2017.8327267

[6] M. Fendt, "Airbus launches advanced indoor inspection drone to reduce

aircraft inspection times and enhance report quality, Commercial Aircraft", Airbus S.A.S , 2018. (Accessed: 10 April 2021)
[7] A. Ward, "Ubisense and MRO Drone launch world's first 'Smart Hangar' solution", Ubisense 18.April 2018 Source: https://ubisense.com/ubisense-and-mro-drone-launch-worlds-first-smart-hangar-solution/
[8] V. Ažaltovič, I. Škvareková, P. Pecho, B. Kandera, "Calculation of the ground casualty risk during aerial work of unmanned aerial vehicles in the urban environment", Transportation Research Procedia, **44**, 271-275, 2020. doi:10.1016/j.trpro.2020.02.043
[9] A.N. Sedláčková, P. Kurdel, J. Labun, "Simulation of unmanned aircraft vehicle flight precision", Transportation Research Procedia, **44**, 313-320, 2020. doi:10.1016/j.trpro.2020.02.037
[10] ST Engenniering: Smart MRO, internet 1.7.2020 https://www.stengg.com/en/aerospace/capabilities/smart-mro/
[11] Uniting Aviation, "The future of MRO: Emerging Technologies in Aircraft Maintenance", Uniting Aviation on August 1, 2019. (Source: https://www.unitingaviation.com/amp/news/capacity-efficiency/the-future-of-mro-emerging-technologies-in-aircraft-maintenance/) (Accessed: 2 April 2021).
[12] Piper AIRCRAFT, "Airplane Maintenance Manual, Card 1 of 5, PA34-220T, SENECA III (ALL) and SENECA V (S/N's 3448038 THRU 3448079), Piper Aircraft Corporation, 2017, April 17, 2007.
[13] M. Catlos, P. Kurdel, A.N. Sedlakova, J. Labun, M. Ceskovic, "Continual monitoring of precision of aerial transport objects", Paper presented at the NTAD 2018 - 13th International Scientific Conference - New Trends in Aviation Development, Proceedings, 76-81, 2018. doi:10.1109/NTAD.2018.8551683
[14] A. Sorin, J. Kjeld, U.Schultz, "Towards Rule-Based Dynamic Safety Monitoring for Mobile Robots", Journal of Software Egnieering for Robotics, 2016. DOI: 10.6092/JOSER_2016_07_01_P120
[15] F. Škultéty, B. Badánik, M. Bartoš, B. Kandera, "Design of controllable unmanned rescue parachute wing", Transportation Research Procedia, **35**, 220-229, 2018. doi:10.1016/j.trpro.2018.12.026
[16] Min Xue "UAV Trajectory Modeling Using Neural Networks", 17th AIAA Aviation Technology, Integration, and Operations Conference. AIAA 2017-3072, 2017. doi: 10.2514/6.2017-3072
[17] H. Zhou, H.L. Xiong, Y. Liu, N.D. Tan, L. Chen, "Trajectory Planning Algorithm of UAV Based on System Positioning Accuracy Constraints. Electronics **9**, 245-250, 2020. https://doi.org/10.3390/electronics9020250
[18] L. Tan, J. Wu,X. Yang, S. Song, "Research on Optimal Landing Trajectory Planning Method between an UAV and a Moving Vessel", Appl. Sci. 2019, **9,** 3708, 2019. https://doi.org/10.3390/app9183708
[19] F. Marcellin, "Positive uses of drones in aviation: UAV changing airports for the better", Available at: https://www.airport-technology.com/features/positive-uses-of-drones-in-aviation/ (Accessed: 4 April 2021).
[20] A. Allouch, A. Koubâa, M. Khalgui, T. Abbes, "Qualitative and Quantitative Risk Analysis and Safety Assessment of Unmanned Aerial Vehicles Missions Over the Internet", IEEE Access, **7**, 53392-53410, 2019. doi: 10.1109/ACCESS.2019.2911980.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 185-202 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

iDRP Framework: An Intelligent Malware Exploration Framework for Big Data and Internet of Things (IoT) Ecosystem

Osaretin Eboya*, Julia Binti Juremi

Asia Pacific University (APU), Faculty of Computing, Engineering, and Technology (FCET), Kuala Lumpur 57000, Malaysia

ARTICLE INFO

Article history:
Received: 24 June, 2021
Accepted: 20 September, 2021
Online: 30 September, 2021

Keywords:
Internet of Things (IoT) Security
iDRP framework Security
Integrated Security framework
Intelligent Network Security
Intelligent Malware Detection
Intrusion Detection System
Big Data Security
IoT iDRP framework
IoT Malware Security
IoT Botnet Security
Anomalies Detection Techniques
Multilayer Perceptron (MLP)
Deep Neural Network (DNN)
Deep Learning (DL)
Machine Learning (ML)
Artificial Neural Network (ANN)

ABSTRACT

The Internet of Things (IoT) is at a face paced growth in the advanced Industrial Revolution (IR) 4.0 in the modern digital world. Considering the current network security challenges and sophistication of attacks in the heavily computerized and interconnected systems, such as an IoT ecosystem, the need for an innovative, robust, intelligent and adaptive malware attacks and threats security solution is becoming predominant in the current cyberspace. An integrated and scalable IoT malware detection framework called iDRP framework with deep learning method was proposed as a solution to current IoT malware attacks that are largely obfuscated. The novel framework utilized systematic pre-processing and post-processing techniques and methods on the BoTNetIoT malware datasets that contains both benign and malicious IoT traffic data infected by modern day IoT attacks such as Mirai and Gafgyt etc. IoT malware variants in an IoT ecosystem. The raw IoT malware binaries were converted to image files (Gray-scaled) and computed statistically with synthesised sparsed and differential evolutionary hidden feature structures techniques, which were cyclically trained, tested, and cross-validated to establish empirical anomalies with precision in the detection, recognizing, and prediction of malware anomalies in a modern IoT ecosystem. Preliminary experiments were conducted with standardized image binary files such as the MNIST (2-D), and NORB (3-D) datasets as sound scientific exploratory experiments with profound results. The comparative results of the performance of our integrated techniques and methods on the BoTNetIoT IoT malware datasets achieved a 99.98% accuracy, 99.99% ROC/AUC, 99.95% precision, and 99.93 recall rate etc. utilizing the integrated iDRP framework mechanisms for effectively detecting IoT malware in an IoT ecosystem.

1. Introduction

This paper is an extension of the work originally presented in *IEEE 8th R10 Humanitarian Technology Conference (R10-HTC) 2020* [1]. The Internet of Things (IoT) has in recent times emerged as ubiquitous technology to everyday lives especially with the advent of the Industrial Revolution (IR) 4.0 in the digital world. The IoT technologies comprises of smart devices and objects interconnected in a heavily computerized network environment that constitutes the backbone of modern innovative critical infrastructures that supports fast-paced technological driven world. The IoT technologies has been disruptive in many industries such as: (1) IoT in Healthcare, (2) IoT in Manufacturing, (3) IoT in Transportation and Mobility, (4) IoT in Buildings, (5) IoT in Cities, (6) IoT in Agriculture, (7) IoT in Energy, (8) IoT in Retail and Marketing, (9) IoT in Logistics and Supply Chain, and (10) IoT in Industries across various sectors of the economy. Similarly, IoT technologies has brought diverse approaches in solving mundane tasks in these sectors by leveraging and amalgamating innovative and intelligent automation systems. Additionally, the IoT technologies at the present time has been able to help connect multiple devices and objects in an expansive interconnected network area while collecting high treasure trove of data – *big data*. The big data collected in an IoT ecosystem has become paramount to making business decisions, analytics, and gaining valuable insights in multidimensional industries in the modern era. Furthermore, this big data that are continuously generated in an a heavily interconnected environment such as the IoT ecosystem has inadvertently set off more attractions and incentives to the cyber squatters and cyber criminals in a digital world – *Cyber hacking*. The security of IoT ecosystem is ever more crucial and concerning to safeguarding critical infrastructures in the smart environments against prolific attacks and state-of-the-art attacks – *Cyber*

*Corresponding Author: Osaretin Eboya, Email: osaretin@ieee.org

https://dx.doi.org/10.25046/aj060521

intrusions. For example, in recent time there have been a number of cyberattacks that caused outages and losses such as the (https://www.msn.com/en-my/news/world/dc-police-suffer-massive-info-leak-after-ransomware-attack/ar-BB1gHTQl) attack and (https://www.msn.com/en-my/news/world/european-hackers-given-us-5-million-for-key-to-reopen-us-pipeline/ar-BB1gIq7K) outage coupled with the public health (https://www.msn.com/en-my/news/world/ireland-tests-cyber-attack-data-fix/ar-AAKe3Qd) compromise disruptions on critical infrastructures that serves an entire community in the modern society. This is one of the most significant attacks in recent time with massive consequences. The cyberattacks on critical infrastructures evidently indicates how vulnerable and compromising the present-day heavily computerized network systems can be, and consequently disruptive and detrimental to everyday lives in a digital world. Moreover, this clearly indicates that most *modern* highly interconnected critical infrastructures serving every day processes in the most developed world can be highly susceptible and vulnerable to being hacked with a myriad of modern malware attacks such as Ransomware attacks and '*Zero Day*' attacks. Theoretically, in the modern digital world, any device or object that can connect to a network system or grid can be hacked.

Considering the general consensus in the cybersecurity space that humans are usually the weakest link in the chain of command in the cybersecurity and network security effort, it is effectively safe to deduce that the current static, dynamic, and automated malware defense approaches against cyberattacks and cyber threats in an IoT ecosystem in a digital world is essentially a futile effort. Particularly in the bountiful cyberattacks and cyber threats warfare in the modern cyber space clearly indicates the fragility and exposure of current interconnected critical infrastructures and its related network systems. This has essentially led to the theory that *detection* is a better approach to combating cyberattacks and cyber threats in a smart ecosystem of massively interconnected computer network systems – *malware detection* [1]. Malwares are typically spread over an interconnected computer network system otherwise known as the internet, which is the core backbone of an IoT ecosystem. The lack of global unified principles and security of IoT protocols present major security and counter-intelligence conundrum on a daily basis in the IoT cyberworld. This inadvertently exposes various valuable resources and assets in an organization to the cybercriminals. The IoT ecosystem can be attacked, infected, and breached by different types of malicious software commonly known as malwares with categorized malware families such as: (a) Ransomwares, (b) Spywares, (c) Scareware, (d) Adware, (e) Viruses, (f) Worms, (g) Rootkits, (h) Botnets, (i) Trojan Horses that all threatens the *integrity*, *privacy*, and *security* of big data in a heavily computerized computer network system like an IoT ecosystem – *tampering of valuable assets; big dataset.* One of the major problems [1] with these malware families and their nefarious activities combined with innovatory technologies is that they have become even harder to detect with traditional and automated malware detection tools and techniques in a heavily computerized network system – *polymorphic malwares.* If and when potent malwares go undetected in a network system such as an IoT ecosystem, this can cause massive damages and disruptions to critical computer network infrastructures [1], loss of personal information, and potentially loss of livelihood and lives in a modern technological-reliant society – *obfuscation and evasion of malware detection.*

We proposed a novel IoT Security framework known as the *intelligent*, *detection*, *recognition*, and *prediction* (iDRP) framework [1] to address these computer network security challenges and problems in the modern digital world. The proposed iDRP framework is an innovative, intelligent, robust and adaptive framework with ultramodern approach to solving malware anomalies in an IoT ecosystem. In the unique approach proposed, a synergy of the subsets of Artificial Intelligence (Ai); Machine Learning (ML) and Deep Learning (DL) will be synthesized with modularity and scalability implementation for a lifelong training and learning capabilities. Eventually, the goal of the research work is to effectively detect, recognize, and predict anomalies in an IoT ecosystem innovatively using a subset of Artificial Neural Network (ANN) with focus on Multilayer Perceptron (MLP); *hierarchical* DL techniques i.e., Deep Neural Network (DNN) to solving the wide range of polymorphic malware anomalies problems in an IoT ecosystem.

2. Related Work

The drawback of heavily depending on the storage of big data set in one or more interconnected storage medium is that adversaries now have the perceived knowledge and understanding of where to target in their quest in wreaking havoc on present-day interconnected network systems. In the event that proper anomaly analysis mechanism for countering the attacks on the treasure trove of big data is not put in place, then, it means that the assailants will in most cases get away with their attacks either undetected or worst-case scenario, reside on the host system while plunging the gigantic and diverse data in the system [1]. The lack of highly-developed, efficient, robust security apparatus with well-established intelligent processes that can accurately measure and predict future occurrences of such attacks could be disastrous to the specific individual, organization, or government alike. The general IoT Security challenges range from Authentication, Access Control, Privacy, Confidentiality, Integrity, Policy Enforcement, Trust, Secure Middleware, Mobile Security [2],[3] [4], and Network Security. The complexity of the interactivity between the gigantic number of smart devices and objects together with smart network systems give rise to a major security gap in the IoT ecosystem. This poses big security risks due to the incessant and sophisticated cyber-attacks in the modern internet age – *IR 4.0*. A common denominator amongst all the aforementioned IoT Security challenges is the Distributed Denial of Service (DDoS) attack. The DDoS attacks occur on major global establishments on a regular basis [5]-[7]. Several peculiar IoT Security constraints such as integrity, availability, authentication, authorization, and privacy have hampered the efforts to address the growing IoT Security challenges posed by malicious actors.

The malicious attacks on large corporations such as Fortune 500 enterprises have clearly shown that the security of internet-connected devices and objects is paramount considering the fact that the particular DDoS Malware attack called the "Mirai Botnet" crippled even the network servers of these largest technology corporations. Some of these large corporations are situated in the United States and major European countries [8], (https://www.theguardian.com/technology/2016/oct/26/ddos-

attack-dyn-mirai-botnet), and (https://www.wired.com/story/reaper-iot-botnet-infected-million-networks/). This type of nefarious malware attacks contributes to the urgent needs for a viable solution to the lingering IoT Security challenges [6], [9], [10] with the ever-growing IoT inter-connected network environment. According to [8], [11], [12]. One of the security challenges of smart devices and objects today is the inability to *detect* when there is an intrusion in the IoT network system. This is largely due to the fact that most smart devices lack the capability to log and report when such malicious attack occurs in the IoT ecosystem. IoT Security needs to be improved and revolutionized to prevent, identify, and neutralize the malicious network traffic threats on the Billions of IoT connected servers. The exponential ramifications of the risks and costs that the malicious attacks on the IoT infrastructures pose [13] to businesses, governments, and individuals who rely on IoT services such as Healthcare, electricity, industrial production, social amenities, communications, and logistics is still unknown. Nevertheless, the researchers discovered that, the two major challenges that an IoT ecosystem currently face are; Low-end Embedded Devices [6], [8], [14] with insecure Operating System (OS), and Mirai-Styled Malware Attacks [10] otherwise known as Mirai Botnet, which is a modern and sophisticated variation of a Distributed Denial-of-Service (DDoS) attack on the current precarious IoT environments and its smart devices and objects.

2.1. IoT Attacks

IoT applications are prone to attacks from different hostile actors. The multidiscipline industries that adopt IoT as a technology have their own variation of attacks on their IoT applications and infrastructures. The concept of sensors connecting to each other and transmitting big data in a two-way traffic in an IoT ecosystem typically creates vulnerabilities [13], [15]-[17] in the IoT network system. In the Smart City ecosystem in Ukraine that utilizes Smart energy grid system to serve the city, 30 of such Smart energy were attacked [8], [16], and (https://www.thenational.ae/uae/smart-cities-open-door-to-cyber-attacks-say-security-experts-1.672214), and 80,000 residents were without power supply for several hours due to the adversarial malicious attack on the Smart Grid system. In the United States, 156 emergency sirens designed to alert residents of the city of Dallas [8], [18], and (https://www.nytimes.com/2017/04/08/us/dallas-emergency-sirens-hacking.html) in the event of critical emergencies such as earthquakes, floods, and nuclear attacks were set off because of an attack [2] on their Smart City system. The impact of such false alarm set off in such a malicious manner can be consequential in the end. There is a high likelihood of the residents of the city to ignore a true alarm warning in the event of an actual emergency, which could potentially lead to loss of lives and livelihood [2]. Similarly, 2,200 residents were without water supply (https://www.wired.com/2011/11/hackers-destroy-water-pump/) when hackers attacked and destroyed the Smart Water Pump Supply system in the municipality of Springfield, Illinois, United States.

The malicious attacks on IoT applications are not limited to Smart Cities and critical infrastructures alone, but also Smart Automobile compute system where attackers targeted a popular Smart Car maker and hijacked its core automation compute system that powers their connected cars for their financial gains. The financial gains, also known as "*Cryptojacking*" [13], [18], [19] where hackers use illegally acquired computational resources to mine the cryptocurrency shows that hackers will launch attacks on IoT applications for monetary reasons. In Smart Healthcare system [13], attackers have successfully hacked [6], [8] a heart monitoring electronics health (e-health) infrastructure that intelligently monitors and analyses patient's heart rate and helps prevent cardiac arrest and heart attacks [20]. Furthermore, the security of a Smart Home system that helps parents monitor their infants was compromised when attackers [6], [8] were able to obtain the login credentials of the Smart Home system and about 700 Webcam feeds of babies were posted on the internet [13], [20] in a malicious attempt by the hackers. The security breaches ultimately led to privacy [20] on the victims of the IoT attacks. All these incessant IoT attacks [6], [13] indicate that IoT devices and objects are more vulnerable and susceptible [19], [21] to various hostile adversaries who are constantly looking for weaknesses [16], [18] in the IoT infrastructures and IoT software applications

2.2. IoT Malwares

Malicious software and worms are very devastating to the IoT applications' ecosystem. Malware such as worms pose greater threats to the security of IoT applications. Some of the known malware worms are Mirai Botnets [22], Ransomware [20], [23], Over The Air (OTA) Worm [21],[24], and Struxnex Worm. According to the case study and analysis conducted by [18], discovered that Mirai Botnet can be directed towards the disruption of the availability of the targeted IoT application and the IoT compute resources in the IoT environment. In addition, [8],[23] classified the Ransomware as primarily Crypto Ransomware, and Locker Ransomware, which encrypts and locks critical files, respectively in an IoT ecosystem have increased by a staggering 670% and 350% between 2015 and 2017 respectively and continue to grow exponentially nowadays. The implications of such malicious software are severe and could disrupt and harm crucial IoT infrastructure and potentially cripple an entire IoT ecosystem as evident [13],[19] in the coordinated attacks by hackers on the key connected internal applications in the city of Atlanta, United States.

The OTA Worm exploits vulnerabilities in the IoT smart devices and objects and spreads in the IoT environment by compromising the integrity, privacy, and security of the interconnected devices and objects therein. The OTA Worm embeds itself into the IoT ecosystem thereby causing a breakdown of the IoT ecosystem referred to as an "epileptic seizure" [6], [16], [24] of the specific IoT smart devices and objects at a close range of 350 meters to the building with the help of a drone or Unmanned Arial Vehicle (UAV) equipped with the malicious worm. Moreover, the mode of launching the OTA Worm attack is also carried out with mobility such as physically getting close or driving cars to a close proximity of the IoT subject. Even though the OTA Drone Worm [6], [16], [24] was successfully launched on a major Smart home appliances manufacturer's Smart bulb product line, it was, however, performed in a controlled environment and the concerned manufacturer of the defected IoT Smart Light Bulb was notified, and the appropriate fix was implemented, which now reduces the possible attack range to 1 meter [6], [16], [24]. However, the risk of exploitation by enemies is present. This shows that malicious adversaries are becoming more sophisticated

in their approach towards unleashing malicious worm attacks with technological advancement. Furthermore, the Struxnex Worm, which is a variant and obfuscation of malicious worm was directed at Smart Industry such as the Uranium plant [6] connected to an IoT infrastructure to intelligently monitor and analyse the increment and decrement of the acceleration of the centrifuge and enrich the Uranium plant as the need arise autonomously. The main goal of the Struxnex worm is to infect and break down the smart devices and objects in a Smart Industrial ecosystem [18] such as a Uranium plant.

2.3. IoT Existing Security Challenges and Gaps

The core existing gap in the IoT Security paradigm is IoT security, IoT privacy [3], [4], [17], [19]. However, this research has gone further with the IoT security and IoT integrity dimension. The importance of anomaly detection in an IoT ecosystem cannot be overemphasized in various IoT industries such as Smart Healthcare, Smart Education, Smart Agriculture, and Smart manufacturing, and Smart Buildings and Smart Cities [13], [25]-[28]. In the IoT Security dimension, we explored and discovered that research has been done on the Classification and Visualization of IoT security [20], [21], [29] in which the researchers performed experiments to visualize and compare different classes of malware in a similarity matrix table based on the images and features of some malware samples such as Worms and Trojans Horses. Their research work can be improved further to increase the accuracy of the detected similarities of the various malware samples and increase the performance of the workload on big data. Moreover, the experiment was conducted using 'mechanical' compute method, which means that their method could possibly detect existing malware at a significant slow pace, coupled with the fact that, their classification and visualization method cannot intelligently predict future occurrences [13], [19], [20] of such IoT malware anomalies autonomously.

3. Methodology

We present the schematic illustration process of the work proposed iDRP framework, including the realization of the universal approximations of Multilayer Perceptron (MLP) with multiple hidden layers by exploiting the capabilities of H-ELM and other related techniques [30]-[33], which have the preeminent underexplored potentials for accelerated speed, rapid feature learning, and improved classification performance [30], [31], [34], [35].

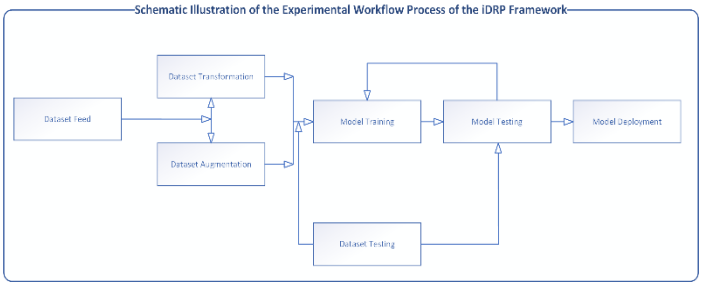


Figure 1: Schematic Illustration of the Experimental Workflow Process of the iDRP Framework

Figure 1 Shows the schematic illustration of the setup of the experimental workflow process of the iDRP framework. The big dataset will be fed into the system, when then gets transformed by various MLP techniques such as augmentation, normalization and classification techniques, and subsequently prepared for the algorithm. Feature extractions will be carried out by augmenting and fine-tuning the big dataset for the generalization and capability of the proposed iDRP framework.

The experimental implementation is to demonstrate the proof-of-concept of the proposed iDRP framework using the generalized MLP techniques for demonstrating how a non-traditional malware detection mechanism [32], [33] can be applied to systematically, dynamically, adaptively, and intelligently detect, recognize, and predict malware anomalies in the generated big data of an IoT ecosystem through converted image file, such as a log file, and binary dataset to create synthesized Machine Learning (ML) and Deep Learning (DL) models. Based upon the premise that attackers are getting more sophisticated with the advancement of technology every day, and considering that attackers have been developing malwares that can dynamically modify their signatures while in the network environment, which approximately makes it more difficult to detect. This fundamentally relegates the notion of applying the traditional and automated Signature-Based Analysis method to solving a potential attacks and threats in an IoT ecosystem. Furthermore, the Heuristic-Based Analysis method, which is otherwise known as Behavioural-Based Analysis technique applied by several Anti-Virus and Anti-Malware vendors on computer network systems has not been sufficient either. Hence, there are various gaps [19],[20],[36] in the field of study to develop a novel approach to solving the inherent and existential security threats in computer network systems, especially in the fast-growing IoT ecosystems that has now become ubiquitous to everyday life. The hypothesis is that nowadays, contaminated and malicious programs in an IoT ecosystem can easily masquerade themselves as benign software applications once in the network space. Therefore, a dynamic, robust, and intelligent Pattern-based model approach to discerning and distinguishing the genuine and malicious program in a system as counter-intelligence is paramount as the building block to proactively combating an unwanted and illegitimate program from executing and consequently causing some irreversible damages and potentially loss of lives in an IoT ecosystem with its critical infrastructures.

To explore the binary classification and analysis techniques, we will be using the standardized Modified National Institute of Standards and Technology (MNIST) dataset (http://yann.lecun.com/exdb/mnist/), which has been demonstrated to have the capability to be embedded with malicious codes (virus-MNIST) (https://arxiv.org/abs/2103.00602) adaptively, NORB (https://cs.nyu.edu/~yann/research/norb/) to perform the experiment by encoding the dataset, which will serve as a standard benchmark for comparative training, testing, evaluation, and analysis of the experiment.

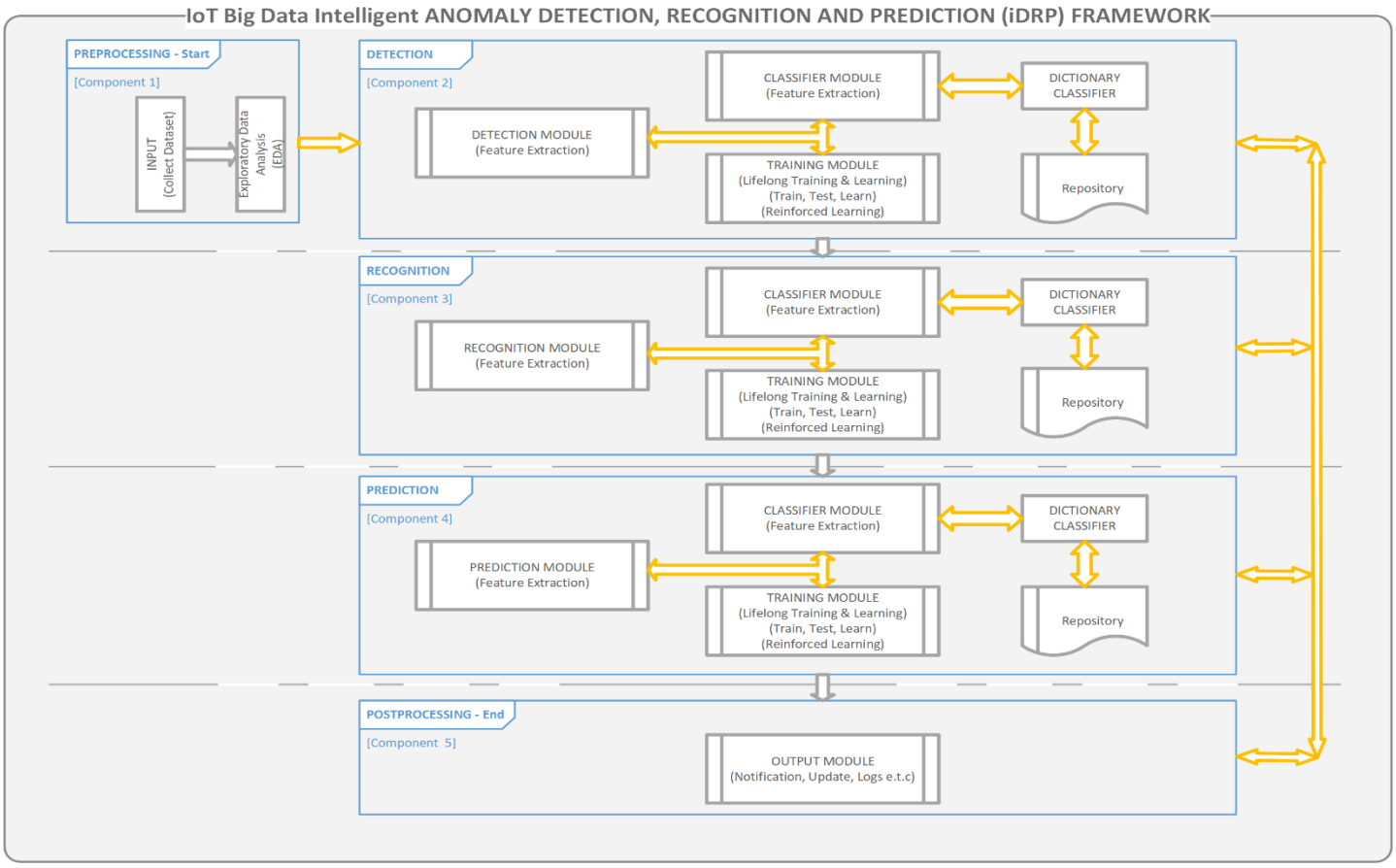


Figure 2: Proposed iDRP framework for Anomaly (Malware) Identification in IoT ecosystem – Overview [1]

4. The Proposed iDRP Framework Architecture

This section provides the analysis of the proposed Intelligent anomalies detection, recognition, and prediction (DRP) also known as iDRP-framework as shown in Figure 2. The conceptualization of the proposed iDRP-framework is based on the relationship between feature extraction, feature classification, precision in the identification of anomalies, and predictions of anomalies in the generated dataset of IoT ecosystem. Table 1 and 2 displays the synthesis of the list of the development techniques comparatively in tandem with the corresponding frameworks' techniques to be synthesized and applied for the development and validation of the proposed intelligent framework. The information gathered in the literatures of this research study has clearly shown that there still exist gaps in effectively tackling the myriad of anomaly detections, identifications, and recognition in complex data sets generated in an IoT ecosystem.

The limitation in the forecasting of anomalies (malwares) in the IoT ecosystem forms the spine of this research study and consequently the creation of an Intelligent framework called iDRP-framework to address the issues of anomalies (malwares) identification in an IoT ecosystem. In particular, the ability to accurately Detect, Recognize, and Predict (DRP) anomalies (malwares) from logging big datasets generated in an IoT ecosystem cannot be overemphasized.

Evidently, the extraction of miniscule features of the big datasets to assist in the '*fine-grained classification*' of features of complex datasets has been demonstrated by other researchers [30],[36] to help in the precision of anomaly identification in an IoT ecosystem.

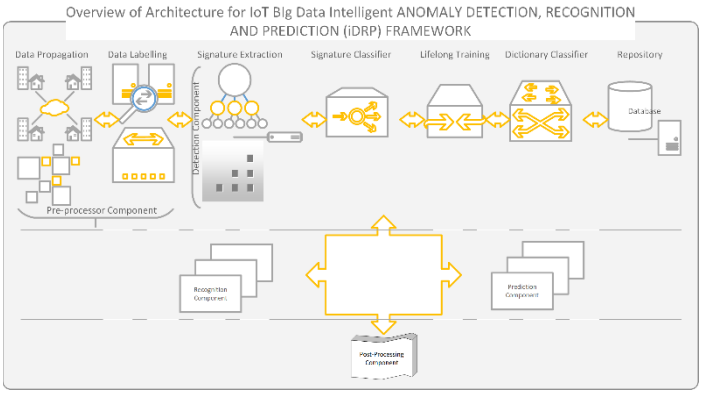


Figure 3: Proposed iDRP framework for Anomaly (Malware) Identification in IoT ecosystem – Architecture Breakdown

Figure 3 illustrates the overview of the proposed iDRP framework architecture with the "*Data Augmentation*" segment consisting of data propagation, data labelling, feature extraction, and feature representation etc. The proposed iDRP framework architecture is flexible, robust, and adaptive in its design to intelligently and effectively extract mapped minuscule information in the stacked layers of the proposed iDRP framework for detecting, recognizing, and predicting malware binaries in an IoT ecosystem.

Table 1: Synthesized Components of Development and Validation of the Proposed iDRP Framework I [1]

No	Comparative Features of Synthesised Frameworks I							
	Features	H-ELM	SAPIM	Smart Weather	Flight Monitoring	eTRIKS	Data Lifecycle	DRIPROM
1	Pre-processor	✓	×	✓	✓	✓	✓	✓
2	Detection	✓	✓	✓	✓	✓	✓	✓
3	Recognition	✓	×	×	×	×	✓	✓
4	Prediction	×	✓	×	×	×	×	×
5	Post-processor	×	×	×	×	×	×	×
6	Lifelong Learning	×	×	×	×	×	×	×
7	Lifelong Training	×	×	×	×	×	✓	×
8	Cyclic Feed	×	×	×	×	×	✓	×

Key: H-ELM (Tang, Deng and Huang, 2016), SAPIM (Wang et al., 2018), Smart Weather (Onal et al., 2017), Flight Monitoring (Li, Ming and Li, 2017), eTRIKS (Oehmichen et al., 2017), Data Lifecycle (Arass, Tikito and Souissi, 2018), DRIPROM (Cuzzocrea and Damiani, 2018).

Table 2: Synthesized Components of Development and Validation of the Proposed iDRP Framework II

No	Comparative Features of Synthesised Frameworks II						
	Features	IoT-HarPSecA Framework	3-Way IoT Framework	IoT-Flock Framework	FIFAC Framework	BiDeL Framework	Combat Intelligence Framework
1	Pre-processor	✓	×	×	×	✓	✓
2	Detection	✓	✓	✓	✓	✓	✓
3	Recognition	✓	×	×	×	×	✓
4	Prediction	×	✓	×	×		
5	Post-processor	×	×	×	×	×	×
6	Lifelong Learning	×	×	×	×	✓	×
7	Lifelong Training	×	×	×	×	×	✓
8	Cyclic Feed	×	×	✓	✓	×	✓

Key: IoT-HarPSecA Framework (Samaila et al., 2020), 3-Way IoT Security Framework (Zeeshan, Reed and Siddiqui, 2019), IoT-Flock Framework (Ghazanfar et al., 2020), FIFAC Framework (Awadelkarim Mohamed and Abdallah M. Hamad, 2020), BiDeL Framework (Otoo-Arthur and van Zyl, 2020), Combat Intelligence Information Monitoring Framework (Jin, Xing and Wang, 2020).

Table 1 and Table 2 lists the synthesis of the development techniques comparatively in tandem with the corresponding frameworks' techniques to be applied for the development and validation of the proposed iDRP framework.

4.1. Overview of Proposed iDRP Framework

Notably, Figure 1 shows the diagram of the proposed iDRP framework with extraction mechanism as a pre-processor and cyclical post-processor apparatus of logged complex dataset in an IoT ecosystem. These form the backbone of the novelty in the proposed iDRP framework. Significantly, from what has been learnt from the literatures of the existing knowledge domain that was conducted through the in-depth investigations, the aforementioned gaps in the anomalies (malwares) identification in an IoT ecosystem will be effectively tackled with the proposed technique for the novel iDRP framework.

4.2. Components of Proposed iDRP Framework

- *Preprocessing* – The pre-processing component, which is one of the major contributing steps in the framework to effectively address the aforementioned identified problems in the anomaly identification in IoT ecosystem. This component 1 involves the incorporation of Pre-Processor of a complex dataset by using ML technique like the enhanced traditional Data Augmentation techniques – blend of Enhanced Data Augmentation techniques and Propagations.
- *Detection* - The signatures extracted in component 1 will be fed forward to the component 2, the detection component for effective classification of signatures and labels i.e., '*fine-grained classification*' mechanism for feature learning to precisely detect Malware anomalies in an IoT ecosystem. Likewise, the Learning and Training of anomalies signatures experiments will be conducted extensively combined with Dictionary Classifiers and the Database of both the new and historical malwares in the detection component.
- *Recognition* - A similar procedure will be performed in the component 3, the recognition component, with the primary aim of precisely recognizing each specific Malware in the logged complex data of IoT ecosystem.
- *Prediction* - Correspondingly, in the component 4, Prediction component, the forecast of anomalies will be rationalized and realized by synthesizing the components 2, 3 and 4 for accurate predictions of the occurrence of anomalies in an IoT ecosystem.
- *Postprocessing* - In component 5, Post-Processing, the assessed anomalies will be fed back into the entire system in a cyclic manner. This '*feed-forward*' and '*feed-backward*' propagation methods that were applied and implemented in the model of the framework will essentially ensure that the framework is well-adjusted, robust, intelligent, and adaptable adequately to address the growing threats and risks in the IoT security paradigm.

4.3. Preliminary Experiment for Proposed iDRP Framework

The preliminarily experiments were conducted using MNIST dataset, a standardized dataset, which is comprised of small 28x28 pixels of randomized hand written number images with corresponding annotated numbers as labels. The MNIST dataset presents a better experimental starting point for developing the MLP for the recognition of patterns in the converted log file image with minimal pre-processing and formatting overhead cost for the ML and DL models. The standardized MNIST dataset used in this experiment contains 34,300 training set samples and 14,700 test images of 28x28 = 784 pixels for each image, which is a subset of original 60,000 training set samples of the standard benchmark MNIST dataset.

4.4. Conceptual Implementation of Proposed Framework

The Data Augmentation techniques and propagation techniques [30],[31],[37], which helped with the '*fine-grained extraction and classification*' of signatures of complex datasets were applied to the model implementation of the proposed novel iDRP framework. The applied research approach for the development of the iDRP framework is implemented using Python Programming language (https://www.python.org/), which is an open-source language with rich and extensive libraries that is easily accessible to the general masses. Keras (https://keras.io/), a high-level library that is built upon a low-level library such as TensorFlow (https://www.tensorflow.org/) is used for implementing the required Neural Network (NN) with simplistic interface that is built on the Python platform as a native Python library. TensorFlow serves as the tensor manipulation library to Keras, as a bridge low-level Application Programming Interface (API) for the robustness of implementing minimalistic and concise NN models. Significantly, Keras has several extensible modules that are suitable and adaptable to advanced scientific research work, which has formed the basis for the informed decision to choose the combination in implementing the complex NN models

especially for the stacking of the various layers; hierarchical models for this experiment.

4.5. Neural Network (NN) of Framework

The NN has been able to provide access and manipulation to the fed benchmarked dataset dynamically while exploring the binary features of the converted logged image files with different categories in the pre-processing of the dataset, while the necessary adjustments were made accordingly. Importantly, different layered NN combinations have been explored while calling upon the Functional model, '*tf.keras*' to provide the building blocks and backend support for the configurable NNs [29] for the training, testing, and prediction in the application. The novelty of the approach in the development of the proposed iDRP framework is to use dynamic Pattern-Based model approach to intelligently and accurately detect, recognize, and predict anomaly in an IoT ecosystem.

4.6. H-ELM Derivative Formula

A formular for the H-ELM was derived from [30],[31] as a result of the amalgamation of the Radial Bias Function (RBF) nodes with the defined Activation function [30] for mapping indiscriminate features of the converted logged datasets of an IoT ecosystem. The primary aim is to precisely approximate the continuous targets with mobility at the evolving RBF nodes adaptively [30],[31],[37]. This ensures the irregular initialization of the hidden nodes H in a given set of big data in an IoT ecosystem. This is expected to significantly reduce the training errors while increasing the precision in the accuracy of the output with better performance and speed.

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \dots & h_L(x_1) \\ \vdots & \vdots & \vdots \\ h_1(x_N) & \vdots & h_L(x_N) \end{bmatrix} \quad (1)$$

where H is the randomized hidden layer output matrix, which denotes the hidden nodes and output weights of the Neural Network, while $x_1 \dots x_N$ is the number of training dataset.

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{t_1^T} \\ \vdots \\ \boldsymbol{t_N^T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{t_{11}} & \dots & \boldsymbol{t_{1m}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{t_{N1}} & \dots & \boldsymbol{t_{Nm}} \end{bmatrix} \quad (2)$$

where T represents the training dataset of the target matrix.

In order to optimize the performance of the ELM feature mapping, the following formula is applied:

$$\beta = \mathrm{H}^{\mathrm{T}} \left(\frac{1}{\lambda} + HH^T\right)^{-1} T \quad (3)$$

The resultant output of the ELM feature mapping will be

$$f(x) = h(x)\beta = h(x)\, \mathrm{H}^{\mathrm{T}} \left(\frac{1}{\lambda} + HH^T\right)^{-1} T \quad (4)$$

Alternatively, the equation is condensed to

$$\beta = \left(\frac{1}{\lambda} + HH^T\right)^{-1} H^T T \quad (5)$$

The derived output function of the ELM feature mapping is

$$f(x) = h(x)\beta = h(x) \left(\frac{1}{\lambda} + HH^T\right)^{-1} H^T T \quad (6)$$

In the hidden nodes H of the training dataset, hidden information can now be exploited and extracted from the hidden layers h to transform T the output weights β. This will ensure minimal fine-tuning of the parameters or weights of the hidden layers with stable minimal overhead cost to achieve maximum output. This randomized approximation and differential evolution technique has been proven [30]-[32] to be effective, which ultimately secures the increment in both the feature learning and classification.

5. The Need to Convert IoT Malware Binaries to Images

The technological advancements in present-day cyberattacks has made the activities of advanced attackers more complex to detect as a result of emerging obfuscation techniques [27], [38], [39] and interactions [40] with stealth variations carried out on IoT ecosystems.

In modern times, the emerging polymorphic malware attacks in the IoT ecosystems have been a major concern [1] due to complex obfuscation code structures that are mostly time based [41], [42]. These IoT malware signatures attacks that are predominantly multivariate [39], [40], [42]-[44] are updated sequentially on a minute-by-minute or hour-by-hour basis by the attackers, thereby inundating and *silencing* any potential alert system, which may cause massive vulnerabilities for exploitations in an IoT ecosystem. The current widespread detection and mitigation mechanisms for these emerging polymorphic IoT malware attacks that are largely obfuscated intricately can be problematic and resource intensive to both the traditional and automated malware detection solutions such as the signature based (e.g., large database), and automated based techniques (insufficient information) etc., adopted by major cybersecurity vendors, practitioners, and researchers in the cross-discipline cybersecurity industries. The domain experts and analysts may write different rules manually or automatedly to detect complex obfuscated malware scripts in an IoT ecosystem as a possible approach to solving the problem, the caveat to such an approach in determining the benignity or maliciousness of such malware binaries in an IoT ecosystem is that it would be near impossible to efficiently and accurately determine either a clean or infected malware script in the IoT ecosystem – *subtle obfuscated IoT malware classification problem.* It certainly would be time and resource intensive to adequately write thousands and/or millions of rules for virtually all current and new malware obfuscation variants in the contemporary IoT ecosystem.

Evidently, the aforementioned complex IoT malware detection approaches are no longer suitable for current growing subtle and complex obfuscated cybersecurity paradigm for five (5) major reasons: (1) the distinct fragments (sections) of converted IoT malware image file can be visualized [27], [38]; (2) the distinct textural patterns of the converted IoT malware image files can be classified (and reclassified), and analyzed using image processing [27], [45] classification techniques commonly applied in computer vision [30], [31]; (3) the advanced attackers typically modify or reuse codes to generate new variants of complex obfuscated IoT malware scripts [38]-[40]; (4) the vector structures of converted IoT malwares image files of the same family are similar [27], which means vectorized representation (i.e., faster speed) of the IoT malware image files will be sufficient to be fed directly as

input into the similar ML-DL algorithms [30], [31] built for image and pattern recognitions [27], [45]-[47]; (5) the captured IoT malware dataset once converted to image files will be immutable (nondestructive and tamper proof), so there is no workaround its state by the attackers. Therefore, an amalgamation and syntheses of the modern ML-DL techniques and methods designed utilizing converted immutable IoT malware binaries to image files is both a logical and pragmatic solution to the current rapidly growing complex IoT malware problems that would scale adaptively.

5.1. The Benefits of Converting IoT Malware Binaries to Images

One of the biggest advantages of converting IoT malware binaries to Gray-scale image files is the fact [27] that the IoT image maintains the structures and patterns with rich information – *structural integrity*. The immutability of the IoT image files guarantees that converted complex IoT malware scripts will always be reliable to work with (ensuring a total control of the collected IoT malware dataset in the environment without the risk of compromise externally) by leveraging the ever-growing advancement of modern technologies such as Artificial Intelligence. Moreover, the structural integrity of the stately converted IoT malware image files facilitate the high-level precision and accuracy of detection of minuscule byte codes of the clean or infected codes with corresponding Cross-Validation techniques of modern ML-DL algorithms – *efficient detection, recognition, and prediction of anomalies in an IoT ecosystem.* Another advantage of converting IoT malware binaries to image file is the speed (time) of processing and computing the thousands and millions of both the benign and malicious IoT malware image file vectors using ML-DL techniques – *computational time*.

5.2. Visualization of IoT Malware Image for Exploitations

The textural and structural anatomy of IoT malware image files reveals the intricacies and patterns of the minuscule information that are sufficient for synthesized spatial and differential exploitations and analysis [27], [30], [31] of the tiniest byte codes (exploiting hidden structural and textural information by utilizing advanced ML-DL techniques and methods i.e., '*fine-grained*' layer-by-layer mapping and representations) of the features of converted complex IoT malware image files with high-level precision and accuracy in an IoT ecosystem – effectively distinguishing the benignity and maliciousness within a few bytes of codes as shown in Figures 4 and 5.

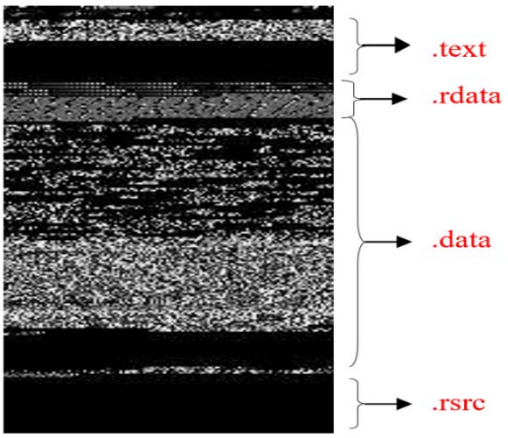


Figure 4: Visual structure of IoT Malware image file

The Figure 4 above shows the visual representation of the sections of a converted IoT malware image file with different segment textures (https://code.google.com/p/pefile/). The above section representations (*.text* = '*fine grained*' section with unique validity of data on the Operating System (OS) level e.g., Windows, Linux, *.rdata* and *.data* = the uninitialized and initialized scripts are the header and section table of the file that distinctively separates the unlaunched and launched codes respectively, *.rsrc* = the resource section that contains mostly padded zeros. These structures and patterns are important because the slightest of modifications (changes) in the variations of IoT malware binary codes can be easily detected stately – often within a few bytes.

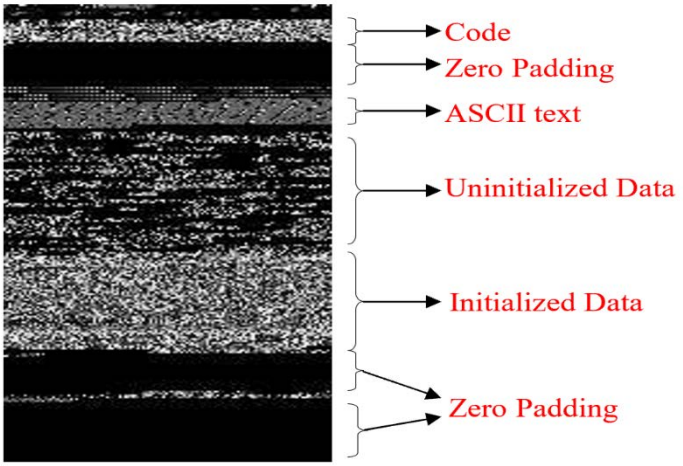


Figure 5: Visual structure of IoT Malware image file (Rich Information)

The Figure 5 above shows the rich visual representation of the section of an IoT malware image file with extensive details for exploitations and analysis. The segmented textures and patterns are essential for the pattern recognition of both benign and malicious IoT malware binaries converted to image files in an IoT ecosystem. The various segments of the IoT malware binary file structure are easily isolated for the synthesis of spatial and evolutionary feature mapping (i.e., spatial and evolutionary data matrix), extractions, representations, classifications, and Cross-Validation of the hidden minuscule features (i.e., the IoT malware signatures in the *header*, *sections 1*, *2*, and *section+n*) of the converted IoT malware image files by utilizing ML-DL techniques and methods with high-level precision and speed of detection of complex obfuscated anomalies in an IoT ecosystem.

5.3. Overview of Converting IoT Malware Binaries to Images

The conversion of typical binary files is based on standardized OS binary file format structures, which is analogous to debugging and homologous to reverse engineering of malware binary programs structurally. This can be simply performed using various *python* libraries such as *pefile* (https://code.google.com/p/pefile/) for fundamental exploration, understanding, and analysis of the layout features of a binary program. However, for a more in-depth exploration, exploitation, and analysis of rich hidden features, a systematic approach has to be applied to gain valuable insight into the various layers and architectures of the IoT malware binary file to image file. These procedures for converting IoT malware binaries to image files consists of the following five (5) fundamental steps: (1) Sort Dataset (i.e., variants, variant name,

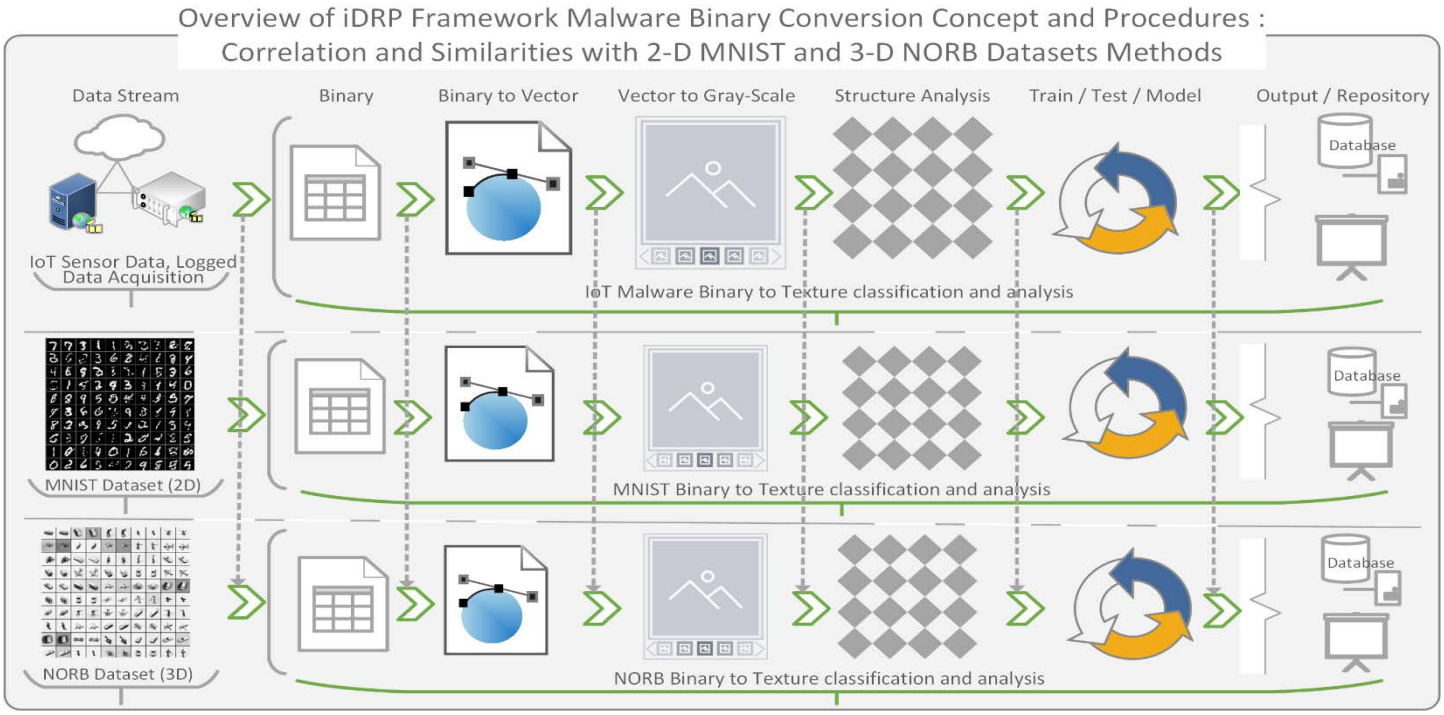


Figure 6: Comparison of similarities between the classification and analysis procedures for IoT malware images and MNIST and NORB datasets

no. of occurrence) retrieved by attributes; (2), Import Libraries (e.g., os, *numpy*, *scipy*, and *array*) to access and extrapolate the IoT malware binaries; (3) Augment IoT Dataset (i.e., *reshape*, *resize*, and *scale* et al) the dataset; (4) Standardize the IoT dataset; (5) Save the converted IoT malware images in a repository. Thereafter, feature selections techniques for consistent feature vectors (i.e., color range, intensity, pattern structure) can be performed on the converted IoT malware image files and processed as input feature classifications – *supervised learning*.

6. Experimental Implementation of iDRP Framework

The approach for the experiment of the proposed novel iDRP framework [1] involves the cleaning, manipulation, and extraction of IoT malware binaries to be preprocessed, and visualized as IoT malware image file for the evolutionary exploitation of hidden minuscule segment features for both supervised and unsupervised structural features representations and classifications with high-level precision to detect, recognize, and predict anomalies in a heavily computerized network system such an IoT ecosystem.

This experimental implementation of the proposed novel iDRP Framework is a proof-of-concept for the research study. The experiments were conducted with several distinct structural different datasets (BoTNet-IoT [43], MNIST, virus-MNIST, and NORB datasets etc.), yet similar conceptual procedures and methods (e.g., Figure 6) in the multiclassification and analysis to extrapolate meaningful hidden information for sound scientific evaluations and validation quantitatively. The Figures 7, 8, 9, 10 represent and demonstrate the correlations between the textual standardization methods of the IoT malware images, MNIST images, and NORB images datasets. This is to enable the systematic implementation and evaluation of the performance of the proposed novel iDRP framework techniques and methods (e.g., Figures 11, 12, 13, 14) for optimal malware binaries' classification and analysis in the present-day IoT ecosystem. The datasets used to perform the experiments were converted to 8-bit gray-scale 2-D array with a channel range between 0 to 255 for consistencies in the datasets. Essentially, this process generates a vectorized 8-bit binaries, which are then converted to the gray-scale 2-D image files for structural texture uniformity for extracting and detecting padded hidden information in a logged IoT malware file.

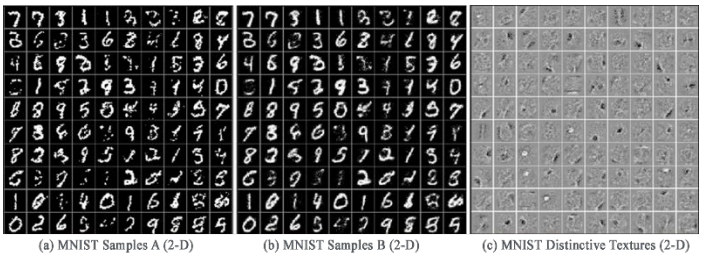


Figure 7: MNIST (2-D) images to Gray-scale images (0 to 255) standard channel

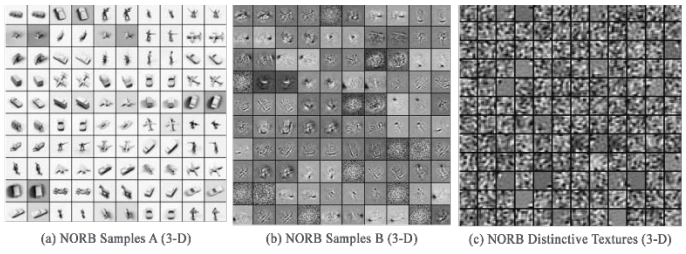


Figure 8: NORB (3-D) images to Gray-scale images (0 to 255) standard channel

The 3-D NORB image files, which has uneven generic object representations such as animals, vehicles, and nature etc. provide similar abundant natural environment and characteristics for representing, visualizing, and detecting certain distinct patterns of structurally complex textures of IoT malware variants that conceals hidden features that modern polymorphic and new malware engines in an IoT environment exhibits. This hypothesis was helpful for the scientific empirical and statistical observations

for the classification and analysis for the detection of IoT malware anomalies with high level precision and performance in the experiment for uniquely solving contemporary IoT malware obfuscation problems by attackers [1].

We performed the experiments on the premise and fact [48] that malware of the same family has similar characteristics and textures, while being distinct from other malware families in nature. Comparatively [48], for visual analysis of IoT malware binaries, the binary fragments of polymorphic and new malware variants are visually similar and consistently fragmented alike by the same token in terms of their image dissections and layouts as with their primitive nature irrespective of the malware variants or obfuscations and their specialized binary complex code structures.

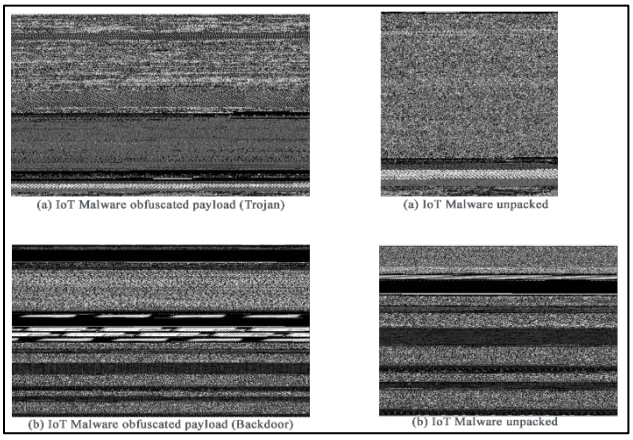


Figure 9: Captured IoT malware images with binaries I

Therefore, we approached the problem by proposing an iDRP framework techniques and methods that leverages both the H-ELM [30] and E-HELN [31] techniques and methods for universal approximation using differential evolution for optimized generation of complex image features to precisely extrapolate hidden features of IoT malware binary images files for expedient training and learning in the DNN. The advantage of these synthesized techniques and methods is that they both provide an optimized classification and characteristics learning and training coupled with minuscule and hidden feature representation modules for accurately *detecting*, *recognizing*, and *predicting* IoT malware binaries speedily.

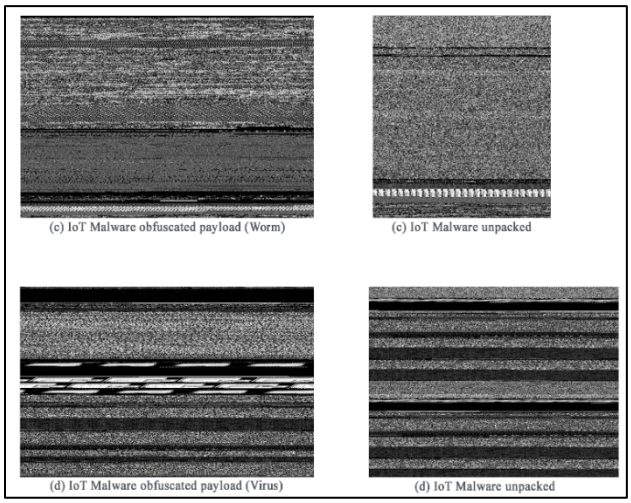


Figure 10: Captured IoT malware images with binaries II

The generative model techniques and methods that were applied cyclically in the iDRP framework to provide better discriminative capabilities for analyzing and evaluating the captured IoT malware binary information and statistics with high level precision and performance.

Overall, we explored and computed (gray-scaled etc.,) the binaries of the engineered BoTNetIoT malware dataset [43] exploits (Botnet attacks e.g., Mirai, Gafygt and their variants, and benign traffic etc. on IoT objects, devices, and environs), split them into 70% (training) and 30% (testing) ratio processed systematically and Cross-validated with a combination of both authentic benign and infected IoT malware traffic dataset with different strains [43], similar techniques and methods as with the statistical exploratory MNIST and NORB datasets experiments performed. The converted IoT malware gray-scaled images were normalized (min-max normalization, anti-aliasing techniques etc.), flattened, reshaped into 1-D and 2-D pixel and arrays for consistency and standardization. These were performed with the structured techniques and methods by applying the evolutionary adversarial training and testing techniques and methods for classification and Cross-validation to accurately estimate the comparative capabilities and skills (e.g., accuracy, precision, recall, and F1-score etc.) of the implemented techniques and methods in the proposed and implemented experiments. These were conducted for effective classification estimation and validation – *pre-processing of data*. These matched aggregated distribution of the train-test split techniques and methods together with the resampling methods, which are both semi-supervised and unsupervised classification clustered enabled the optimized resampling of the utilized datasets quantitatively. Specifically, with the categorical distributions mapped against data distributions of the IoT malware binaries and structures, empirical and statistical analysis of the results were scientifically established.

Evidently, captured IoT malware binary classification and analysis using the proposed integrated novel iDRP framework techniques and methods contributed provides treasure trove of hidden information exploited with both 2-D and 3-D images synthesized based on their distinct structural compositions and textures unpacked in comparison to static, dynamic, and automated analysis techniques and methods that are prone to present-day complex and unpacked code obfuscation and time consuming.

6.1. Pre-Processing Component Technique Overview

The data Pre-processing component in the proposed iDRP framework involved the augmentation, transformation, and classification of the raw dataset that are fed and parsed into the network. Feature extraction and mapping of the sparse data feed for layer representation of random features were performed to fully exploit the minuscule information [1] in the converted logged image files for universal approximation (i.e., MD5 Hashes and malware binaries) of the mapped datasets in the IoT ecosystem.

Figure 11 shows the pre-processing component of the proposed iDRP framework and its integration into the overall architecture. The pre-processor 'feed-forward' the systematically processed dataset (i.e., 'fine-grained' gray-scaled converted logged file) directly into the detection component of the proposed iDRP framework to ensure the universal approximation and classification of the precision and accuracy of detection of

anomaly is optimized for high-performance with detailed description for exploits – preparation of dataset for standardization.

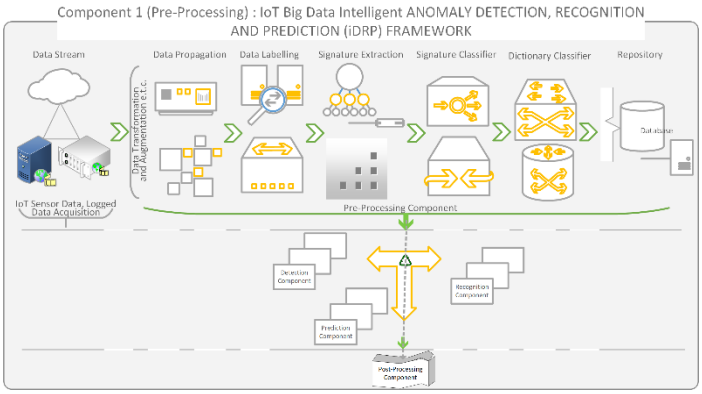


Figure 11: Pre-processing Component of iDRP Framework Overview

The systematic data manipulation techniques such as the systematic extrapolation of feature representations to form the vectorized structures applied in the experiment substantially helped in the distinction of the effective assignment of randomized and differentiated hidden layer weightages synthesized. The feature immutability techniques to stately capture complex code implemented in effect ensures that the feature extractions and data augmentations performed are compact and dynamic to materially eliminate unnecessary redundancies in the generated data. These scalable techniques intrinsically set the approximation numbers of the hidden nodes in the converted logged file of the original input dataset to achieve compact features for exploitations. Uniquely, these innovative and intelligent techniques implemented radically extrapolates rich hidden features for universal approximation competence on both the raw datasets and converted logged file to ensure that the delay in the binary classification analysis is effectively mitigated for optimal outputs in the proposed iDRP framework. Expediently, the compact features extracted and delivered to the next 'following' components iteratively, which is the detection components in the proposed iDRP framework.

6.2. Detection Component Technique Overview

The detection component involves the identification, determination, and classification of converted logged file dataset, whether such dataset is benign or infected. Considering this is a binary classification problem, several techniques [1] were deployed to achieve the expected results by conducting several systematic experimentations to discover the best suitable techniques. In particular, the systematic cross-validation splits, and weightage allocation techniques were beneficial for the classification precision and accuracy in the experiment of the detection component of the proposed iDRP framework.

Figure 12 shows the detection component of the proposed iDRP framework that is strategically the responsible component for determining whether the pre-processed data feed in the network is benign or infected with optimized classification precision and accuracy with maintained high-performance in an IoT ecosystem. Mainly, the mapped converted logged files are systematically rescaled and gray-scaled before they are dynamically fed 'feed forward' into the DNN of the IoT ecosystem.

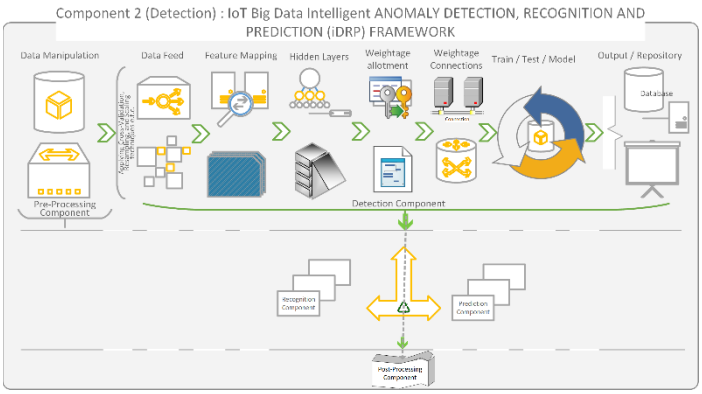


Figure 12: Detection Component of iDRP Framework Overview

The primary focus of the detection component is to intelligently identify the presence of polymorphic malware anomalies [1] with high level of accuracy in the converted logged dataset, which creates the unique separation of concern in its architecture – design principle. The detection component is robustly and innovatively designed in its structure to encapsulate the interaction between the dynamic and static states of the dataset in the converted logged file of an IoT ecosystem. By and large, the detection components feed and receives immutable mapped hidden information dynamically to exploit with other modular components in the proposed iDRP framework. The 'Loose Coupling' and yet cohesion of the detection component [1] essentially promotes the dynamic and robust interaction of the various components directly or indirectly connected to it in the proposed iDRP framework while explicitly performing the primary function of detecting anomalies (randomly and differential mapping hidden features) in the network system.

6.3. Recognition Component Technique Overview

The recognition component is inherently responsible for identifying what class of anomaly the detected malware family is associated with or lack thereof [1]. The recognition component explicitly curates the type of anomalies present if found in the network by precisely defining the exact variants or obfuscation of the IoT malware anomalies for launching attacks such as zero-day attacks. This unique technique involves the heuristic classification of variants of IoT malware anomalies in the network system [1], which are universally approximated in the random and differential feature mapped converted logged file that are rescaled and represented as gray-scale images in an IoT ecosystem.

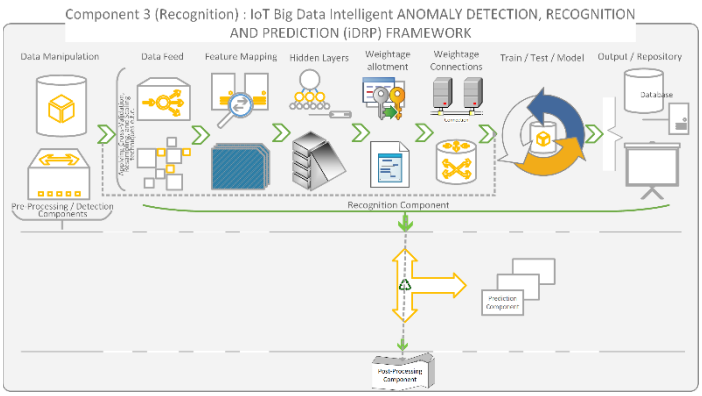


Figure 13: Recognition Component of iDRP Framework Overview

Figure 13 shows the recognition component of the proposed iDRP framework. The recognition component in the architecture of the proposed iDRP framework facilitates the structural distinction of the universally approximated mapped converted logged file in the DNN of the IoT ecosystem. Thereby determining and identifying the particular class of malware anomalies that each existential or inherent malware family belongs to while intelligently discovering new variants or obfuscation of such IoT malware anomalies in the process from the repository that could potentially result in a zero-day attack in an IoT ecosystem. This unique technique provides the capability for the proposed iDRP framework to be able to adaptively and robustly perform the required intelligent classification precision and accuracy [1] of recognizing specific types of anomalies in IoT ecosystem with high performance. The systematic integration of data from both the Pre-Processor and detection components ensures that there is both shared characteristics and cohesion amongst the interconnected components of the proposed iDRP framework. Constructively, the arbitrary boundaries established between the multi-directional connections that are 'Loosely coupled' essentially eliminates constraints in the shared information such as hidden layer features to and from 'fro' the recognition components to the preceding and forwarding components virtually in the proposed iDRP framework – architecture intelligence.

6.4. Prediction Component Technique Overview

The prediction component is practically responsible for predominantly predicting the occurrence of captured malware anomalies in the IoT ecosystem. The systematic elimination of unnecessary '*checkpointing*', '*overlearning*', and '*overfitting*' etc. techniques applied in the implementation of the prediction component effectively helped in the classification precision and accuracy of the predicted malware binaries innovatively in the experimental implementation of the proposed iDRP framework.

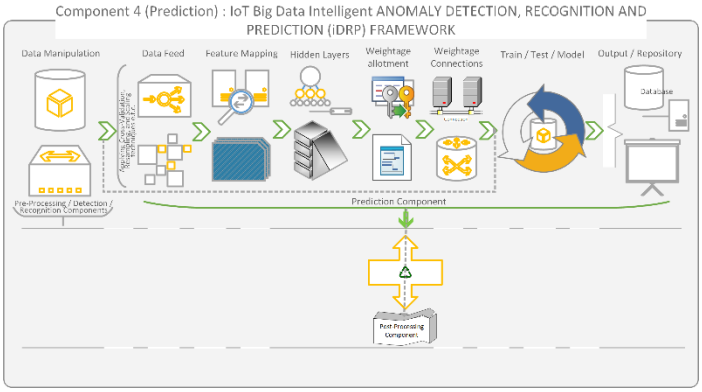


Figure 14: Prediction Component of iDRP Framework Overview

Figure 14 shows the prediction component of the proposed iDRP framework. One of the main features of the prediction components is that it can be configured to predict specialized family of malware binaries that are captured in the IoT ecosystem. Likewise, the systematic assignment of numerical score in the prediction component implementation ensures that the '*learning curve*' of the prediction classification precision and accuracy is properly monitored and dynamically optimized for high-performance. Crucially, the loss or error in the prediction component is systematically calibrated to ensure that the network is neither '*under-fitting*' nor '*under-learning*' to achieve maximum probabilistic outcomes in detection, recognition and prediction of malware binaries in an IoT ecosystem.

The prediction component employs the network weights to accurately predict the occurrence of malware binaries in the network system. The cyclical and infinite interconnection of the prediction component [1] to other components with accessibility to the repositories in the network ensures that the continuous learning of rapidly evolving malware families is adaptively and intelligently tackled in the growing threats and attacks on the technologically driven IoT ecosystem. In the prediction component, rigorous systematic estimate of the capability of the DNN model were performed exclusively. The prediction component of the proposed iDRP framework is highly effective predominantly due to the capability to systematically stabilize the universal approximation of the mapped hidden features that are fully exploited in the network – intelligent classification prediction technique.

6.5. Post-Processing Component Technique Overview

The Post-Processing component ensures the cyclical and infinite enhancement of the information distributed in the entire proposed iDRP framework [1]. The optimization techniques implemented guarantees that the loss i.e., 'validation loss' in the network model will always be minimized while fostering the continuous and lifelong learning in the network. The substantially reintroduction of the refined information dynamically and innovatively into the network forms the unsupervised building block technique in the overall implementation of the proposed iDRP framework. Broadly, the 'autoencoder' combined with the 'weightage reloading' concepts [1] are integral aspects of the unique features of the Post-Processing component in effectively maximizing the classification precision and accuracy support for the other interconnected components in the proposed iDRP framework.

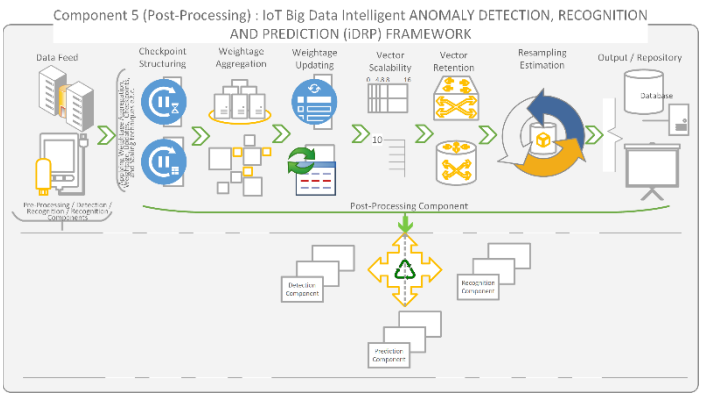


Figure 15: Post-Processing Component of iDRP Framework Overview

Figure 15 shows the Post-Processing component of the proposed iDRP framework. The systematic implementation techniques implemented in the Post-Processing component provide the comparable performance for the convergence of both the training and testing models for stochastic consistencies in the network while storing the models for exploration capabilities in the IoT ecosystem [1]. The Post-Processing component adequately helped stabilize the performance behaviour in the operations of the

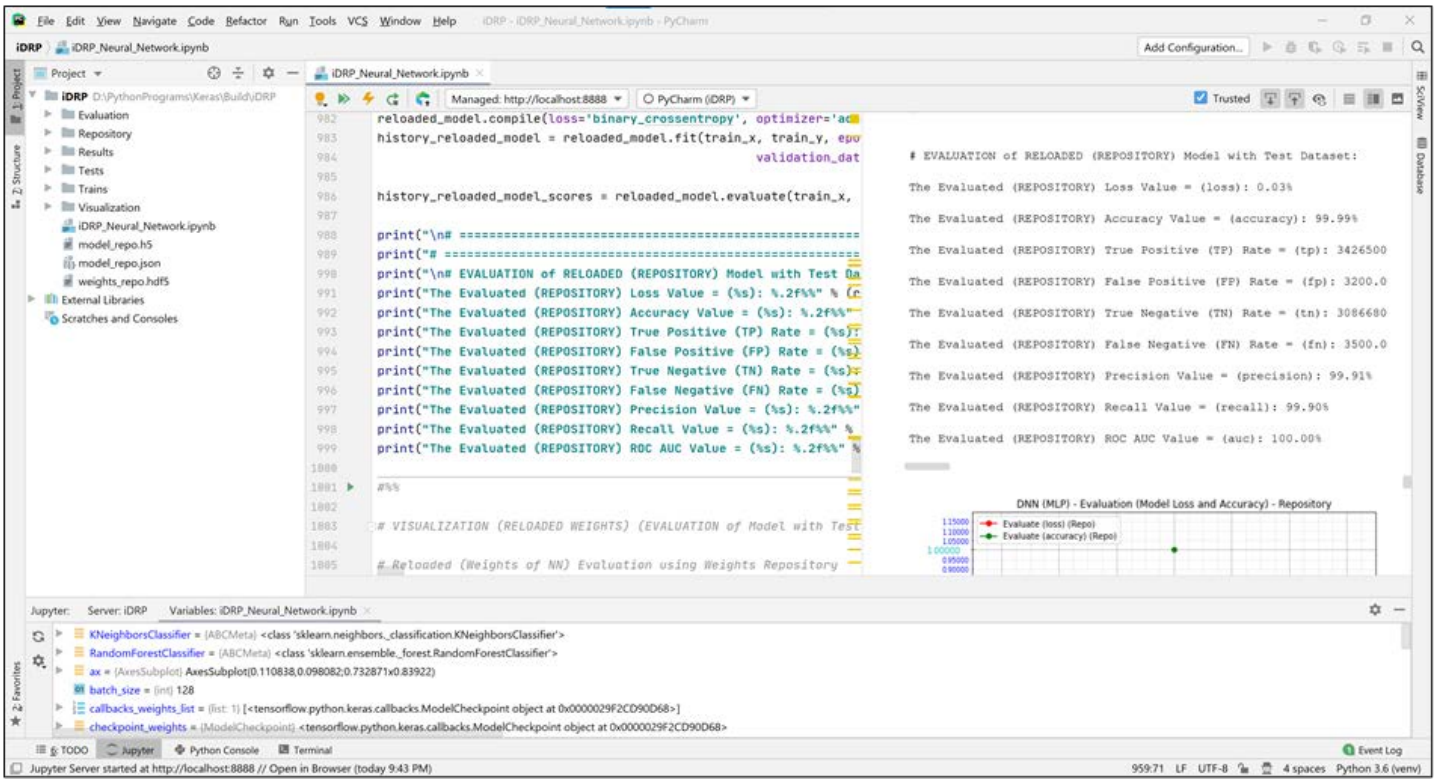


Figure 16: Results of implemented experiment of iDRP Framework with the corresponding snippet of metrics for evaluation

overall proposed iDRP framework via the *checkpoint structuring*, *weightage aggregation*, *weightage updating*, and *vector scaling and retention* techniques etc. With minimal 'hyper-parameterization', 'regularization', and 'resampling' techniques etc., the robust and innovative Separation of Concerns technique effectively helped to reduce redundancies in the implementation of the network for optimal competence and high-performance in detecting, recognizing, and predicting malware binaries and anomalies in a massively computerized network system such as an IoT ecosystem.

Specifically, the Post-Processing component is pivotal in the construction of the defined weights of the hidden features exploited for the lifelong learning of the overall proposed iDRP framework. With the Post-Processing component in the architecture of the proposed iDRP framework, minimal 'fine-tuning' is required while the other interconnected components can fully explore compact features in the universally approximated hidden features in the network [1]. Essentially, these specialized innovative and dynamic techniques implemented in the experiment further help to minimize the training, testing and learning time and overhead cost for both the existing and new datasets in the network [1]. The historical information aggregation technique in the network of the proposed iDRP framework are used to systematically discover and generate visualization and analysis of the classification precision and accuracy together with its overall high competency and performance.

6.6. Summary of Components Techniques

The generalized Artificial Neural Network (ANN) applied in the experimental implementation of this proposed iDRP framework is systematic and innovative Multilayer Perceptron (MLP) with Deep Neural Network (DNN) with integrated modular components to expedite the cross-validation methods for universal approximation and classification towards the effective learning, training, and testing capabilities for optimal outcome. The amalgamations of the unique innovative and intelligent random and differential feature extraction, augmentation, transformation, scaling, universal approximation, compactness, mapping, auto encoding, weightage saving, and weightage reloading, vector scaling and retention, cross-validation, resampling etc., techniques were successfully designed and implemented in the interconnected components of proposed iDRP framework to achieve core optimal results. The specialized immutable randomized and differentiated projections techniques implemented in the experiment formed the core of the feature representations of the complex malware binaries for classification precision, accuracy and analysis in the implementation of the proposed iDRP framework – *binary classification problem with complex structure system*.

The '*checkpoints*' of the network structures were set up to save model weights architecture intelligently and automatically whenever there is any improvement attained in the model classification system. The best part of some of the amalgamated specialized techniques implemented is that if there is no improvement in the weight classification of the model, then, the last highest improved weights will be applied in the model. Thereby maximizing the capability to accurately detect, recognize, and predict captured and generated malware binaries anomalies in an IoT ecosystem. The singularity of each of the components in the proposed iDRP framework ensured an easy to manage and maintain network. This generalized Separation of Concern technique ensures that each component in the proposed iDRP framework remains consistent and true to its primary function in the architecture.

The dynamic and innovative abstraction and segregation of the different components the implemented proposed iDRP framework makes it adaptable to the growing threats and attacks in the heavily computerized systems and smart technologies such as an IoT ecosystem. This fundamentally delineates specialized responsibilities amongst the various components in the network with flexible boundaries. Essentially, malware binaries data are analyzed and shared in the interconnected components while at its core separated in execution to perform specified operations in the proposed iDRP framework. The strategy deployed in the experimental implementation of the proposed iDRP framework has demonstrated the intelligent capabilities to determine the slightest of improvements in the network model weightage's structure and only save the best weights and architecture model together for future usage. This simply means that the overall system has been designed to continually learn, adapt, improve and perform better with time in combating the ever-growing malware threats such as the prevalent polymorphic IoT malware threats and attacks in the IoT ecosystem.

The overall procedures involve (a). pre-processing of dataset, (b). loading of dataset, (c). augmentation of dataset (d). creating of model architecture (e). compiling of model (f). fitting of model (g). evaluation of model (h). saving of model weights (i). loading of saved model weights (j). evaluation of saved model (k). post-processing of outcome. These approaches were conducted in multiple hidden layers Neural Network 'feed-forward' propagation and cyclic method using Deep Learning (DL) models to precisely, efficiently and optimally and intelligently detect, recognize, and predict (iDRP) anomalies in an IoT ecosystem. Cumulatively, several ANN techniques were leveraged innovatively and intelligently to significantly improve the classification precision, accuracy, and performance in the implementation of the proposed iDRP framework. Notably, the dynamic and innovative extensibility and maintainability of the proposed iDRP framework has made it highly adaptable and adoptable in effectively safeguarding and protecting critical infrastructures in multi-industries that depends on heavily computerized interconnected network systems and smart technologies such as an IoT ecosystem in the digital space.

7. Results, Observations and Discussions

A 100% accuracy rate, 100% ROC/AUC rate together with a precision rate of 99.94%, and recall rate of 99.90%, and merely a loss rate of 0.02% were achieved in the preliminary implementation of the iDRP framework experiment. Please see Figure 16 screenshot for the output of results and metrics. These results were attained based on the syntheses of the various techniques and procedures applied and discussed in the previous sections.

In like manner, a 99.98% accuracy rate, 99.99% ROC/AUC ROC/AUC rate together with a precision rate of 99.95%, and recall rate of 99.93 %, and a mere loss rate of 0.03% were achieved in the actual implementation of the integrated iDRP framework experiment using real IoT malware dataset – BoTNetIoT dataset.

The metrics used in evaluating the experimental implementation of the proposed iDRP framework are: Accuracy rate, Precision rate, Recall rate, Receiver Operator Characteristic (ROC) and Area Under the Curve (AUC) rate, Loss rate, True Positive (TP) rate, False Positive (FP) rate, True Negative (TN) rate, False Negative (FN), and F1-Score respectively.

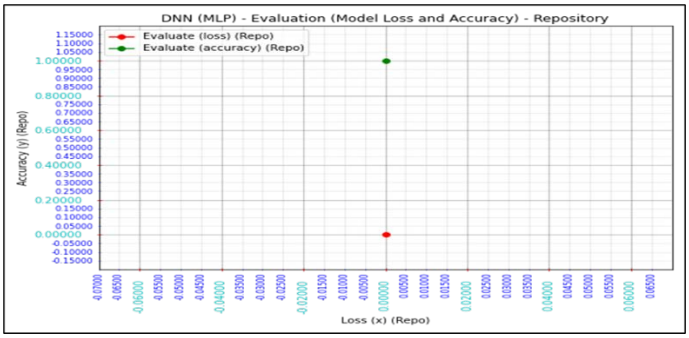


Figure 17: Graph of Accuracy and Loss Results of implemented experiment of proposed iDRP Framework

Figure 17 shows the result of the evaluated models of the experimental implementation together with the corresponding rates and scores. The 100% accuracy value, 100% ROC/AUC value with a precision value of 99.90%, and crucially with a much lower False Positive (FP) rate. that were attained in the experimental implementation of the proposed iDRP framework shows that there is a correlation in the accuracy, ROC/AUC, and precision rate in the Deep Learning (DL) classifiers in tandem with the hypothesis of the research study. It further proves that the experimental implementation of the proposed iDRP framework in this research study is highly innovative, applicable, useful, relevant and efficient; accurate, precise, robust, dynamic, adaptable, and scalable in detecting, recognizing, and predicting anomalies in the form of polymorphic malwares in an IoT ecosystem.

The visualization of the results of the models evaluated were represented in plotted Deep Neural Network graphs and charts to gain better insights and understanding of the complex binary classification matrix problems and solutions in the experimental implementation of the proposed iDRP framework. Figure 10 displays the optimal efficiency and accuracy correspondingly with insignificant loss rates in relations to the performance of the evaluated model for detecting, recognizing, and predicting the anomalies in an IoT ecosystem. Minimal training was required.

Logically, it is understandably that 50 epochs would likely fare better than a 5 epochs of model training and testing of datasets. At the same time, with a direct comparison of the regular 50 epochs with the 50 epochs' '*Checkpointing*' Deep Neural Network model technique of intelligently and systematically saving and reloading the best performance weightage in the repository of the overall system, a 100% accuracy rate and a 100% ROC/AUC rate together with a 99.94% precision rate and 99.90% recall rate of detection, recognition, and prediction of the anomalies in an IoT ecosystem was achieved in the experiment.

7.1. Cross-Validation and Evaluation Performance of Predictor Model Results

The results of the various output attained in the experiment on the proposed iDRP framework network has been reported in Table 3. The output of the different techniques towards the innovative synthesis, augmentations permutations, adjustments, and tuning of parameters in the Neural Networks together with various optimization techniques in the experiment has been extensively

investigated and compared to show the selected outcome in the project.

Table 3: Comparison of the Experimental Results

Cross-Validation Classification Parameter Tuning Metrics of iDRP (MLP) Model using *Keras* Library with Tensorflow Backend

Dataset (MNIST)	NN Layers	Epoch (Iterations)	Optimizer	Learning Rate	TPR	FPR	Precision	Recall	MAE	F-Score	Accuracy (%)	Training Time (s)
✓	5	5	ADAM	0.02	0.924	0.924	0.924	0.924	0.200	0.947	92.43	1.00
✓	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
✓	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
✓	50	50	ADAM	0.02	0.996	0.996	0.996	0.996	0.159	0.979	99.60	10.00
✓	50	50	ADAM	0.01	0.999	0.998	0.999	0.998	0.010	0.998	100.00	0.03

Legend: MLP=**Multilayer Perceptron,** NN=**Neural Network,** Learning Rate=**Dropout Rate,** TPR=**True Positive Rate,** FPR=**False Positive Rate,** MAE=**Mean Absolute Error,** F-Score=**Measure of Test Accuracy**

Table 3 displays the comparison in the various parameter tuning and optimizations performed in the experiments together with their corresponding metrics and results. It is evident that the continuous tuning process has significant impact on the overall output, especially with the classification precision and accuracy of the predictors. Likewise, the optimization parameter ADAM plays a vital role in minimizing loss function while increasing the speed of the convergence of the Neurons in the system. The dynamic and cyclic hyperparameter tuning approach in the Neural Network construct of this experiment will ultimately facilitate an incremental and better outcome.

7.2. Comparison of Generalized Artificial Neural Network (ANN) Results

The selection of the appropriate NN model for the implementation of the proposed iDRP framework involved the systematic experimentations of various streamlined and focused ANNs such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Multilayer Perceptron (MLP), and Hybrid Network (synergy of RNN and CNN) to effectively determine the best possible NN model for the binary classification problem of this nature.

Table 4: Comparison of Artificial Neural Network (ANN) - RNN, CNN, MLP, and Hybrid Models Results

No.	Metric \ ANN	RNN	CNN	MLP	Hybrid (RNN + CNN)
1	Detection	98.53%	99.71%	**100.00%**	99.83%
2	Recognition	98.24%	99.64%	**100.00%**	99.84%
3	Prediction	97.32%	99.52%	**100.00%**	98.91%

Table 4 shows the comparison between the various NNs in determining the chosen NN for the implementation of the proposed iDRP framework. Evidently, MLP performed best for the binary classification of precision and accuracy in the experiment – binary classification problem.

The MLP NN mainly performed better with model flexibility in terms of training, testing, and development. The optimal capability for an MLP NN to be abstracted in its architecture design and robustly extract fine-grained features with universal approximation made it a perfect candidate for the implementation NN model of the proposed iDRP framework. Chiefly, the MLP NN allows the Separation of Concerns in its architecture by creating boundaries (methods, and objects) for the several interconnected components in the proposed iDRP framework while providing the layers and tiers the independent capability for each component to maintain their core operations dynamically.

7.3. Comparison of Classification Techniques on Standardized Datasets Results – Preliminary Experiment

The amalgamations of innovative MLP techniques applied in the implementation of the proposed iDRP framework network demonstrated to be advantageous in the specialized classification precision and accuracy on accurately detecting, recognizing, and predicting malware binaries in the standardized dataset in an IoT ecosystem.

Table 5 shows the comparison of the different classification results of experiments on standardized datasets. The synthesized innovative techniques in the proposed iDRP framework network technique that have been successfully implemented have proven to be more efficient and competent in tackling polymorphic IoT malware attacks and their variants or obfuscation that are typically responsible for zero-day attacks in a heavily computerized network system and smart technologies such as the IoT ecosystem. The learning capabilities and accuracy of a specific predictor problem of this nature was significantly amplified through the innovative classification precision and accuracy techniques that were proposed and implemented in the systematic experiments iDRP framework network.

Table 5: Comparison of Classification Experimentations Techniques with Standardized Datasets and Results – Preliminary Experiment

No.	Technique \ Dataset	MNIST	NORB
		Accuracy (%)	Accuracy (%)
1	SAE	98.60	86.28
2	SDA	98.72	87.62
3	DBN	98.87	88.47
4	DBM	99.05	89.65
5	MLP-BP	97.39	84.20
6	ML-ELM	99.04	88.91
7	H-ELM	99.13	91.28
8	iDRP	**100.00**	**100.00**

Key: SAE (Stacked Auto Encoders), SDA (Stacked Autoencoders), DBN (Deep Belief Networks), DBM (Deep Boltzmann Machines), MLP-BP (Multilayer Perceptron Back Propagation), ML-ELM (Machine Learning Extreme Learning Machine, H-ELM (Hierarchical Extreme Learning Machine), iDRP (Intelligent Detection Recognition Prediction)

7.4. Comparison of Classification Techniques on Standardized Datasets (BoTNetIoT) Results

The standardized IoT malware generic traffic metric and evaluation were used as parameters to compare the effectiveness of the learning and training techniques in correlations with the validation of the techniques applied to detecting benign and malicious network traffic in an IoT ecosystem. Table 6 shows the attained results.

Table 6 indicates that the proposed novel iDRP framework techniques and methods perform better with other notable techniques for IoT malware detections. Significantly, the optimal results were achieved with the synthesis of both the discriminative features and differential evolution techniques for the universal approximation of the extrapolated hidden information together with connected structures in the integrated iDRP framework.

Table 6: Comparison of Classification Techniques on Standardized IoT Malware Datasets (BoTNetIoT) Results

No.	Techniques	Accuracy	Precision	Recall	F-score
1	Resnet34 [41]	92.39%	93.57%	64.55%	76.40%
2	Resnet50 [41]	94.50%	95.78%	94.02%	94.90%
3	MobileNet [41]	91.32%	91.67%	91.03%	91.35%
4	SOINN [41]	91.75%	89.68%	95.52%	92.50%
5	iDRP	**99.98%**	**99.95%**	**99.93%**	**99.97%**

Evidently, the innovative iDRP framework network techniques performed better with even the state-of-the-art MLP NN techniques and methods etc. The flexibility, extensibility, stability, and maintainability in the architecture design of the innovative iDRP framework network that were developed have demonstrated to be optimal in helping achieve the overall outcome. Through the leveraged innovative and systematic techniques implemented in the experiments of the proposed iDRP framework network, it is safe to mention that the optimal security and protection of heavily computerized network systems and smart technologies such as an IoT ecosystem against evolving attacks and threats can be better managed with improved innovation and intelligence in the digital ecosystem.

8. Conclusion

The novelty of this work are the systematic integrations, techniques and methods of the adaptive and scalable *pre-processing* and *post-processing* components designed and developed for multi-class classifications and capturing the immutable vector state structures from categorical distributions to mapped data distributions in order to estimate and exploit hidden textural features of converted polymorphic and new IoT malware binary images in correlations with their aggregated distributions in a network. Together with, the modular *detection*, *recognition*, and *prediction* components DL approach provided new concepts through the syntheses of the three (3) key components with the *multi-class classification*, *data augmentation*, *vector scaling and retention*, *clustering*, *cross-validation*, and *resampling* techniques and methods etc., synthesized and interconnected evolutionarily by the *post-processing* techniques in the proposed novel iDRP framework, which can be multidisciplinary. In particular, considering the textural and structural changes in polymorphic and new malware variants or obfuscation, the applied techniques and methods that were designed and developed ensured the improved and accurate *detection*, *recognition*, and *prediction* of the complex code malwares that specifically changes their binary structures on execution in an IoT ecosystem. A higher level of feature extraction and classification estimation of the modular hidden layer information were achieved to intelligently and efficiently detect the present-day obfuscation of malware attacks and infections in an IoT ecosystem. Overall, the derived intelligent immutability of the stately vectorized mapped data estimation techniques and methods implemented ensured optimal results in the experiments.

We conducted and tested the proposed implemented experiments with sound scientific procedures and processes to provide the proof-of-concept by conceptualizing and applying conventional scientific processes such as conducted in the engineering, medical, animal science fields etc., to establish empirical and statistical analysis in the performed experiments i.e., structurally complex, yet similar vector images; MNIST (2-D) and NORB (3-D) datasets prior to the comparative real IoT malware (BoTNetIoT) dataset converted image files as benchmark, which is analogous to animal to human experiments in the pure science and engineering fields systematically. The premise of the research experiments is to discover conclusive scientific approaches, techniques, and methods to solving polymorphic and new malware anomalies problems in an IoT ecosystem. These were achieved by applying the aforementioned procedures and techniques to solving the IoT malware problems with new ideas in this work. Extensive analysis was conducted by synthesizing the various techniques on the big datasets utilized in the proposed experiments. The syntheses of the integrated procedures and techniques in the proposed novel iDRP framework, which consists of five (5) components; *pre-processor*, *detection*, *recognition*, *prediction*, and *post-processor* components provided the seamless evolutionary information distributions for a high-level accuracy result in the experiment. Standardized metrics such as *accuracy* rate, *ROC/AUC* rate, *precision* rate, *True Positive* (TP) rate, *False Positive* (FP) rate, *True Negative* (TN) rate, *False Negative* (FN) rate, *recall* rate, and *F1-Score* were calculated and recorded to effectively determine the performance of the finalized generative DNN models in the proposed and implemented systematic experiments of the iDRP framework research.

In the end, the resultant output of the proposed experiments was tested and evaluated to achieve the best performance and consistency, accuracy, and precision etc., of the comparative aggregated model results achieved. As a result, a 100% accuracy rate, 100% ROC/AUC rate, 99.94% precision rate, and 99.90% recall rate were successfully achieved with the preliminary experiment, while with a real IoT malware data, BoTNetIoT, a 99.98% accuracy rate, 99.99% ROC/AUC rate, 99.95% precision rate, and 99.93 recall rate were successfully attained as solution with the synthesized techniques and methods to the regression and binary classification problems for intelligent malware anomalies detection in an IoT ecosystem.

Future Work

The future work would be to evaluate the systematic '*checkpointing*' of various weightages and best performing weights measurement together with the corresponding network topography and architecture in IoT security models, which will be verified, validated and explored further for better security solution in expanding heavily computerized interconnected systems.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors received no specific funding for this research work.

References

[1] O. Eboya, J. Juremi and M. Shahpasand, "An Intelligent Framework for Malware Detection in Internet of Things (IoT) Ecosystem," in 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), 1 - 6, 2020, doi: 10.1109/r10-htc49770.2020.9356961.

[2] R. AL MOGBIL, M. AL ASQAH and S. EL KHEDIRI, "IoT: Security Challenges and Issues of Smart Homes/Cities," in 2020 International Conference on Computing and Information Technology (ICCIT-1441), 1 - 6, 2020, doi: 10.1109/iccit-144147971.2020.9213827.

[3] E. Tabane and T. Zuva, "Is there a room for security and privacy in IoT?," in 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE), 1-6, 2016, doi: 10.1109/icacce.2016.8073758.

[4] C. Vorakulpipat, E. Rattanalerdnusorn, P. Thaenkaew and H. Dang Hai, "Recent challenges, trends, and concerns related to IoT security: An evolutionary study," in 2018 20th International Conference on Advanced Communication Technology (ICACT), 1 - 6, 2018, doi: 10.23919/icact.2018.8323773.

[5] Z. Zheng, S. Xie, H. Dai, X. Chen and H. Wang, "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," in 2017 IEEE International Congress on Big Data (BigData Congress), 1 - 6, 2017, doi: 10.1109/bigdatacongress.2017.85.

[6] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum and N. Ghani, "Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on Internet-Scale IoT Exploitations," in 2019 IEEE Communications Surveys & Tutorials, 2702-2733, 2019, doi: 10.1109/comst.2019.2910750.

[7] Y. Seralathan et al., "IoT security vulnerability: A case study of a Web camera," in 2018 20th International Conference on Advanced Communication Technology (ICACT), 1 - 6, 2018, doi: 10.23919/icact.2018.8323685.

[8] N. Karie, N. Sahri and P. Haskell-Dowland, "IoT Threat Detection Advances, Challenges and Future Directions," in 2020 Workshop on Emerging Technologies for Security in IoT (ETSecIoT), 1 - 6, 2020, doi: 10.1109/etseciot50046.2020.00009.

[9] O. Georgiana Dorobantu and S. Halunga, "Security threats in IoT," in 2020 International Symposium on Electronics and Telecommunications (ISETC), 1 - 6, 2020, doi: 10.1109/isetc50328.2020.9301127.

[10] M. Antonakakis et al., "Understanding the mirai botnet," 1st ed. Vancouver, BC, Canada: Proceedings of the 26th USENIX Security Symposium, USENIX Association, 2017.

[11] K. Sha, R. Errabelly, W. Wei, T. Yang and Z. Wang, "EdgeSec: Design of an Edge Layer Security Service to Enhance IoT Security," in 2017 IEEE 1st International Conference on Fog and Edge Computing (ICFEC), 1 - 6, 2017, doi: 10.1109/icfec.2017.7.

[12] I. Waz, M. Sobh and A. Bahaa-Eldin, "Internet of Things (IoT) security platforms," in 2017 12th International Conference on Computer Engineering and Systems (ICCES), 1 - 5, 2017, doi: 10.1109/icces.2017.8275359.

[13] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf and Y. Bangash, "An In-Depth Analysis of IoT Security Requirements, Challenges, and Their Countermeasures via Software-Defined Security," in IEEE Internet of Things Journal, 10250-10276, 2020, doi: 10.1109/jiot.2020.2997651.

[14] E. Baccelli et al., "RIOT: An Open Source Operating System for Low-End Embedded Devices in the IoT," in IEEE Internet of Things Journal, 4428-4440, 2018, doi: 10.1109/jiot.2018.2815038.

[15] R. Johari, N. Gaurav, S. Chaudhary and A. Pramanik, "START: Smart Stick based on TLC Algorithm in IoT Network for Visually Challenged Persons," in 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 1 - 6, 2020, doi: 10.1109/i-smac49090.2020.9243517.

[16] A. Gupta and R. Johari, "IOT based Electrical Device Surveillance and Control System," in 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), 1 - 5, 2019, doi: 10.1109/iot-siu.2019.8777342.

[17] L. Tawalbeh, H. Tawalbeh, H. Song and Y. Jararweh, "Intrusion and attacks over mobile networks and cloud health systems," in 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 1 - 6, 2017, doi: 10.1109/infcomw.2017.8116345.

[18] Y. Zheng, A. Pal, S. Abuadbba, S. Pokhrel, S. Nepal and H. Janicke, "Towards IoT Security Automation and Orchestration," in 2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), 1 - 6, 2020, doi: 10.1109/tps-isa50397.2020.00018.

[19] M. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali and M. Guizani, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," in IEEE Communications Surveys & Tutorials, 1646-1685, 2020, doi: 10.1109/comst.2020.2988293.

[20] F. Hussain, R. Hussain, S. Hassan and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," in IEEE Communications Surveys & Tutorials, 1686-1721, 2020, doi: 10.1109/comst.2020.2986444.

[21] M. Bettayeb, O. Waraga, M. Talib, Q. Nasir and O. Einea, "IoT Testbed Security: Smart Socket and Smart Thermostat," in 2019 IEEE Conference on Application, Information and Network Security (AINS), 1 - 6, 2019, doi: 10.1109/ains47559.2019.8968694.

[22] J. Moos, "IoT, Malware and Security," in 2017 ITNOW, 28-29, 2017, doi: 10.1093/itnow/bwx013.

[23] A. Zahra and M. Shah, "IoT based ransomware growth rate evaluation and detection using command and control blacklisting," in 2017 23rd International Conference on Automation and Computing (ICAC), 1 - 5, 2017, doi: 10.23919/iconac.2017.8082013.

[24] E. Ronen, A. Shamir, A. Weingarten and C. O'Flynn, "IoT Goes Nuclear: Creating a Zigbee Chain Reaction," in 2018 IEEE Security & Privacy, 54-62, 2018, doi: 10.1109/msp.2018.1331033.

[25] S. Sonune, D. Kalbande, A. Yeole and S. Oak, "Issues in IoT healthcare platforms: A critical study and review," in 2017 International Conference on Intelligent Computing and Control (I2C2), 1 - 5, 2017, doi: 10.1109/i2c2.2017.8321898.

[26] M. Mekala and P. Viswanathan, "A Survey: Smart agriculture IoT with cloud computing," in 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS), 1 - 6, 2017, doi: 10.1109/icmdcs.2017.8211551.

[27] P. Rughoobur and L. Nagowah, "A lightweight replay attack detection framework for battery depended IoT devices designed for healthcare," in 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), 1 - 7, 2017, doi: 10.1109/ictus.2017.8286118.

[28] A. Ukil, S. Bandyoapdhyay, C. Puri and A. Pal, "IoT Healthcare Analytics: The Importance of Anomaly Detection," in 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), 1 - 4, 2016, doi: 10.1109/aina.2016.158.

[29] S. Venkatraman and M. Alazab, "Classification of Malware Using Visualisation of Similarity Matrices," in 2017 Cybersecurity and Cyberforensics Conference (CCC), 1 - 5, 2017, doi: 10.1109/ccc.2017.11.

[30] J. Tang, C. Deng and G. Huang, "Extreme Learning Machine for Multilayer Perceptron," in 2016 IEEE Transactions on Neural Networks and Learning Systems, 809-821, 2016, doi: 10.1109/tnnls.2015.2424995.

[31] Y. Zeng, L. Qian and J. Ren, "Evolutionary Hierarchical Sparse Extreme Learning Autoencoder Network for Object Recognition," in 2018 Symmetry, 1 - 11, 2018, doi: 10.3390/sym10100474.

[32] M. Shahini, R. Farhanian and M. Ellis, "Machine Learning to Predict the Likelihood of a Personal Computer to Be Infected with Malware," in 2019 SMU Data Science Review, 1 - 24, 2019, doi: https://scholar.smu.edu/datasciencereview/vol2/iss2/9/.

[33] J. Villanueva, R. Juanatas and L. Lacatan, "Malware Predictor using Machine Learning Techniques," in 2020 Research Gate, 5665 - 5674, 2020, doi: https://www.researchgate.net/publication/339935591_Malware_Predictor_using_Machine_Learning_Techniques.

[34] A. Onal, O. Berat Sezer, M. Ozbayoglu and E. Dogdu, "Weather data analysis and sensor fault detection using an extended IoT framework with semantics, big data, and machine learning," in 2017 IEEE International Conference on Big Data (Big Data), 2037-2046, 2017, doi: 10.1109/bigdata.2017.8258150.

[35] J. Wang, C. Liu, M. Zhu, P. Guo and Y. Hu, "Sensor Data Based System-Level Anomaly Prediction for Smart Manufacturing," in 2018 IEEE International Congress on Big Data (BigData Congress), 158-165, 2018, doi: 10.1109/bigdatacongress.2018.00028.

[36] N. Zeeshan, M. Reed and Z. Siddiqui, "Three-way Security Framework for Cloud based IoT Network," in 2019 International Conference on Computing, Electronics & Communications Engineering (iCCECE), 183-186, 2019, doi: 10.1109/iccece46942.2019.8941877.

[37] B. Li, X. Ming and G. Li, "Big data analytics platform for flight safety monitoring," in 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), 1 - 3, 2017, doi: 10.1109/icbda.2017.8078837.

[38] Y. Xing, H. Shu, H. Zhao, D. Li and L. Guo, "Survey on Botnet Detection Techniques: Classification, Methods, and Evaluation," in 2021 Mathematical Problems in Engineering, 1-24, 2021, doi: 10.1155/2021/6640499.

[39] Y. Song, S. Hyun and Y. Cheong, "Analysis of Autoencoders for Network Intrusion Detection," in 2021 Sensors, 4294, 2021, doi: 10.3390/s21134294.

[40] T. Palla and S. Tayeb, "Intelligent Mirai Malware Detection for IoT Nodes," in 2021 Electronics, 1241, 2021, doi: 10.3390/electronics10111241.

[41] G. Bendiab, S. Shiaeles, A. Alruban and N. Kolokotronis, "IoT Malware Network Traffic Classification using Visual Representation and Deep Learning," in 2020 6th IEEE Conference on Network Softwarization (NetSoft), 1 - 9, 2020, doi: 10.1109/netsoft48620.2020.9165381.

[42] R. Kozik, M. Pawlicki and M. Choraś, "A new method of hybrid time window embedding with transformer-based traffic data classification in IoT-networked environment," in 2021 Springer, 1 - 9, 2021, doi: 10.1007/s10044-021-00980-2.

[43] A. Alhowaide, I. Alsmadi and J. Tang, "Towards the design of real-time autonomous IoT NIDS," in 2021 Cluster Computing, 1 - 14, 2021, doi: 10.1007/s10586-021-03231-5.

[44] K. Sudheera, D. Divakaran, R. Singh and M. Gurusamy, "ADEPT: Detection and Identification of Correlated Attack Stages in IoT Networks," in 2021 IEEE Internet of Things Journal, 6591-6607, 2021, doi: 10.1109/jiot.2021.3055937.

[45] Q. Ngo, H. Nguyen, V. Le and D. Nguyen, "A survey of IoT malware and detection methods based on static features," in 2020 ICT Express, 280-286, 2020, doi: 10.1016/j.icte.2020.04.005.

[46] P. Sudhakaran, C. Malathy, T. Vardhan and T. Sainadh, "Detection of Malware from IOT Devices Using Deep Learning Techniques," in 2021 Journal of Physics: Conference Series, 1 - 7, 2021, doi: 10.1088/1742-6596/1818/1/012219.

[47] B. Khammas, "The Performance of IoT Malware Detection Technique Using Feature Selection and Feature Reduction in Fog Layer," in 2020 IOP Conference Series: Materials Science and Engineering, 1 - 11, 2020, doi: 10.1088/1757-899x/928/2/022047.

[48] K. Han, J. Lim, B. Kang and E. Im, "Malware analysis using visualized images and entropy graphs," in 2015 International Journal of Information Security, 1-14, 2015, doi: 10.1007/s10207-014-0242-0.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 189-204 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Survey of Big Data Techniques for Extracting Information from Social Media Data

Carla Blank[1], Matthew McBurney[1], Maria Morgan[1], Raed Seetan[2,*]

[1]*Department of Mathematics and Statistics, Slippery Rock University, Slippery Rock, 16057, United States*

[2]*Department of Computer Science, Slippery Rock University, Slippery Rock, 16057, United States*

ARTICLE INFO

Article history:
Received: 10 February, 2021
Accepted: 11 May, 2021
Online: 23 May, 2021

Keywords:
Data Mining
Social Media
Big Data

ABSTRACT

Data mined from social media can be used in a variety of methods. The goal of this paper is to explore some of the various methods of mining data from social media and the different areas of its applications. From the analysis of other studies, it was clear that methods such as text analysis, classification, clustering, mapping, testing/validity methods, regression, and research methods were the overarching themes of the previously done research. Pros, cons, and possible extensions were examined for the current research evaluated in the social media data mining area. At the conclusion of this survey, our research team found that text analysis, sentiment analysis, and support vector machine classifiers were among the most common themes of the research methods in this field. In most cases, multiple methods were attempted for each topic to be able to cross compare results.

1. Introduction

In this study, we examined the various techniques that have been used in data mining for social media data. With many social media platforms used by millions, such as Facebook, Twitter, Instagram, etc., all posts by users are collected and comprise the "big data" of social media. Social media data can be used in a variety of ways from predicting buying habits of users to generating information about emergency events.

The idea of using social media to make predictions about individuals is more relevant than ever. Across all types of individuals, social media usage has shown to be on the rise. For example, Figure 1a-c shows the social media growth over three variables (age, race and income). In all cases, more social media use is being experienced allowing us to collect data on a wide range of individuals for a variety of applications. Similarly, Figure 2 shows the frequency that these applications are being used. As seen, for the majority of the individuals, daily data is being obtained which can help better explain the behaviors showcased by people [1].

Selecting efficient methods in the task of social media data mining is a critical consideration for researchers because the methods used need to be executed efficiently when there is a large amount of data. Researchers must also consider the various types of data that are present when analyzing social media data.

A previous survey of data mining techniques in social media analyzed 66 reports on data mining in social media [2]. The authors of this paper mention the important fact that establishing successful data mining techniques in social media is difficult because of the amount of data and the velocity of that data; all social media data is collected in real time so there is constantly more data being added to the repositories where this data is stored. In this study, the authors identified 19 different data mining techniques that were used by researchers to analyze social media data and compared the effectiveness of these techniques. The authors noted that many of the researchers did not back up their methods with statistical significance testing, which presents an issue with the reliability of the methods used.

For the purpose of this research, we analyzed 12 previous studies that have been conducted on data mining social media data. While conducting this research, the following research questions were kept in mind to guide the focus of our analysis: what topics (or areas of interest) have been explored in regards to social media data mining, what methods have been used for data mining in social media data in these areas, are certain data mining methods applied to specific areas/topics, how do each of these data mining methods perform and can we make suggestions on how to approve future analysis projects on social media data mining.

*Corresponding Author: Raed Seetan, Computer Science Department, Slippery Rock University, 16057, USA, Email: raed.seetan@sru.edu Tel: 724-738-2940

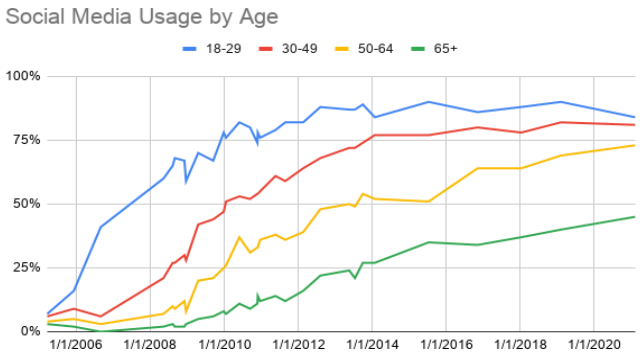


Figure 1a: Social Media Use by Age

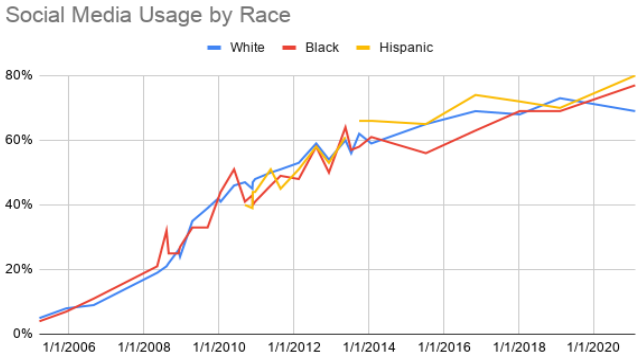


Figure 1b: Social Media Use by Race

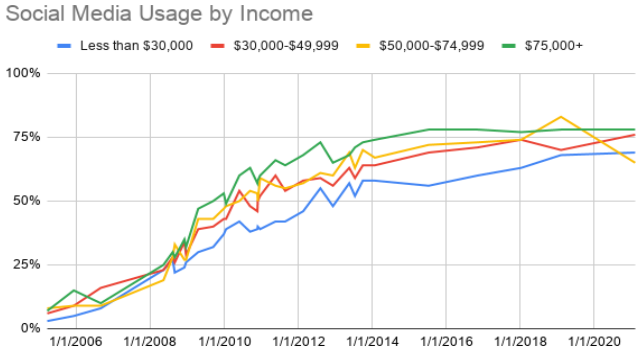


Figure 1c: Social Media Use by Income

The remainder of this paper will be structured as follows: we will first discuss the various methods that have been used in data mining social media data, we will then provide summaries on the 12 papers selected for this research followed by an overview of the findings from each paper, and finally we will provide a conclusion on the research that has been conducted for this paper.

2. Methods

There are many machine learning methods used in analyzing big data. The methods discussed below are used in the studies observed in this research paper.

2.1. Text Analysis

Text analysis is used to gather important information from unstructured text. Patterns of specific text can be collected, or common words and phrases can be determined using text analysis. Many text analysis methods use natural language processing (NLP). Text analysis is frequently used in social media studies as users can post about anything they chose. Valuable insights can be derived from these data, such as companies collecting feedback about their brand or a specific product.

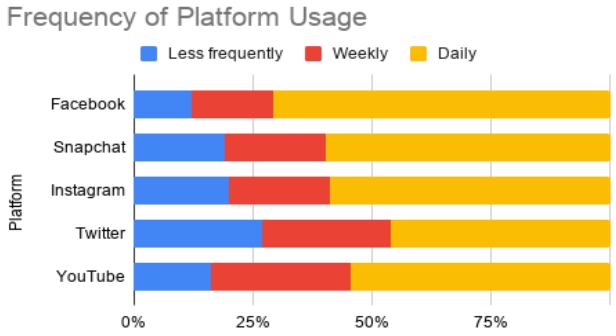


Figure 2: Frequency of platform usage

1) Latent Dirichlet Allocation (LDA): Latent Dirichlet Allocation (LDA) is a topic model in natural language processing. LDA allows sets of observations to be explained by unobserved groups that explain the similarity in some parts of the data. In LDA, both word distribution for each topic and topic proportion for each document are modeled using Dirichlet distributions [3]. Common topics and words within the text data can be ranked by their resulting probabilities to determine which occur most frequently.

2) Normalized Discounted Cumulative Gain: Normalized Discounted Cumulative Gain (NDCG) is a method for measuring the quality of a set of search results. The result of NDCG is a scaled ranking system which tells the user which results are relevant using a numbered ranking system. For example, a score of 0 meaning least relevant and a score of 2 meaning most relevant. In NDCG, very relevant results are more important than partly relevant or irrelevant results, which is called cumulative gain. Discounting is when relevant results are more important when they appear earlier in the search results. The ranking of the result should not depend on the query performed; this is normalization [4]. In discounted cumulative gain, results that appear lower are discounted, i.e., they are not as valuable as results that appear higher in the search results. Normalization scales the results to the best result seen.

3) Sentiment Analysis: Sentiment analysis is the classification of textual data to uncover classifications of emotions that can be seen in the text. The method is either rule based, governed by manually determined rules (like looking for specific words or phrases), automatic, governed by machine learning techniques, or a combination of the two. Naïve Bayes, Linear Regression, Support Vector Machines, and Deep Learning are commonly used classification algorithms to perform sentiment analysis [5]. Given the variety of machine learning techniques that can be used to perform sentiment analysis, the user is able to select which technique best suits their needs based on the type of data being used and the desired format of the result.

4) Text Analysis/Classification: Text analysis and classification have a primary goal of taking something which may be unstructured and trying to impose a structure on it. From a classification standpoint, this involves taking an unsorted list of items and attempting to assign them to appropriate groups. Text analysis, similarly, takes text and makes it understandable and easier to manage. This can be helpful while analyzing masses of text data such as social media, comments, and surveys [6]. Like sentiment analysis, the user has options to choose from when deciding how to perform text classification. Users can use a rule-based system, which classifies text into certain groups based on a set of linguistic rules that are pre-determined by the user. This method requires more work from the user to determine the rules to apply when classifying. Machine learning-based systems, such as Naive Bayes and Support Vector Machines, may also be used in text classification. It is up to the user to decide which method will perform best based on the size of their dataset and the result they wish to achieve.

5) TF-IDF: Term frequency - inverse document frequency (TF-IDF) serves as a statistical measure which determines the relevancy of a specific word in a defined collection of documents. TF-IDF is a widely accepted statistic because it not only accounts for how often a word appears, but also how many documents they appear in. Therefore, words such as "and" are not deemed inaccurately as relevant. This method is used for text vectorization (turning text into numbers) to further be processed by machine learning algorithms for various applications [7].

6) Twitter API Search: Twitter standard search API returns a set of tweets considered relevant by whatever the specified query was [8]. This method is used to gather results that can later be used in text analysis and sentiment analysis. An API search allows for "tweets" to be pulled from as early as 2006. This serves as helpful data when needing to examine text from a specific period of time (or comparison of times) for the analysis.

7) Active semi-supervised Clustering based Two-stage text classification (ACTC): ACTC is a method that turns text classification into a supervised problem [9]. This algorithm is done in two stages. The first stage is a clustering stage in which the data is clustered using a method called SemiCCAc; this algorithm was created to address that cluster-bias problem that often exists when clustering text data. The second step is a classification method using discriminative classifiers. In the first part of the algorithm, the unlabeled text data is clustered to create tentative labels for the clustered data. In the second part, the discriminative classifier is used to assign the final labels to each group of the text data. The resulting output after the classification stage is complete may require additional processing using supervised learning techniques. The idea behind the method of combining clustering with classification is to attempt to minimize the amount of text data that would be mis-labeled by using either clustering or classification alone; this combination should result in a reduction of the further processing of the output that would be needed when compared to only using one of the methods alone.

8) Example Adaption for Text categorization (EAT): Traditional text categorization algorithms work by categorizing text data based on negative and positive labels [10]. For example, if we have defined a category to gather all text related to plants and we wish to analyze a new set of text data to determine if this data is related to this category, the new data set may not be easily identified as either positive or negative because there are a variety of terms that may be used to identify text related to the plant category. The EAT algorithm works by teaching the classifier given the new dataset that is to be used. The authors who proposed this method gave the example of categorizing text data about an election. In analyzing a new text dataset, we also want to look out for terms such as "voters" to be associated with the election dataset. This new method works by identifying new terms that should be considered by the classifier in order to assign the text data a positive or negative label.

9) General Architecture for Text Engineering (GATE): GATE is a set of tools written in Java which can be used for analysis projects involving things like text mining and information extraction [11]. There are various tools that make up the GATE system, such as ANNIE, which is used for information extraction given a text dataset as input. Users of GATE can perform a wide variety of analytical tasks on text data, and there are various software that offer user-friendly versions of the system in a GUI environment. The GATE system accepts text datasets in various formats, such as txt, html, pdf, etc. The various inputs that are allowed in this system make it an ideal system in a field such as social media data mining because the text data associated with these tasks can often come in various formats. Because the user has the option of a GUI environment, results may be interpreted in a simpler manner as opposed to requiring additional processing.

2.2. Classification

Classification can be either supervised or unsupervised machine learning models, although they are commonly supervised learning. For unsupervised learning, the software determines the characteristics of the data to define classes. For supervised learning, a user can select the classes or use training sets with classes defined in order to train the software. These are used to predict the class of new data points.

1) Discriminant Analysis: Linear Discriminant Analysis (LDA) is both a classifier and a dimensionality reduction technique. LDA is the preferred method of analysis when there are more than two classes [12]. LDA is commonly used in the pre-processing step of machine learning and pattern classification projects. LDA uses Bayes' Theorem to estimate the probability that a new set of input values belongs to every class. The output class is the class with the highest probability. LDA assumes that the data are normally distributed and also assumes a common covariance matrix. Quadratic Discriminant Analysis (QDA) is a variant of LDA that allows for non-linear separation of data. There is not an assumption that the covariance of each of the classes is identical in QDA as with LDA [12].

2) Naïve Bayes: Naïve Bayes is a probabilistic classification method based on Bayes' theorem. It makes decisions using the Maximum A Posteriori decision rule in a Bayesian setting [2]. The goal is to determine the probability of the features occurring in each class and to return to the most likely class. The only thing that must be done before prediction is finding the parameters for the features' individual probability distributions. A naïve Bayes classifier assumes that the presence or absence of a particular feature of a class is unrelated to the presence or absence of other features. Naïve Bayes classifiers are popular for text classification. A classic use of naïve Bayes classifiers is spam filtering.

3) Support Vector Machines: Support Vector Machines are extremely popular supervised learning models. This method of classification creates a decision boundary which is designed by training points to then classify new data points based on which side of the boundary they fall on. Support vector machines offer a large advantage because they can create nonlinear boundaries easier than some other methods. This is done using a kernel trick which replaces its predictors with a kernel function [13].

4) Decision Trees: Decision trees are a way of performing classification on datasets [14]. Decision trees are popular in machine learning because they are easy to understand. The idea behind decision trees is that similar data points are grouped together in leaf nodes based on decisions that are made. These decisions are represented by the branches of the tree. Each lead node represents a label for a class. Data points are assigned to these classes by following the structure of the tree and going through the tests represented by each branch. Data points are assigned to a leaf node based on the result of the test at each branch that it goes through. So, for unlabeled text data, which is often what data sets consist of in social media data mining, the text data would go through the various tests along the tree structure in order to be assigned a class label.

5) K-Nearest Neighbor (KNN): KNN is a classification algorithm used in machine learning [15]. This algorithm works by analyzing the raw data points specifically by what other data points in the dataset are close to each other. A value K is selected by the researcher; this determines the number of "neighbors" each data point should have. The distance between a raw data point and it's neighbors is pre-determined by the researcher, meaning for a data point to be a neighbor it must be within a certain distance. Class labels are formed by the groups that are formed between each raw data point and its neighbors. This algorithm can be trained on a training dataset and tested on a test dataset in order to select the optimum distance and optimum value of K.

6) Logistic Regression: Logistic regression is a powerful classification algorithm; however, it is limited to two classes or binary classification problems [2]. While it can be extended for additional classes, it is rarely used this way. Logistic regression can be used to predict the likelihood of an outcome based on the input variables.

2.3. Clustering

Clustering is an unsupervised machine learning technique in which data are grouped together in order to make assumptions about the data. The data in each group, or cluster, are most similar to the other data in the same group as opposed to the other groups. It is useful in determining characteristics of data, such as clustering customer orders to determine which types of customers purchase certain products. Outliers are also more evident when using clustering. There are many types of clustering algorithms. Some of the more commonly used algorithms are K-means clustering, DBSCAN clustering, Affinity Propagation clustering, BIRCH clustering, and Mean-Shift clustering. The clustering methods discussed below are the methods used in the studies reviewed in this paper.

1) Positive examples and Negative examples Labeling Heuristics (PNLH): the purpose of Positive examples and Negative examples Labeling Heuristics (PNLH) is to extract reliable positive and negative examples [16]. Reliable negatives are extracted using features that occur more frequently in positive data. Then the sets of positives and negatives are enlarged by clustering the reliable negatives. If an example is close to a positive but not close to a negative example, it is added to the reliable positives. If an example is close to a negative example but not close to a positive example, it is added to the reliable negatives.

2) Similarity Matrix: a similarity matrix is a method of expressing the closeness of relationship between two data points. There are various methods of calculating a similarity score. A similarity matrix has many applications in clustering. Additionally, methods exist based on the similarity matrix such as the clustering-based similarity partitioning algorithm. This matrix provides a convenient way to showcase how points are related and showcase the clustering rationale [17].

3) Parallel Density-based Spatial Clustering with Noise (PDBSCAN): DBSCAN is a clustering method that utilizes a density method to cluster data points, it is highly used in the machine learning field [18]. For each data point within a cluster, there needs to be a minimum number of other data points within a predetermined radius of that data point. So, data points are grouped based on two predetermined parameters. The first parameter is "eps" which indicates how close the data points are to be to each other to be considered as being in the same cluster. The second is "minPoints" which tells the algorithm how many points must be within the radius of each data point. This traditional model of DBSCAN can be computationally expensive with large datasets, thus an updated version of this method called PDBSCAN has been proposed for big data mining, such as social media data mining [19]. This parallel version of this algorithm works by first creating local clusters within different pieces of the data. This is done in parallel, so the data is divided into sections and these local clusters are formed at the same time. Then, these local clusters for each section are combined to form the final clusters. Researchers have found that the parallel adjustment to the DBSCAN algorithm has greatly improved its performance with large datasets.

4) K-means Clustering: K-means clustering is an unsupervised learning algorithm that is used to cluster similar data points together [20]. The raw dataset is analyzed so that centroids are formed based on a given number K, K being the number of clusters that are to be formed from the data. Data points are assigned to clusters based on their similarity to the cluster's centroid. This algorithm will typically involve a lot of testing to determine the optimal value of the parameter K. K-means clustering is often performed by building the model on a subset of the raw data, called a training dataset, and then testing the algorithm on the remaining data, called the test dataset. Using this methodology of testing and training the algorithm allows researchers to more accurately cluster their data using the optimal value of K.

2.4. Mapping

Mapping is a machine learning technique in which input data is mapped to a category or class. It is a type of Artificial Neural Network (ANN) and can be either supervised or unsupervised.

1) Self Organizing Maps: Self organizing maps, while coming with a set of challenges, offer the great advantage of not needing humans to help them learn. This method is based on unsupervised learning and is a cooperative learning network. Self-organizing maps take high dimensional data and make it low dimensional. The map units are typically thought of as points on a two-dimensional lattice. Neighborhoods are created by identifying sets of points which can be classified into the same class, based on traits and attributes deemed best by the unsupervised learning technique [21].

2) Supervised Mapping: Supervised learning techniques require the use of training data as examples of input - output pairs. These examples allow the method to learn and appropriately classify new inputted data to the appropriate category. Such a technique uses a learning algorithm such as support vector machines, linear or logistic regression, naïve bayes, or linear discriminant analysis to name a few. The performance of the supervised learning technique can be very dependent on which algorithm is chosen to do the classification task [22].

2.5. Statistical Test/Model Validity Methods

Statistical tests use mathematics to determine the relationship between random and nonrandom variables. The random variables do not have a statistical influence on the dataset. Model validation shows the accuracy of the model. There are many methods that can measure the accuracy of a model, such as k-fold cross validation, Wilcoxon signed-rank test, and train/test split. The tests and methods used below were used in the studies observed for this paper.

1) Goodness of Fit: Goodness of fit is a measure between the observed values and the expected values of a model [23]. Goodness of fit describes how well a statistical model fits the dataset presented. It tells how well the sample data represents data in the actual population.

2) Kernel Density Estimation (KDE): The Kernel Density Estimation is a mathematical process of finding an estimated probability density function of a random variable [24]. KDE attempts to make inferences of a population based on a representative data set. The technique creates a smooth curve given a set of data. The estimation can also be used to generate points that look like they were from a dataset, which can help power simulations [24]. The KDE is created by plotting the data and creating a curve of the distribution. The curve is calculated by weighing points at a specific location. When more points are grouped at a location, the estimation is higher because the probability of a point being at that location is higher. The kernel function is used to weigh the points.

3) Precision at Rank k: Precision at k is the proportion of recommended items in the top k-set that are relevant to the number of recommended items [25]. A downfall of precision at k is that it fails to take the position of the recommended items in the top k into consideration.

4) Analysis of Variance (ANOVA): ANOVA is a statistical method that is used to examine differences among variables across two or more categories [26]. In simple terms, ANOVA allows analysts to identify the differences among different groups of variables at the same time. This is a very important technique in the field of social media data mining because it allows researchers to identify the unique relationships that may exist between certain social media users, for example. Furthermore, it allows researchers to then perform further analysis to find the reasonings behind these differences. The result of an ANOVA test is the F-statistic, which allows researchers to analyze the variance that exists among the various groups of data. Using the F-statistic, researchers can see if there is a real variance among groups of data, if true variance does not exist then the F-ratio resulting from the ANOVA test should be close to 1.

2.6. Regression

Regression is a statistical technique used to describe relationships between dependent and independent variables. There are various types of regression, including linear regression, logistic regression and multiple regression. The techniques described below are those used in the studies discussed in this paper.

1) Multiple Regression: Multiple regression is designed to create regression models based on multiple independent variables and a single dependent variable [27]. The model that is created provides the impact that the independent variables have on the dependent variable. It is important to keep the number of variables low because the model becomes more complex as variables are added.

2) Time Series Analysis: Time series analysis consists of taking time series data and trying to find the underlying pattern in the data to extract meaningful statistics or characteristics that would be helpful in forecasting or predicting future values. Various prediction methods exist, and many depend on the acknowledgement of cyclical, periodic, and seasonal behavior that exists in the data [28].

2.7. Research Methods

The research methods discussed in this section are methods used for analysis for various research projects. These methods did

not fit into the previous categories of methods defined. Only methods discussed in the observed studies are listed.

1) 5W Method (What, Where, When, Why, Who): The 5W method (What, Where, When, Why, Who) is a general model that can be used to approach a variety of analysis problems [29]. This method works by identifying the main themes related to the problem to be solved. In the field of data mining in social media, we can describe this model in general terms. The What aspect identifies what the problem is, what kind of data are we looking for? The Where aspect identifies where the data is located that we need for analysis. The When aspect identifies when the data of interest is to be collected. The Why aspect identifies why the event of interest occurred. The Who identifies who the social media data is being retrieved from (what users' posts are we interested in).

2) Inductive Content Analysis: Inductive content analysis is a technique used by researchers when the data of interest is not numeric, such as text data often seen in social media data mining tasks [30]. Researchers use inductive content analysis to identify features and develop theories about their data of interest. The process is inductive in that researchers analyze their dataset to come up with themes and conclusions by repeatedly analyzing their data of interest through a series of repeated steps. A process of inductive content analysis typically begins with the researchers organizing their data through some method such as classification and identifying the unique groups that are formed. These processes are then repeated until the researchers have come up with theories associated with their data that can either be further tested or explained for the purpose of the analysis task at hand.

3) Speeded-Up Robust Features (SURF): SURF is an advanced feature descriptor used to represent photo content [31]. It is a fast and robust algorithm used for feature extraction, feature description and feature matching.

3. Study Summaries

3.1. Crowdsourcing based Timeline Description of Urban Emergency Events using Social Media

A previous study utilized the 5W model in order to crowdsource information on urban emergency events from social media posts [29]. The 5W model consists of describing the what, where, when, who, and why information about the event. This study uses posts from Weibo users to test the effectiveness of their model in providing information on urban emergency events. Weibo is a social media platform, similar to Twitter, which has approximately 500 million users in China. The components of the 5W method are described and generated as follows:

1) What: What is the event that occurred? Generated by analyzing keywords from Weibo user posts, the authors refer to these keywords as "concepts". For example, if the event we are interested in is a fire, "fire" would be one of the concepts used to identify posts related to that event. Posts that contain these concepts are referred to as "positive samples".

2) Where: Once the positive samples are identified from the "what" element, spatial information can be extracted from those posts. For example, a post like "I see a fire on Huaihai Road" indicates that there is a fire and where that fire is located. Geographic Information System (GIS) based information from check-in locations can also be utilized to determine the "where". Two location-based service applications in China, Jiepang and Dianping, allow Weibo users to post their check-in location.

3) When: Determining when the event occurred can be done by determining the earliest time stamp for the posts in the positive sample and the latest time stamp. The first timestamp is referred to as the starting timestamp (ST) and the last timestamp is referred to as the ending timestamp (ET).

4) Who: Weibo posts can identify witnesses of the event and potentially the participator of the event. The witness would be the poster of the Weibo event, and the participator may be revealed from the witness' post.

5) Why: Posts by Weibo users may reveal why the event occurred, this will often come from posts by news outlets or police departments.

The authors present two case studies, a fire and a hijack situation, to test how the 5W model can be used to gain information on urban emergency events by crowdsourcing social media posts. The authors demonstrated that this model was beneficial in retrieving the needed information concerning each event from Weibo posts. The authors were able to identify relevant posts for each situation, which they called positive samples. These posts were retrieved by searching Weibo posts in the locations of the two events on the days in which the events occurred. From these positive samples, the authors were able to identify what happened (the event of interest), where the event occurred (from location information attached to the posts), when the even occurred (by looking at time stamps related to the posts in the positive samples), who the event affected, and why the event occurred.

The authors conclude that the 5W model was successful in extracting information about the urban emergency events presented in the two case studies. Future applications of this model may involve comparing the effectiveness when applying it to Weibo posts and posts from a different social media platform, such as Twitter or Facebook.

3.2. Mining Social Media Data for Understanding Students' Learning Experiences

In [32], the study conducted focused on analyzing Twitter posts from engineering students to understand the issues they face in regard to their educational experiences [32]. The model developed was then used to detect student issues in Twitter posts from students at Purdue university. The two main research goals of this study were to analyze social media data for educational purposes using various data mining techniques and qualitative analysis and to analyze Twitter posts from engineering students to understand the issues faced by these students in regard to their educational experience.

The authors began by collecting Twitter posts that used the hashtag "#engineeringProblems" over a time period of 14 months to create their first dataset. The second dataset was created by collecting all Twitter posts that are geo-tagged with the location

of Purdue University through a search of the Twitter search API. The Tweets gathered by searching for the hashtag showed that the problems faced by engineering students fall into distinct categories. The authors used a Naïve Bayes classifier to classify the Tweets into their categories. This classifier was then used to classify the Tweets collected from Purdue University to arrange the engineering students' Tweets into multiple categories based on the issues they were facing. The authors' multi-label classification model allowed for one Tweet to fall into multiple categories. The researchers found that the main issues faced by engineering students were things like sleep deprivation and lack of study-life balance.

Because the Twitter posts obtained from the hashtag search did not contain pre-defined categories, the authors used inductive content analysis to identify categories within the dataset. The authors were interested in developing categories based on the major issues faced by engineering students. From this inductive content analysis, the authors developed six major classes: heavy study load, lack of social engagement, negative emotion, sleep problems, diversity issues, and others. The authors used multi-label classification so that one Twitter post could fall into more than one of these classes, with the exception of the others class. The Twitter posts were organized into these classes by two researchers, and F1 scores were generated based on how each researcher labeled each post in order to decide on the finalized class assignments. The authors found that the Naïve Bayes classifier was the most effective classification method for their datasets. This classification method was then used to classify the Twitter posts from the Purdue University dataset into the six classes. The authors believe that the methods they have presented here can be applied to other social media platforms, such as Facebook, to further analyze issues faced by students.

3.3. Investigating the impact of social media advertising features on customer purchase intention

Social media ads have become increasingly popular, with companies investing a large amount of money for them. The biggest rationale of choosing social media ads is the idea that they can be customized to the specific individual. For example, based on what users post about, like, and share, advertisers can predict the most relevant ads for the user to see that will result in a sale or a click.

The study lists out two primary questions,

1) What is a suitable conceptual model that could be adopted to provide a clear picture covering the main aspects related to social media advertising?
2) What are the main factors associated with social media advertising that could predict the customer's purchase intention?

In order to figure out what data need mined, a variety of previous studies were analyzed looking at various factors such as attitude, irritation, informativeness, creativity, and privacy concerns were targeted using platforms such as Facebook, Twitter, and Myspace. Based on customer researchers, they found that there was a strong correlation between the performance expectancy of ads (how well they were geared to individuals) and people's views on social media advertising [33].

Structural equation modeling was the technique used for analyzing a proposed model based on Cronbach's alpha value. Some methods used in the analysis of the model was Goodness of Fit Index, the Adjusted Goodness of Fit Index, Normed-Fit Index, Comparative Fit Index, and Root Mean Square Error of Approximation.

A hypothesis test was run on each of the factors to see if they were significant in contributing to the prediction of proper advertisements leading to increased chance of a consumer purchasing using regression analysis. Multicollinearity tests ensured for the multiple regression that the variance inflation factors were not too high. Average variance extracted (AVE) and composite reliability (CR) were additionally tested to make sure they were in their recommended levels.

The research done in this study was able to confirm that being able to identify these characteristics can heavily improve the ability to target customers with appropriate ads. Hence, when mining social media posts in real time these factors are the important aspects to be able to filter out and utilize [33].

3.4. Corpus for Customer Purchase Behavior Prediction in Social Media

One application of making these models was done in a data mining study that analyzed how well iPhone sales could be predicted based on social media data. By analyzing tweet's such as "I bought an iPhone" and "Checking out the new iPhone", the corpus collected English tweets related to purchase behavior and then manually recorded if a user purchased a product [34]. After defining key phrases, extracting product details from eBay pages, and analyzing tweets. The proposed methodology by researchers in this type of situation would be to use machine learning to train a classifier on whether a phrasing results in a buy or not.

The corpus created for data mining tasks like this can be automated as much as possible. Examples of automated labeling tasks were provided such as finding the location of users, frequency, and past analysis of the same user. All of these items are believed to enhance a machine learning corpus to provide much more accurate and meaningful results [34].

3.5. Predicting Postpartum Changes in Emotion and Behavior via Social Media

As opposed to being able to help predict marketing and purchasing information, the researchers utilized user data from social media to predict the effect that giving birth would have on a mother. In order to accomplish this, tweets were analyzed for dimensions of social engagement, emotion, social network, and linguistic style [35].

In order to develop a predictive model and sort through the tweets, a two-stage approach was used to construct the sample of new mothers. Twitter being public presents a wide range of challenges because of the number of public posts and ones that may be irrelevant to the predictions being made. To condense the data set, the first strategy was to analyze tweets to find users that appeared to have a recent childbirth in their lives. These remaining individuals served as possible individuals for the sample.

The various factors of engagement, ego-network, emotion, and linguistic style were quantified so they could be rated based on the analysis of the tweets. Since Twitter maintains all of the Tweets from the user since the beginning of the account, the differences in these factors were able to be determined from before the child birth to after the child birth. With the calculations of these differences, groups were formed (extreme-changing mothers and standard-changing mothers) to use in supervised learning techniques.

To avoid overfitting the data, principal component analysis and randomized forest procedures were utilized. Methods such as linear, quadratic, discriminant classifiers, naïve Bayes, k-nearest neighbor, decision trees, and SVM with radial kernels were used in the classification process. Out of all of these, the SVM had the greatest accuracy with using a five-fold cross validation method. Using this methodology, there was a highest point of 80% accuracy on a testing data set of being able to classify new mothers who would experience postpartum depression after giving child birth [35].

3.6. Big data analytics for disaster response and recovery through sentiment T analysis

Big data from Twitter and other social media has been used for many applications, such as healthcare, multi-channel, finance, log analysis, traffic control, homeland security, telecommunications and retail marketing but is only in the early stages for disaster response and recovery [36]. Emergency help related requests during disaster situations rarely use social networks [36]. Because crisis situations are chaotic and disorganized, big data analysis of these data is a good fit for handling the environment. Typically, because of lack of information for the affected team, disaster management teams rely on incomplete or inaccurate information. Big data analytics can assist in getting the right information from large amounts of data, analyzing it and providing results to the disaster management or rescue teams to enable them to make the best course of action.

The authors of this study propose a method to identify a sentiment towards the philanthropic aids received by the people during and after a disaster [36]. While governments typically provide relief, the needs of people are not always met as the need is not always identified. This study uses tweets collected during three disasters to help build a model to meet the needs of those affected. First, the data are analyzed and categorized by the various needs of the people during and after the disaster. Second, various features, such as bag of words, parts of speech-based features, and various lexicon-based features are analyzed and the best performing algorithm in each category is identified [36]. Finally, a visualization method is proposed for the sentiment of the basic needs which would help the emergency responders. In order to provide help to emergency responders, a solution must be built that uses real-time data and analysis to determine the needs of the people during a disaster. The objective of this research is to build a conceptual model for disaster response and recovery by using the best features that classify the disaster data with highest accuracy [36].

The data obtained for this study were obtained for the following disasters: the India-Pakistan floods in September 2014, a seven cyclonic storm named HUDHUD in October 2014 and another seven cyclonic storms named Nilofar. These data were collected from Twitter and contained 70,817 tweets. Part of these tweets were collected from Streaming Twitter API. This API only allows users to collect the past seven days' data. Since historical data cannot be obtained through this API, the rest of the data was collected using a third-party vendor 'Followthehashtag'. Keywords used to collect data are as follows: HUDHUD, Vizag flood, Nilofar, Kashmir floods, India-Pakstan floods, Pakistan floods [36]. The data collected for each disaster was collected for a specific time frame: HUHUD and Nilofar were collected from September 5, 2014 to November 15, 2014; Kashmir floods were collected from September 1, 2014 to October 15, 2014.

The preprocessing step is completed so that the machine learning algorithm can interpret the data. Since Twitter data contains Uniform Resource Locators (URLs), numbers, abbreviations, etc., it is important to remove these before performing further steps. The disaster area tweets are categorized by keyword filtering technique, which is common in Twitter analytics. The keywords are selected by identifying words that are used more than five times and also relevant to the category. Two steps are used in the text classification stage, the subjective sentences segregation and feature vector generation. The subjective sentences provide the sentiment related information, while features convert the subjective sentences to feature vectors to be used in the machine learning algorithm [36]. Text is separated into subjective and objective sentences based on the usage of words. The sentiment of people is identified in the subjective sentences. Since separating subjective and objectives is a manual task, tools can be used to automate the process. Publicly available sentiment-based lexicons are used to categorize the subjective sentences. To categorize disaster data, three lexicons were considered for the model, one from Twitter domain (AFINN) and the other lexicons which contain slang words, misspelled words, and morphological variants [36]. The disaster data is filtered with the list of positive and negative words from the lexicons. Anything not classified as positive or negative is removed. Parts of speech (POS) tagging is carried out to filter the adjectives and adverbs from entire streams of tweets. Bag of words (BOW) feature is where the entire text is represented as a list of words. The occurrence of each word is used as a feature in training a classifier [36]. Another feature is combining words. The adjacent two words are the bigram feature; the adjacent three words are the trigram feature. This is extended up to n numbers, so n-gram features to train the classifier. As n increases, the accuracy decreases. While training, the input is limited to the unigram model. Next a machine learning algorithm, support vector machine (SVM) is applied, as previous studies show SVM is the best algorithm for text classification. During preprocessing, the text is converted into a vector which consists of a set of features that represents the corresponding disaster related data [36]. The SVM algorithm then calculates and plots a hyperplane through supervised learning that divides positive and negative texts. The linear kernel was found to provide the best results. The results are evaluated by calculating the precision, recall and F-measure which are the three main metrics for measuring the performance of a classification system [36]. Precision is the percentage of classified text that is relevant. Recall is the fraction of classified text that is retrieved. F-measure is the ratio of the combination of precision and recall [36]. After the data has been

classified by the SVM algorithm, the positive and negative words are used to identify the opinion towards each category [36]. Sentiment strength is calculated from these results. In this study, the combination of the subjective phrase and machine learning algorithm with bigram feature presents the best classification accuracy of disaster related data [36].

This study used social media to determine the needs of the affected people and then using sentiment analysis determines if those needs are being met. Various visualizations were used in this paper, including box plots. The visualizations make it easier for emergency management personnel to understand the needs of those affected in order to provide the best assistance. The study also helps connect people who are willing to help with information regarding how they can help. Many times, an unaffected person will offer help, such as clothing or food, but is unsure how to get these to the affected people. The proposed method helps to bring those people together.

3.7. World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets

It is indicated in recent social media research that sports fans use social media for a variety of reasons, including emotional response [37]. Present analysis uses big data analysis techniques to analyze sports fans' emotions and behaviors. The authors propose to use social media responses to measure the emotions and responses to sporting events or television shows [37]. If the results prove to be true, this would indicate that text-analysis is valid based on specific criteria. The authors also examine whether the tweets change with response to the gameplay. This study focused on Twitter data with a location stamp from the United States, although the World Cup took place in Brazil. This ensured the tweets were from users who were not in attendance at the World Cup.

The affective disposition toward a team can range from intense liking to intense disliking [37]. Typically, when a team is doing well, their fans are more likely to show joy or positive emotions. When a team is doing poorly, their fans are more likely to show negative emotions. If a viewer is not a fan of either team, they are typically not emotionally invested in the team and do not show the same emotions as a fan. Research in social media for emotional reasons is limited. Part of the literature review for this study found a survey method that analyzed the why sports fans used social media during sports and found a significant but weak relationship with sports fans' game enjoyment and their use of social media during mediated games [37]. Typically, emotion behaviors in data are coded by humans. Since social media data is so large, it would be impossible to analyze tweets manually for the number of fans watching a sporting event.

Sentiment analysis has increased over the last decade. Existing methods are based on linguistic resources or machine learning. The more common approach is to use linguistic resources, which uses lists of positive and negative words. This method counts the number of times a word appears which allows it to recognize words with positive polarity (favorable sentiment), negative polarity (unfavorable sentiment) or no polarity (neutral) [37].

This study used a natural experiment approach to monitor U.S. fans' emotions and reactions during the games when the U.S. National Soccer team competed in the FIFA World Cup 2014. Sentiment analysis was used to examine the emotions manifested in the tweets [37]. Tweets were retrieved using Twitter's Search API during the 2014 FIFA World Cup. A web crawler was designed to collect and parse English tweets in real time using the following predefined hashtags: #FIFA, #Football, #WorldCup, or #Soccer, ignoring letter case.

Two analysis methods were used for sentiment analysis. First, NRC Word-Emotion Association lexicon was used to measure specific emotions, such as fear, anger, joy, sadness, disgust, surprise, trust and anticipation [37]. The authors used the lexicon words' frequency to measure the strength of the specific dimension [37]. An R application was used to extract the features from a tweet and create a score by counting the words that matched the eight categories or emotions listed above. Second, the tweets were then analyzed for emoticons expressing the mood with either a :(meaning sadness or :) meaning joy [37].

The findings of this study show that fanship of a team enhances one's worries and involvement in the game but being a fan of a sport is mainly tied to enjoyment and emotional release (anticipation). The use of big data analysis of the sentiments was as expected and can be reasonably explained [37].

3.8. A context-aware personalized travel recommendation system based on geotagged social media data mining

A previous study focused on using social media data to build a travel recommendation system [38]. The goal of this study was to build a system to recommend travel points of interest in different locations based on data obtained from social media posts previously made by users during previous travels. The idea being that the system can recommend a point of interest to a user in a new travel city, based on the previous points of interest they have visited in other cities. The authors built a system with a variety of functions in order to accomplish their goal of building this travel recommendation system. Their method was tested on the public API of the social media platform Flickr to collect the metadata of 736,383 photos that were taken between January 2001 and July 2011 in six different cities within China.

Geotag information available on the photos allowed the authors to cluster photos together based on location using the P-DBSCAN clustering algorithm. This method worked by clustering photos by location based on a predetermined maximum radius. Once the clusters were created, the authors used term frequency-inverse document frequency to score the location tag of each photo. The higher the score indicates a tag for a more distinctive location. Based on these results, the authors used Google Places to find points of interest associated with each location. The authors then used the timestamps of users' photos to determine the average time a visitor would spend at each point of interest. Matrices were built to examine the interests of users in each location. The Pearson correlation metric was then used to build a similarity matrix to find similarities among users' traveling patterns. The authors conclude that the method they have proposed was successful in examining user preferences and correctly suggesting travel points of interest in new cities based on their previous posts and their similarities to other users.

3.9. A Big Analytics Method for Tourist Behaviour Analysis

Big data analysis is lacking in the social media field in relation to decision making, specifically in Tourist Destination (TD) Management. This study focuses on supporting strategic decision-making in this sector because big data analytics has not provided use cases for strategic decision support [31]. TDs are areas of interest within a locality that tourists may visit or participate in sightseeing. Tourists share data via multiple sites: Flickr for photo sharing, YouTube for video sharing, Twitter for immediate response sharing, and Facebook for photo and comment sharing and discussion. These data could provide important information to tourism authorities, although they rarely collect the data. This study uses publicly available geotagged photo data on the photo-sharing site, Flickr.

This study evaluates a method based on unstructured data with meaningful content that is tourist focused. The method used in this study combines text processing, geographical data clustering, visual content processing and time series modeling to address the DMOs decision support needs. The results may provide DMOs with information in forecasting future and seasonal demands of tourism development, management and planning [31].

Textual metadata often contains keywords. These data are usually unstructured, which requires some pre-processing. The authors employed a text processing tool, General Architecture for Text Engineering (GATE). The geographical data clustering step is the next step in the process. Here, the authors aim to identify popular locations for each of the identified tourist interests. A clustering technique, P-DBSCAN, is used to identify popular areas of interest. It includes both the number of tourists that visit the location and the number of photos. This ensures that the identified locations have many visitors.

For each location, tourism managers are interested in the most representative photos for each tourist interest [32]. Representative photos identification was completed in two steps: Visual Content Representation and Kernel Density Estimation. For the visual content representation step, the authors adopted Speeded-Up Robust Features (SURF). A popular approach to represent photo content using local region descriptors is to represent each image as a visual bag of words. The descriptors are first extracted for a large set of local regions extracted from a set of random photos. K-means clustering is then applied to create a visual word vocabulary. Visual words are defined as the center of clusters and the value of k determines the number of visual words available [31]. Kernel Density Estimation (KDA) is then applied. The probability density function is calculated for each photo. The photos with the highest probability density function are considered the representative photos. The theme for each photo collection can be identified by examining the small selection of representative photos.

In order to determine how frequently and how many tourists can be expected to visit in certain months, a time series model is created. Popular fitting functions include linear, exponential, and quadratic types [31]. The mean absolute error (MAE) is used to select the most appropriate model for trend estimation rather than predict the actual value of the time series. A seasonal component is removed from the time series. The trend was modeled using a quadratic function and the seasonal means are calculated by the average of seasonal components for each month.

Those four techniques are combined in the solution artefact to process and analyze different types of data (textual, geographic, photos and time) in order to offer insight into tourist behavior and perceptions. The authors propose using the time-series model to predict tourism demand. The DMO will have insight into tourist interests but also time series trends, including any seasonal trends, which will help with strategic planning and decision-making.

The authors chose to use a representative sample of tourists' social media data for Melbourne, Australia. Geotagged data was extracted and processed using the techniques already described. These data were extracted from Flickr using its API using the coordinates of a bounding box around Melbourne, Australia (based on the longitude and latitude from Google maps).

Time series models were then built for forecasting future tourist demand in Melbourne. Seasonal mean models were created to determine seasonal patterns. These models help determine seasonal patterns of travel which can be used to determine tourist arrival. This will aid the tourism managers in strategic planning and decision making. These data showed no seasonal pattern for local Australian tourists. The pattern of tourists was different among the groups of data (Australia, Asia, Europe and North America), however was the same within the groups. For example, in North America, the high peak season was between January and March and in November; the low peak season was April through September.

This study automatically detects tourists interests in places and objects [31]. The authors were able to extract meaningful information from the geotagged photos, thus creating a solution that provides DMOs with information to make meaningful decisions based on key patterns and trends in tourist social media data. The full artefact in this study was presented to both academic and industry audiences, which helped improve the design [31]. The authors were also able to test the results of the artefact using not only the Melbourne data but data from other locations. Only data from Melbourne was presented in this study. If a sufficient number of photos are collected, this method proves to be effective for any city.

3.10. Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care

In [39], the study conducted aimed to analyze forum posts related to cancer treatment to identify positive and negative feedback to the treatments and side effects [39]. This study also aims to identify "influential users" within the forums of interest. Pharmaceutical companies have been tasking their IT departments with using social media data mining techniques to retrieve reviews from patients that have used their medication in order to easily retrieve feedback on their product. Social media data mining in reporting feedback of care can also be used by doctors and patients to improve care. Doctors may use such a system to recommend more effective treatment options and patients can use such a system to see previous feedback on treatment options that have been proposed by their doctors. This task is often completed using patient surveys; however, the authors indicate that social

media data mining for this purpose allows for access to more readily available information.

The authors of this study were interested in user feedback of the drug Erlotinib, which is used in the treatment of lung cancer. The authors began this experiment by searching the web for popular cancer forums, specifically those with a significant amount of posts related to lung cancer, four online forums were identified in this step. Within these forums identified, the authors searched for posts specifically related to lung cancer medication and noted that Erlotinib was the most frequently discussed drug. The authors then searched all of the forum posts for common terms related to positive and negative feedback, this was done using a Rapidminer data collection and processing process. The result of this step was a final list of words associated with negative and positive feedback; each word was given a term-frequency-inverse document frequency (TF-IDF) score. The words with the highest TF-IDF scores were tagged using a modified NLTK toolkit in order to add positive tags to negative words and vice versa. For example, for a post that says, "I do not feel great", a negative tag was added to the word "great". The authors then removed similar words and words that appeared less than 10 times, the result was a list of 110 words, half of which were positive and the other half being negative.

Another word list was created based on searching all forum posts related to side effects of Erlotinib. In order to identify terms associated with side effects, the authors referred to a vocabulary list from the National Library of Medicine's Medical Subject Heading. A list of the words related to side effects were then fed into the Radidminer processing tool in order to assign TF-IDF scores to each word. Words with the highest score, that appeared at least 10 times, were kept for the final list of words relating to side effects. The result of these two processes was a dataset of user forum posts which were each transformed into two vectors, one with 110 variables for the TF-IDF scores for the positive/negative terms and another with 10 variables containing the TF-IDF scores for the side effect terms. The vectors were then analyzed for clusters using Self Organizing Maps (SOM). The forum posts were first manually labeled as negative or positive before they were fed into the SOM. The purpose of the SOM step was to analyze the clusters formed, and the weights associated with those clusters, to see if there was a relationship between the clusters identified and the classification of the post being negative or positive.

The next step was to identify those users in the forums that are considered "influential." This was a topic of interest because these influential users can often drive the opinions of other users within the forum. This process was done using a combination of network-based analysis with a Markov chain modeling the various connections between nodes of the network. The opinions of these influential users were then analyzed to identify them as being negative or positive using a module average opinion (MAO) approach. This formula takes the sum of the TF-IDF scores associated with positive terms, minus the TF-IDF scores of the negative terms, and then divides this difference value by the sum of all TF-IDF scores. Lastly, forum posts related to side effects were identified using the list of side effect terms previously developed. The authors found that most feedback related to the drug was positive. The authors demonstrated that they were successful in identifying "influential users" within the forums. These users were found to interact with other users on the forums very often and that users looked to these influential ones for further information. The authors used t-tests to score the posts associated with side effects, finding that rash and itching were the most frequent side effects reported. The authors summarize that methods similar to these can be used in the future to improve patient care and reduce costs.

3.11. Filtering big data from social media –Building an early warning system for adverse drug reactions

Adverse drug reactions (ADR) are unintended reactions or side effects for a medication. Usually, these are determined during clinical drug testing trials. The World Health Organization (WHO) uses "early warning" systems to monitor ADRs. Usually, the early warning systems are based on feedback from pharmacists and physicians rather than the patient. A growing number of pharmaceutical companies are now allowing consumers to report symptoms directly to them in order to ease legal and monetary implications [40]. Not all countries use consumer reports and even the ones that do have experience a delay in getting consumer reports of serious ADRs. Because of this, a different approach is needed.

Patients have taken to social media to discuss their medication use and experience. This information could be used for healthcare research. The user-generated content (UGC) is an emerging trend of data for early detection of adverse disease events [40]. Users tend to share their perceived risk of an ADR on a social media platform. These data provide an opportunity to mine the data for a relationship between drugs and their interactions. The challenge to this type of data is the messages are usually sparse and highly distributed. It is easy to obtain large amounts of unlabeled social media data, but it is costly and time consuming to manually classify and label the ADR messages to build an early warning system.

The objective of this study was to develop a process to scan large amounts of text-based posts collected from drug-related Web forms. The system will integrate both text and data-mining techniques to extract important text features from the posts and then classify the posts into positive/negative examples based on a few pre-identified ADR related posts. The classification process is based on a partially supervised learning method, which uses a small number of known positive posts to identify other posts of similar text features from a large corpus of unlabeled posts [40]. The results can be used to build an early warning system to assist the Food and Drug Administration and pharmaceutical companies.

Data collection consisted of crawling of Web forums to gather patient posts from social media. This study focused solely on discussion boards, in particular Medhelp, however the method could be expanded to include social media platforms such as Twitter and Facebook. The authors used domain experts to tag a large sample of posts related to the drugs in this study. If any ADR wordings were identified, it was labeled as a positive example; otherwise, it was a negative example.

The Latent Dirichlet allocation (LDA) is applied to construct a topic space over the corpus. It uses a small number of topics to

describe a collection of documents. The LDA model reduces the dimensions while maintaining the structure of the document. The authors then used the Gibbs sampling inference algorithm for approximation. During classifier construction, the authors selected reliable positive and negative examples to fit the distribution of the positive and negative classes, respectively.

In order to reduce the diversity of topics, the negative dataset is partitioned into smaller clusters. Here, they used the k-means clustering method. After the negative examples were extracted, they moved to the positive examples. The authors used Support Vector Machines as the text classifier.

Four benchmark methods, Example Adaptation for Text categorization (EAT), Positive examples and Negative examples Labeling Heuristics (PNLH), Active semi-supervised Clustering based Two-stage text Classification (ACTC) and Laplacian SVM, were compared with the method used in this study. The paired-sample Wilcox signed-rank test was applied to assess the statistical significance with respect to the best benchmark method. The approach in this study outperformed the benchmark techniques in most cases.

3.12. Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality

In dealing with data mining, text analysis plays a vital role in filtering through reviews that people write for products to determine if they are useful or not. In this study, a method was developed to model the large amounts of reviews that people post on a daily basis (from sites such as Amazon) to determine which of those reviews are actually helpful.

In order to collect the data, reviews were collected and multiple regression was used to classify reviews as helpful or not. Text analysis revealed that extremely high and extremely low reviews were often over inflated, being based on very individualized experiences that would not help a wider audience. ANOVA served as a statistical measure to analyze the differences between the various categories of review (such as by star rating). Text analysis found that in the majority of cases, the most helpful and detailed information was being found in the 4-2 star rating [41].

Further filtering was able to be performed to determine truly useful features of reviews. Further modeling was performed to be able to factor in product type, which had different factors that impact what reviews are deemed to display quality information. Further considerations in this work can be to put important reviews more accessible to customers or identify fake reviews that are being posted simply to increase or decrease a specific product rating for an online site [41].

4. Overview of Findings

4.1. Social Media Studies

In the paper describing urban emergency event description through social media data mining, the authors developed a simple 5W model to describe the event of interest [29]. This method is very simple in that it only requires searching through public posts within a given location on a certain date range to describe an event that has already occurred. This paper showed how easily information can be extracted from social media posts using basic words to search posts. While this paper develops a user-friendly way to search for social media posts of interest, there are some issues that could arise from this method. For example, if a researcher were to use this information to search for posts related to a fire that occurred in a certain area on a given date, searching for the word “fire” in Weibo posts may result in posts being returned that had this word but that were not related to the event that occurred. While this is still a good method to identify posts of interest, one should be aware that a method would need to be developed to filter-out posts that are irrelevant to the topic of interest. In the future, researchers should come up with an algorithm that can be employed to filter-out posts that are irrelevant to the current topic of interest. Future work involving these methods should also consider that size of the dataset that will need to be searched for relevant posts. In this example, the event was in a specific location. If this model were to be applied to future research that was concerned with a larger geographic location, for example an entire country, additional methods, such as programing, may need to be deployed to efficiently search through social media posts to find those that are of interest.

In the paper which described a system for identifying student experience, the authors used Twitter API search and Naïve Bayes classification to examine the problems faced by engineering students [32]. An easy method of searching for a relevant hashtag within Twitter was implemented. This approach works well for a system like Twitter in which users’ posts very often include hashtags to identify the main themes in the post. The authors showed that the Naive Bayes classifier performed best with their dataset of interest. It is important to test multiple classification methods when analyzing datasets such as this one to find the best method. Future research that involves a similar dataset could use the methods described by these authors to compare if such methods would apply to all Twitter post datasets searching for hashtags of interest, or if the best classification method varies depending on what the resulting dataset is after a Twitter API search. One limitation to this specific approach was that the authors only searched for problems faced by engineering students but did not develop a method to identify positive feelings from engineering students. A collection of data like this may imply that only negative experiences were faced by engineering students. It would’ve been beneficial to compare positive posts from engineering students to learn more about the engineering student experience as a whole. In the future, this same algorithm could be applied to include other social media platforms, such as Facebook, because not all students actively use Twitter. An addition of this experiment should also involve coming up with positive hashtags or terms that would be used to identify posts of engineering students.

While studying “Investigating the impact of social media advertising features on customer purchase intention” it was clear that sites such as Facebook and Twitter performed well for the model [33]. However, the applications of this research were not tested in other social media environments to see if the same results would hold up. One argument for this is the reasoning of why certain personalities gravitate towards specific types of social media platforms. Since much time was devoted to model validity, it would have been worthwhile to explore how well the model did

analyzing photo captions on a site such as Twitter, or tying in photo analysis itself. As opposed to just text, video and images on social media could provide an unseen angle as to how customers look at products [33]. As viewed in the introductory section, with all social media platforms experiencing a rise in recent years, further work should be done to view correlations and associations between the various sites to see what behaviors can be seen as a commonality. As the data was collected from a survey of participants, it would be beneficial to anyone using this model to test it against a randomly selected group of individuals from other areas.

Corpus for Customer Purchase Behavior Prediction in Social Media provided insightful methodology of how social media posts could be used to determine if an individual would buy a product based on what they posted [34]. The downfall of this article was that it did not actually create a workable model that could be tested. Therefore, all proposed methodology were based off of other studies and what should theoretically" work and produce the best results (i.e., appropriate classifiers, sentiment analysis procedures, learning methods) without demonstrating that these claims were in fact true. This work does however provide an easy point for new researchers to build off of to help develop a working model. While the article mentions twitter extensively to obtain the data, no reference is mentioned as to how this would actually fully be implemented. Based on other studies, it seems reasonable that Twitter API searches could be implemented to gather tweets that could be fed into this proposed model once it is created [34].

The paper "Predicting Postpartum Changes in Emotion and Behavior via Social Media" provided several different methods of how to classify the data once text analysis was performed. This idea was an extremely highlighting feature of this paper since it involved linear and Quadratic discriminant classifiers, Naïve Bayes, k-nearest neighbor, decision trees, and support vector machines as tested methods. This allowed the reader to confirm the results of several other studies that SVM appeared to be the classification method which yielded the highest success and reliability rate [35]. The paper also shows social media can apply to make predictions about important aspects of people's lives. While the study focused primarily on Twitter, it would have been beneficial to see if the performance held up over other platforms as well [35]. Extensions in the research could also examine commonalities between the various social media platforms to make extensions on how a model for one social media platform may perform or be adjusted for any of the others. Based on the analysis performed and the more reliable results from the SVM classification method, this should form a good method choice for testing the analysis of text on sites that allow users to be more expansive with their word limits in posts. Additional analysis of the SVM method's reliability correlated with character length would provide great insight into if SVM is an overall choice for this type of analysis, or if it is the most effective when a stricter character limit is imposed on a user.

In the paper describing a system to perform sentiment analysis in response to disaster responses and recovery, sentiment analysis and SVM were used to create a classification system for social media. These results were presented to emergency responders in a visual form to help understand the needs of those affected [36]. The combination of both sentiment and SVM proved to be an effective method of classification. The method also included bag of words, where the entire text is shown as a list of words. The downfall to this is when taken out of context, the words may seem unimportant but in the full context of a sentence or tweet the word may be very important. Improvements to this study would be to include data from other sources, instead of only Twitter data, as well as considering sentences rather than individual words within tweets to ensure the context is captured properly. Also, particular keywords were provided to extract the appropriate data. Another improvement would be to create an algorithm that monitors Twitter data for an increase in use of a particular word or sentence, as is done with data that is trending in Twitter. The authors suggested future studies could include data from various disasters to improve the accuracy of the model. With multiple social media platforms, extracting data from multiple sources, while complex, could also improve the disaster response models.

In the paper describing a system to analyze sentiments of U.S. sports fans using tweets, sentiment analysis was performed using two methods: NRC Word-Emotion Association and analysis of tweets for emoticons expressing the mood. This study presented results that were expected: a fan's emotional state during a game is directly related to how well or how poorly their team is performing [37]. A limitation of this study is that a very specific group of fans was analyzed. The study used tweets from US fans that watched the World Cup 2014 from the United States, not in person. The authors assumed that when fans were happy it was because their team was doing well, so their assumption and hypothesis were essentially identical [37]. Data was collected only during the games. An improvement to this study would be to collect tweets from a longer time frame to determine if the users are truly fans of the team or only fans during the particular game evaluated. Additional statistical techniques could be applied, such as SVM, which works well with text classification. Individual words within the tweets were examined, rather than examining the context of a word within a tweet. To improve this method, future research could involve examining the entire tweet to ensure the proper context is examined.

4.2. Photo Sharing Studies

In the paper describing the creation of a travel recommendation system, the authors used a combination of P-DBSCAN clustering and TF-IDF to create a travel recommendation system based on the user's previous travel data [38]. This system was based on making recommendations for new travel locations based on previous locations the user has traveled to. The authors proved that their method of combining P-DBSCAN clustering and TF-IDF was efficient, and successful, in handling large datasets. With the size of the dataset that this method was tested on, one can assume that the same methods would be successful if an application of this kind were to be available to all social media users when traveling. It is likely these methods would be successful in other recommendation applications that used social media, such as an application to recommend restaurant options in a user's location. While the methods presented here provide a reliable system to follow to make future recommendations, there is some room for improvement in that the researchers only consider making

recommendations for users in which previous data has been supplied. A system such as this would not work for new users because there would be no previous data to go off of. A system such as this one should include an additional feature that new users can use to answer questions based on previous travel that has not been tracked in the travel recommendation system. A questionnaire which identifies previous places the user has traveled to and points of interest visited in those locations would improve the function of the travel system to be more beneficial to new users.

In the paper describing a system defining tourist behavior, the authors used Flickr's API to export photos and related metadata for analysis. The methods used were General Architecture for Text Engineering (GATE) for text processing, and P-DBSCAN for clustering [31]. P-DBSCAN is used to identify popular tourist areas. Speeded-Up Robust Features (SURF) and Kernel Density Estimation are applied. Time series models were used to determine the future tourist demand and seasonality demands. This system proved to be effective for not only the location analyzed in this study, Melbourne, Australia, but also for other locations if sufficient data is collected [31]. The strength of this study was that the authors incorporated tourist photos, which had not been done in previous studies. One limitation was the metadata for the photos. Complete metadata was not available for all photos extracted from Flickr. Another limitation was only one site was used to extract data. This limits the information to only photos and does not include emotional responses that may be found in other social media, such as tweets. A future recommendation would be to include data from some of the other social media platforms, such as Instagram, Snapchat, Facebook and Twitter to improve recommendations. Multiple types of data could also be analyzed, including photos, videos and status updates. Some users may tag themselves into a location without sharing a photo, making it valuable to collect other types of data. These additional data would likely require adjustments to the model; however, the model would provide a good foundation.

4.3. Medical Forum Studies

In the paper which described a system for improving medical care, the authors use a combination of TF-IDF and self-organizing maps to analyze forum posts within various medical forums [39]. The authors' model was careful to add negative and positive tags so that these posts were categorized correctly. The authors also identified influential users that could potentially sway the opinions of new users. As shown in the other papers mentioned here, TF-IDF is a common method used when analyzing social media posts, and it has shown to perform well for extracting information of interest. The authors show that the use of self-organizing maps is successful when the goal is not only to find social media posts relevant to your topic of interest, but also to find which users are considered to have influence over other users. One item that the authors did not consider was how long users of the forums may have been taking the drug of interest. It is well known that there is a period of time when a person begins taking a new drug in which their body is adapting to the drug, so some side-effects may be more intense for a period of time in the beginning. An improvement to this system may be to retrieve information on how a users' post may change over time after they have been taking the drug of interest for a longer period. The same logic would apply to medical treatments that don't specifically involve drugs, such as a new kind of therapy that a patient is trying.

In the paper describing an early warning system for adverse drug reactions, the authors used Latent Dirichlet Allocation and Support Vector Machines to classify text from Medhelp [40]. Domain experts were used to tag sample posts related to the drugs in this study. If any ADRs were identified, the posts were marked as positive examples; otherwise, they were marked as negative examples. Support Vector Machines are particularly effective in text analysis. LDA performed well against the benchmarks from other studies that were presented in this paper. One limitation to this study is the need to use domain experts to classify posts. When human interaction is needed with big data, the sheer number of posts that can be evaluated is limited. In the future, an improvement could be creating an automated system to evaluate the social media posts rather than relying on a human to interact with the data. This would provide a larger sample of data, and allow the system to react in real-time, or close to real-time. As the authors discussed, an improvement to this model would be to collect from other sources, such as social media posts, rather than just web forums. While this method appeared successful compared with the benchmark methods, it is hard to imagine doing large scale analysis while relying on domain experts to manually review data in order to build a model. Other methods considered as well, since SVM and sentiment analysis were successful in several of the studies observed.

4.4. Product Review Studies

In the paper which described how to build a function to determine the helpfulness of user-generated reviews as a function of review sentiment, product type and information quality, text analysis was relied heavily on to be able to examine reviews [41]. Text analysis is a very big need when collecting data based on what people right. With the number of variables required to take into consideration, multiple regression was selected as the method of analysis with ANOVA being used to analyze the differences in the types of review classifications (such as 1-star, 2-star, etc.). In this text analysis situation, sentiment analysis did not appear. Since the text was already broken down into categories based on how a user rated them, it is interesting to wonder if this is the reason why sentiment analysis never made an appearance. From a classification standpoint, sentiment analysis would likely need to be used to classify them into their rating category, if this step was not already completed. Hence, it would have been an interesting extension if the researchers could have applied this to social media (as other papers) to analyze the mentions of products and then analyze if posts are useful [41]. As the authors only analyzed three product searches and three experience searches, the fact that the helpfulness and quality varied greatly over the review sentiment and product type raises questions of how this methodology would perform if used to model more diverse products.

5. Conclusion

In this paper, we evaluated twelve research papers that explored various data mining techniques for social media data.

While many methods were used for analysis, common themes were text analysis, SVM and sentiment analysis. The previous researchers used a variety of methods to extract data, such as using the API for the social media platform. Both the Twitter Search API and the Flickr API were used to extract data in some of the studies. Future recommendations were made for each individual study, sometimes recommending methods that worked particularly well in one of the other papers evaluated. Several of the studies evaluated were similar in theme but did not necessarily use the same types of analysis. All of the methods provided sufficient results for the studies, but it was interesting to look at different methods used for similar studies.

In the study that utilized social media posts to describe emergency events, the 5W method was used. Another study that also used social media for emergency events used sentiment analysis to determine the needs of the affected people. The authors used several algorithms, including SVM and sentiment analysis. Both of these studies had similar goals but very different and effective approaches. This may show that while a particular big data technique may work best in a situation, there are multiple research methods worth testing.

Two studies used proposed methods using Flickr data to provide recommendations to travelers and Travel Destination Management. Both of these studies used P-DBSCAN, but one expanded the recommendations further by analyzing seasonal timeframes of destinations. P-DBSCAN was effective in both of these studies.

The studies that used social media for improving medical care used different methods. The study monitoring feedback of the drug Erlotinib used a combination of TF-IDF and SOM. The study monitoring ADRs used LDA, k-means clustering and SVM as a classifier. Since the methods tested provided sufficient results, it would be worth further study to compare both methods to determine the best approach for using social media to analyze medical data.

Three studies evaluated social media for products. All three used different techniques in data mining. The Goodness of Fit, structured equation modeling and other similar models were used when investigating the impact on purchase intent. The study that focused on predicting iPhone sales proposed using machine learning to train a classifier based on phrasing in social media posts. The final product study focused on the helpfulness of product reviews. Multiple regression and text analysis were used in classifying reviews.

One study used Twitter data to evaluate students' learning experiences. Naïve Bayes was found to be the most effective classification method.

While all of these studies were based on social media data, they show there many diverse ways to effectively classify and analyze the data.

References

[1] Pew Research Center. (2021, April 26). Demographics of Social Media Users and Adoption in the United States. Pew Research Center: Internet, Science & Tech https://www.pewresearch.org/internet/fact-sheet/social-media/?menuItem=ad42e188-04e8-4a3c-87fb-e101714f1651

[2] M. N. Injadat, F. Salo, and A. B. Nassif, "Data mining techniques in social media: a survey." Neurocomputing, **214**, 2016, doi: 10.1016/j.neucom.2016.06.045.

[3] Data science & big data analytics: discovering, analyzing, visualizing and presenting data. Wiley, 2015.

[4] V. Mhaskar, "Measure Search Relevance using NDCG - Know More: T/DG Blog - Digital Thoughts," Measure Search Relevance using NDCG - Know More | T/DG Blog - Digital Thoughts, 28-Jun-2015. [Online]. Available: https://blog.thedigitalgroup.com/measuring-search-relevance-using-ndcg. [Accessed: 05-Apr-2020].

[5] S. Gupta, Sentiment Analysis: Concept, Analysis and Applications. 19-Jan-2018. [Online]. Available https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17 . [Accessed: 15-Apr-2020].

[6] "Text Classification," MonkeyLearn, 27-Jan-2020. [Online]. Available: https://monkeylearn.com/text-classification/. [Accessed: 15-Apr-2020].

[7] H. C. Wu, R. W. P. Luk, K. F. Wong, and K. L. Kwok, "Interpreting TF-IDF term weights as making relevance decisions," ACM Transactions on Information Systems,. **26**(3), 1–37, 2008, doi: 10.1145/1361684.1361686..

[8] "Standard search API - Twitter Developers," Twitter. [Online]. Available: https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets. [Accessed: 15-Apr-2020].

[9] X. Zhang and W.-X. Xiao, "Clustering based two-stage text classification requiring minimal training data," in 2012 International Conference on Systems and Informatics (ICSAI2012), 2012, doi: 10.1109/ICSAI.2012.6223496.

[10] S. Shimojo, Web and communication technologies and internet related social issues: proceedings, Springer, 2005.

[11] "Get Started with GATE (General Architecture for Text Engineering)," WhoIsHostingThis.com, 12-Dec-2018. [Online]. Available: https://www.whoishostingthis.com/resources/gate/. [Accessed: 20-Apr-2020].

[12] "Linear, Quadratic, and Regularized Discriminant Analysis," datascienceblog.net: R for Data Science, 30-Nov-2018. [Online]. Available: https://www.datascienceblog.net/post/machine-learning/linear-discriminant-analysis. [Accessed: 05-Apr-2020].

[13] R. Gandhi, "Support Vector Machine - Introduction to Machine Learning Algorithms," Medium, 05-Jul-2018. [Online]. Available: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47. [Accessed: 15-Apr-2020].

[14] R. S. Brid, "Decision Trees - A simple way to visualize a decision," Medium, 26-Oct-2018. [Online]. Available: https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb. [Accessed: 20-Apr-2020].

[15] O. Harrison, "Machine Learning Basics with the K-Nearest Neighbors Algorithm," Medium, 14-Jul-2019. [Online]. Available: https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761. [Accessed: 20-Apr-2020].

[16] J. Bekker and J. Davis, "Learning from positive and unlabeled data: a survey," Machine Learning, 109**,** 719-760, Apr. 2020, doi: 10.1007/s10994-020-05877-5.

[17] K. Baxter, "Similarity Matrix," Similarity Matrix - an overview | ScienceDirect Topics. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/similarity-matrix. [Accessed: 15-Apr-2020].

[18] K. S. do Prado, "How DBSCAN works and why should we use it?," Medium, 03-Jun-2019. [Online]. Available: https://towardsdatascience.com/how-dbscan-works-and-why-should-i-use-it-443b4a191c80. [Accessed: 20-Apr-2020].

[19] M. M. A. Patwary, D. Palsetia, A. Agrawal, W.-K. Liao, F. Manne, and A. Choudhary, "A new scalable parallel DBSCAN algorithm using the disjoint-set data structure," 2012 International Conference fo1`r High Performance Computing, Networking, Storage and Analysis, 2012, doi: 10.1109/SC.2012.9.

[20] M. J. Garbade, "Understanding K-means Clustering in Machine Learning," Medium, 12-Sep-2018. [Online]. Available: https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1. [Accessed: 20-Apr-2020].

[21] J. Hollmen, Self-Organizing Map (SOM), 08-Mar-1996. [Online]. Available: https://users.ics.aalto.fi/jhollmen/dippa/node9.html. [Accessed: 15-Apr-2020].

[22] D. Soni, "Supervised vs. Unsupervised Learning," Medium, 16-Jul-2019. [Online]. Available: https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d. [Accessed: 15-Apr-2020].

[23] W. Kenton, "Goodness-Of-Fit Definition," Investopedia, 29-Jan-2020. [Online]. Available: https://www.investopedia.com/terms/g/goodness-of-

fit.asp. [Accessed: 05-Apr-2020].
[24] "Kernel Density Estimation," DeepAI, 17-May-2019. [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/kernel-density-estimation. [Accessed: 05-Apr-2020].
[25] M. Malaeb, "Recall and Precision at k for Recommender Systems," Medium, 04-Aug-2019. [Online]. Available: https://medium.com/@m_n_malaeb/recall-and-precision-at-k-for-recommender-systems-618483226c54. [Accessed: 05-Apr-2020].
[26] W. Kenton, "How Analysis of Variance (ANOVA) Works," Investopedia, 29-Jan-2020. [Online]. Available: https://www.investopedia.com/terms/a/anova.asp. [Accessed: 20-Apr-2020].
[27] P. Grant, "Understanding Multiple Regression," Medium, 23-Mar-2020. [Online]. Available: https://towardsdatascience.com/understanding-multiple-regression-249b16bde83e. [Accessed: 05-Apr-2020].
[28] "Time Series Analysis," Statistics Solutions. [Online]. Available: https://www.statisticssolutions.com/time-series-analysis/. [Accessed: 15-Apr-2020].
[29] Z. Xu et al, "Crowdsourcing based Timeline Description of Urban Emergency Events using Social Media," International Journal of Ad Hoc and Ubiquitous Computing, **25**(1/2), 1, 2017, doi: 10.1504/IJAHUC.2017.083481.
[30] S. Hall, "What Is Inductive Content Analysis?," Small Business - Chron.com, 26-Oct-2016. [Online]. Available: https://smallbusiness.chron.com/inductive-content-analysis-24666.html. [Accessed: 20-Apr-2020].
[31] S. J. Miah, H. Q. Vu, J. Gammack and M McGrath, "A Big Data Analytics Method for Tourist Behavior Analysis," Information & Management, **54**(6), 771-785, 2017, doi: 10.1016/j.im.2016.11.011.
[32] X. Chen, M. Vorvoreanu, and K. P. Madhavan, "Mining Social Media Data for Understanding Students' Learning Experiences," IEEE Transactions on Learning Technologies, **7**(3), 246–259, 2014, doi: 10.1109/TLT.2013.2296520.
[33] A. A. Alalwan, "Investigating the impact of social media advertising features on customer purchase intention," International Journal of Information Management, **42**, 65–77, 2018, doi: 10.1016/j.ijinfomgt.2018.06.001
[34] S. Sakaki, F. Chen, M. Korpusik, & Y.-Y. Chen, "Corpus for customer purchase behavior prediction in social media," in Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 2976–2980, 2016.
[35] M. D. Choudhury, S. Counts, and E. Horvitz, "Predicting postpartum changes in emotion and behavior via social media," in CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3267-3276, 2013, doi: 10.1145/2470654.2466447
[36] J. R. Ragini, P. R. Anand and V. Bhaskar, "Big data analytics for disaster response and recovery through sentiment analysis," International Journal of Information Management, **2018**(42), 13-24, 2018, doi: 10.1016/j.ijinfomgt.2018.05.004.
[37] Y. Yu and X. Wang, "World cup 2014 in the Twitter world: a big data analysis of sentiments in U.S. sports fans' tweets," Computers in Human Behavior, **48**, 392-400, 2015, doi: 10.1016/j.chb.2015.01.075
[38] A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, and J. Woodward, "A context-aware personalized travel recommendation system based on geotagged social media data mining," International Journal of Geographical Information Science, **27**(4), 662–684, 2013, doi: 10.1080/13658816.2012.696649.
[39] A. Akay, A. Dragomir, and B.-E. Erlandsson, "Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care," IEEE Journal of Biomedical and Health Informatics, **19**(1), 210–218, 2015, doi: 10.1109/JBHI.2014.2336251.
[40] M. Yang, M. Kiang and W. Shang, "Filtering big data from social media – Building an early warning system for adverse drug reactions," Journal of Biomedical Informatics, **54**, 230-240, 2015, doi: 10.1016/j.jbi.2015.01.011.
[41] A. Y. Chua, & S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," Computers in Human Behavior, **54**, 547–554, 2016, doi: 10.1016/j.chb.2015.08.057.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 189-195 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Mitigation of Nitrous Oxide Emission for Green Growth: An Empirical Approach using ARDL

Hanan Naser*, Fatema Alaali

College of Business and Management, American University of Bahrain, Riffa, 959, Kingdom of Bahrain

ARTICLE INFO

Article history:
Received: 02 March, 2021
Accepted: 01 July, 2021
Online: 20 July, 2021

Keywords:
N_2O emissions
Environmental Kuznets Curve
FDI
Economic Growth
Financial Development
Electric Power Consumption

ABSTRACT

Although the perception of Environmental Kuznets Curve (EKC) has been thoroughly investigated, but there is inconsistency in the results. The relationship between nitrous oxide (N_2O) emissions, financial development, economic growth, foreign direct investment, and electric power consumption in Bahrain over the period 1980 – 2012 is examined in this paper. The autoregressive distributed lags (ARDL) technique is employed to test for the cointegration in the long run. The results reveal a reversed U-shape long run relationship between N_2O emissions and economic growth for Bahrain. Moreover, electric power consumption affects N_2O emissions positively in the short and long run. Whereas foreign direct investments and financial development affects the emissions of N_2O negatively. Therefore, Bahrain should assist households in installing solar cells to generate clean energy and enhance its financial sector.

1. Introduction

One of the most considerable problems that the whole world is experiencing is the environmental degradation, which leads to irreparable damages to the natural world and human society. Greenhouse gas emissions are the main pollutants of the environment, and the major source of these pollutants is burning fossil fuels. When greenhouse gas emissions are mentioned, it is usually referring to carbon dioxide (CO_2). However, greenhouse gas story involves other types of gases. Nitrous oxide (N_2O) is partially responsible for the greenhouse gas diffusions and its effect on the atmosphere from the point of global warming is greater than that of CO_2. One ton of N_2O corresponds to almost 298 tons of CO_2.[1] Moreover, the procedure followed to eliminate the N_2O from the atmosphere contributes into depleting ozone layer. Therefore, N_2O is considered as an ozone eradicator as well as being greenhouse gas.

Bahrain that is an archipelago made up of 33 islands has a total area of 780 km^2. The current population in the year 2019 is about 1.6 million with a population density of 2155 per km^2 and a population growth of 4.9% in year 2018.There is an increase in total energy consumption by 2% per year on average over the period 2010 to 2017 with a record of 14.5 mega ton of oil equivalent (Mtoe) in 2017. The fundamental factor behind this increase in energy consumption is the increase in population. Rise in population means greater local demand for energy, which increases the greenhouse gas emissions including the N_2O emissions.

One of the most substantial origins of N_2O emissions is burning fossil fuels for energy and transport. Compared to the quantity of CO_2 in the air, the amount of N_2O is much less, but the warming impact of each molecule of N_2O is nearly 300 times that of Co_2.[2] Accordingly, a number of papers have examined the soundness of the Environmental Kuznets curve (EKC) hypothesis using N_2O diffusions as a measure for the environmental deterioration. In [1], the authors scrutinized the theoretical and empirical basis of the EKC using a panel of 156 countries and three different pollutants including N_2O emissions. The results of their paper indicate the presence of a reversed U-curve for all the three pollutants but at different levels of income. In [2], the researchers reviewed the available literature to explain the EKC phenomenon for atmospheric variation and decide the association with the economic evolution of Bangladesh in regard to the EKC.

*Corresponding Author: Hanan Naser, E-mail: hanan.naser@aubh.edu.bh

[1] http://theconversation.com/meet-n2o-the-greenhouse-gas-300-times-worse-than-co2-35204

[2] https://www.carbonbrief.org/nitrous-oxide-emissions-could-double-by-2050-study-finds

Their results show that the EKC is valid only for developed countries which have low-income turning point. The paper of [3] confirm the presence of the EKC assumption for Germany by employing time series data between 1970 – 2012. The study of [4] employed a panel of 36 developing and developed nations between 1995-2013 and the EKC assumption is confirmed for N_2O and CO emissions. Therefore, the contribution of this study is to examine the short- and long-term effects of GDP growth, foreign direct investment, financial development, and electric power consumption on Bahrain N_2O diffusions over the period 1980 – 2012.

The remainder of the paper is formatted as follows: section 2 highlights the review of literature; the following section describes the data and models employed in this study followed by the discussion of results in section 4. Finally, the last section covers the conclusions and policy implications.

2. Literature Review

EKC affirms that environmental pollutants increase as the national output increases just before a certain threshold of output after which emissions decrease as output increases [5]. Particularly, EKC assumption indicates a reversed U-shape relationship between national output and environmental deterioration. This aspect was first called EKC by [6].

A substantial consideration is given to the relationship between economic development and environmental diffusions in the last decades. Various practical studies have scrutinized this hypothesis in distinct countries all over the world. In general, the results of these studies are mixed. One part validates the EKC hypothesis and approves the presence of a reversed U-shape (for example, [7]-[10] among others). The other part of these studies did not succeed to prove the EKC hypothesis but found either a linear or N-shaped relationship (such as [11]-[14] among others).

The presence of the difference in the results of empirical studies is apparent. This difference in the results can be illustrated by the following components. The first component is explained by the various measures of air pollutants used. For example, one part of the studies utilized air contamination index such as greenhouse gas emissions, CO_2, CH_4, SO_2 and NO_x emissions (such as [11]; [15]; [16]). However, the other group of research has assessed the EKC hypothesis by employing other environmental measures such as ecological footprint [17], deforestation [18] and [19], and hazardous waste [20]. The second component is pertained to the different models and methodologies utilized. To illustrate, the studies that employ a panel data for a set of countries examined the EKC theory using panel cointegration [21] and [22] or fixed effects regression [23]. On the contrary, studies that employ an individual country apply time series techniques such as Vector Autoregressive (VAR) models [24] and [25] cointegration approaches of Granger and Johanson and the ARDL bounds techniques (such as [26]-[30], among others). The third component is associated with the different variables incorporated in the model. Some studies examined the EKC hypothesis by incorporating the GDP per capita and its square only. Other set of research papers have expanded the primary model with other explanatory variables like energy consumption, foreign direct investment, financial development, urbanization and trade openness. The last component is related to the various countries included in the study and the chosen period.

The amount of research that examines the EKC theory in the Gulf Cooperation Council GCC area is limited. For example, in [30], the author uses the data of Saudi Arabia to investigate the long-run relationship between the emissions of CO_2, economic development, energy consumption and urbanization and found that the EKC does not exist. This conclusion is supported by the study of [31], which scrutinized the relationship between CO_2 emissions from transports, energy consumption by road transports and economic development in Saudi Arabia. On the contrary, the study of [32] proved the presence of EKC assumption by studying the effect of trade and income level on CO_2 emissions. In the UAE, the presence of EKC relationship was approved [33] and [34]. The empirical results of [35] indicate that inverted U-shaped relationship is held when using the ecological footprint as a measure of environmental deterioration but not for CO_2 emissions. This paper adds to the existing literature by studying the relationship between N_2O emissions, income level, electricity use, FDI and financial development in Bahrain.

3. Material and Method

3.1. Data

To attain the target of this research paper, yearly N_2O emissions data (thousand metric tons of CO_2 equivalent) are used as the environmental pollutant. To check the reliability of EKC theory, GDP per capita at constant 2010 US$ and its square are employed.[3] As asserted by the study of [17] that one of the factors that leads to the divergence in the results of the EKC testing studies is the variables included in the estimated equation. One of the sources of N_2O emissions is fossil fuel combustion that is used to generate electricity. Therefore, this paper augments its model with electric power consumption (KWh per capita) to examine its impact of N_2O emissions.

In [36] and [37], the authors argue that foreign investors and international corporations prefer investing in countries that have loose environmental policies and standards. Most of these investments sponsor forms of production that are environmentally inefficient [38]. Accordingly, this paper augments the basic model with additional variables such as the foreign direct investment, net inflows (% of GDP). Furthermore, domestic credit provided by financial sector (% of GDP) is utilized as a measure of financial development.

[3] GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2010 U.S. dollars.

The World Development Indicators (WDI) database is utilized to collect annual data for all the variables between 1980 – 2012.[4] In order to reduce heteroscedasticity and stabilize variances all the variables are transformed using the natural logarithm except the financial development indicator and FDI as they are measured as a percentage of GDP. Table 1 summarizes the descriptive statistics of all the variables under concern.

Table 1: Descriptive Statistics

Variables Description	**N_2O emissions** (Thousand metric tons of CO_2 equivalent)	**GDP per capita** (constant 2010 US$)	**Elec Electric Power Consumption** (KWh per capita)	**Fin Domestic credit provided by Financial Sector** (% of GDP)	**FDI Foreign direct investment, net inflows** (% of GDP)
Mean	86.39	20487.51	16972.75	30.42	4.57
Std. deviation	26.90	1995.28	5003.57	21.68	7.79
Minimum	39.684	16571.4	4612.55	-5.82	-13.61
Maximum	131.257	22955.1	21644.38	72.22	33.57
Skewness	0.142	-0.61	-1.53	0.36	1.39
Kurtosis	1.928	1.89	4.32	2.41	7.79
obs	33	33	33	33	33

3.2. Technical Tool and Econometric Model

On the basis of the pioneering work by [39] that is followed by the study of [40], this paper uses a single equation model that allows to examine the emission –growth relationship for Bahrain. Basically, it is suggested that the degradation of environment in Bahrain can be briefly written as follows:

$$N_2O = f\ (GDP, GDP^2, Elec, Fin, FDI) \quad (1)$$

where N_2O represents the nitrous oxide emissions in Bahrain. The above function shows that the explanatory variables for nitrogen oxide emissions are the economic growth (GDP), square of economic growth (GDP^2), electricity consumption (Elec), financial developments (Fin) and foreign direct investments (FDI).

Looking at the main objective of this study, it is suggested to derive a natural log form equation from the above linear equation, so it allows for testing the hypothesis of the EKC. Having natural log model ensure the production of reliable and effective results. Furthermore, in [41], the authors have reported that a natural log model will help in satisfying the stationary condition of the variance-covariance matrix. Accordingly, re-writing the model above can be represented by the following equation:

$$ln(N_2O)_t = \beta_0 + \beta_1\ ln\, GDP_t + \beta_2\ (ln\, GDP_t)^2 + \beta_3\ ln\, Elec_t + \beta_4\ ln\, Fin_t + \beta_5\ ln\, FDI_t + \varepsilon_t \quad (2)$$

where ε_t is the error term associated with the estimation while β_0 is the intercept. The coefficients $\beta_1, \beta_2\ \beta_3, \beta_4\ and \beta_5$ represent, respectively, the impacts of GDP per capita and its quadratic term, electricity consumption per capita, financial developments, and the percentage of foreign direct investments. Based on the theory of EKC, it is likely that β_1 has positive sign while β_2 has a negative sign.

Having a glance at (2) shown above, it is revealed that none of GDP term or its square term can be excluded. The aim of adding both terms in the same equation is to investigate the authenticity of the EKC, which can be verified only if the relationship among environmental emissions and economic growth is expressed by an inverted U-shape curve. In another words, if the relationship among the environmental emissions and economic growth is an inverted U-shape, EKC hypothesis cannot be rejected. Therefore, the environmental emissions will tend to increase as much as there is an increase in output per capita, until it reaches to a specific level (peak). Accordingly, the quality of the environment will be better.

In a study conducted by [6], it is reported that if there is no significant change in technology in an economy, it is expected that an inverted U-shape curve represents the linkage between GDP per capita and the emissions of the environment. Precisely, an expansion of an economy causes negative impacts on environment at early stages. However, as much as it grows, structural change is experienced by the economy due to many factors including information intensive industries and services, advances in technology, increase in environmental consciousness and implementation of environmental protocols, which may have positive influence on reducing environmental emissions.

3.3. ARDL Approach

The purpose of this study is to investigate the hypothesis of having an environmental Kuznets curve, that be represents a curve of an inverted U-shape to express the linkage between the emissions of environment and the growth of an economy. Accordingly, long- and short-term dynamics of the model can be investigated using a cointegration approach named Autoregressive Distributed Lag Model (ARDL). Among others, the authors of [42], [43], [44], and [45] have used ARDL, which starts as a general model and then moves into more specific one to capture the characteristics of the data included in the regression using a sufficient number of lags. As literature has discussed many advantages for using ARDL in cointegration models, it is worth to shed the light on some key advantages that have affected the selection of the model for this study. First, in [42], the author has reported that the ARDL approach is suitable for variables with different integration orders that could be either fractional or at I (0) or I (1). In addition, the equilibrium features of both short- and long-term dynamics are captured by the error correction model (ECM) which is obtained from a transformation that is done linearly for the ARDL model.

Considering the size of the sample, the authors of [44] have claimed that when the investigation is done on a small sample, then the approach of ARDL is more appropriate in comparison to that of [46] cointegration. They have also admitted that ARDL has minimal residual correlation and thus considered as a superior approach against serial correlation problems attributed to endogeneity issue. Lastly, being able to proceed with the

[4] The most available data for N_2O were till the year 2012.

estimation even if the explanatory variables are endogenous is another key advantage of ARDL model [42], [45].

3.4. Unit Root Testing

As time series data exhibit trending behavior and thus the null hypothesis of non-stationary cannot be rejected, unit root testing is applied using three different approaches called Augmented Dickey & Fuller (ADF) [47], Phillips and Perron (PP) [48] and KPSS [49] to test the order of integration of each variable. In fact, both ADF and PP tests suggest that the null hypothesis is non-stationary where KPSS claims a stationary null hypothesis. Since the original Dickey & Fuller test [47] cannot accommodate complex models, the ADF approach is a customized version of the test that can satisfy the need of having an appropriate stationary test for complex models. The ADF for models with unknown orders can be represented as shown below:

$$\Delta y_t = \theta_0 + \alpha_0 t + \alpha_1 y_{t-1} + \sum_{i=0}^{p} \theta_i \Delta y_{t-1} + \mu_t \quad (3)$$

where y_t is the variable in period t; Δy_{t-1} is the y_{t-1}-y_{t-2}; the disturbance term represented by μ_t is i.i.d and has zero mean where the variance is 1; t the linear time trend and p is the lag order. It is worth to note that the ADF test has been developed to examine univariate time series for the presence of unit root. In another word, any time series with presence of unit root should be treated to ensure that the variable is stationary (if it is required for modelling).

Since $\boldsymbol{\alpha_0}$ and $\boldsymbol{\alpha_1}$ are the coefficients of the time trend term (t) and the variable $\boldsymbol{y_{t-1}}$ in previous period (t-1), it is important to investigate the order of integration of the variable y_t. This is done by examining whether or not $\alpha_1 = 0$ in (3). Accordingly, if α_1 is not significantly less than zero, the null hypothesis of a unit root cannot be rejected. Otherwise, both level and first differenced variables are tested against unit root. Since in [49], the KPSS stationary test was introduced which assumes that stationary is the null hypothesis, this study has also applied this test to make sure that the results are superior.

3.5. Bound Testing Approach

The next step after getting the order of integration for each variable under concern, the ARDL bound testing [45] is implemented to investigate the existence of long run association between the variables. The ARDL bound testing depends on assessing the joint significance of the lagged variables coefficients using F-Wald test, which has a null hypothesis of H0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, whereas the alternative is that at least one coefficient (β) is not equal to zero. There are upper and lower critical values that should be compared with the results obtained from the F-statistics, where all are tabulated by [45]. There will be a proof of cointegration with a presence of long run association between the variables if the estimated F-statistic is more than the upper limit. If the computed F-statistic falls between the two critical limits, no conclusion can be provided by the test. However, the null hypothesis cannot be rejected if the values of F-statistic is less than the lower limit. In addition, this applies that there is no existence of cointegration among the tested variables.

If the estimated F-statistic shows the presence of long run relationship between the variables, the next stage of the ARDL specification is estimating the long run coefficients in (2) [50]. The long run impact on the N_2O emissions is measured by the estimated values of β_i . The Akaike Information Criteria (AIC) [51] is utilized to decide on the optimal lag length for each variable.

The residuals obtained from estimating (2) are used to approximate the error correction term (ECT), which shows the speed that the variables restore to their equilibrium levels in the long run from the short run. Therefore, the value of the coefficient of ECT should be negative and less than or equal to one besides being highly significant. Hence, the specifications of the error correction term of the ARDL approach can be estimated as follows:

$$\Delta ln\,(N_2O)_t = \delta_0 + \sum_{i=0}^{n} \delta_{1i}\,\Delta ln\,N_2O_{t-i} + \sum_{k=0}^{q} \delta_{1k}\,\Delta ln\,GDP_{t-k} + \sum_{j=0}^{d} \delta_{2j}\,\Delta\,(ln\,GDP_{t-j})^2 + \sum_{l=0}^{b} \delta_{3l}\,\Delta\,ln\,Elec_{t-l} + \sum_{w=0}^{y} \delta_{4w}\,\Delta ln\,Fin_{t-w} + \sum_{m=0}^{r} \delta_{5m}\,\Delta\,ln\,FDI_{t-m} + \theta ECT_{t-1} + \varepsilon_t \quad (4)$$

where Δ represents the growth or changes in N_2O emissions ($\boldsymbol{N_2O}$), GDP per capita ($\boldsymbol{GDP}$) and its quadratic term, electricity consumption per capita ($\boldsymbol{Elec}$), financial developments ($\boldsymbol{Fin}$), and the percentage of foreign direct investments ($\boldsymbol{FDI}$). The term of $\boldsymbol{ECT}$ reflects the speed of adjustment when a deviation take place. The value of the $\boldsymbol{ECT}$ is negative.

4. Empirical Results and Discussion

Before estimating any time series model, it is essential to examine the integration order of the variables and identify their order of integration. This study employs ADF test, KPSS test and PP test. The results of the three tests are reported in Table 2 which reveals that all the variables except the FDI are having unit root at level however a first difference convert them to stationary variables. Hence, the integration order is 1; i.e. I (1).

Table 2: Unit root results

Variable	ADF		KPSS		PPerron	
	Constant	Constant and Trend	Constant	Constant and Trend	Constant	Constant and Trend
lnN_2O	-2.33	-3.93**	0.48**	0.14*	-3.01**	-4.10*
lnGDPpc	-1.07	-2.33	0.31	0.11	-1.29	-2.55
lnElec	-3.23**	-2.55	0.37	0.15**	-3.65**	-2.57
Fin	-0.11	-3.04	0.46**	0.15**	0.22	-2.99
FDI	-5.51***	-5.48***	0.29	0.15**	-5.61***	-5.55***
ΔlnN2O	-7.49***	-8.35***	0.32	0.11	-7.37***	-8.45***
ΔlnGDpc	-4.74***	-4.63***	0.14	0.11	-4.76***	-4.62***
ΔlnElec	-5.24***	-5.69***	0.35	0.15**	-5.23***	-5.73***
ΔFin	-5.83***	-5.74***	0.29	0.13	-6.02***	-5.92***
ΔFDI	-8.05***	-7.94***	0.31	0.17**	-11.2***	-11.1***

Notes: ADF is the Augmented Dickey Fuller Unit root test, KPSS is the Kwiatkowski, Phillips, Schmidt & Shin stationarity test. PP is Phillips & Perron unit root test. lnN_2O is the natural logarithm of nitrous oxide emissions, lnGDPpc is the natural logarithm of GDP per capita, lnElec is the natural logarithm of electric power consumption, Fin is the Domestic credit provided by Financial Sector (% of GDP) and FDI is Foreign direct investment, net inflows (% of GDP). Δ is the first difference. *, ** and *** show 10%, 5% and 1% level of significance, respectively.

The findings obtained from stationary tests serve as the basics to implement the following steps of estimation. In order to explore the presence of long run relationship among the N_2O emissions and its determinants, the bound testing method of ARDL that aims for exploring cointegration is employed in (4) and the results are shown in Table 3. AIC is utilized in selecting the optimal lag structure for all the variables and the results are as (1,0,0,0,2,0) for the function $N_2O = f\ (GDP, GDP^2, Elec, Fin, FDI)$. The estimated F statistic is 10.052 which is greater than the value of the upper critical limit developed by [52]. Therefore, it is possible that the null hypothesis of no cointegration is rejected.

Table 3: Results of ARDL bound testing to cointegration

Model	Optimal lag structure	F - value	t - statistics
$N_2O = f\ (GDP, GDP^2, Elec, Fin, FDI)$	(1,0,0,0,2,0)	10.052	-6.77***

Table 4: Estimated Coefficients from ARDL (1,0,0,0,2,0) for Model $N_2O = f\ (GDP, GDP^2, Elec, Fin, FDI)$

	Variable	Coefficients	t-statistics
Long run estimates: lnN_2O as **dependent variable**	GDP_t	4.5	2.18**
	${GDP^2}_t$	-2.3	-2.20**
	$Elec_t$	0.66	2.74**
	Fin_t	-0.20	-4.64***
	FDI_t	-0.40	-1.33*
Short run estimates: ΔlnN_2O as **dependent variable**	GDP_{t-1}	2.30	2.42**
	${GDP^2}_{t-1}$	-4.6	-2.44**
	$Elec_{t-1}$	0.72	2.36**
	Fin_{t-1}	-0.10	-1.29
	Fin_{t-2}	0.40	3.60**
	FDI_{t-1}	-0.20	-0.33
	ECT_{t-1}	-0.76	-6.77***
	Constant	14.43	1.54
	trend	-0.44	-6.97***

Table 5: Diagnostic Tests

Test	Coefficient
R^2	0.76
Adjusted R^2	0.66
F- statistics	10.052(0.023)
Jarque-Bera normality test	1.062(0.334)
Heteroscedasticity Test: ARCH	2.83(0.984)
Breusch-Godfrey Serial Correlation LM Test	0.718(0.315)
Ramsey RESET test	0.723(0.413)

Table 4 presents the findings of ARDL estimation. The upper part of Table 4 illustrates the coefficients of the long run relationship. All the estimated coefficients appear to have significant impacts on N_2O emission at 1% or 5% level of significance except the coefficient of FDI, which is only significant at 10% level. Specifically, the GDP per capita elasticity is statistically significant and has a positive sign in both long and short run relationships. Moreover, the coefficient of the squared GDP per capita turned out to be negative and significant in both timeframes. The negative coefficient on the squared value of GDP per capita confirms the presence of the Environmental Kuznets Curve (EKC), which indicates the relationship between Bahrain economic growth and the level of N_2O emissions follows the inverted U-shape curve in the long run.

The findings of Table 4 demonstrate the coefficients of the other variables under concern. A 1% increase in electricity consumption causes 0.66% surge in N_2O diffusion. However, foreign direct investments and financial development variables are negatively related to the N_2O emissions as they cause a decrease in N_2O emissions of 0.4% and 0.2%, respectively. This is a good sign for policy makers to consider the improvement in the financial sector and draw the attention of foreign direct investments, which may help boosting air quality in Bahrain. Furthermore, the electricity consumption variable occurs to have a negative significant impact on N_2O emissions level.

The lower part of Table 4 reveals the findings of the short run estimations of the dynamic effects on N_2O diffusions by its determinants. The Δ sign implies that the variables are in their first differences. Briefly, for the error correction term (ECT_{t-1}), the coefficient is not only negative (as expected) but also significant at 5% level. This reinforces the hypothesis of cointegration and gives a measure for the how fast is the adjustment to equilibrium in the short run, which is around 76% in a year.

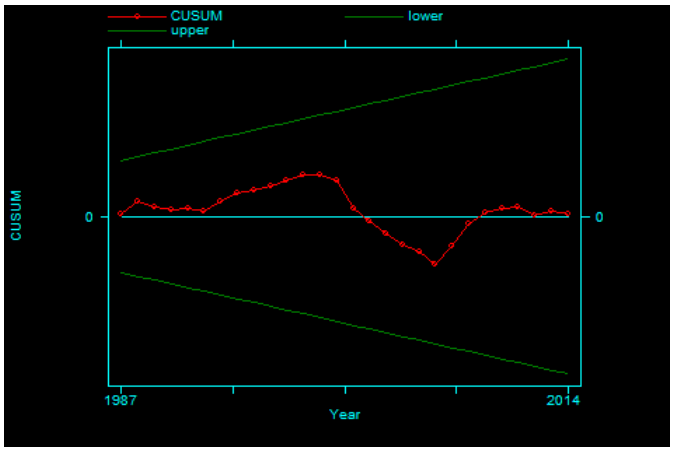


Figure 1: Plot of Cumulative Sum of Recursive Residuals

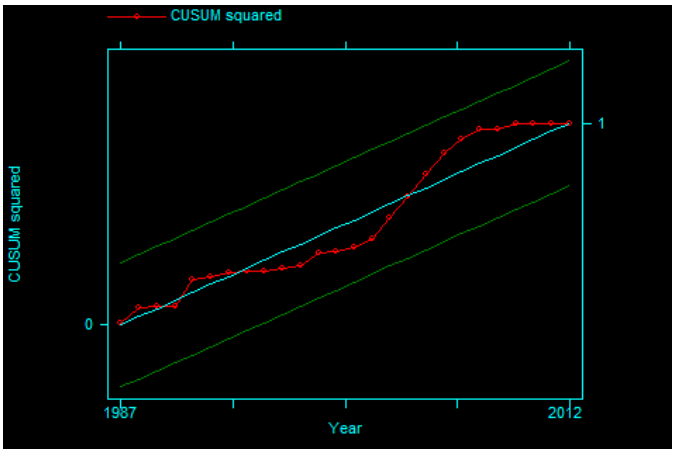


Figure 2: Plot of Cumulative Squared Sum of Recursive Residuals

To reassure the stability of the obtained findings, Table 5 reports the results obtained from some diagnostic tests that are applied for ARDL estimations. The findings prove that there is no serial correlation in the residuals, and the distribution is normal. Moreover, the variance of the errors is constant (homoscedastic).

Furthermore, the cumulative sum of recursive residuals (CUSUM) and the cumulative squared sum of recursive residuals (CUSUMSQ) established by [53] are plotted in Figures 1 and 2 to examine the model's parameters stability. Figures 1 and 2 illustrate the stability of the estimated model as both plots occur within the limits of the 5% confidence interval. It is important to check the stability of the model for it to be strong enough for forecasting issues and accordingly applicable for policies implementations.

5. Conclusion and Recommendations

This paper examines the validity of the EKC hypothesis in Bahrain over the period 1980 – 2012 with the implication of nitrous oxide (N_2O) emissions as an environmental pollutant. The estimated equation is augmented with electric power consumption, foreign direct investment and financial development indicator. ARDL approach is employed to examine the short and long run impact of the variables under interest on N_2O emissions. The obtained results from the ARDL estimation indicate the validity of the EKC hypothesis. Moreover, electric power consumption has a positive impact on N_2O emissions in the short and long run. However, foreign direct investments and financial development have negative impact on the emissions of N_2O.

In order to decrease the impact of electric power consumption on N_2O emissions, Bahrain should continue its efforts in decreasing the environmental emissions such as assisting households to install solar cells on the top of their houses and use solar energy in generating their own need for electricity, which may help in reducing the electricity production using fossil fuel combustion.

Although the authors of [36] and [37] argue that foreign investors and international corporations prefer investing in countries that have loose environmental policies and standards as most of these investments sponsor forms of production that are environmentally inefficient [38], this paper found that FDI reduces N_2O emissions. Our results are in line with that of [54] who found that FDI can be good for the environment. They explain this relationship by referring to the possibility of transferring the foreign firms' green technologies to their domestic counterparts which may have low environmental-friendly technologies. As the FDI helps decreasing N_2O emissions, Bahrain should increase its focus on the quality of FDI to be technology oriented FDI. Moreover, the financial development indicator reduces the N_2O emissions. Therefore, Bahrain should enhance its financial sector by developing bond and securities markets. This will improve the financial services and provide more funds to be invested in research and development on new and advanced techniques to generate clean energy.

Acknowledgment

Authors would like to thank the referees for their useful comments and support.

References

[1] R.J. Hill, E. Magnani, "An exploration of the conceptual and empirical basis of the environmental Kuznets curve," Australian Economic Papers, **41**(2), 239–254, 2002. doi: 10.1111/1467-8454.00162

[2] M.D. Miah, M.F.H. Masum, M. Koike, "Global observation of EKC hypothesis for CO2, SOx and NOx emission: A policy understanding for climate change mitigation in Bangladesh," Energy Policy, **38**(8), 4643–4651, 2010. doi: 10.1016/j.enpol.2010.04.022

[3] M.A. Zambrano-Monserrate, M.A. Fernandez, "An environmental Kuznets curve for N2O emissions in Germany: an ARDL approach," in Natural resources forum, Wiley Online Library: 119–127, 2017. doi: 10.1111/1477-8947.12122

[4] A.M. Rasli, M.I. Qureshi, A. Isah-Chikaji, K. Zaman, M. Ahmad, "New toxics, race to the bottom and revised environmental Kuznets curve: The case of local and global pollutants," Renewable and Sustainable Energy Reviews, **81**, 3120–3130, 2018. doi: 10.1016/j.rser.2017.08.092

[5] G.M. Grossman, A.B. Krueger, Environmental impacts of a North American free trade agreement, National Bureau of economic research Cambridge, Mass., USA, 1991.

[6] T. Panayotou, Empirical tests and policy analysis of environmental degradation at different stages of economic development, International Labour Organization, 1993.

[7] S. Farhani, M. Shahbaz, R. Sbia, A. Chaibi, "What does MENA region initially need: grow output or mitigate CO2 emissions?," Economic Modelling, **38**, 270–281, 2014. doi: 10.1016/j.econmod.2014.01.001

[8] L.-S. Lau, C.-K. Choong, Y.-K. Eng, "Investigation of the environmental Kuznets curve for carbon emissions in Malaysia: do foreign direct investment and trade matter?," Energy Policy, **68**, 490–497, 2014. doi: 10.1016/j.enpol.2014.01.002

[9] S. Oshin, A. Ogundipe, "An empirical examination of environmental Kuznets curve (EKC) in West Africa," Euro-Asia Journal of Economics and Finance, **3**(1), 2014. https://ssrn.com/abstract=2512136

[10] N. Apergis, "Environmental Kuznets curves: New evidence on both panel and country-level CO2 emissions," Energy Economics, **54**, 263–271, 2016. doi: 10.1016/j.eneco.2015.12.007

[11] S. Park, Y. Lee, "Regional model of EKC for air pollution: Evidence from the Republic of Korea," Energy Policy, **39**(10), 5840–5849, 2011. doi: 10.1016/j.enpol.2011.06.028

[12] O.A. Onafowora, O. Owoye, "Bounds testing approach to analysis of the environment Kuznets curve hypothesis," Energy Economics, **44**, 47–62, 2014. doi: 10.1016/j.eneco.2014.03.025

[13] S. Özokcu, Ö. Özdemir, "Economic growth, energy, and environmental Kuznets curve," Renewable and Sustainable Energy Reviews, **72**, 639–647, 2017. doi: 10.1016/j.rser.2017.01.059

[14] E.L. Effiong, A.O. Iriabije, "Let the data speak: semiparametric evidence on the environmental Kuznets curve in Africa," Quality & Quantity, **52**(2), 771–782, 2018. doi: 10.1007/s11135-017-0487-6

[15] M. Fodha, O. Zaghdoud, "Economic growth and pollutant emissions in Tunisia: an empirical analysis of the environmental Kuznets curve," Energy Policy, **38**(2), 1150–1156, 2010. doi: 10.1016/j.enpol.2009.11.002

[16] C.-H. Cho, Y.-P. Chu, H.-Y. Yang, "An environment Kuznets curve for GHG emissions: a panel cointegration analysis," Energy Sources, Part B: Economics, Planning, and Policy, **9**(2), 120–129, 2014. doi: 10.1080/15567241003773192

[17] L. Charfeddine, Z. Mrabet, "The impact of economic development and social-political factors on ecological footprint: A panel data analysis for 15 MENA countries," Renewable and Sustainable Energy Reviews, **76**, 138–154, 2017. doi: 10.1016/j.rser.2017.03.031

[18] Y. Chiu, "Deforestation and the environmental Kuznets curve in developing countries: A panel smooth transition regression approach," Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie, **60**(2), 177–194, 2012. doi: 10.1111/j.1744-7976.2012.01251.x

[19] K. Ahmed, M. Shahbaz, A. Qasim, W. Long, "The linkages between deforestation, energy and growth for environmental degradation in Pakistan," Ecological Indicators, **49**, 95–103, 2015.

[20] M. Mazzanti, A. Montini, R. Zoboli, "Municipal waste generation and the EKC hypothesis new evidence exploiting province-based panel data," Applied Economics Letters, **16**(7), 719–725, 2009. doi: 10.1080/13504850701221824

[21] P.K. Narayan, S. Narayan, "Carbon dioxide emissions and economic growth: Panel data evidence from developing countries," Energy Policy, **38**(1), 661–666, 2010. doi: 10.1016/j.enpol.2009.09.005

[22] Z. Zoundi, "CO2 emissions, renewable energy and the Environmental

Kuznets Curve, a panel cointegration approach," Renewable and Sustainable Energy Reviews, **72**, 1067–1075, 2017. doi: 10.1016/j.rser.2016.10.018

[23] S. Sinha Babu, S.K. Datta, "The relevance of environmental Kuznets curve (EKC) in a framework of broad-based environmental degradation and modified measure of growth–a pooled data analysis," International Journal of Sustainable Development & World Ecology, **20**(4), 309–316, 2013. doi: 10.1080/13504509.2013.795505

[24] F. Abbasi, K. Riaz, "CO2 emissions and financial development in an emerging economy: an augmented VAR approach," Energy Policy, **90**, 102–114, 2016. doi: 10.1016/j.enpol.2015.12.017

[25] S. khoshnevis Yazdi, B. Shakouri, "The renewable energy, CO2 emissions, and economic growth: VAR model," Energy Sources, Part B: Economics, Planning, and Policy, **13**(1), 53–59, 2018. doi: 10.1080/15567249.2017.1403499

[26] U. Al-Mulali, S.A. Solarin, I. Ozturk, "Investigating the presence of the environmental Kuznets curve (EKC) hypothesis in Kenya: an autoregressive distributed lag (ARDL) approach," Natural Hazards, **80**(3), 1729–1747, 2016. doi: 10.1007/s11069-015-2050-x

[27] E. Dogan, B. Turkekul, "CO 2 emissions, real output, energy consumption, trade, urbanization and financial development: testing the EKC hypothesis for the USA," Environmental Science and Pollution Research, **23**(2), 1203–1213, 2016. doi: 10.1007/s11356-015-5323-8

[28] W. Ali, A. Abdullah, M. Azam, "Re-visiting the environmental Kuznets curve hypothesis for Malaysia: fresh evidence from ARDL bounds testing approach," Renewable and Sustainable Energy Reviews, **77**, 990–1000, 2017. doi: 10.1016/j.rser.2016.11.236

[29] U.K. Pata, "Renewable energy consumption, urbanization, financial development, income and CO2 emissions in Turkey: testing EKC hypothesis with structural breaks," Journal of Cleaner Production, **187**, 770–779, 2018. doi: 10.1016/j.jclepro.2018.03.236

[30] B. Raggad, "Carbon dioxide emissions, economic growth, energy use, and urbanization in Saudi Arabia: evidence from the ARDL approach and impulse saturation break tests," Environmental Science and Pollution Research, **25**(15), 14882–14898, 2018. doi: 10.1007/s11356-018-1698-7

[31] A.S. Alshehry, M. Belloumi, "Study of the environmental Kuznets curve for transport carbon dioxide emissions in Saudi Arabia," Renewable and Sustainable Energy Reviews, **75**, 1339–1347, 2017. doi: 10.1016/j.rser.2016.11.122

[32] H. Mahmood, T.T.Y. Alkhateeb, "Trade and environment nexus in Saudi Arabia: An environmental Kuznets curve hypothesis," International Journal of Energy Economics and Policy, **7**(5), 291–295, 2017. https://EconPapers.repec.org/RePEc:eco:journ2:2018-03-5

[33] M. Shahbaz, R. Sbia, H. Hamdi, I. Ozturk, "Economic growth, electricity consumption, urbanization and environmental degradation relationship in United Arab Emirates," Ecological Indicators, **45**, 622–631, 2014. doi: 10.1016/j.ecolind.2014.05.022

[34] L. Charfeddine, K. Ben Khediri, "Financial development and environmental quality in UAE: Cointegration with structural breaks," Renewable and Sustainable Energy Reviews, **55**, 1322–1335, 2016. doi: 10.1016/j.rser.2015.07.059

[35] Z. Mrabet, M. Alsamara, "Testing the Kuznets Curve hypothesis for Qatar: A comparison between carbon dioxide and ecological footprint," Renewable and Sustainable Energy Reviews, **70**, 1366–1375, 2017. doi: 10.1016/j.rser.2016.12.039

[36] N.D. Woods, "Interstate competition and environmental regulation: a test of the race-to-the-bottom thesis," Social Science Quarterly, **87**(1), 174–189, 2006. doi: 10.1111/j.0038-4941.2006.00375.x

[37] C. Dick, A.K. Jorgenson, "Sectoral foreign investment and nitrous oxide emissions: A quantitative investigation," Society & Natural Resources, **23**(1), 71–82, 2009. doi: 10.1080/08941920802392690

[38] P. Grimes, J. Kentor, "Exporting the greenhouse: foreign capital penetration and CO? Emissions 1980 1996," Journal of World-Systems Research, 261–275, 2003.

[39] U. Soytas, R. Sari, B.T. Ewing, "Energy consumption, income, and carbon emissions in the United States," Ecological Economics, **62**(3–4), 482–489, 2007. doi: 10.1016/j.ecolecon.2006.07.009

[40] M. Shahbaz, M.M. Rahman, "Foreign capital inflows-growth nexus and role of domestic financial sector: an ARDL co-integration approach for Pakistan," Journal of Economic Research, **15**(3), 207–231, 2010.

[41] T. Chang, W. Fang, L.-F. Wen, "Energy consumption, employment, output, and temporal causality: evidence from Taiwan based on cointegration and error-correction modelling techniques," Applied Economics, **33**(8), 1045–1056, 2001. 10.1080/00036840122484

[42] M.H. Pesaran, "The role of economic theory in modelling the long run," The Economic Journal, **107**(440), 178–191, 1997. https://www.jstor.org/stable/2235280

[43] M.H. Pesaran, R.P. Smith, "Structural analysis of cointegrating VARs," Journal of Economic Surveys, **12**(5), 471–505, 1998. doi: 10.1111/1467-6419.00065

[44] H.H. Pesaran, Y. Shin, "Generalized impulse response analysis in linear multivariate models," Economics Letters, **58**(1), 17–29, 1998. doi: 10.1016/S0165-1765(97)00214-0

[45] M.H. Pesaran, Y. Shin, R.J. Smith, "Bounds testing approaches to the analysis of level relationships," Journal of Applied Econometrics, **16**(3), 289–326, 2001. doi: 10.1002/jae.616

[46] S. Johansen, K. Juselius, "Maximum likelihood estimation and inference on cointegration—with appucations to the demand for money," Oxford Bulletin of Economics and Statistics, **52**(2), 169–210, 1990. doi: 10.1111/j.1468-0084.1990.mp52002003.x

[47] D.A. Dickey, W.A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," Journal of the American Statistical Association, **74**(366a), 427–431, 1979. doi: 10.1080/01621459.1979.10482531

[48] P.C.B. Phillips, P. Perron, "Testing for a unit root in time series regression," Biometrika, **75**(2), 335–346, 1988. doi: 10.1093/biomet/75.2.335

[49] D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," Journal of Econometrics, **54**(1–3), 159–178, 1992. doi: 10.1016/0304-4076(92)90104-Y

[50] M. Bouznit, M. del P. Pablo-Romero, "CO2 emission and economic growth in Algeria," Energy Policy, **96**, 93–104, 2016. doi: 10.1016/j.enpol.2016.05.036

[51] Y. Sakamoto, M. Ishiguro, G. Kitagawa, "Akaike information criterion statistics," Dordrecht, The Netherlands: D. Reidel, **81**(10.5555), 26853, 1986. doi: 10.1080/01621459.1988.10478680

[52] P.K. Narayan, "The saving and investment nexus for China: evidence from cointegration tests," Applied Economics, **37**(17), 1979–1990, 2005. doi: 10.1080/00036840500278103

[53] R.L. Brown, J. Durbin, J.M. Evans, "Techniques for testing the constancy of regression relationships over time," Journal of the Royal Statistical Society: Series B (Methodological), **37**(2), 149–163, 1975. doi: 10.1111/j.2517-6161.1975.tb01532.x

[54] B.A. Demena, S.K. Afesorgbor, "The effect of FDI on environmental emissions: Evidence from a meta-analysis," Energy Policy, **138**, 111192, 2020. doi: 10.1016/j.enpol.2019.111192

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 4, 196-211 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Boltzmann-Based Distributed Control Method: An Evolutionary Approach Using Neighboring Population Constraints

Gustavo Alonso Chica Pedraza[1*], Eduardo Alirio Mojica Nava[2], Ernesto Cadena Muñoz[3]

[1]*School of Telecommunications Engineering, Universidad Santo Tomás, Bogotá D.C., 10111, Colombia*

[2]*Department of Electric and Electronic Engineering, Universidad Nacional de Colombia, Bogotá D.C., 10111, Colombia*

[3]*Department of Systems and Industrial Engineering, Universidad Nacional de Colombia, Bogotá D.C., 10111, Colombia*

ARTICLE INFO

Article history:
Received: 30 April, 2021
Accepted: 06 July, 2021
Online: 27 July, 2021

Keywords:
Distributed control
Entropy
Learning
Population dynamics
Selection-Mutation

ABSTRACT

In control systems, several optimization problems have been overcome using Multi-Agent Systems (MAS). Interactions of agents and the complexity of the system can be understood by using MAS. As a result, functional models are generated, which are closer to reality. Nevertheless, the use of models with permanent availability of information between agents is assumed in these systems. In this sense, some strategies have been developed to deal with scenarios of information limitations. Game theory emerges as a convenient framework that employs concepts of strategy to understand interactions between agents and maximize their outcomes. This paper proposes a learning method of distributed control that uses concepts from game theory and reinforcement learning (RL) to regulate the behavior of agents in MAS. Specifically, Q-learning is used in the dynamics found to incorporate the exploration concept in the classic equation of Replicator Dynamics (RD). Afterward, through the use of the Boltzmann distribution and concepts of biological evolution from Evolutionary Game Theory (EGT), the Boltzmann-Based Distributed Replicator Dynamics are introduced as an instrument to control the behavior of agents. Numerous engineering applications can use this approach, especially those with limitations in communications between agents. The performance of the method developed is validated in cases of optimization problems, classic games, and with a smart grid application. Despite the information limitations in the system, results obtained evidence that tuning some parameters of the distributed method allows obtaining an analogous behavior to that of the conventional centralized schemes

1 Introduction

This original research paper is an extension of the work initially presented in the Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI) 2020 [1]. In this version, readers can find a full view of the proposed learning distributed method, which uses concepts from Game Theory (GT) to control complex systems. This paper also presents an evaluation of the method from an evolutionary perspective of the obtained equations. This work also simulates a modified version of the case study presented in the conference, which includes different communication constraints and attributes of the generators employed in the power grid. Moreover, some additional cases in the context of classic games and maximization problems are introduced to make clearer the incidence of some control parameters in the behavior of agents.

The idea to model and control complex systems has increased over time. In this context, Engineering applications have received special interest due to their affinity with the use of mathematical techniques to prove new models and concepts on applications closer to reality [2]. In recent decades, research has been focused on the study of distributed systems with large-scale control. Numerous models and techniques have been developed to overcome issues such as the expensive computational requirements, the structure of the communication, and the calculation of the data required to complete a task in large-scale systems. These issues can be managed by using Multi-Agent Systems (MAS) and concepts from game theory [3]. In this sense, the interactions of agents have been thoroughly studied, as some strategies can help agents maximize their outcomes. For example, [4] establishes relations among games, learning, and

*Corresponding Author: Gustavo Alonso Chica Pedraza, Carrera 9 No 51-11, Bogotá D.C., Colombia, +5715878797 Ext 1654, gustavochica@usantotomas.edu.co

optimization in networks. Other studies have focused on games and learning [5] or on algorithms for distributed computation in topologies of dynamic networks [6]. Authors in [7] studied the main applications of power control in the frameworks of distributed and centralized game theory. Regarding smart grid control applications, please refer to [8]. Other research has concentrated on cases with issues in coordination and negotiation that guide the study of the interactions of agents[9]. For further studies on applications of power control using game theory, please refer to [10]. Research on game theory considers three types of games. First, continuous games consider the way an agent can have a pure strategy looking for maximum profit. Second, in matrix games, agents are regarded as individuals and can take only one shot to play simultaneously. Finally, dynamic games suppose that players can learn in some way about the environment, that is, their actions and states. This assumption means agents can learn and correct their behavior based on the outcomes of their actions [11]. Dynamic games must deal with the following challenges: modeling the environment for agents interaction, modeling the agents goals, the prioritization of the agents actions, and the estimation of the amount of information owned by a player [12].

The study of dynamics of agents changing over time is a concept of dynamic games introduced by Evolutionary game theory (EGT) [13]. The concept of the evolutionary stable strategy popularized EGT thanks to the analogy with biology concepts and the comparison with natural behaviors [14]. Some real-life control applications have employed EGT, whose understanding serves as a basis for the replicator dynamics (RD) approach. The revision protocols describe the way agents choose and modify their strategies, while population games determine the agents' interactions. The combination of both revision protocols and population games produces the concept of evolutionary game dynamics [14]. This perspective of evolution is often used to model large-scale systems because its mathematical background helps to describe this process with differential equations [13]. Many areas of Engineering have applied EGT, for example, optimization problems, control of communication access, systems of microgrids, etc. [11]. The use of EGT to model engineering problems has revealed the following benefits: ease to relate a game to an engineering problem, where payoff functions can be defined with the objective function and the strategies, and the relationship between the optimization concept and Nash Equilibrium, which is enabled under particular conditions that met the conditions of the first-order optimization of the Karush–Kuhn–Tucker. Last but not least, EGT uses local information to achieve solutions. In this sense, distributed approaches emerge to tackle engineering problems, which is useful when considering the implementation cost of centralized schemes and their complexity [11]. Distributed schemes of population dynamics have outstanding features over techniques like the method of dual decomposition, which requires a centralized coordinator [15]. This characteristic reduces the associated cost with the structure of communication. Additionally, in comparison with distributed learning algorithms in normal-form-games, there are no failures in distributed population dynamics when all the variables involved in decision-making have limitations[16]. This makes Distributed Population Dynamics suitable for solving issues regarding allocation of resources like in a smart city design [17]. For these purposes, the distributed power generation needs to be integrated so that electric grids be more reliable, robust, efficient, and flexible. Nevertheless, modeling a grid using a distributed approach instead of the classic centralized, is an option to consider due to its realism and flexibility, according to microgrids constraints [18]. In this sense, control operations are considered individually in microgrids, as they make a distinction among the power generation, the secondary frequency, and the economic dispatch [19]. Static optimization concepts are employed to manage the economic dispatch [20] or even methods like the offline direct search [21]. The analysis may be more complicated if it includes loads, the generator, and power line losses in the distributed model. Other approaches cannot consider the dynamic conditions like the economic dispatch time dependence [22]. Some approaches have been developed to face these challenges. For instance, [23] presents a management system for a microgrid with centralized energy and stand-alone mode to study its static behavior. Other research employs a distributed control strategy considering power line signaling for energy storage systems [24]. The employment of the MAS framework in economic problems using a distributed approach was gathered in [25], taking into account the delays in the communication system. Microgrid architectures have also been proposed considering distributed systems like the microgrid hierarchical control [26].

This paper presents an approach to overcome some of the issues identified in the literature review. The aim is to show how to develop a control method of learning to study the influence of the exploration concept in MAS, that is, interaction between agents. RD was developed from simple learning models [27], so this research seeks to bring the exploration concept into the traditional exploration-less expression of RD, using the Q-learning dynamics. As a result, the combination of these frameworks opens up a path to tackle dynamics in a scenario where the feedback of each agent is determined by the agent itself and by other agents, and where interaction between them is limited. For the analysis, the Boltzmann distribution includes a distributed perspective of the Replicator Dynamics as a way to regulate the agents' behavior in a determined scenario. The developed method employs a temperature parameter and the presence of entropy terms, to modify the learning agents' behavior and link the selection-mutation process from EGT and the exploration-exploitation concept from RL. This attribute complies with the traditional positive condition of EGT techniques (modeling agents' interaction). Nevertheless, In the control area, the employment of these techniques has to be understood more on the normative side of things. To explain these features, this approach employs theory of RL, EGT, and decision-making to solve some cases in the context of classic games and maximization problems using a novel distributed model of learning. It also uses experimental data to tackle an economic dispatch problem, which is a common problem in smart grids. The results obtained by the proposed approach are contrasted with the classical centralized framework of RD.

The remainder of this paper is organized as follows. Section 2 presents, a short synopsis of game theory and reinforcement learning, as well as the relationship between EGT and Q-learning using the Boltzmann distribution. Section 3 explains a distributed neighboring concept used for the Boltzmann control method, considering the behavior of replicator dynamics. Section 4 introduces important concepts from the previous Section, related to evolutionary game theory and reinforcement learning. In Section 5, the employment of

the learning method on traditional cases of GT and maximization problems presents the background to analyze the application of the Boltzmann model behavior on a smart grid real-life case. Finally, the main conclusions of the study are summarized in Section 6.

2 Preliminaries

Game theory includes a group of equations and concepts to study the background in decentralized control issues. Most of the time, a game comprises a group of players (agents) with similar population behavior that choose the best way to execute actions. The strategy of a player can decrease rewards after performing a wrong action or increase rewards when the action was correct [28]. The theory of learning is used to understand this behavior. In this sense, the scheme of RL explains the relationship among the environment, signals, states, and actions. In the interaction, at each step, each player gets a notification with the current state of the environment and a reinforcement signal, then, the player chooses a strategy. Each player of the game aims to find the policy that produces the best rewards after recognizing the consequence of its actions, that is, reward or punishment. A structure of estimated value functions is characteristic of traditional RL methods [29]. The total reward that a player can obtain is usually a pair state-action or a state value. This means that the optimal value function is needed to find the policy that correctly fulfills payoffs. The Markov decision process and value iteration algorithm can be employed for this purpose [30] when the scenario is familiar. In other cases, Q-learning can be used as an adaptable method of value iteration where the model of the scenario does not require to be specific. Equation (1) depicts the Q-learning interaction process [31]:

$$Q_{t+1}(s,a) \leftarrow (1-\alpha)Q_t(s,a) + \alpha(\Gamma + \gamma max_{a'} Q_t(s',a')) \quad (1)$$

The whole process begins at time Q_{t+1} with an initial pair of action-state (s,a), then, after performing action a achieves the $Q_t(s',a')$, where (s',a') represents the newest values of s and a, respectively. $max_{a'}$ obtains the uppermost value of Q from s' by selecting the action that increases its value. α represents the general step size parameter, Γ is the instant reinforcement, and γ is a deduction parameter. When players have complete access to the game information and there are no communication limitations, the theory of learning and games are valuable instruments to deal with control applications that use a centralized approach. Nevertheless, these models aim to provide a close description of optimal circumstances, but they have some drawbacks when dealing with more realistic conditions, communication constraints, and the individuals rationality. In this vein, EGT tries to loosen the idea of rationality, by substituting it with biological notions like evolution, mutation, and natural selection [32, 33]. In EGT, there is a genetic encoding of the strategies of the players, which are called genotypes and represent the conduct of every player employed to calculate its outcomes. The quantity of other types of agents in the scenario determines the payoff of the genotype of each player genotype. In EGT, the population strategies begin to evolve employing a dynamic process that allows finding the expected value of this process through the use of the Replicator dynamics equation. An evolutionary system often returns to two concepts: mutation and selection. On the one hand, mutation provides variety to the population. On the other hand, selection provides priority to some varieties where every genotype is a pure strategy $Q_j(n)$, where the RD offspring expresses this behavior. The general equation of RD [27] is presented in Equation (2).

$$\frac{dx_i}{dt} = [(Ax)_i - x \cdot Ax]x_i \quad (2)$$

where x_i is the portion of a population that plays the i-th strategy. The payoff matrix is written as A and it owns diverse payoff values that each replicator obtains from other agents. The vector of probability $x = (x_1, x_2, ..., x_J)$ often defines the population state (x), and evidence the diverse density values of each type of replicator. Consequently, $(Ax)_i$ is the payoff obtained by the i-th player with x state. Then, the average payoff would be written as $x \cdot Ax$. Similarly, $\frac{dx_i}{dt}$ symbolizes the growth rate of the population playing the i-th strategy, which is calculated using the obtained payoff value after playing the i-th strategy and its difference with the average population payoff. [34].

2.1 Relating EGT and Q-Learning

In [35], the frameworks of RD and Q-learning are related in the context of two-player games, where players have different strategies. This relationship is conceivable as players can also be considered Q-learners. For modelling this case, a differential equation is needed for player R (rows) and another one for player C (columns). When $A = B^t$, the standard RD Equation (2) is employed, where x_i is substituted by r_i or c_i. Thence, A or B, and the change in state (x) for r or c determine the payoff matrix for a specific player. Therefore, $(Ax)_i$ switches to $(Ac)_i$ or $(Br)_i$ and is the reward obtained by the i-th player with a r or c state. Likewise, for players R and C, the growth rate $\frac{dx_i}{dt}$ switches to $\frac{dr_i}{dt}$ or $\frac{dc_i}{dt}$, respectively. This behavior is explained using the following system of differential equations [27] below:

$$\frac{dr_i}{dt} = [(Ac)_i - r \cdot Ac]r_i \quad (3)$$

$$\frac{dc_i}{dt} = [(Br)_i - c \cdot Br]c_i \quad (4)$$

Equations (3) and (4) denote the group of replicator dynamics equations used to model the behavior of two populations. Each population has a growth rate determined by the other populations. For example, A and B denote two payoff matrices that are needed to estimate the rate of change for two different current players in the problem using this group of differential equations. To find the relationship between the Q-learning framework and the RD equations, Equation (5) is introduced:

$$x_i(\delta) = \frac{e^{\tau Q_{a_i}(\delta)}}{\sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta)}} \quad (5)$$

where the notation $x_i(\delta)$ means the prospect of using strategy i at time δ, and τ symbolizes the temperature. Equation (5) is well-known as the Boltzmann distribution and is used in [35] to obtain the continuous time model of Q-Learning in the context of a game played by two players, as shown in Equation (6), where $\frac{dx_i}{dt}$ is written as $\dot{x}_i$.

$$\frac{\dot{x}_i}{x_i} = \tau\Big[\frac{dQ_{a_i}}{dt} - \sum_{j=1}^{n} \frac{dQ_{a_i}}{dt} x_j\Big] \quad (6)$$

The expression $\frac{dQ_{a_i(t)}}{dt}$ in Equation (6) can be solved by using Equation (1) to represent the Q-learner update rule. Equation (7) presents the equation of difference for the function Q.

$$\Delta Q_{a_i}(\delta) = \alpha\Big[\Gamma_{a_i}(\delta+1) + \gamma \max Q - Q_{a_i}(\delta)\Big] \tag{7}$$

The term σ expresses the time spent between two repetitions of the Q-values updates, where $0 < \sigma \leq 1$, while $Q_{a_i}(\delta\sigma)$ denotes the Q-values at time $k\sigma$. Then, by assuming an infinitesimal scheme of this expression, Equation (7) converts to Equation (8) after taking the limit $\sigma \to 0$.

$$\frac{\dot{x}_i}{x_i} = \tau\alpha\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j} + \sum_{j=1}^{n} x_j(Q_{a_j} - Q_{a_i})\Big] \tag{8}$$

As $\frac{x_j}{x_i}$ comes to $\frac{e^{\tau}\Delta Q_{a_j}}{e^{\tau}\Delta Q_{a_i}}$, the part after the sum in Equation (8) can be written in logarithm terms:

$$\alpha\Big[\tau\sum_{j} x_j(Q_{a_j} - Q_{a_i})\Big] = \alpha\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{9}$$

The last expression in Equation (8) is reorganized and replaced, so it converts to Equation (10).

$$\frac{\dot{x}_i}{x_i} = \alpha\tau\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j}\Big] + \alpha\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{10}$$

For using payoff matrices in games with two players, Γ_{a_i} as $\sum_j a_{ij}y_j$ can be written, then, the expressions for players 1 and 2 are expressed as shown in Equations (11) and (12), respectively:

$$\dot{x}_i = x_i\alpha\tau\Big[(A\mathbf{y})_i - \mathbf{x}\cdot A\mathbf{y}\Big] + x_i\alpha\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{11}$$

$$\dot{y}_i = y_i\alpha\tau\Big[(B\mathbf{x})_i - \mathbf{y}\cdot B\mathbf{x}\Big] + y_i\alpha\Big[\sum_{j=1}^{n} y_j \ln\left(\frac{y_j}{y_i}\right)\Big] \tag{12}$$

These expressions denote the derivation of the continuous-time model for Q-learning. For the full process of the derivation, see Annex A. The Equations (11) and (12) can be considered as a centralized perspective, analogous to the Equations (3) and (4) that represent the standard RD form to model actions of players R and C, in a game of 2 players. However, the Boltzmann model produces the main differences with the introduction of α and τ parameters, and the emergence of an additional term. This approach has been applied in some scenarios such as multiple state games, multiple player games, and in the context of 2×2 games [27]. Nevertheless, research is still needed to use this approach in real-life problems.

The following Section presents our approach, which is a learning method that uses a distributed population perspective to control agents' behaviors. This proposal uses some of the principles stated in [35] to introduce the Boltzmann-based distributed replicator dynamics approach. This paper also uses the concept of population dynamics but employing constraints in the agents communications and assuming players should use neighboring strategies, thus, having a scenario where players have no full information of the system.

3 The Boltzmann-based distributed replicator dynamics method

In the following paragraphs, we describe the Boltzmann-based distributed replicator dynamics method. The starting point needed to perform the development of this method is the Equation (11). This formalism is useful since it employs the Boltzmann concept and its first term has the classic form of the RD when modeling games that use payoff matrices. Considering the idea to have an analogous and more general form to express the RD expression, Equation (11) can be written as Equation (13):

$$\dot{x}_i = \alpha x_i\tau\Big[f_i(x) - \overline{f}(x)\Big] + \alpha x_i\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{13}$$

In this equation, a fraction of a determined population can augment or diminish depending on the higher/lower fitness values of its individuals with respect to the population average. The population is represented by the state vector $x = (x_1, x_2, ..., x_n)^n$ with $0 \leq x_i \leq 1, \forall_i$ and $\sum_{i=1}^{n} x_i = 1$, which denotes the portions that belong to each of the n-types. In $f_i(x)$, i denotes the fitness type. Consequently, the fitness average of the population is expressed by $f(x) = \sum_j x_j f_j(x)$. Using these assumptions, this expression becomes:

$$\dot{x}_i = \alpha x_i\tau\Big[f_i(x) - \sum_{j=1}^{n} x_j f_j(x)\Big] + \alpha x_i\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{14}$$

The first term of Equation (14) is written as the centralized equation for the RD. We propose to adapt it to a decentralized form, to compute the local information of the players to tackle limitations in communication. The decentralized expression of this step is written in Equation (15):

$$\dot{x}_i = \underbrace{\alpha x_i\tau\Big[f_i(x) - \sum_{j=1}^{n} x_j f_j(x)\Big]}_{Centralized} = \underbrace{\alpha x_i\tau\Big[f_i(x)\sum_{j=1}^{n} x_j - \sum_{j=1}^{n} x_j f_j(x)\Big]}_{Decentralized} \tag{15}$$

where $\sum_{j=1}^{n} x_j$ is equivalent to the unit, since the term x_j of the operation denotes the probabilities of selecting the *jth* strategy. Likewise, when using logarithms rules, the second term in Equation (14) becomes:

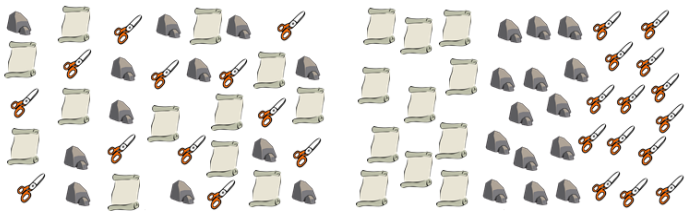
$$\underbrace{\alpha x_i\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big]}_{centralized} = \underbrace{-\alpha x_i\Big[\ln x_i - \sum_{j=1}^{n} x_j \ln x_j\Big]}_{Decentralized} \tag{16}$$

Finally, substituting Equations (15) and (16) in Equation (14), becomes Equation (17), that expresses the Decentralized form of the Replicator Dynamics equation in connection with Boltzmann probabilities.

$$\dot{x}_i = \alpha x_i\tau\Big[f_i(x)\sum_{j=1}^{n} x_j - \sum_{j=1}^{n} x_j f_j(x)\Big] - \alpha x_i\Big[\ln x_i - \sum_{j=1}^{n} x_j \ln x_j\Big] \tag{17}$$

This equation is studied in detail in Section 4 considering EGT with the selection-mutation concept and the exploration-exploitation

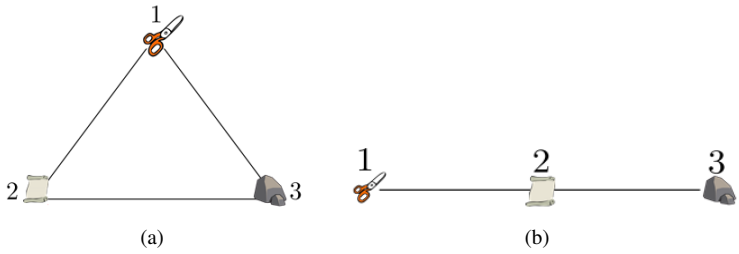
approaches with their influence on MAS. in Equation (17), the first parenthesis corresponds to alterations in the proportion of players that are using the i-th strategy and require complete information about the state of the whole population and the payoff functions. Consequently, complete information of the system is required so that population dynamics evolve. However, since this work aims to control scenarios where agents cannot access the complete information of the system, there should not be dependence on complete information, for example, in scenarios with limitations in communication infrastructure, big systems, or privacy matters that obstruct the process of sharing information. Since the population structure determines the features that explain players behaviors, the population structure in the classic approach owns a complete and well-mixed structure, which means that players can choose any strategy with the same probability as the others. Figure 1a illustrates this concept with some players in a game. We use element shapes such as scissors, paper, or stone to represent the chosen strategies of each agent.



(a) (b)

Figure 1: (a) Population Structure Without Constrains, (b) Constrained Population Structure

Considering EGT, each player can equally obtain a revision opportunity. When receiving this opportunity, players choose arbitrarily one of their neighbors and switch their chosen strategy to one of their neighbors based on the selected revision protocol. As players are supposed to have a full and well-mixed structure, any opponent has the same possibility of selecting and playing any strategy of the structure (Figure 1a). On the contrary, Figure 1b shows a case where constraints in the structure limit the capacity of an agent to select some strategies, which is also an approach closer to reality. In this case, all agents are equally likely to be given a revision opportunity, but a neighbor does not have the same probability to choose and play a particular strategy. For instance, when a player obtains a revision opportunity with a paper strategy, there is no opportunity of choosing an opponent with scissors. The reason is that no papers are close to any scissors. Nevertheless, in this player case, the prospect of choosing an adversary with a paper or stone plan is higher than in the scissors situation. The graph $G = (T, L, M)$ establishes a mathematical way to represent the behavior of agents and their dynamics. The set T symbolizes the strategies an agent can choose. Set L is the meeting probability between strategies. For contextualizing, the notation $M = [a_{ij}], a_{ij} = 1$ suggests that strategy j and i can find each other, but $a_{ij} = 0$ indicates that these strategies cannot meet. Thence, it is possible to define Ni as the set of neighbors of agent i. Full and well-mixed and constrained mixed populations can be represented by two types of graphs. Figure 2a depicts a complete graph for the full and well-mixed structures, while Figure 2b illustrates the case with constraints in the structure. The form of the graph is determined by the particular structure of the population. In this research, undirected graphs are employed, which means that the probability that strategies j and i find each other are the same as in strategies i and j.



(a) (b)

Figure 2: Graphs topology for (**a**) full and well-mixed structure and (**b**) constrained structure.

Now, for regulating agents interactions, the limitation of incomplete information dependency in the population structure of Equation (17) must be overcome. For this purpose, the work proposed by [2] is considered. Therefore, to incorporate the neighboring concept, we use the pairwise proportional imitation protocol, as expressed in Equation (18):

$$p_{ij} = p_j[f_j(p_{Ni}) - f_i(p_{Ni})]_+ \tag{18}$$

where the calculation of p_i only requires knowing the portions of the population that are playing neighboring strategies. Then, the following expression is assumed:

Assumption 1 *Operations that update behaviors of agents by employing the pairwise proportional imitation protocol use the neighboring concept, which means that the iterations in the sums and the payoff function are determined by those neighbors communicating effectively with the i-th player.*

In this vein, Equation (19) denotes the obtained distributed replicator dynamics that fulfill the limitations of the population structure and enable agents to regulate the calculation of incomplete information:

$$\dot{x}_i = \alpha x_i \tau\Big[f_i(x_{Ni}) \sum_{j\in Ni}^{n} x_j - \sum_{j\in Ni}^{n} x_j f_j(x_{Nj})\Big] \tag{19}$$

where $f_{i/j}(x_{Ni/j})$ is the payoff function for the ith or jth player, estimated by the proportion of population that effectively communicates with neighbors, and $\sum_{j\in Ni}^{n} x_j$ is a sum that just considers those neighbors who communicate effectively. As our statement about the neighboring concept was implemented just in the first part of Equation (17), the second part of the equation (second parenthesis) including this concept is written as follows:

$$-\alpha x_i\Big[\ln x_i - \sum_{k\in Ni}^{n} x_k \ln x_k\Big] \tag{20}$$

In this equation, k represents i-th neighbor with an active communication link that employs strategy j. The end of the equation expresses the way the i-th player behaves regarding the proposed method using the Boltzmann concept. Equation (21) denotes in a complete manner the Boltzmann-Based Distributed Replicator

Dynamics (BBDRD) which includes both concepts: the distributed and the neighboring.

$$\dot{x}_i = \alpha x_i \tau \underbrace{\left[f_i(x_{Ni}) \sum_{j \in Ni}^{n} x_j - \sum_{j \in Ni}^{n} x_j f_j(x_{Nj}) \right]}_{Exploitation} - \alpha x_i \underbrace{\left[\ln x_i - \sum_{k \in Ni}^{n} x_k \ln x_k \right]}_{Exploration} \tag{21}$$

As stated above, the BBDRD equation evidences the implementation of the exploration and the exploitation notions of RL, and the selection-mutation approach of EGT, as explained in the next section. The implementation of this approach and examples of its application in the context of classic games, maximization problems, and for a smart grid control are developed in Section 5.

4 Evolutionary Approximation

This section presents the control method stated in Equation (21) from the perspective of RL and in an evolutionary approximation, which is helpful to comprehend the introduction of the notion of exploration in the classic RD expression.

4.1 Evolutionary Perspective

The traditional structure of RD is represented in the first part of the dynamics of Equation (21). This allows approximating to the Q-learner dynamics from EGT, because the mechanism for selection is contained in it. Then, the mechanism for mutation is found in the complementary part of the expression, which means:

$$x_i \alpha \left(\sum_{k \in Ni}^{n} x_k \ln(x_k) - \ln(x_i) \right) \tag{22}$$

In Equation (22), there are two recognizable entropy values: the distribution of probability x and the value of the strategy x_i. The expressions for entropy can be written as:

$$E_i = -x_i \ln(x_i) \tag{23}$$

and

$$E_n = -\sum_{k \in Ni}^{n} x_k \ln(x_k) \tag{24}$$

where E_i represents the available information regarding strategy i, while E_n is the information of the complete distribution. Consequently, the mutation equation can be expressed now as:

$$-(\alpha x_i E_n - \alpha E_i) \tag{25}$$

The following expression is the mutation equation derived, considering the difference between old and new states of x_i.

$$\sum_{k \in Ni}^{n} \epsilon_{ik} x_k - x_i \tag{26}$$

In Equation (26), ϵ_{ik} expresses the rate of mutation of agents that employ the i-th strategy and select another strategy from the pool of the k neighbors, for example, strategy j. When k is higher or equal to 1, ϵ_{ik} becomes bigger than or equal to zero. Considering EGT, in the framework of Q-Learning dynamics, mutation is directly connected with entropy that expresses the strategy state. However, this connection already existed, since it has been evidenced that entropy augments with mutation [36]. This connection is described in [37] from the perspective of thermodynamics, taking into account the trend of mutation to augment to increase entropy. Additionally, the Q-learning dynamics evidence that RD is the basis for the development of the selection concept. In RD, the resulting payoff can favor or be independent of a strategy, and the behavior of its opponent is strongly related to the resulting payoff. The concept of mutation can be found too. This fact is estimated by comparing the value of the entropy strategy with the value of entropy of the entire population.

4.2 Reinforcement Learning Perspective

Reinforcement learning aims to compensate the exploration and exploitation mechanisms. For gaining the maximum profit, a player must execute an action. Commonly, the player chooses actions that paid a high compensation before. Nevertheless, if the player wants to identify these actions, it must choose actions that were not chosen before. The notion from RL of exploitation-exploration is understood from a biological perspective by establishing connections between exploitation/exploration and mutation/selection. For clarity purposes, the first term of Equation (21) always chooses the best courses of actions, which matches the exploitation concept. Likewise, the exploration term is introduced into the RD expression due to its direct connection with the terms of entropy in Equation (22). Note that high values of entropy produce a high level of uncertainty in choosing one course of action. Therefore, the term of exploration augments entropy and gives diversity all at once. Consequently, the exploration and mutation concepts are strongly related, as both of them give variety, and a feature of heterogeneity to the environment. Being in control of particular scenarios like heterogeneity and communication limitations in a system is a demanding task when addressing real-life cases. In this sense, the compensation of the exploration-exploitation mechanisms can be quite difficult since a fine adjustment is often required for the parameters involved in the learning process. This adjustment must be performed to regulate the behavior of players in the process of decision-making. This problem can be solved by using the BBDRD control method as demonstrated in Equation (21) and explained in the following section.

5 Illustrative cases

5.1 Rock-Paper-Scissors as a classic game

In this part of the document, the concept introduced in Equation (21) is implemented in one of the classic games for excellence, the rock-paper-scissors game. For this purpose, a single population with three strategies has been considered, where $x = [x_1, x_2, x_3]^\top$ represent each of them respectively. In the same sense, the expression $F(x) = Ax$ denotes the fitness function, where A represents the payoff matrix showed in Equation (27). It is worth noting that the classic payoff matrix has been modified to guarantee positive values

of the payoffs in all cases.

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix} \tag{27}$$

To start running the simulation, a time of 30 units was considered with 300 agents and 5000 iterations. Additionally, the following initial conditions were stated $x_0 = [0.2, 0.7, 0.1]^\top$. The classic behavior of the rock-paper-scissors game proposes that every single strategy has the same probability to be selected, which means the absence of a dominant strategy. This behavior can be evidenced using Δ representation, which is defined as follows:

Definition 1 *Let Δ be the representation of a triangle of n-dimensions known as a Simplex.*

Using a simplex helps in the understanding of the implicit dynamics. Since the simplex is composed of three vertices, each of them represents a strategy e.g. rock, paper, scissors, then, the classic expected simulation of this situation is depicted as shown in Figure (3)a. Similarly, Figure (3)b, shows how the evolution of the population strategies is completely symmetrical, which means they keep constant along the time.

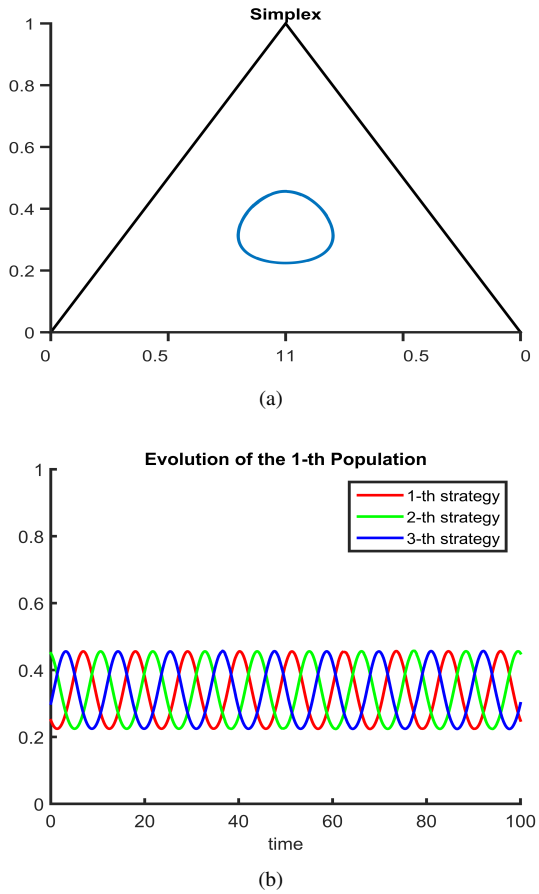


Figure 3: (a) Classic Rock-Paper-Scissors Behavior in a Simplex. (b) Evolution of the Population Behavior.

We also consider simulating a general distributed case to further compare it with the results of the BBDRD method. In both cases, the same simulation parameters were considered, but communication between agents was limited in the following way: agents playing strategy 1 were not allowed to communicate with agents playing strategy 3 and vice versa. Figure 4a shows the behavior of the distributed case, where the graphic seems to be an oval. This means that the interaction between strategies using only the first part of Equation (21) (general distributed case without entropy) tends to have a similar behavior to the one found in Figure 3a. Additionally, results depicted in Figure 4b show the evolution of the population strategies under the distributed case, where the symmetry is altered by the constraints in the communication of agents.

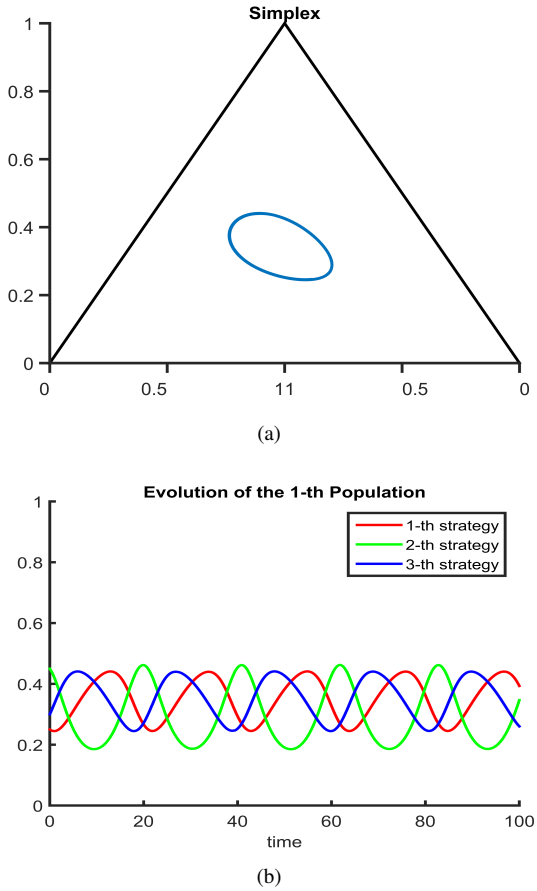
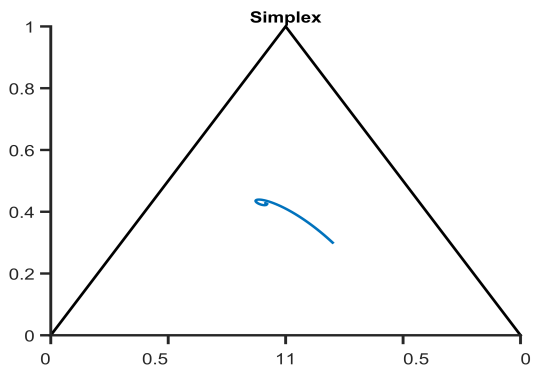
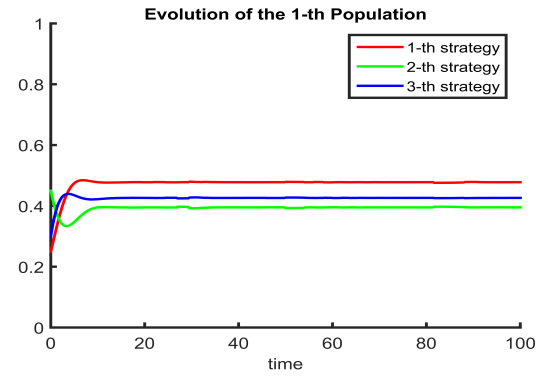


Figure 4: (a) Distributed Rock-Paper-Scissors Behavior in a Simplex. (b) Evolution of the Distributed Behavior.

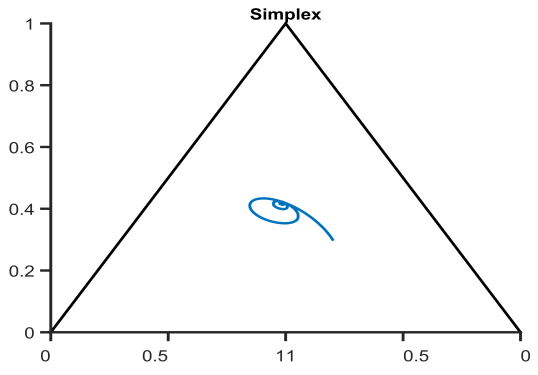
As mentioned previously, to compare these results with those obtained using the BBDRD method of Equation (21) (Distributed + Entropy case), Figure 5a shows that using τ= 1 the blue line depicts just one part of the oval (in contrast to 4a). Additionally, Figures 5c and 5d show the behavior of the model using τ values of 10 and 100 respectively. As evidenced, the bigger the term τ is, the more similar the behavior is to that obtained in the distributed case i.e. the contour of the oval seems to be equal to that obtained in Figure 4b), which at the same time is similar to the classic case. Finally, In Figure 5b, the evolution of the strategies population seems to be stable in all cases. This can be understood due to the introduction of the entropy term.



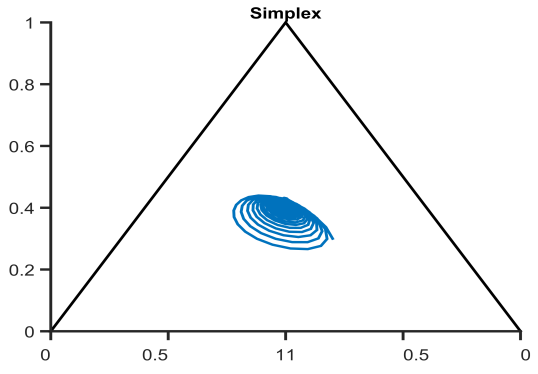
(a)



(b)



(c)



(d)

Figure 5: (a) Distributed + Entropy Rock-Paper-Scissors Behavior in a Simplex. (b) Evolution of the Population Behavior. (c) Simulation with τ =10. (d) Simulation with τ =100.

5.2 Solving maximization Problems

In this part of the document, we propose an application of the proposed method by understanding how it works under single and multi-population cases to solve maximization problems.

5.2.1 Single Population Case

This case considers a population where each agent can choose one of the $n+1$ strategies. In this case, the first n strategy corresponds to one variable of the objective function and the $n+1$ th strategy can be seen as a slack variable. Thus, x_k is the proportion of agents that use the *kth* strategy, and it corresponds to the *kth* variable, i.e., $x_k = z_k$. The fitness function of the *kth* strategy F_k is defined as the derivative of the objective function with respect to the *kth* variable, thus,

$$F_k(x) \equiv \frac{\partial}{\partial x_k} f(x)$$

Note that if $f(x)$ is a concave function, then its gradient is a decreasing function. As mentioned previously, users attempt to increase their fitness by adopting the most profitable strategy in the population, e.g. the *kth* strategy. This lead to an increase of x_k, which in turns decrease the fitness $F_k(x)$. Furthermore, the equilibrium is reached when all agents that belong to the same population have the same fitness. Thus, at equilibrium $F_i(x) = F_j(x)$, where $i, j \in \{1, \ldots, n\}$. If we define $F_{n+1}(x) = 0$, then, at equilibrium $F_i(x) = 0$ for every strategy $i \in \{1, \ldots, n\}$. Since the fitness function decreases with the action of users, it can be concluded that the strategy of the population evolves to make the gradient of the objective function equal to zero (or as close as possible). This resembles a gradient method to solve optimization problems. Recall that the evolution of the strategies lies in the simplex, that is, $\sum_{i \in S^p} z_i = m$, thence this implementation solves the following optimization problem:

$$\begin{aligned} \underset{z}{\text{maximize}} \quad & f(z) \\ \text{subject to} \quad & \sum_{i=1}^{n} z_i \leq m, \end{aligned} \tag{28}$$

where m is the total mass of the population.

Figure 6 shows an example of the setting described above for the function

$$f(z) = -(z_1 - 5)^2 - (z_2 - 5)^2. \tag{29}$$

Figure 6a shows the classic behavior to solve the maximization problem using a centralized approach. The simulation is executed during 0.6 time units. The black line founds the maximum with a very short deviation. Figure 6b depicts the case using a decentralized maximization approach. Once again, the maximum is reached but the deviation is bigger than the centralized approach. Finally, Figures 6c, 6d, 6e and 6f show the behavior of the Boltzmann-Based Distributed Replicator Dynamics, i.e. communication between agents is limited (Equation 21). In these cases, values of τ of 0.1, 0.5, 1, and 10 were used, respectively. Using the obtained model, it can be observed that as τ grows, the behavior of the simulation tends to be very similar to that of the centralized approach. Conversely, the shorter the τ value, the farther it is from the maximization point.

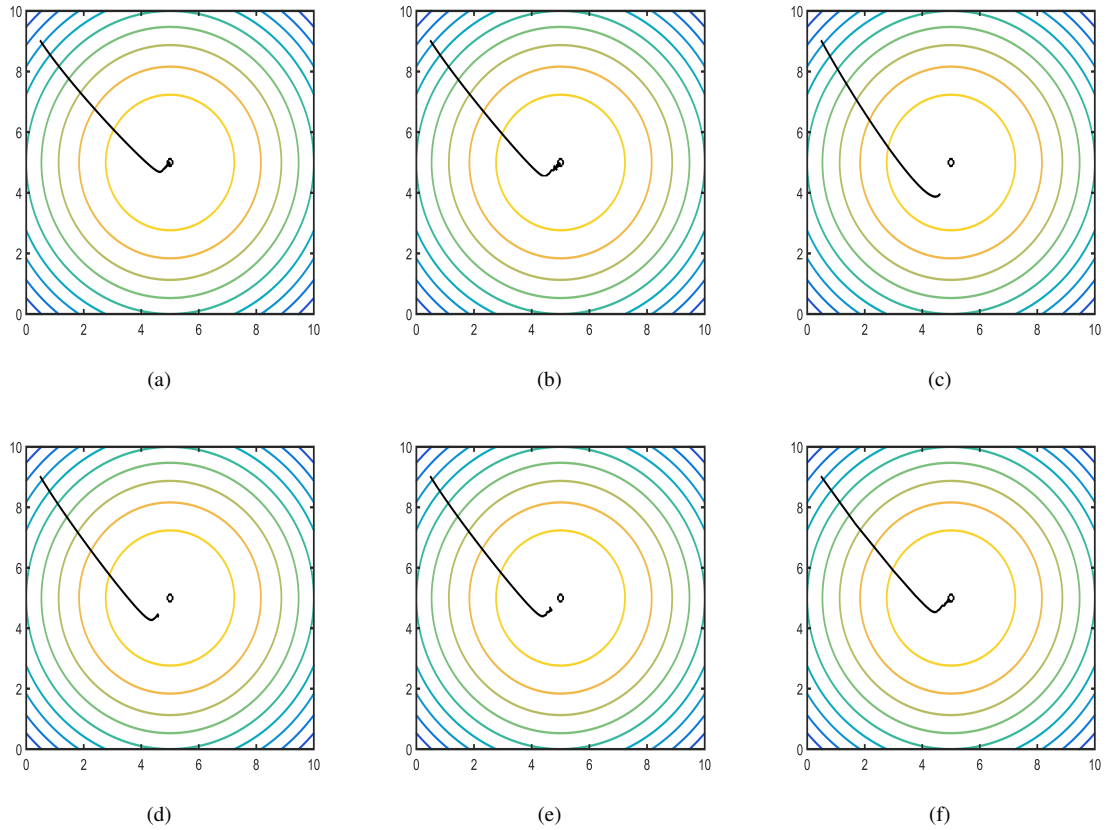


Figure 6: (a) Centralized Maximization Approach. (b) Decentralized Maximization Approach. (c) Distributed Maximization + Entropy Approach for τ =0.1. (d) Distributed Maximization + Entropy Approach for τ =0.5. (e) Distributed Maximization + Entropy Approach for τ =1. (f) Distributed Maximization + Entropy Approach for τ =10

5.2.2 *Multi Population Case*

Consider n populations where each agent can choose one out of two strategies. One population is defined per each variable of the maximization problem and also n additional strategies that resemble slack variables. Thus, x_i^p is the proportion of agents that use the *ith* strategy in the *pth* population. In this case x_1^k corresponds to the *kth* variable, that is, $x_1^k = z_k$, while x_2^k is a slack variable. The fitness function F_1^k of the *kth* population is defined as the derivative of the objective function with respect to the *kth* variable, that is, $F_1^k(x) \equiv \frac{\partial}{\partial x_1^k} f(x)$. Additionally, $F_2^k(x) = 0$. This implementation solves the following optimization problem:

$$\begin{aligned} &\underset{z}{\text{maximize}} && f(z) \\ &\text{subject to} && z_i \leq m^i, i = \{1, \ldots, n\}. \end{aligned} \tag{30}$$

Figure 7a shows the way the system gets to the maximum point using the centralized approach. Using a multi-population, the plotted line is made almost without deviations. Similarly, Figure 7b depicts the result for the classical distributed approach, where the multiple populations reach the maximum, but the following form has some deviations before reaching it. Figures 7c, 7d, 7e and 7f show the behavior of the Extended Distributed Replicator Dynamics (see full model of Equation (21)). In these cases, values of τ of 0.1, 0.5, 1, and 10 were used respectively. Results show once again, that using the Boltzmann-Based Distributed Replicator Dynamics method evidences that as τ grows, the behavior of the simulation tends to be very similar to that of the centralized approach (where full information is assumed within agents). Conversely, the shorter the φ value, the farther and the more deviant it is from the maximization point.

5.3 *Smart Grids Application*

This part of the paper presents how the use of the Boltzmann-based distributed replicator dynamics can be developed in a power grid. Some of the main issues to solve in these kinds of applications are cases of the economic dispatch problem (EDP). In these problems, first, it is necessary to reduce the global value of the power generation and, second, to maximize the overall effectiveness of the power generators, thus fulfilling the limitations of generation capacity and power balance simultaneously [38]. In this sense, traditional approaches to EDP have employed offline direct-search methods [21, 15], or static optimization algorithms [20]. One of the first works that introduced a different approach to deal with EDP is [39], where the authors proposed changing the resource allocation as a solution to this issue. Our work takes into account this approach and complements it with the introduction of the Boltzmann-based distributed replicator dynamics as a way to find the place to execute the dispatch algorithm at a microgrid, by using distributed population dynamics. Our work also assumes that loads, generators, and other devices in the grid share information in the system and have

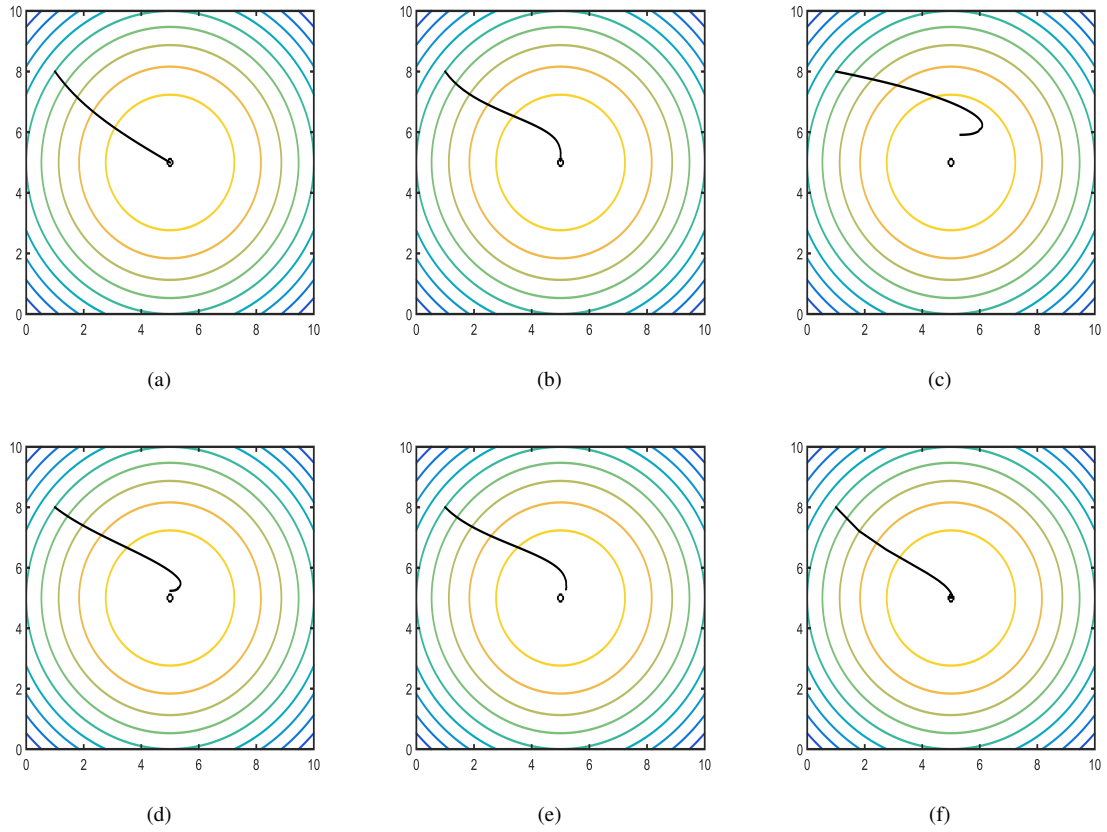


Figure 7: (a) Centralized Maximization Approach. (b) Decentralized Maximization Approach. (c) Distributed Maximization + Entropy Approach for τ =0.1. (d) Distributed Maximization + Entropy Approach for τ =0.5. (e) Distributed Maximization + Entropy Approach for τ =1. (f) Distributed Maximization + Entropy Approach for τ =10

a cooperative role with other controllable devices in the grid. The general case of the microgrid is explained in [40], where authors formulate a grid with two different control levels. At the lowest level, an inverter attaches loads to a source of voltage comprised of seven distributed generators (DGs). The output voltage and the operation frequency are controlled by a drop-gain regulator. Figure 8 depicts the distribution of the microgrid.

The uppermost level employs a strategy that can dynamically dispatch setpoints of power. The economic limitations, like load demands and power production costs, come from the inferior level of control and are directed to the central controller of the microgrid. Therefore, a classic RD is implemented. The controller obtains dynamic values of load demands and costs, which means that it is possible to include renewable energy resources. As a result, the dispatch is carried out , that is, the uppermost control level. The expression of the EDP is written as follows:

$$\begin{aligned} &\text{maximize} \quad J(\varphi) = \sum_{i=i}^{n} J_i(\tau_i), \\ &\text{subject to} \sum_{i=1}^{n} \varphi_i = \sum_{i=1}^{n} \psi_i = \varphi_D \end{aligned} \tag{31}$$

In Equation (31), $0 \leq \varphi_i \leq \varphi_{\max i}, \forall i \in \mathbb{Z}$, n represents the quantity of distributed generators, φ_i denotes the the i-th DG set-point of power, ψ_i simbolizes the loads, φ_D represents the total load that the grid requires, $\varphi_{\max}$ establishes the i-th DG maximum capacity of generation, and $J_i(\varphi_i)$ represents the utility function of every DG. The criterion of the economic dispatch determines the utility function [38], which in turn settles the performance of all the generation units with the same marginal utilities stated in Equation (32)

$$\frac{dJ_1}{d\varphi_1} = \frac{dJ_2}{d\varphi_2} = ... = \frac{dJ_n}{d\varphi_n} = \delta, \tag{32}$$

Consider $\delta > 0$, so that $\sum_{i=1}^{n} \varphi_i = \varphi_D$. According to the EDP criterion expressed in Equation (32), it is possible that the EDP of Equation (31) obtain a solution by employing utility functions with quadratic form for every DG [39].

5.3.1 *The Economic Dispatch Problem Using a Population Games Perspective*

From the Population Games Perspective, the EDP can be managed using the Replicator Dynamics approach. For the simulation purposes, we limit the communications constraints among agents at random, which allows us to have another point of view to compare results with those obtained in [1]. Using the population games approach, n represents the quantity of DGs in the grid. Consider the selection of a DG as the i-th strategy, then, φ_i would be the amount of power allocated to each DG, which is associated with the number of players that choose the i-th strategy in S. The term φ_D represents the sum of every power set-point, which means $\sum_{i=1}^{n} \varphi_i = \varphi_D$ to obtain an appropriate steady-state performance. Likewise, to accomplish

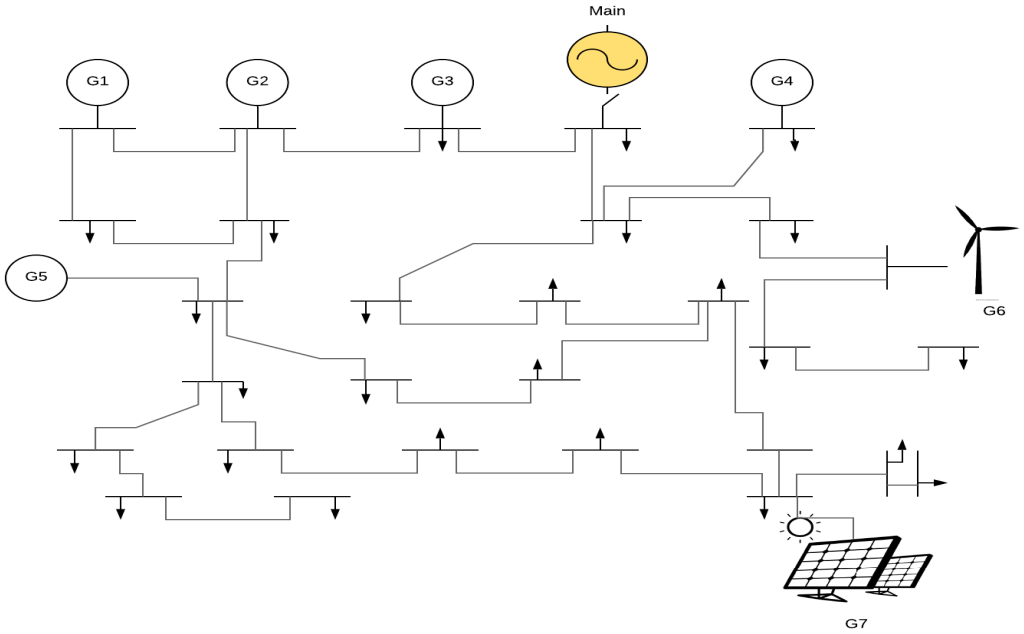


Figure 8: Distribution of a Power Grid. Adapted from [11].

the power balance, $\bar{f} = (1/\varphi_D) \sum_{i=1}^{n} \varphi_i f_i$ must be implemented to enable the invariance of set $\Delta = [\varphi \epsilon^t \geq 0 : \sum_{i \epsilon S} \varphi i = m]$ [41]. This equation ensures that in case $\varphi(0) \epsilon \Delta$, then $\varphi(t) \epsilon \Delta, \forall t \geq 0$, that is, the strategy of control have to determine set-points to guarantee the correct equilibrium between the demanded and generated power by generators. This behavior makes it possible to perform a proper control of frequency. To include in the control strategy economic and technical criteria, the capacity of power generation and the associated cost are relevant factors for determining the final power dispatched to every DG. RD seems to be suitable, since its stationary state is achieved once the average outcome equals all the outcome functions. This characteristic relates RD to EDP, as it is the same as the economic dispatch criterion of Equation (32) when the outcome function is chosen as follows:

$$f_i(\varphi_i) = \frac{dJ_i}{d\varphi_i}, \forall i = 1, 2, ..., n, \tag{33}$$

It is worth noting that the EDP approach in Equation (32) ensures an optimal solution of the system if constraints are satisfied. These kinds of optimization issues may be tackled using marginal utilities for the outcome functions. This is possible because the outcome functions are equal to $\bar{f}$. The outcome chosen can be modeled as an expression whose growth/reduction depends on the distance of the desired set-point from/to the power. In this vein, RD allocate resources to generators according to the average result. The following function [42] can illustrate this phenomenon:

$$f(\varpi) = r\varpi\left(1 - \frac{\varpi}{k}\right) \tag{34}$$

where k represents the carrying capacity so that the independent variable $\varpi \epsilon (0, K)$. Here, parameters such as the carrying capacity and a cost factor of generation, among other parameters, are used by the outcome functions. As a result, the outcome function of each DG can be expressed as:

$$f_i(\varphi_i) = \frac{dJ_i}{d\varphi_i} = \frac{2}{ci}\left(1 - \frac{\varphi i}{\varphi_{max}}\right), \forall i = 1, 2, ..., n, \tag{35}$$

The population game can transform into a potential game by the addition of marginal utilities to the outcome functions [43]. The outcome functions in Equation (35) become functions of quadratic utility for every DG in the optimal EDP [39]. This outcome function has been implemented in other research, e.g. [39, 40, 44].

$$J_i(\varphi_i) = \frac{1}{ci}\left(2\varphi_i - \frac{\varphi i^2}{\varphi_{maxi}}\right), \quad \forall i = 1, 2, ..., n, \tag{36}$$

5.4 Simulation Results

The BBDRD control model presented in this paper is validated in a study case that considered a low voltage smart grid comprising seven DGs. The system used $\varphi_D = 9$ kW as the overall power demand in the network; DG 4 had the lowest cost and DG 7 the highest. DGs 1, 3, 5, and 6 had no significant differences in cost, and DG 2 had their lowest cost. The system employs 60 Hz and a nominal capacity of 3.6 kW for all generators, except for DGs 2 and 6 that employ 1.5 KW and 4 KW, respectively. Note that these initial conditions differ from those used in [1], where DG 3 had the lowest cost and the nominal capacity of DGs 2 and DG 6 was 3.6 KW and 2 KW, respectively.

For comparison purposes, first, the classic centralized case was simulated, taking into account the availability of full information. Figure 9a shows the results of this step. There is an unexpected rise in the load of 3 KW and various values for each generator. The frequency was stable, except for t = 0.8, where there is a variation of approximately 0.2 Hz produced by an increase in the load,

however, it returns to stability right after it. Figure 9a also depicts the quantity of power delivered to each DG. First, generator DG 7 transmits a minimum power when there is low demand, owing to its costly behavior. On the contrary, DG 3 approximates to its maximum capacity and remains near this value without affectations by changes in the load. In case that the demand augments, DG 7 augments its capacity too, intending to counter-weigh the demand. DG 6 approximated to its maximum capacity just after changes in the load. DGs 1, 4, 5, and 6 evidence a comparable behavior, since they present analogous conditions. Finally, DG 2 approximates to its maximum performance thanks to its low-cost performance. With the results of the classic centralized case, the BBDRD control method was employed to contrast its behavior. Figure 9b–d present the outcomes of this step for different values of τ. For simulation purposes, we use constraints in the communication of agents at random. When the τ value augments, the system imitates the centralized approach. Low values of τ (Figure 9b) produced the biggest differences, as DGs need more time to achieve their working level. On the contrary, high values of τ (Figure 9d) allow DGs to achieve their working levels faster. Concerning the results obtained in [1], we observe a similar behavior of the microgrid. Despite using limitations in the communications of the generators at random, after employing the BBDRD method with high values of τ, once more the behavior tends to be equal to that of the centralized approach of the classic RD. The main difference evidenced is the behavior of DG4 in comparison with the centralized approach, that is, when the demand augments, it delivers more power as a result of the communication limitations topology and the effect of the exploration concept (second term in Equation (21). This effect was also evidenced in [1] with the behavior of DG5.

6 Conclusions

The Boltzmann-based distributed replicator dynamics shown in Equation (21) might be defined as a learning method of distributed control that includes the exploration scheme from RL in the classic equation of RD. In this sense, exploration can be related to the mutation concept of EGT, and involves a method for measuring variety in the system with the entropy approach. The Boltzmann-based distributed replicator dynamics also employs the scheme of the Boltzmann distribution to include the τ parameter for controlling purposes. An appropriate temperature function can be chosen using methodological search and reliably set to fulfill an anticipated convergence distribution. Regarding stability, Section 3 presents a derivation process that has low or no significant variations in the presence of multiple agents. This behavior is explained with the inclusion of the population approach in the BBDRD method. The neighboring approach provides the missing piece to prevent centralized schemes from happening and compels players to consider just the available information of other players before performing an action. The method was validated in the context of classic games, maximization problems, and in a smart grid that allowed initializing parameters beforehand, and providing evidence that behavior using the BBDRD approach tends to be similar to cases using centralized schemes.

Engineering problems represent real scenarios whose complexity can be simulated using MAS, through the analysis of the communication between the agents. EGT presents some helpful tools to tackle communication between players and control them. This paper evidences the advantages of applying a distributed control approach of EGT to a real-life smart grid. The BBDRD performance is presented using experiments that include limitations in communication, therefore, it emerges as a helpful tool for developing more realistic control strategies in Engineering problems with distributed schemes. This advantage becomes particularly relevant because it offers the opportunity to deal with complex systems using local information of the agents, taking into account communication limitations without the need of a centralized coordinator and evading expensive implementation costs, as in classic approaches, like the dual decomposition method. The distributed control concept proposed to tackle cases of classic games, maximization issues, and the Economic Dispatch Problem can be further applied to other real-life situations, including some other problems in the smart grid context, like as the physical limits of power-flow, the presence of power losses, and the inconsistency of renewable generation, among others.

Despite using incomplete information, results demonstrated that the system can imitate the performance of a centralized approach when the τ value increases. conversely, when τ takes values lower than the unit, the behavior was distant from outcomes obtained under the optimal communication scenario of a centralized approach. The possibility of adjusting the behavior and parameters of the method using communications limitations between players proved to be successful. This can also logically be extended to any number of players or populations. Results also evidenced that the Boltzmann-based distributed method has adequate performance for solving some cases of maximization problems, including the economic dispatch problem in a smart grid. This is possible since the features of the DGs were coherent with their power capacity and operation cost.

7 Future Work

For future work, the optimization of wireless sensor networks can be an option for the building automation field. Various critical issues may be tackled by implementing the distributed replicator dynamics approach to solve the EDP in a smart grid scenario, for example, considering power losses, the limitations of physical power-flow, or the uncertainty of renewable generation. Open issues should be considered with respect to the control strategies developed that must include decentralization, scalability, and robustness. In consequence, novel methods should incorporate economic incentives and the information necessary to ensure that more elements can be included in the system without a reconfiguration of the whole system. The Boltzmann-based distributed concept that tackles EDP will be expanded to other problems in the context of a smart grid framework, for example, the inconsistency in renewable generation, physical limitations of power-flow, and incorporation of power loss. In summary, distributed techniques used to manage open problems represent a suitable option for modeling the complexity of these scenarios. Innovative approaches are still required to include scalable solutions and features closer to reality.

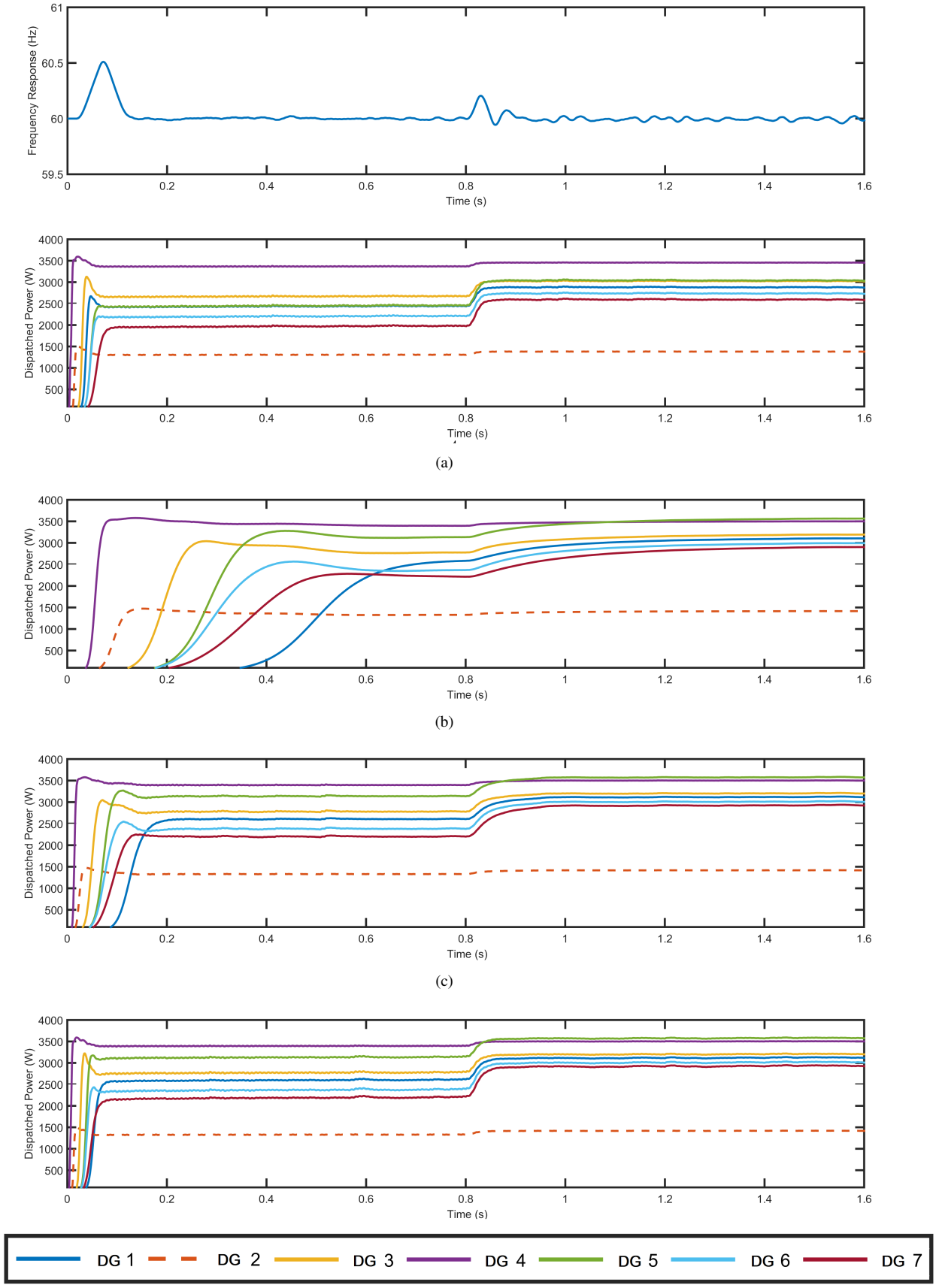


Figure 9: Results for a microgrid system. (**a**) Frequency response and active power response of DGs for the classic RD. The analysis of the performance of the Boltzmann-based distributed replicator dynamics for different values of τ: (**b**) $\tau = 0.4$ (**c**) $\tau = 2.5$, (**d**) $\tau = 7$.

Conflict of Interest The authors declare no conflict of interest.

Funding This research was funded by Colciencias, grant Doctorado Nacional number 727.

Acknowledgment We express our gratitude to Colciencias for the founding of this Project. We thank Universidad Nacional de Colombia and Universidad Santo Tomás for allowing us to use laboratories, hardware and software for the development of this work.

Annex A

In this part of the document, we reconstruct the full process of derivation necessary to have a continuous-time limit for the model of Q-learning, where the Q-values are considered as Boltzmann probabilities for action-selection mechanisms. For clarity purposes in the construction of the learning model, this analysis starts considering an extended version of the equations obtained in [35], where dynamics for the Q-learners in two-players games were defined.

To find the relationship between the Q-learning framework and the RD equations, the use of the Equation (1) that describes the Boltzmann probabilities is done.

$$x_i(\delta) = \frac{e^{\tau Q_{a_i}(\delta)}}{\sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta)}} \tag{1}$$

Here, $x_i(\delta)$ represents the prospect of selecting the i strategy at δ step time, and τ symbolizes the temperature. From the Boltzmann distribution, it is easy to find the expression for $x_i(\delta+1)$ as follows:

$$x_i(\delta+1) = \frac{e^{\tau Q_{a_i}(\delta+1)}}{\sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta+1)}}$$

now dividing $x_i(\delta+1)$ into $x_i(\delta)$:

$$\frac{x_i(\delta+1)}{x_i(\delta)} = \frac{e^{\tau Q_{a_i}(\delta+1)} \sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta)}}{e^{\tau Q_{a_i}(\delta)} \sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta+1)}}$$

after organizing terms it gets to:

$$\frac{x_i(\delta+1)}{x_i(\delta)} = \frac{e^{\tau Q_{a_j}(\delta+1)} e^{-\tau Q_{a_i}(\delta)}}{\sum_{j=1}^{n} e^{\tau Q_{a_j}(\delta+1)} \sum_{j=1}^{n} e^{-\tau Q_{a_j}(\delta)}}$$

Then, using Δ to denote a small difference between operations it takes the following form:

$$\frac{x_i(\delta+1)}{x_i(\delta)} = \frac{e^{\tau \Delta Q_{a_i}(\delta)}}{\sum_{j=1}^{n} e^{\tau \Delta Q_{a_j}(\delta)}}$$

This result can be rewritten in the following way:

$$x_i(\delta+1) = x_i(\delta) \frac{e^{\tau \Delta Q_{a_i}(\delta)}}{\sum_{j=1}^{n} x_j e^{\tau \Delta Q_{a_j}(\delta)}}$$

Now, considering the difference equation for x_i:

$$\begin{aligned} x_i(\delta+1) - x_i(\delta) &= \frac{x_i(\delta) e^{\tau \Delta Q_{a_i}(\delta)}}{\sum_{j=1}^{n} x_j(\delta) e^{\tau \Delta Q_{a_j}(\delta)}} - x_i(\delta) \\ &= x_i(\delta) \Big[\frac{e^{\tau \Delta Q_{a_i}(\delta)} - \sum_{j=1}^{n} x_j(\delta) e^{\tau \Delta Q_{a_j}(\delta)}}{\sum_{j=1}^{n} x_j(\delta) e^{\tau \Delta Q_{a_j}(\delta)}} \Big] \end{aligned}$$

At this point, to describe the continuous time version, it is assumed that σ, with $0 < \sigma \leq 1$, describes the time amount spent between game repetitions. In the case of $x_i(\delta\sigma)$, it represents the x-values at time $k\sigma = t$. Under these premises, the expression takes the following form:

$$\begin{aligned} \frac{x_i(\delta\sigma+\sigma) - x_i(\delta\sigma)}{\sigma} =& \Big[\frac{x_i(\delta\sigma)}{\sigma \sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}} \Big] * \\ & \Big[e^{\tau \Delta Q_{a_j}(\delta\sigma)} - \sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)} \Big] \end{aligned}$$

Nevertheless, the main interest is finding the limit of $x_i(\delta\sigma)$, given $\sigma \to 0$, $\delta\sigma \to t$ and $t \geq 0$, then:

$$\begin{aligned} \lim_{\sigma\to\infty} \frac{\Delta x_i(\delta\sigma)}{\sigma} = \lim_{\sigma\to 0} \Big[& \Big(\frac{x_i(\delta\sigma)}{\sigma \sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}} \Big) * \\ & \Big(e^{\tau \Delta Q_{a_j}(\delta\sigma)} - \sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)} \Big) \Big] \end{aligned}$$

This expression can be rewritten as follows:

$$\begin{aligned} \lim_{\sigma\to\infty} \frac{\Delta x_i(\delta\sigma)}{\sigma} = & \lim_{\sigma\to 0} \Big[\frac{x_i(\delta\sigma)}{\sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}} \Big] * \\ & \lim_{\sigma\to 0} \Big[\frac{e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} - \frac{\sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} \Big] \end{aligned}$$

In the first limit, the expression $e^{\tau \Delta Q_{a_j}(\delta\sigma)}$ is equal to 0, and the summation becomes 1 because it is referred to the sum of all the probabilities. This means that the first limit becomes x_i.

$$\lim_{\sigma\to\infty} \frac{\Delta x_i(\delta\sigma)}{\sigma} = x_i * \underbrace{\lim_{\sigma\to 0} \Big[\frac{e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} - \frac{\sum_{j=1}^{n} x_j(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} \Big]}_{L2}$$

In the second limit, an undefined situation is presented; the numerator and denominator become zero, therefore, after using l'hopital rule, this limit equals (for short *L2*):

$$L2 = \lim_{\sigma\to\infty} \Big[\frac{\tau \Delta Q_{a_j}(\delta\sigma) e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} \Big] - \sum_{j=1}^{n} x_j(\delta\sigma) * \lim_{\sigma\to\infty} \Big[\tau \Delta Q_{a_j}(\delta\sigma) \frac{e^{\tau \Delta Q_{a_j}(\delta\sigma)}}{\sigma} \Big]$$

Which allow finding the following expression:

$$L2 = \tau \frac{dQ_{a_i}(t)}{dt} - \sum_{j=1}^{n} x_j(t) \frac{dQ_{a_j}(t)}{dt}$$

Now, it is possible to find the total limit, that is, the Q-Learning continuous time model derived as shown in Equation (2):

$$\frac{\frac{dx_i}{d_t}}{x_i} = \tau \Big[\frac{dQ_{a_i}}{dt} - \sum_{j=1}^{n} \frac{dQ_{a_j}}{dt} x_j \Big] \tag{2}$$

To solve the expression $\frac{dQ_{a_i(t)}}{dt}$, the first player takes the following update rule:

$$Q_{a_i}(\delta+1) = Q_{a_i}(\delta) + \alpha\Big[\Gamma_{a_i}(\delta+1) + \gamma \max_{ai} Q - Q_{a_i}(\delta)\Big]$$

Therefore, the last expression represents the equation of difference for the Q-function and can be rewritten as follows:

$$\Delta Q_{a_i}(\delta) = \alpha\Big[\Gamma_{a_i}(\delta+1) + \gamma \max_{ai} Q - Q_{a_i}(\delta)\Big] \tag{3}$$

if Equation (3) takes an infinitesimal scheme, it is supposed that the amount of time spent performing two update iterations of the Q-values is given by σ with $0 < \sigma \leq 1$. Additionally, $Q_{a_i}(\delta\sigma)$ symbolizes the Q-values at time $\delta\sigma$. Applying these assumptions, Equation (3) gets to:

$$\Delta Q_{a_i}(\delta\sigma) = \Big[\alpha(\Gamma_{a_i}((\delta+1)\sigma) + \gamma \max_{ai} Q - Q_{a_i}(\delta\sigma)\Big] * \Big[(\delta+1)\sigma - \delta\sigma\Big]$$

which is equal to:

$$\Delta Q_{a_i}(\delta\sigma) = \alpha\sigma\Big[\Gamma_{a_i}((\delta+1)\sigma) + \gamma \max_{ai} Q - Q_{a_i}(\delta\sigma)\Big]$$

Once again, the limit $\sigma \to 0$ is the state sought. Taking the limit of $Q_{a_i}(\delta\sigma)$, it gets to Equation (4):

$$\frac{dQ_{a_i}}{dt} = \alpha\Big[\Gamma_{a_i} + \gamma \max_{ai} Q - Q_{a_i}\Big] \tag{4}$$

Now, substituting Equation (4) on Equation (2):

$$\frac{\frac{dx_i}{dt}}{x_i} = \tau\Big[\alpha\Gamma_{a_i} + \alpha\gamma \max_{ai} Q - \alpha Q_{a_i} - \sum_{j}^{n} x_j \alpha(\Gamma_{a_j} + \gamma \max_{ai} Q_{a_i} - Q_{a_j})\Big]$$

$$= \tau\alpha\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j} - Q_{a_i} + \sum_{j=1}^{n} Q_{a_j}x_j\Big]$$

Taking into account that $\sum_j^n x_j = 1$ and using $\dot{x}_i$ to denote $\frac{dx_i}{dt}$, it is obtained:

$$\frac{\dot{x}_i}{x_i} = \tau\alpha\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j} - Q_{a_i}\sum_{j=1}^{n} x_j + \sum_{j=1}^{n} Q_{a_j}x_j\Big]$$

$$\frac{\dot{x}_i}{x_i} = \tau\alpha\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j} + \sum_{j=1}^{n} x_j(Q_{a_j} - Q_{a_i})\Big]$$

since $\frac{x_j}{x_i}$ equals $\frac{e^{\tau}\Delta Q_{a_j}}{e^{\tau}\Delta Q_{a_i}}$, the second part of the last expression can be expressed in logarithm terms:

$$\alpha\Big[\sum_{j=1}^{n} x_j ln(\frac{x_j}{x_i})\Big] = \alpha\Big[\tau\sum_{j} x_j(Q_{a_j} - Q_{a_i})\Big]$$

After reorganizing and substituting, the result is:

$$\frac{\dot{x}_i}{x_i} = \alpha\tau\Big[\Gamma_{a_i} - \sum_{j=1}^{n} x_j\Gamma_{a_j}\Big] + \alpha\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big]$$

To bring the concept of the payoff matrices into a 2 x 2 game, it can be expressed r_{a_i} as $\sum_j a_{ij}y_j$, thus obtaining the Equation (5) which represents the behavior for the first player as follows:

$$\dot{x}_i = x_i\alpha\tau\Big[(A\mathbf{y})_i - \mathbf{x}\cdot A\mathbf{y}\Big] + x_i\alpha\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big] \tag{5}$$

Similarly, for the second player, the expression is:

$$\dot{y}_i = y_i\alpha\tau\Big[(B\mathbf{x})_i - \mathbf{y}\cdot B\mathbf{x}\Big] + y_i\alpha\Big[\sum_{j=1}^{n} y_j \ln\left(\frac{y_j}{y_i}\right)\Big] \tag{6}$$

Since the approach of the classic RD can be used at this point, Equation (5) may be stated as shown in the following expression [27]:

$$\dot{x}_i = \alpha x_i\tau\Big[f_i(x) - \overline{f}(x)\Big] + \alpha x_i\Big[\sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{x_i}\right)\Big]$$

It should be noted that depending on the value obtained from the fitness of a specific type of population, this value may increase or decrease depending on the average value obtained by the entire population.

References

[1] G. Chica, E. Mojica, E. Cadena, "Boltzmann-Based Distributed Replicator Dynamics: A Smart Grid Application," in 2020 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI), 1–6, IEEE, 2020, doi: 10.1109/CONIITI51147.2020.9240335.

[2] J. Barreiro-Gomez, G. Obando, N. Quijano, "Distributed population dynamics: Optimization and control applications," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **47**(2), 304–314, 2016, doi:10.1109/TSMC.2016.2523934.

[3] G. Chica-Pedraza, E. Mojica-Nava, E. Cadena-Muñoz, "Boltzmann Distributed Replicator Dynamics: Population Games in a Microgrid Context," Games, **12**(1), 1–1, 2021, doi:10.3390/g12010008.

[4] G. Bacci, S. Lasaulce, W. Saad, L. Sanguinetti, "Game theory for networks: A tutorial on game-theoretic tools for emerging signal processing applications," IEEE Signal Processing Magazine, **33**(1), 94–119, 2015, doi: 10.1109/MSP.2015.2451994.

[5] C. Mu, K. Wang, "Approximate-optimal control algorithm for constrained zero-sum differential games through event-triggering mechanism," Nonlinear Dynamics, **95**(4), 2639–2657, 2019, doi:10.1007/s11071-018-4713-0.

[6] M. Zhu, E. Frazzoli, "Distributed robust adaptive equilibrium computation for generalized convex games," Automatica, **63**, 82–91, 2016, doi: 10.1016/j.automatica.2015.10.012.

[7] S. Najeh, A. Bouallegue, "Distributed vs centralized game theory-based mode selection and power control for D2D communications," Physical Communication, **38**, 100962, 2020, doi:https://doi.org/10.1016/j.phycom.2019.100962.

[8] R. Tang, S. Wang, H. Li, "Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids," Applied Energy, **250**, 118–130, 2019, doi:10.1016/j.apenergy.2019.04.177.

[9] K. Främling, "Decision theory meets explainable ai," in International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems, 57–74, Springer, 2020, doi:https://link.springer.com/chapter/10.1007/978-3-030-51924-7_4.

[10] A. Navon, G. Ben Yosef, R. Machlev, S. Shapira, N. Roy Chowdhury, J. Belikov, A. Orda, Y. Levron, "Applications of Game Theory to Design and Operation of Modern Power Systems: A Comprehensive Review," Energies, **13**(15), 3982, 2020, doi:10.3390/en13153982.

[11] N. Quijano, C. Ocampo-Martinez, J. Barreiro-Gomez, G. Obando, A. Pantoja, E. Mojica-Nava, "The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory," IEEE Control Systems Magazine, **37**(1), 70–97, 2017, doi: 10.1109/MCS.2016.2621479.

[12] M. Lanctot, E. Lockhart, J.-B. Lespiau, V. Zambaldi, S. Upadhyay, J. Pérolat, S. Srinivasan, F. Timbers, K. Tuyls, S. Omidshafiei, et al., "OpenSpiel: A framework for reinforcement learning in games," arXiv preprint arXiv:1908.09453, 2019.

[13] W. H. Sandholm, Population games and evolutionary dynamics, MIT press, 2010.

[14] L. Hindersin, B. Wu, A. Traulsen, J. García, "Computation and simulation of evolutionary Game Dynamics in Finite populations," Scientific reports, **9**(1), 1–21, 2019, doi:https://doi.org/10.1038/s41598-019-43102-z.

[15] D. P. Palomar, M. Chiang, "A tutorial on decomposition methods for network utility maximization," IEEE Journal on Selected Areas in Communications, **24**(8), 1439–1451, 2006, doi:10.1109/JSAC.2006.879350.

[16] J. R. Marden, "State based potential games," Automatica, **48**(12), 3075–3088, 2012, doi:10.1016/j.automatica.2012.08.037.

[17] L. Zhao, J. Wang, J. Liu, N. Kato, "Optimal edge resource allocation in IoT-based smart cities," IEEE Network, **33**(2), 30–35, 2019, doi:10.1109/MNET.2019.1800221.

[18] A. Cagnano, E. De Tuglie, P. Mancarella, "Microgrids: Overview and guidelines for practical implementations and operation," Applied Energy, **258**, 114039, 2020, doi:10.1016/j.apenergy.2019.114039.

[19] J. P. Lopes, C. Moreira, A. Madureira, "Defining control strategies for microgrids islanded operation," IEEE Transactions on power systems, **21**(2), 916–924, 2006, doi:10.1109/TPWRS.2006.873018.

[20] T. Ibaraki, N. Katoh, Resource allocation problems: algorithmic approaches, MIT press, 1988.

[21] S.-J. Ahn, S.-I. Moon, "Economic scheduling of distributed generators in a microgrid considering various constraints," in 2009 IEEE Power & Energy Society General Meeting, 1–6, IEEE, 2009, doi:10.1109/PES.2009.5275938.

[22] G. Strbac, "Demand side management: Benefits and challenges," Energy policy, **36**(12), 4419–4426, 2008, doi:10.1016/j.enpol.2008.09.030.

[23] D. E. Olivares, C. A. Cañizares, M. Kazerani, "A centralized optimal energy management system for microgrids," in 2011 IEEE Power and Energy Society General Meeting, 1–6, IEEE, 2011, doi:10.1109/PES.2011.6039527.

[24] P. Quintana-Barcia, T. Dragicevic, J. Garcia, J. Ribas, J. M. Guerrero, "A distributed control strategy for islanded single-phase microgrids with hybrid energy storage systems based on power line signaling," Energies, **12**(1), 85, 2019, doi:10.3390/en12010085.

[25] B. Huang, L. Liu, H. Zhang, Y. Li, Q. Sun, "Distributed optimal economic dispatch for microgrids considering communication delays," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **49**(8), 1634–1642, 2019, doi:10.1109/TSMC.2019.2900722.

[26] J. C. Vasquez, J. M. Guerrero, J. Miret, M. Castilla, L. G. De Vicuna, "Hierarchical control of intelligent microgrids," IEEE Industrial Electronics Magazine, **4**(4), 23–29, 2010, doi:10.1109/MIE.2010.938720.

[27] D. Bloembergen, K. Tuyls, D. Hennes, M. Kaisers, "Evolutionary dynamics of multi-agent learning: A survey," Journal of Artificial Intelligence Research, **53**, 659–697, 2015, doi:10.1613/jair.4818.

[28] H. Peters, Game theory: A Multi-leveled approach, Springer, 2015.

[29] R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction, MIT press, 2018.

[30] W. Ertel, "Reinforcement Learning," in Introduction to Artificial Intelligence, 289–311, Springer, 2017.

[31] F. L. Da Silva, A. H. R. Costa, "A survey on transfer learning for multiagent reinforcement learning systems," Journal of Artificial Intelligence Research, **64**, 645–703, 2019, doi:10.1613/jair.1.11396.

[32] T. Başar, G. Zaccour, Handbook of Dynamic Game Theory, Springer, 2018.

[33] J. Newton, "Evolutionary game theory: A renaissance," Games, **9**(2), 31, 2018, doi:10.3390/g9020031.

[34] J. W. Weibull, Evolutionary game theory, MIT press, 1997.

[35] K. Tuyls, K. Verbeeck, T. Lenaerts, "A selection-mutation model for q-learning in multi-agent systems," in Proceedings of the second international joint conference on Autonomous agents and multiagent systems, 693–700, 2003, doi: 10.1145/860575.860687.

[36] A. E. Eiben, J. E. Smith, Introduction to Evolutionary Computing, Springer, 2015.

[37] D. Stauffer, "Life, love and death: Models of biological reproduction and aging," Institute for Theoretical physics, Köln, Euroland, 1999.

[38] W. Aj, B. Wollenberg, "Power generation, operation and control," New York: John Wiley & Sons, 592, 1996.

[39] A. Pantoja, N. Quijano, "A population dynamics approach for the dispatch of distributed generators," IEEE Transactions on Industrial Electronics, **58**(10), 4559–4567, 2011, doi:10.1109/TIE.2011.2107714.

[40] E. Mojica-Nava, C. A. Macana, N. Quijano, "Dynamic population games for optimal dispatch on hierarchical microgrid control," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **44**(3), 306–317, 2013, doi: 10.1109/TSMCC.2013.2266117.

[41] J. Hofbauer, K. Sigmund, et al., Evolutionary games and population dynamics, Cambridge university press, 1998.

[42] N. F. Britton, Essential mathematical biology, Springer Science & Business Media, 2012.

[43] H. P. Young, S. Zamir, "Handbook of Game Theory with Economic Applications," Technical report, Elsevier, 2015.

[44] E. Mojica-Nava, C. Barreto, N. Quijano, "Population games methods for distributed control of microgrids," IEEE Transactions on Smart Grid, **6**(6), 2586–2595, 2015, doi:10.1109/TSG.2015.2444399.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 203-212 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Extraction of Psychological Symptoms and Instantaneous Respiratory Frequency as Indicators of Internet Addiction Using Rule-Based Machine Learning

Hung-Ming Chi[1], Liang-Yu Chen[2], Tzu-Chien Hsiao[1,3,*]

[1]*Department of Computer Science, College of Computer Science, National Yang Ming Chiao Tung University, Hsinchu, 300, Taiwan*

[2]*Institute of Computer Science and Engineering, College of Computer Science, National Yang Ming Chiao Tung University, Hsinchu, 300, Taiwan*

[3]*Institute of Biomedical Engineering, College of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu, 300, Taiwan*

ARTICLE INFO

Article history:
Received: 20 July, 2021
Accepted: 25 September, 2021
Online: 30 September, 2021

Keywords:
Extended Classifier System
Behavioral Addiction
Psychological Symptom
Instantaneous Frequency

ABSTRACT

Internet addiction (IA) has adverse effects on psychophysiological responses, interpersonal relationships, and academic and occupational performance. IA detection has received increasing attention. Although questionnaires enable long-term assessment (over 6 months) and physiological measurements to aid the short-term evaluation (over 2 min) of IA, the lack of algorithms results in an inability to detect IA in real time. A computer-aided system can address this problem. This study used the extended classifier system with continuous real-coded variables (XCSR) for rule-based machine learning to classify IA risk. Chen Internet Addiction Scale (CIAS) items were verified and instantaneous respiratory features of IA were extracted with "don't care" attribute values. The result demonstrated that the XCSR model achieved more than 95% classification accuracy. Using the "don't care" attribute values, the CIAS items were reduced from 26 to 19, and the instantaneous frequency (IF) of respiratory muscle contractions, respiratory wall movements, and body movements were extracted as IA-related features. These findings suggested that the XCSR model is a potentially useful system for detecting IA. The modified 19-item CIAS and IF of respiration can be adopted to assist in the real-time detection of IA and explore the psychophysiological developments of IA users. In future studies, more samples must be collected to validate these findings and instantaneous physiological responses investigated with different window sizes while participants with IA engage in active online gameplay.

1. Introduction

This paper constitutes a continuation of the work originally presented at the 2019 Genetic and Evolutionary Computation Conference [1]. The Internet has changed our daily lives profoundly over the last three decades. People now have easy access to the Internet by using desktops, smartphones, tablets, laptops, and other digital devices, through which they interact with friends, purchase products, play games, search for information, and send emails. Despite the convenience the Internet lends to our lives, some serious concerns remain.

Some individuals intensely indulge in and depend on the Internet, which can lead to negative ramifications, such as estrangement from family and friends, impaired mental and physical health, and low academic and work performance [2, 3]. This condition has been termed Internet addiction (IA) [4], and it has various subtypes, including problematic online game addiction [5], problematic cellphone use [6], and social network mental disorder [7]. Scholars have reported that problematic game use in particular was associated with severe clinical impairment, and Internet gaming disorder (IGD) has been preliminarily listed in the *Diagnostic and Statistical Manual of Mental Disorders (DSM), Fifth Edition* [8]. A 2019 announcement by the World Health Organization recognized gaming disorder as a mental disorder [9]. The IA prevalence rate in Asia [3], Europe [10], and the United States [11] was 5.6%, 7.97%, and 8.1%, respectively, with the risk of developing IA increasing as a result of the 2019 coronavirus

*Corresponding Author: Tzu-Chien Hsiao, 705R, MIRC, 1001 University Road, Hsinchu, Taiwan, labview@g2.nctu.edu.tw

disease pandemic [12, 13]. IA is a common behavioral addiction globally and continues to receive considerable attention.

Psychological symptoms and physiological responses are key indicators for detecting addiction behavior. IA-related questionnaires are used to evaluate the symptoms of IA, but they have some limitations. First, different questionnaires must be designed that correspond to various cultures and environments worldwide. The question of how many psychological symptoms must be incorporated into an IA questionnaire remains debated. Second, questionnaires are typically adopted to assess Internet experiences over a long-term period (over 6 months) [2, 3, 5, 8, 9, 14–18] and are not ideal for examining short-term or current experiences. Third, Internet users must spend time completing these questionnaires. Fourth, subjective experiences can affect the assessment of IA. Researchers can potentially observe the physiological responses of Internet users alongside questionnaires to assist in the assessment of IA risk. In such assessments, the major physiological signals are analyzed or calculated using descriptive statistics or generic mathematics [10, 19–23], although these algorithms may be insufficient to estimate the dynamical and instantaneous nature of physiological responses. Some studies have collected user behavior or social media and online gaming records to input into machine learning systems to identify addiction behaviors [7, 24–28]. However, this type of data collection is time consuming, and the real-time detection of IA presents a considerable challenge.

The extended classifier system (XCS) method is a rule-based machine learning system that assumes an agent role to interact with the environment and produce the optimum policy [29]. The XCS method generates rules with each interaction, each in the form of an "IF condition THEN action." The conditions in the rules employ coded symbols {0, 1, #}; the symbol # involves ignoring some values and is also referred to as the "don't care" attribute. The XCS method is applied to manage binary problems, and when combined with the continuous real-coded variables (XCSR) method, it can solve real-value problems [30]. Unlike other machine learning systems, the learning process of the XCSR method can be observed directly, and the "don't care" attribute can extract the main features for classification. Therefore, this study adopted the XCSR method as a computer-aided system to detect the risk of IA using Chen Internet Addiction Scale (CIAS) items [3] as input data. The XCSR method was then applied to verify the major psychological symptoms of IA from the CIAS data. To investigate the instantaneous physiological responses of IA, the ensemble empirical mode decomposition (EEMD) method [31] was used to decompose a physiological signal into instantaneous components, and the XCSR method was employed to extract the IA-related physiological features from these components. This study has two main contributions, namely the use of a rule-based machine learning system to classify the risk of IA, and the extraction of the major symptoms and instantaneous physiological features of IA from a questionnaire and from physiological components, respectively. These developments may be used to design a new short-form IA questionnaire and instantaneous physiological model to detect IA in real time. These extracted psychological symptoms and instantaneous responses can assist in exploring the developmental process of IA behavior.

The organization of this paper is as follows: The Literature Review section outlines the psychological symptoms covered in the IA questionnaire, physiological responses of the IA group, and detection of behavioral addiction using machine learning systems. The Research Methodology section presents the overview of the XCSR and EEMD methods, and the Experiment section describes the design of experiment and data analysis. The Results section reveals the classification accuracy of the XCSR using the CIAS items and instantaneous physiological waveforms as input data, with the rate of non-"don't care" values calculated for these input data. In the Discussion section, the study findings and limitations are discussed, and the Conclusion section summarizes the study.

2. Literature Review

2.1. IA questionnaires

The design of IA questionnaires is based on substance abuse, substance dependence, and pathological gambling criteria from the *DSM, Fourth Edition*, IGD criteria from the *DSM, Fifth Edition*, and gaming disorder criteria from the *International Classification of Diseases, Eleventh Revision*. IA questionnaires cover various psychological symptoms (Table 1), including abandoning other activities (S_A), craving (S_C), compulsive symptoms (S_{CS}), deceit or cover-up (S_{DC}), health-related problems (S_H), interpersonal problems (S_I), mood modification (S_{MM}), preoccupation or salience (S_{PS}), school or occupational problems (S_{SO}), tolerance (S_T), time management problems (S_{TM}), and withdrawal symptoms (S_{WS}). Researchers have debated how many psychological symptoms can be incorporated into an IA questionnaire. Young presented an eight-item IA diagnostic questionnaire that included S_{CS}, S_{DC}, S_I/S_{SO}, S_{MM}, S_{PS}, S_{TM}, S_T, and S_{WS} [2]. Griffiths proposed that S_I/S_H, S_{MM}, S_{CS}, S_{PS}, S_T, and S_{WS} were the core components of addiction [32]. The IGD questionnaire comprises S_A, S_{CS}, S_{DC}, S_I, S_{MM}, S_{PS}, S_{SO}, S_T, and S_{WS} [14], and the CIAS consists of S_{CS}, S_I/S_H, S_T, S_{TM}, and S_{WS} [3]. Authors suggested that S_I/S_H, S_{MM}, S_{CS}, S_{SO}, S_T, S_{PS}, and S_{WS} be used to design a questionnaire [5]. Authors structured their IA questionnaire by using the psychological symptoms of S_A, S_C, S_{CS}, S_{DC}, S_H/S_{SO}, S_I, S_{MM}, S_{PS}, S_T, and S_{WS} [9].

IA questionnaires are typically translated into various languages or modified into suitable structures for application in different countries or areas. For example, Young's IA Test is a well-known questionnaire for assessing IA risk [15]. Researchers in Germany converted this 20-item test into a 12-item short-form version [16]. Researchers also modified the nine IGD criteria outlined in the *DSM, Fifth Edition*, into a new nine-item short version for UK Internet users [17]. The CIAS is a common Chinese-language tool for evaluating IA symptoms [3]. Researchers in Hong Kong reduced the 26 items of the CIAS to 19 [18]. The aforementioned studies have adopted factor analysis to extract the psychological symptoms of IA and confirm the reliability and validity of their new IA questionnaire versions [16–18]. However, the factor analysis employed for extracting these symptoms used single-dimension calculations, and thus may be unable to analyze non-Markovian problems. Non-Markovian indicates that the optimum decision cannot be determined only from the current state, and memory is needed to assist decision-making. In addition, current IA questionnaires are limited to assessing prolonged Internet use (over 6 months) after browsing Internet (also called narrative surveys). Few questionnaires

evaluate the current experiences of the Internet user during browsing Internet (also termed the experience sampling method). Therefore, the design of an IA questionnaire for the immediate detection of IA risk is a substantial challenge.

Table 1: Psychological symptoms of IA

Symptom	Explanation
S_A	stop or reduce take part in other activities due to using the Internet
S_C	desire to use the Internet
S_{CS}	try to stop or reduce Internet use repeatedly without success
S_{DC}	lie about their amount of using the Internet
S_H	produce psychosocial and physical problems due to overusing the Internet
S_I	lose relationships due to overusing the Internet
S_{MM}	use the Internet to escape from stress or relieve negative mood
S_{PS}	always think about Internet activities or expect the next Internet activities
S_{SO}	risk the educations or occupations due to overusing the Internet
S_T	need to increase time spent on the Internet to achieve the same level of satisfaction as before
S_{TM}	reduce sleep and other activities to use the Internet
S_{WS}	feel displeasure when stop or reduce using the Internet

2.2. Physiological response of IA

Individuals with IA experience dynamic emotional stimulation on the Internet and then regulate psychophysiological responses instantaneously. Cardiovascular, respiratory, body temperature, skin conductance, and other responses in IA have been observed for short periods. The aim of such monitoring is to provide additional data that can be analyzed with IA questionnaires. Relevant results indicated that the risk of IA was positively associated with respiratory responses during 6 min of Internet browsing but negatively related to peripheral temperature [19]. Compared with individuals with low-risk Internet addiction (LIA), individuals with high-risk Internet addiction (HIA) showed higher skin conductance responses after 15 min of Internet use [10]. The respiratory sinus arrhythmia responses of users with HIA while watching negative and positive videos of 2 to 3 min were lower than that of users with LIA [20]. The heart rate and systolic blood pressure of individuals with problematic Internet use both increased after 2 min of stopping Internet use [21]. When playing an online game for 5 min, the high-frequency heart rate variability and cardiopulmonary synchrony for problematic Internet users in an excessive online gaming group were lower than those of the healthy group [22]. Compared with 5 min of baseline values, the natural logarithm of high-frequency heart rate variability of gamers with IGD was lower during 5 min of online gaming with attention [23].

Among the physiological signals, respiration is the only one that reflects emotional responses and affects emotions, motivation, and other physiological signals. Patients with IA and anxiety disorder obtained alleviation for their addiction and anxiety symptoms after 10 weeks of cognitive-behavioral therapy with breathing exercises [33]. Mindful breathing has also been employed to assist gamers with IGD [34] and individuals with a gambling addiction [35] to relieve their desire to play. For physiological regulation, respiration is related to the vagus nerve, which modulates sympathetic and parasympathetic nervous system activities. Slow respiratory rates reduce sympathetic nervous activity or increase parasympathetic nervous activity. Inhalation and exhalation change the intrathoracic pressure, which in turn affects the heart rate and blood pressure. Individuals with high-risk IGD reduced their respiratory rates by using slow and deep breathing exercises while watching negative films [31]. Therefore, the respiration response is a key feature for detecting IA risk, and respiratory adjustments constitute a valuable means of affecting psychophysiological responses.

The major IA-related features are generally extracted from physiological signals using descriptive statistics and generic mathematics methods. However, these methods are limited in analyzing physiological signals with dynamic and instantaneous changes and solving non-Markovian problems. The extraction of instantaneous physiological features and investigation of the relationship between each instantaneous physiological component are crucial for determining IA.

2.3. IA detection using machine learning

Some researchers have proposed machine learning models as computer-aided systems for detecting IA and extracting IA-related psychological and physiological features. Impulsive and compulsive symptoms were adopted to construct prediction models of problematic Internet use by employing naïve Bayes, random forest, and logistic regression methods [24]. Authors used a decision tree to determine IGD and extract the features of gaming cost, gaming time on weekdays and weekends, and offline gaming social, marriage, and self-perceptions [25]. In supervised machine learning, demographic information, gamer statistics, and self-esteem scores were input to classify IGD [26]. Authors combined long short-term memory with an adversarial autoencoder to predict the gamer's IGD risk using duration, costs, feelings of depressive responses, and contextual information in games as input data [27]. Social data from Facebook and Instagram were applied to detect problematic social network use by employing a support vector machine framework [7]. The aforementioned studies were required to collect data over a long period of time. Authors proposed that 10- or 25-s biosignals, such as heart and respiratory rates, eye blinks, and saccadic movements, be input as variables into the support vector machine to determine the urge of an IGD group for gaming [28]. Although long-term and short-term information has been used to predict addiction, the lack of algorithms impedes the use of instantaneous physiological response measurements in IA detection.

3. Research Methodology

3.1. Empirical mode decomposition

Empirical mode decomposition (EMD) [36] is an adaptive filter method applied to decompose nonlinear and nonstationary signals into oscillatory components from high to low frequencies; these components are also termed intrinsic mode functions (IMFs). The decomposition process of EMD is expressed in Algorithm 1.

Algorithm 1: Decomposition process of EMD

Result:
$s_1(t) \leftarrow s(t)$
$c(t) \leftarrow s(t)$
while $c_i(t)$, i =1, 2, …, n **do**
  \\ Sifting Process
  if standard_deviation ($c(t)$) $\geqslant$ 0.02 **then**; % stop criterion
    $p_{max} \leftarrow$aidentify_local_maxima($c(t)$)
    $p_{min} \leftarrow$iidentify_local_minima($c(t)$)
    $e_{max}(t)$ xncompute_upper_envelop(p_{max})
    $e_{min}(t)$ npcompute_lower_envelop(p_{min})
    $m(t) \leftarrow (e_{max}(t)+ e_{min}(t))/2$; % mean envelop
    $c(t) \leftarrow c(t) – m(t)$
  end
  $IMF_i(t) \leftarrow c(t)$
  $s_{i+1}(t)$ 1 $s_i(t) – IMF_i(t)$
end

EMD, however, has boundary-effect and mode-mixing problems. Boundary-effect indicates IMF endpoint distortions. Mode mixing refers to different oscillatory waveforms being present in a single IMF. These two problems lead to difficult or erroneous interpretation of the physical meaning of the IMFs but can be resolved using the EEMD method [37]. The analysis process of EEMD involves the addition of Gaussian noise to the source signal; this noise-added signal is then decomposed using the sifting process of the EMD method.

3.2. XCS

The XCS is a powerful machine learning model with high interpretability and flexibility. The interpretability of the XCS stems from the use of a set of collaborative rules known as a population set [P] to express the solution of the problem domain. The interpretation of each rule is the "IF condition THEN action," with the condition part of each rule corresponding to the detected environmental state and coded with a ternary character {0, 1, #}. The # symbol is termed "don't care," which indicates that the corresponding environmental state can be ignored regardless of whether it is 0 or 1. The action part of each rule is the decision-making that the XCS executes. If the output is a binary classification problem, the action can be encoded as a binary string {0, 1}. Furthermore, each rule contains the following three parameters to indicate the prediction and quality of each rule: 1) *prediction*, which is used to predict how much feedback can be obtained from the environment after the action is executed; 2) *prediction error*, which is used to calculate the error between the prediction and actual feedback; and 3) *fitness*, which is used to evaluate the accuracy and quality of the prediction. A modular architecture allows XCS to address many problem areas flexibly, including performance components, exploration components, and credit assignment.

The performance component is responsible for detecting the state of the environment and making decisions. First, the XCS converts the detected environmental state into a binary string. Second, the condition part of each rule in the [P] is compared with the environmental state, and the rule is moved to the matching set [M] only if the conditions of the rule match the environmental state. If [M] is empty, the XCS initiates the *covering* mechanism to generate new rules that satisfy the environmental state. Third, the XCS must select an action to execute according to the rules in [M]. The system uses rules with the same action in [M] to calculate the expected feedback of the action and form a prediction array [PA]. The XCS adopts two mechanisms, exploration and exploitation, to select the action. In the exploration mode, the expected feedback calculated by each action in [PA] is used as the probability of random selection; in the exploitation mode, the system directly selects the action with the greatest expected feedback. Fourth, the selected action is executed through the effector, and the rules in [M] that match the action move to the action set [A] to update the parameters. The XCS repeats these four steps until convergence or the maximum number of iterations has been reached.

The exploration components are used to assist the XCS in generating new rules to search for the optimum solution in the problem domain. Evolutionary computation (EC) introduces small random factors to gradually optimize the solution and is suitable for the exploration components of the XCS. Typically, the XCS adopts the genetic algorithm (GA) in EC as its exploration components. GA solves the optimization problem by simulating natural selection, heredity, crossover, and mutation mechanisms in evolutionary biology. In the XCS, the GA randomly selects two rules from [P] for inheritance based on the fitness of each rule as the probability. Subsequently, the GA performs crossover on each part of the condition in accordance with the two rules based on the probability χ and mutates each part with probability μ. The GA generates two new rules to replace the rules with lower fitness in [P] to assist the XCS in identifying the global optimum in the problem domain.

The credit assignment assigns and updates the parameters of the rules in [A]. The XCS uses the feedback from the environment to modify the prediction of each rule in [A] as a reference parameter for the next iteration. The modified technique uses the Widrow–Hoff delta rule, as described in Eq. (1), where p is the *prediction* of the rule, P is the actual feedback obtained from the environment, t is the current iteration number, and β is the learning rate ($0 < \beta \leq 1$). The XCS can then use the absolute error to correct the *prediction error* of the rule, as detailed in Eq. (2), where ε denotes the prediction error. The last parameter *fitness* is used to evaluate the quality of the rule. The calculation of *fitness* is based on the relative accuracy of the rule in [A]. First, the XCS must calculate the accuracy k_t of each rule in [A], as expressed in Eq. (3), where α controls the attenuation of the exponential function ($0 < \alpha < 1$), ε_t is the *prediction error* of the rule at time t, and ε_0 is the tolerable error. Subsequently, the XCS calculates the relative accuracy k_t' of the rule compared with the other rules in [A], as described in Eq. (4). The system utilizes the relative accuracy k_t' as the *fitness* of the rule and employs the Widrow–Hoff delta rule to update, as detailed in Eq. (5), where F is the *fitness* of the rule.

$$p_{t+1} = p_t + \beta(P_t - p_t) \tag{1}$$

$$\varepsilon_{t+1} = \varepsilon_{t+1} + \beta(|P_t - p_t| - \varepsilon_t) \tag{2}$$

$$k_t = \begin{cases} \exp\left[\dfrac{(\ln \alpha)(\varepsilon_t - \varepsilon_0)}{\varepsilon_0}\right], & \varepsilon_t > \varepsilon_0 \\ 1, & \text{otherwise} \end{cases} \tag{3}$$

$$k_t^{'} = \frac{k_t}{\sum_{k_t \in [A]} k_t} \quad (4)$$

$$F_{t+1} = F_t + \beta(k_t' - F_t) \quad (5)$$

The XCS provides the rule with a *numerosity* parameter to record how many of the same rules are in [P] and uses a rule as a representative of these rules. Therefore, the same rules must only be matched once to maximize the computation speed. The *subsumption* mechanism is also added to assist in generating the most accurate and general rules. In the *subsumption* mechanism, one rule can subsume another rule, and the rule therefore undergoes a certain amount of training that strengthens its generalizability and accuracy. If rule r_a can subsume another rule r_b, r_b is removed from [P], and the numerosity of r_a increases by one. The *subsumption* mechanism usually acts on the GA and [A].

3.3. XCSR

The encoding method used in the *condition* of the rule can only process discrete variables. Wilson proposed to modify the ternary-alphabet encoding of the *condition* into interval predicates to enable the XCS to process continuous variables [30]. The expression of interval predicates is $int_i = (c_i, s_i)$, where int_i is the rule condition that the ith interval predicate corresponds to the ith variable, c_i is the central value of the range, and s_i refers to the size of the coverage boundary. Changing the expression of the condition modifies the mechanism of components, such as *matching*, *covering*, and the GA. For *matching*, assuming the ith variable is x_i, if $c_i - s_i \leq x_i < c_i + s_i$, then the int_i matches x_i. For *covering*, the XCS generates a rule whose interval predicates cover all input variables. The generation method is to set c_i to x_i, and s_i is set to a random value between 0 and s_r, where s_r is defined by the user. For the GA, *crossover* exchanges the int_i with probability χ; *mutation* adds a random value from 0 to s_m to the range of int_i with probability μ, where s_m is defined by user. The remaining mechanism is the same as the XCS operation, with this modified version known as the XCSR. The modular structure allows the XCS to be flexibly modified to adapt to various problem domains.

4. Experiment

4.1. Design of experiment

(1) Objective of this experiment

The CIAS scores of participants with HIA and LIA were collected, and their respiratory signals were observed while they watched online game videos. This study was granted ethical approval by National Chiao Tung University (NCTU) Research Ethics Committee for Human Subject Protection (Approval Number: NCTU-REC-102-009-e).

(2) Method

To assess the risk of IA, we adopted the CIAS, which contains 26 items, with each item ranked from 1 (strong disagreement) to 4 (strong agreement) [3]. The total score of the CIAS ranged from 26 to 104, with a score of 26 to 63 and 64 to 104 reflecting LIA and HIA, respectively. This study also adopted the IGD questionnaire to evaluate IGD risk; this questionnaire consists of nine items requiring yes or no responses [14]. The IGD questionnaire score ranged from 0 to 9, with 0 to 4 and 5 to 9 representing non-IGD and IGD, respectively. The internal consistency of the CIAS and IGD questionnaires, as estimated using Cronbach's α, was 0.94 and 0.81, respectively. The extent of participants' experience using computers and the Internet were recorded, and the Self-Assessment Manikin (SAM) was adopted to evaluate the emotions of participants. The SAM measures emotional valence and arousal on a 9-point Likert scale. A 5-point score on valence represents neutral emotion, and anything higher indicates positive emotions. Conversely, a score lower than 5 is indicative of negative emotions. Arousal was scored from 1 (no feeling) to 9 (strong feeling). To arouse a craving for gaming, 2 min of online game videos were selected from League of Legends (LOL; Riot Games, Los Angeles, CA, USA), MapleStory (MS; Nexon, Seoul, South Korea), and Resident Evil (RE; Capcom, Osaka, Japan). Compared with the active stimuli of playing games, watching clips as passive stimuli may induce more physiological stability, potentially lowering the difference in psychological responses between participants.

(3) Indices

The 26-item scores of CIAS and instantaneous frequency (IF) of respiratory signal were used to detect the risk of IA.

(4) Participant

Participants aged 20–40 years and with no symptoms of depression and anxiety, and cardiovascular diseases were included in the study. Each participant provided written informed consent. A total of 50 participants (36 men and 14 women) were recruited from NCTU.

(5) Procedure

Each participant stared at a gray picture to familiarize themselves with the environment and relax physically at the beginning of the experiment. To induce a feeling of calm, the participant looked at a gray picture for 2 min before each video was played. For emotional stimuli, participants watched three videos (LOL, MS, and RE) in a random order. After each video, participants completed the SAM questionnaire, after which time the gray picture was displayed for another 2 min to allow the participant to recover from emotional arousal. During the experimental procedure, respiratory inductance plethysmography (RIPmate Inductance Belt, Ambu, Ballerup, Denmark) and electrocardiography (Ag/AgCl spot electrodes attached to the left and right subclavian and left lumber regions) were used to measure abdominal wall movement and electrocardiography signals, respectively. These signals were acquired through a DAQCard (USB 6218, NI, Austin, TX, USA) with a 1000 Hz sampling rate. To investigate instantaneous respiratory responses, this study only analyzed abdominal wall movement signals.

4.2. Data analysis

(1) Method

The XCSR model was used to detect the risk of IA by using the CIAS data and the IF of respiration as the input features (Figure 1). The abdominal movement signal was analyzed as the IF according to the signal processing step described in a previous study [36], including integral calculation, signal baseline correction, downsampling, signal decomposition, and transformation.

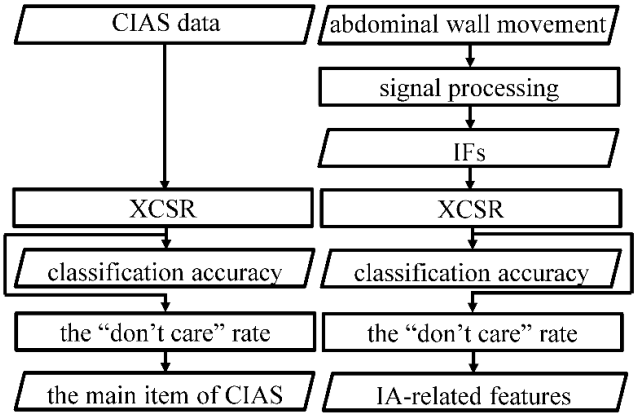


Figure 1: Data analysis procedure.

For statistical calculation, the Mann–Whitney U test was performed to compare differences in the questionnaire data and average IF between the HIA with IGD and LIA with non-IGD groups. The statistical significance was set as $p < 0.05$. To assess the performance of the average IF of respiration in discriminating between HIA with IGD and LIA with non-IGD, the receiver operating characteristic curve (ROC) was calculated, and the area under the curve (AUC) of the ROC indicated the predictive accuracy. The AUC ranged from zero (reflecting incorrect prediction) to one (representing correct prediction); the AUC was 0.7, indicating acceptable predictive power.

(2) Procedure

The integral calculation converted the abdominal movement signal into the volume change of the abdomen by using Simpson's rule. The volume changes of the abdomen were then used to perform baseline correction using the linear fit. To investigate the frequency band of respiratory muscle contractions, the corrected signal was downsampled from a rate of 1000 to 50 Hz. Subsequently, the EEMD method was used to decompose the downsampled signal into eight IMFs. With reference to the frequency bandwidth [38], IMF_1–IMF_2, IMF_3–IMF_4, IMF_5–IMF_6, and IMF_7–IMF_8 represented the diaphragm, respiratory muscle cluster on the abdomen, respiratory wall movement on the abdomen, and body movement components, respectively. In the transformation step, each IMF was transformed into IF by using the normalized direct quadrature [39]. The average value of IF during the viewing of the LOL, MS, and RE videos was treated as the probability of crossover operation (χ) = 0.8, the probability of the instantaneous psychological features that were input into the XCSR model.

Parameters in the XCSR were set as follows [29, 30, 40]: the population size (N) constituted 350 classifiers, probability of # (P#) = 0.33, initial rule ε = 0, initial rule p = 10, initial rule F = 10, learning rate (β) = 0.2, accuracy function (ε_0) = 10, ν = 5, ε_0 = 10, α = 0.1, ϕ = 0.1, mutation operation (μ) = 0.04, threshold parameter of the GA = 50, deletion thresholds of the GA = 20, and subsumption thresholds of the GA = 20; moreover, the reward setup for accurate prediction of HIA with IGD and LIA with non-IGD was set as 1000, or otherwise 0. The iteration number of the XCSR method was 1950 for both the 26-item score in the CIAS and 24 IF values of respiration as input features. The moving average per 50 exploitations was used to calculate the classification accuracy of the XCSR. To verify the items in the CIAS and extract the crucial instantaneous respiratory components for detecting the IA risk, rules with a reward of 1000 in [P] were selected, and then the rate of non-"don't care" values ("non-#") in these rules was calculated. To effectively evaluate the training quality of the XCSR and avoid bias, 10-fold cross-validation was used. To eliminate the unevenness of the IF at different frequency bandwidths, the IF underwent min–max normalization. The XCSR does not update any parameters nor generate new rules for the testing data and only uses the exploitation mode when selecting and performing actions on this data. All data management, respiratory signal processing, and XCSR execution operation were conducted using LabVIEW (v.2021, NI, Austin, TX, USA).

5. Results

Of all participants, 34% were from the College of Electrical and Computer Engineering, 28% from the College of Computer Science, 14% from the College of Management, 8% from the College of Humanities and Social Sciences, 6% from the College of Hakka Studies, 6% from the College of Engineering, and 4% were from other departments. Participants were stratified into four groups according to their CIAS and IGD questionnaire scores, namely the LIA with non-IGD, LIA with IGD, HIA with non-IGD, and HIA with IGD groups. This study focused on investigating the psychological symptoms and physiological responses of the LIA with non-IGD and HIA with IGD groups only. Table 2 lists the demographic information and questionnaire responses of the two groups. No significant differences in age, computer and Internet use experience, or emotional valence and arousal were observed between the two groups. The total scores of the CIAS and IGD questionnaires of the HIA with IGD group were significantly higher than those of the LIA with non-IGD group. The emotional valence results indicated that the LOL and MS videos evoked positive emotions, whereas the RE video prompted negative emotions.

Table 2: Demographic information and questionnaire responses

		LIA with non-IGD	HIGD with IGD
Gender (men, women)		12, 9	14, 4
Age (years old)		22.62 ± 1.79	23.67 ± 5.27
Computer experience (year)		13.19 ± 2.71	12.44 ± 4.50
Internet experience (year)		11.71 ± 2.63	10.89 ± 4.19
CIAS		52.10 ± 8.03	79.33 ± 8.88**
IGD questionnaire		1.48 ± 1.25	6.58 ± 1.54**
SAM	valence_LOL	5.95 ± 1.32	6.61 ± 1.33
	arousal_LOL	4.33 ± 2.01	5.28 ± 1.53
	valence_MS	5.76 ± 1.79	5.56 ± 1.46
	arousal_MS	4.33 ± 1.88	4.22 ± 1.73
	valence_RE	3.90 ± 1.87	4.39 ± 1.61
	arousal_RE	6.14 ± 1.80	5.78 ± 1.63

** $p < 0.01$ in the Mann–Whitney U test

The results of classification accuracy of the XCSR with an average of 30 repetitions using 26-item CIAS scores and 24 IFs of respiration in response to online video stimuli are illustrated in Figure 2 and Figure 3, respectively. The classification accuracy of the XCSR model was over 95%. Table 3 presents the mean ± standard deviation scores of each item in the CIAS for the two

groups and the "non-#" rate of each item. The score of each item for the HIA with IGD group was significantly higher than that of the LIA with non-IGD group. The "non-#" rate of the 11th, 14th, 19th, and 22nd items in S_{CS}; 12th, 13th, 15th, 17th, 18th, and 21st items in S_I/S_H; 6th item in S_T; 1st, 8th, 23rd, and 25th items in S_{TM}; and 2nd, 4th, 10th, and 16th items in S_W were higher than the 0.50 value for the CIAS. The 10-fold cross-validation of the XCSR using the 26-item CIAS score and 24 IFs of respiration in response to online video stimuli is described in Table 4 and Table 5, respectively; the average accuracy of 10-fold cross-validation for the XCSR was 87% and 78%, respectively. Table 6 presents the mean ± standard deviation of the IFs of respiration of the two groups while watching the LOL, MS, and RE videos as well as the "non-#" rate of each IF of respiration. The IF_2 of the HIA with IGD group while watching the RE video stimuli was significantly higher than that of the LIA with non-IGD group. Except for IF_5 during the MS video and IF_5 and IF_7 during the RE video, the "non-#" rate of all IFs was higher than the 0.50 value. The ability to predict the risk of IA using IF values, as calculated with the AUC of the ROC, is detailed in Table 7.

6. Discussion

In this study, the XCSR method as a rule-based machine learning system was used to aid in detecting the risk of IA. The "don't care" attribute of the XCSR was employed to verify the items of the CIAS and extract the instantaneous respiratory features for the classification of IA. The results demonstrated that the XCSR model could identify HIA with IGD group or LIA with non-IGD group with a classification accuracy of over 95%. The average accuracy of the 10-fold cross-validation was also over 78%. Researchers have proposed that self-esteem scores [26], impulsive and compulsive symptoms [24], and physiological signals [28] can serve as the input variables of machine learning systems to predict behavioral addictions. Our findings suggested that the XCSR model is an appropriate computer-aided tool for detecting IA. The psychological symptoms in the CIAS and instantaneous respiratory responses are vital indexes for assessing risky Internet use.

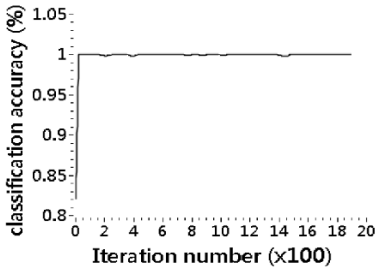


Figure 2: Classification accuracy of the XCSR with the CIAS data as the input feature.

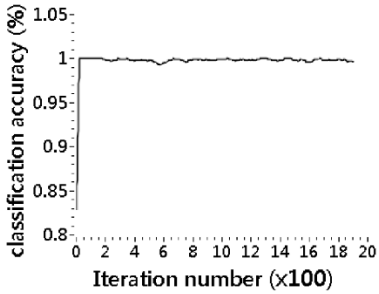


Figure 3: Classification accuracy of the XCSR with the IF of respiration data as the input feature.

Table 3: Psychological symptoms in the CIAS and their corresponding items (modified from [1])

Symptom	item	LIA with non-IGD	HIA with IGD	not # rate
S_{CS}	11. I unable control my impulse of Internet use	1.81 ± 0.68	2.89 ± 0.74**	0.58
	14. When waking up every morning, my first thought is to use the Internet	2.00 ± 0.89	3.05 ± 0.78**	0.75
	19. After stopping Internet use, I crave it again	2.33 ± 0.73	3.32 ± 0.58**	0.67
	20. Without the Internet, my life would be joyless	1.76 ± 0.77	2.79 ± 0.85**	0.42
	22. I try to spend less time on the Internet, but failure	2.14 ± 0.57	3.05 ± 0.71**	0.67
S_I/S_H	7. Even if the Internet negatively affects my interpersonal relationships, my use remains unreduced	1.43 ± 0.60	3.21 ± 0.71**	0.42
	12. I find that I reduce the time spent with friends due to the Internet	1.81 ± 0.51	2.95 ± 0.78**	0.67
	13. I feel aches and soreness in the back or other discomfort due to Internet use	2.00 ± 0.84	3.11 ± 0.94**	0.92
	15. Using Internet has negative effects on my education or work	1.86 ± 0.73	2.84 ± 0.60**	0.75
	17. I reduced interaction with my family due to Internet use	1.76 ± 0.77	3.05 ± 0.85**	0.58
	18. I reduce my recreational activities to use the Internet	2.00 ± 0.63	3.16 ± 0.83**	0.67
	21. Using the Internet has negative effects on my health	2.10 ± 0.70	2.95 ± 0.78**	0.83
S_T	3. I perceive my Internet use as getting longer and longer	2.43 ± 0.51	3.05 ± 0.71**	0.50
	6. I spend more time on the Internet than I originally intended	2.71 ± 0.64	3.53 ± 0.51**	0.58
	9. I have spent more time on the Internet since last semester	2.00 ± 0.63	2.95 ± 0.71**	0.33
	24. I need to increase the amount of time I use the Internet to achieve the same satisfaction as before	1.67 ± 0.48	2.63 ± 0.83**	0.42
S_{TM}	1. I have been told more than once that I spend too much time on the Internet	2.48 ± 0.75	3.21 ± 0.79**	0.58
	8. My sleep time is less than 4 hours from using the Internet	1.76 ± 0.89	3.16 ± 0.96**	0.67
	23. I am used to reducing my sleep time to use the Internet	1.52 ± 0.75	2.79 ± 0.79**	0.67
	25. I do not eat on time due to using the Internet	2.05 ± 0.80	2.84 ± 0.90**	0.58
	26. I use the Internet all night, which causes daytime tiredness	2.33 ± 1.02	3.00 ± 0.58**	0.50
S_{WS}	2. I feel displeasure when I stop using the Internet for a period of time	2.05 ± 0.67	3.26 ± 0.73**	0.58
	4. I feel restless and irritable when the Internet is unavailable	2.52 ± 0.87	3.37 ± 0.68**	0.75
	5. I feel energetic upon using the Internet regardless of a fatiguing experience	1.62 ± 0.59	2.84 ± 0.76**	0.50
	10. I feel distressed when stopping the Internet use for a period of time	1.57 ± 0.51	2.89 ± 0.74**	0.83
	16. I feel like I am missing something when stopping the Internet use for a period of time	2.38 ± 0.97	3.37 ± 0.60*	0.75

* $p < 0.05$ and ** $p < 0.01$ in the Mann–Whitney U test

The results indicated that the 11th, 14th, 19th, and 22nd items in S_{CS}; 12th, 13th, 15th, 17th, 18th, and 21st items in S_I/S_H; 6th item in S_T; 1st, 8th, 23rd, and 25th items in S_{TM}; and 2nd, 4th, 10th, and 16th items in S_W were maintained in the CIAS item verification. This result, however, is inconsistent with that of other works. Authors modified the items in the CIAS for grade 7-to-13 Hong Kong students, retaining the 14th and 20th items in S_{CS}; 12th, 13th, 15th, 17th, 18th, and 21st items in S_I/S_H; 3rd, 6th, 9th, and 24th items in S_T; 1st, 8th, 23rd, and 26th items in S_{TM}; and 2nd, 5th, and 16th items in S_W [18]. Authors changed the 26-item CIAS into a 10-item smartphone addiction inventory, retaining the 22nd item in S_{CS}; 7th, 13th, and 15th items in S_I/S_H; 3rd, 6th, and 9th items in S_T; 23rd item in S_{TM}; and 2nd and 4th items in S_W for Taiwan students aged 18 to 31 years [41]. These inconsistent results are likely attributable to differences in age, country, analysis method, or assessment of psychological symptoms between studies. Our findings must be validated with more data.

In this study, the "don't care" rate of IF_1–IF_8 of respiration during LOL video stimuli; IF_1–IF_4 and IF_6–IF_8 of respiration during MS video stimuli; and IF_1–IF_4, IF_6, and IF_8 of respiration during RE video stimuli were lower than 0.50. IMF_1–IMF_2, IMF_3–IMF_4, IMF_5–IMF_6, and IMF_7–IMF_8 represented the diaphragm, respiratory muscle cluster on the abdomen, respiratory wall movement on the abdomen, and body movement components, respectively [38]. The LOL and MS videos aroused positive emotions in the participants, whereas the RE video evoked negative emotions.

Table 4: The 10-fold cross-validation of the XCSR with the CIAS data as the input feature

Testing	1	2	3	4	5
Accuracy	1.00	1.00	1.00	0.50	1.00
Testing	6	7	8	9	10
Accuracy	1.00	0.75	0.75	1.00	0.67
Average accuracy			0.87		

Table 5: The 10-fold cross-validation of the XCSR with the IF of respiration data as the input feature

Testing	1	2	3	4	5
Accuracy	1.00	0.25	0.5	0.75	0.75
Testing	6	7	8	9	10
Accuracy	1.00	0.75	1.00	0.75	1.00
Average accuracy			0.78		

Emotions affect respiratory responses, and respiratory muscular contraction and relaxation drive the thoracic and abdominal wall movements. Compared with the users with LIA, those with HIA exhibited a higher abdominal muscle contraction amplitude and lower abdominal wall movements whilestaring at positive or negative pictures of 3- to 12-s [42]. The respiratory responses positively related to the scores of the CIAS [19], and the respiratory rate was used to predict the craving of individuals with IGD to play online games [28]. Those with HIA are on the Internet excessively, which can cause aches and soreness in the back [2, 3] or hands; they may change their posture to relieve this pain. Therefore, we determined that the instantaneous responses of respiratory muscle clusters, respiratory wall movements, and body movements constitute potential biomedical indexes through which to observe the psychological reaction of the IA group, with the various emotional stimuli eliciting different respiratory responses. These responses can assist in detecting IA symptoms. We also demonstrated that IF_2 during the RE video stimuli was an acceptable predictor of IA risk, which was calculated using the ROC. However, this finding was inconsistent with the feature extraction of the XCSR model and can be attributed to several possible explanations. The ROC employs conventional mathematical calculation, but the XCSR method observes the relationship between features using multiple dimension attributes. In addition, the sampling size and parameter setup of the XCSR may influence feature extraction. Therefore, future works should employ more data samples and test different parameters of the XCSR method.

Table 6: Mean ± standard deviation of IF of respiration

Feature		LIA with non-IGD	HIGD with IGD	not # rate
LOL	IF_1	12.21 ± 0.15	12.19 ± 0.19	0.61
	IF_2	7.42 ± 0.27	7.32 ± 0.38	0.67
	IF_3	3.64 ± 0.21	3.53 ± 0.33	0.64
	IF_4	1.18 ± 0.20	1.16 ± 0.20	0.59
	IF_5	0.35 ± 0.08	0.34 ± 0.07	0.64
	IF_6	0.18 ± 0.03	0.18 ± 0.04	0.57
	IF_7	0.08 ± 0.02	0.07 ± 0.02	0.59
	IF_8	0.04 ± 0.01	0.04 ± 0.01	0.58
MS	IF_1	12.21 ± 0.20	12.24 ± 0.15	0.72
	IF_2	7.33 ± 0.29	7.33 ± 0.40	0.56
	IF_3	3.55 ± 0.27	3.57 ± 0.38	0.58
	IF_4	1.22 ± 0.28	1.31 ± 0.30	0.58
	IF_5	0.35 ± 0.11	0.36 ± 0.13	0.49
	IF_6	0.18 ± 0.04	0.18 ± 0.05	0.56
	IF_7	0.08 ± 0.02	0.08 ± 0.03	0.54
	IF_8	0.04 ± 0.01	0.04 ± 0.01	0.55
RE	IF_1	12.31 ± 0.17	12.29 ± 0.14	0.63
	IF_2	7.74 ± 0.16	7.61 ± 0.28*	0.53
	IF_3	3.93 ± 0.16	3.81 ± 0.32	0.56
	IF_4	1.54 ± 0.28	1.45 ± 0.35	0.55
	IF_5	0.38 ± 0.09	0.41 ± 0.17	0.37
	IF_6	0.25 ± 0.06	0.21 ± 0.04	0.56
	IF_7	0.10 ± 0.02	0.10 ± 0.03	0.44
	IF_8	0.05 ± 0.01	0.05 ± 0.02	0.66

* $p < 0.05$ in the Mann–Whitney U test

This study has some limitations. First, the small sample size may lead to overfitting and affect the verification of items in a questionnaire or the feature extraction of physiological responses. Second, the parameter setup of the XCSR was based on empirical studies and must be tested with different values. Third, the physiological responses of participants while watching online game videos may differ from those actively playing online games. Fourth, we observed respiratory responses during 2 min of emotional stimuli. Future studies can investigate second-by-second psychological responses under varying window sizes. Fifth, the unequal sample of men and women may affect the analysis of the XCSR. Notwithstanding these limitations, the XCSR model provides researchers with a computer-aided detection method for IA. The verification of items in a questionnaire can assist in the

design of a questionnaire for experience sampling method and the real-time detection of IA. The extraction of instantaneous respiratory features can provide psychological insight into the development of IA behavior.

Table 7: Predictive accuracy for the IF value in individuals watching online game videos, calculated with the area under the curve (AUC) of the ROC

Feature		AUC	*p*-value	95% confidence interval	
				lower bound	upper bound
LOL	IF_1	0.52	0.82	0.33	0.71
	IF_2	0.52	0.80	0.33	0.71
	IF_3	0.58	0.41	0.39	0.76
	IF_4	0.53	0.74	0.35	0.72
	IF_5	0.56	0.55	0.37	0.74
	IF_6	0.49	0.96	0.31	0.68
	IF_7	0.63	0.17	0.45	0.81
	IF_8	0.51	0.96	0.32	0.69
MS	IF_1	0.56	0.52	0.33	0.71
	IF_2	0.51	0.91	0.38	0.74
	IF_3	0.53	0.78	0.32	0.70
	IF_4	0.59	0.35	0.33	0.72
	IF_5	0.52	0.84	0.41	0.77
	IF_6	0.53	0.74	0.33	0.70
	IF_7	0.55	0.61	0.34	0.72
	IF_8	0.56	0.55	0.36	0.74
RE	IF_1	0.54	0.63	0.36	0.73
	IF_2	0.70	0.03	0.51	0.89
	IF_3	0.67	0.07	0.48	0.86
	IF_4	0.64	0.14	0.45	0.82
	IF_5	0.52	0.87	0.32	0.71
	IF_6	0.67	0.07	0.50	0.84
	IF_7	0.58	0.41	0.39	0.76
	IF_8	0.52	0.82	0.33	0.71

7. Conclusion

This study aimed to detect IA using the XCSR model. This model was also applied to verify the items of the CIAS and extract instantaneous respiratory features. The XCSR achieved over 95% classification accuracy for both the CIAS data and IF of respiration during online game video stimuli. The average accuracy of the 10-fold cross-validation was also higher than 75% for both input data. Furthermore, using the XCSR model, the CIAS was modified from a 26-item to 19-item questionnaire and the IFs of respiratory muscle contractions on the abdomen, respiratory wall movements on the abdomen, and body movements were extracted as the IA-related features. The current findings indicated the suitability of the XCSR method as a computer-aided system for real-time IA detection. The extracted psychological symptoms and instantaneous respiratory responses are key features in determining the development of IA. Further studies should collect more samples to validate these findings; design experiments incorporating active gameplay stimuli; and measure the instantaneous psychophysiological responses per second under varying window sizes.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was fully supported by the Taiwan Ministry of Science and Technology (Grant Numbers: MOST 105-2221-E-009-159, MOST 105-2634-E-009-003, and MOST 107-2221-E-009-153). As well, all participants in this study were thanked for their time and contributions.

References

[1] H.-M. Ji, L.-Y. Chen, T.-C. Hsiao, "Real-Time detection of Internet addiction using reinforcement learning system," in 2019 Genetic and Evolutionary Computation Conference (GECCO), 1–9, 2019, doi:10.1145/3319619.3326882.

[2] K.S. Young, "Internet addiction: The emergence of a new clinical disorder," CyberPsychology and Behavior, **1**(3), 237–244, 1998, doi: 10.1089/cpb.1998.1.237.

[3] S.-H. Chen, L.-J. Weng, Y.-J. Su, H.-M. Wu, P.-F. Yang, "Development of Chinese Internet addiction scale and its psychometric study," Chinese Journal of Psychology, **45**(3), 279–294, 2003, doi: 10.1037/t44491-000.

[4] C. Chou, L. Condron, J.C. Belland, "A review of the research on Internet addiction," Educational Psychology Review, **17**(4), 363–388, 2005, doi: 10.1007/s10648-005-8138-1.

[5] J.S. Lemmens, P.M. Valkenburg, J. Peter, "Development and validation of a game addiction scale for adolescents," Media Psychology, **12**, 77–95, 2009, doi: 10.1080/15213260802669458.

[6] M. Beranuy, U. Oberst, X. Carbonell, A. Chamarro, "Problematic internet and mobile phone use and clinical symptoms in college students: The role of emotional intelligence," Computers in Human Behavior, **25**, 1182–1187, 2009, doi:10.1016/j.chb.2009.03.001.

[7] H.-H. Shuai, C.-Y. Shen, D.-N. Yang, Y.-F. C. Lan, W.-C. Lee, P. S. Yu, M.-S. Chen, " A comprehensive study on social network mental disorders detection via online social media mining," IEEE Transactions on Knowledge and Data Engineering, **30**(7), 1212–1225, 2018, doi: 10.1109/TKDE.2017.2786695.

[8] American Psychiatric Association, Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5), American Psychiatric Association Press Inc., Arlington, VA, 2013.

[9] Y.-S. Jo, S.-Y. Bhang, J.-S. Choi, H.-K. Lee, S.-Y. Lee, Y.-S. Kweon, "Clinical characteristics of diagnosis for Internet gaming disorder: comparison of DSM-5 IGD and ICD-11 GD diagnosis," Journal of Clinical Medicine, **8**, 945, 2019, doi:10.3390/jcm8070945.

[10] M. Romano, A. Roaro, F. Re, L.A. Osborne, R. Truzoli, P. Reed, "Problematic internet users' skin conductance and anxiety increase after exposure to the internet," Addictive Behaviors. **75**, 70–74, 2017, doi: 10.1016/j.addbeh.2017.07.003.

[11] J. Morahan-Martin, P. Schumacher, "Incidence and correlates of pathological Internet use among college students," Computers in Human Behavior, **16**(1), 13–29, 2000, doi: 10.1016/S0747-5632(99)00049-7.

[12] Y. Sun, Y.-Y. Li, Y.-P. Bao, S.-Q. Meng, Y.-K. Sun, G. Schumann, T. Kosten, J. Strang, L. Lu, J. Shi, "Brief report: Increased addictive Internet and substance use behavior during the COVID-19 pandemic in China," American Journal on Addictions, **29**, 268–270, 2020, doi: 10.1111/ajad.13066.

[13] M.C. Zara, L.H. A. Monteiro, "The negative impact of technological advancements on mental health: An epidemiological approach," Applied Mathematics and Computation, **396**, 125905, 2021, doi: 10.1016/j.amc.2020.125905.

[14] N.M. Petry, F. Rehbein, D.A. Gentile, J.S. Lemmens, Hans-Jürgen Rumpf, T. Mößle, G. Bischof, R. Tao, D.S.S. Fung, G. Borges, et al., "An international consensus for assessing Internet gaming disorder using the new DSM-5 approach," Addiction, **109**, 1399–1406, 2014, doi: 10.1111/add.12457.

[15] K.S. Young, Caught in the net: How to recognize the signs of internet addiction--and a winning strategy for recovery: John Wiley & Sons, 1998.

[16] M. Pawlikowski, C. Altstötter-Gleich, M. Brand, "Validation and psychometric properties of a short version of Young's Internet addiction test," Computers in Human Behavior, **29**(3), 1212–1223, 2013, doi: 10.1016/j.chb.2012.10.014.

[17] H. M. Pontes, M.D. Griffiths, "Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale," Computers in Human Behavior, **45**, 137–143, 2015, doi: 10.1016/j.chb.2014.12.006.

[18] K.-K. Mak, C.-M. Lai, C.-H. Ko, C. Chou, D.-I. Kim, H. Watanabe, and

R.C.M. Ho, "Psychometric properties of the revised Chen Internet addiction scale (CIAS-R) in Chinese adolescents," Journal of Abnormal Child Psychology, **42**, 1237–1245, 2014, doi: 10.1007/s10802-014-9851-3.

[19] D.-W. Lu, J.-W. Wang, A.C.W Huang, "Differentiation of Internet addiction risk level based on autonomic nervous responses: The Internet-addiction hypothesis of autonomic activity," Cyberpsychology, Behavior, and Social Networking, **13**, 371–378, 2010, doi: 10.1089/cyber.2009.0254.

[20] D.-L. Hsieh, T.-C. Hsiao, "Respiratory sinus arrhythmia reactivity of internet addiction addicts in negative and positive emotional states using film clips stimulation," BioMedical Engineering OnLine, **15**, 69, 2016, doi: 1186/s12938-016-0201-2.

[21] P. Reed, M. Romano, F. Re, A. Roaro, L.A. Osborne, C. Viganò, R. Truzoli, "Differential physiological changes following internet exposure in higher and lower problematic internet users," PLoS ONE, **12**(5), e e0178480, 2017, doi: 10.1371/journal.pone.0178480.

[22] J.-S. Chang, E.-Y. Kim, D. Jung, S.-H. Jeong, Y. Kim, M.-S. Roh, Y.-M. Ahn, B.-J. Hahm, "Altered cardiorespiratory coupling in young male adults with excessive online gaming," Biological Psychology, **110**, 159–166, 2015, doi: 10.1016/j.biopsycho.2015.07.016.

[23] S.-J. Hong, D. Lee, J. Park, K. Namkoong, J. Lee, D.P. Jang, J.E. Lee, Y.-C. Jung, and I.Y. Kim, "Altered heart rate variability during gameplay in Internet gaming disorder: The impact of situations during the game," Frontiers in Psychoiatry, **9**, 429, 2018, doi: 10.3389/fpsyt.2018.00429.

[24] K. Ioannidis, S.R. Chamberlain, M.S. Treder, F. Kiraly, E.W. Leppink, S.A. Redden, D.J. Stein, C. Lochner, J.E. Grant, "Problematic internet use (PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry," Journal of Psychiatric Research, **83**, 94–102, 2016, doi: 10.1016/j.jpsychires.2016.08.010.

[25] M.J. Rho, J.-E. Jeong, J.-W. Chun, H. Cho, D.J. Jung, I.Y. Choi, D.-J. Kim, " Predictors and patterns of problematic Internet game use using a decision tree model," Journal of Behavioral Addictions, **5**(3), 500–509, 2016, doi: 10.1556/2006.5.2016.051.

[26] S. Aggarwal, S. Saluja, V. Gambhir, S. Gupta, S.P.S. Satia, "Predicting likelihood of psychological disorders in PlayerUnknown's Battlegrounds (PUBG) players from Asian countries using supervised machine learning," Addictive Behaviors, **101**, 106132, 2020, doi: 10.1016/j.addbeh.2019.106132.

[27] D. Seth, S. Eswaran, T. Mukherjee, M. Sachdeva, "A deep learning framework for ensuring responsible play in skill-based cash gaming," in 2020 IEEE International Conference on Machine Learning and Applications (ICMLA), 454–459, 2020, doi: 10.1109/ICMLA51294.2020.00077.

[28] H. Kim, J. Ha, W.-D. Chang, W. Park, L. Kim, C.-H. Im, "Detection of craving for gaming in adolescents with Internet gaming disorder using multimodal biosignals," Sensors, **18**, 102, 2018, doi:10.3390/s18010102.

[29] S.W. Wilson, "Classifier fitness based on accuracy," Evolutionary Computation, **3**, 149–175, 1995, doi: 10.1162/evco.1995.3.2.149.

[30] S.W. Wilson, Get real! XCS with continuous-valued inputs. Lanzi et al. (eds) Learning Classifier Systems: From Foundations to Applications. Springer-Verlag, Berlin, 1813 of Lecture Notes in Artificial Intelligence, 209–219, 2000.

[31] H.-M. Ji, T.-C. Hsiao, "A novel cue-induced abdominal reaction analysis for Internet gaming disorder," Journal of Medical Systems, **43**, 94, 2019, doi: 10.1007/s10916-019-1221-9.

[32] M. Griffiths, "Nicotine, tobacco and addiction," Nature, **384**, 18, 1996, doi: 10.1038/384018a0.

[33] V.A. Santos, R. Freire, M. Zugliani, P. Cirillo, H.H. Santos, A.E. Nardi, A.L. King, "Treatment of Internet addiction with anxiety Disorders: Treatment protocol and preliminary before-after results involving pharmacotherapy and modified cognitive behavioral therapy," JMIR Research Protocols, **5**, e46, 2016, doi: 10.2196/resprot.5278.

[34] W. Li, E.L. Garland, J.E. OBrien, C. Tronnier, P. McGovern, B. Anthony, M.O. Howard, "Mindfulness-oriented recovery enhancement for video game addiction in emerging adults: preliminary findings from case reports," International Journal of Mental Health Addiciton, **16**, 928–945, 2018, doi: 10.1007/s11469-017-9765-8.

[35] A.R. Melero Ventola, J.R. Yela, A. Crego, M. Cortés-Rodríguez, "Effectiveness of a mindfulness-based cognitive therapy group intervention in reducing gambling-related craving," **20**(1), 107–134, 2020, doi: 10.24193/jebp.2020.1.7.

[36] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proceedings of the Royal Society A, **454**(1971), 903–995, 1998, doi: 10.1098/rspa.1998.0193.

[37] Z. Wu, N.E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," Advances in Adaptive Data Analysis, **1**, 1–41, 2009, doi: 10.1142/S1793536909000047.

[38] Y.-C. Chen, T.-C. Hsiao, "Towards estimation of respiratory muscle effort with respiratory inductance plethysmography signals and complementary ensemble empirical mode decomposition," Medical & Biological Engineering & Computing, **56**(7), 1293–1303, 2018. doi: 10.1007/s11517-017-1766-z.

[39] N.E. Huang, Z. Wu, S.-R. Long, K.C. Arnold, X. Chen, K. Blank, "On instantaneous frequency," Advances in Adaptive Data Analysis, **1**, 177–229, 2009, doi: 10.1142/S1793536909000096.

[40] M.V. Butz, S.W. Wilson, An algorithmic description of XCS. In Lanzi, P., Stolzmann, W., and Wilson, S., Ed., Advances in Learning Classifier Systems. IWLCS 2000. Lecture Notes in Computer Science, vol 1996. Springer, Berlin, Heidelberg, 253–272, 2001, doi: 10.1007/3-540-44640-0_15.

[41] Y.-H. Lin, Y.-C. Pan, S.-H. Lin, S.-H. Chen, "Development of short-form and screening cutoff point of the smartphone addiction inventory (SPAI-SF)," International Journal of Methods in Psychiatric Research, **26**(2), e1525, 2017, doi: 10.1002/mpr.1525.

[42] D.-L. Hsieh, H.-M. Ji, T.-C. Hsiao, B.-S. Yip, "Respiratory feature extraction in emotion of Internet addiction abusers using complementary ensemble empirical mode decomposition," Journal of Medical Imaging and Health Informatics, **5**, 1–9, 2015, doi: 10.1166/jmihi.2015.1405.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 212-222 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Initial Experiments using Game-based Learning Applied in a Classical Knowledge Robotics in In-Person and Distance Learning Classroom

Márcio Mendonça[*,1], Rodrigo Henrique Cunha Palácios[1,2], Ivan Rossato Chrun[3], Diene Eire de Mello[4], Henrique Cavalieri Agonilha[5], Elpiniki Papageorgiou[6], Konstantinos Papageorgiou[6]

[1]*Programa de Pós-Graduação em Engenharia Mecânica (PPGEM-CP), Universidade Tecnológica Federal do Paraná, Av. Alberto Carazzai, 1640, Cornélio Procópio, Brazil*

[2]*Programa de Pós-Graduação em Informática (PPGI-CP), Universidade Tecnológica Federal do Paraná, Av. Alberto Carazzai 1640, Cornélio Procópio, Brazil*

[3]*Departamento de Engenharia Química, Universidade Estadual de Maringá, Av. Colombo 5790, Maringá, Brazil*

[4]*Programa de Pós-Graduação em Educação (PPEDU- UEL), Universidade Estadual de Londrina, Av. Colombo 5790, Maringá, Brazil*

[5]*Departamento de Acadêmico de Computação (DACOM), Universidade Tecnológica Federal do Paraná, Av. Alberto Carazzai 1640, Cornélio Procópio, Brazil*

[7]*Department of Energy Systems, University of Thessaly, Geopolis Campus, 41500 Larissa, Greece*

ARTICLE INFO

Article history:
Received: 16 March, 2021
Accepted: 13 July, 2021
Online: 27 July, 2021

Keywords:
Digital games
Game-based learning
Mobile robotics
Robotic manipulators
Autonomous robotics

ABSTRACT

This paper addresses experiments with Scratch-developed games in the robotics introduction course at the Federal University of Technology – Paraná. It aims at assisting learning of classical and initial robotics concepts. This proposal, similar to the classic 80s war tanks game on Atari 2600, was developed using an autonomous vehicle. In the first experiment, applied to the class 2019/2, the students (players) had to battle against another autonomous tank developed (in two different ways, in Person and Distance Learning), using keyboard inputs to control their tank. In this game, the students were asked to create states machine models while were being introduced to fundamental concepts such as pose, other basic notions concerning controlled and autonomous robots, and the hierarchy of actions. At the end of the games, a questionnaire answered by the students extracted valuable findings of the examined concepts. In 2021/1 class, the second and third experiments were applied. The former was an extension of the first experiment, using autonomous parking cars. The latter was inspired by the classic Pong game, with the addition of more degrees of freedom (DOF). In this case, the player attempts to reach and catch a ball through the operation of a robotic arm with two rotating joints, using keyboard inputs. Each block or scenario will become more complex, and the student has time to perform a task. In the case of the third experiment, the concepts including 2-D workspace, multiple solutions, inverse, and direct kinematics were explored. Delivery rates for the first and second experiments were 90% and 80%, respectively. Even though three individual experiments were investigated, the single objective was achieved: the implementation of modern didactic tools to deliver critical pedagogical concepts to students in the robotics class.

1. Introduction

The first 20 years of 2000 have been highlighted by the development and creation of computer technologies. This study is devoted to the integration of certain aspects of these technologies to either the working or leisure areas of everyday human life. In the field of leisure, games have been transformed into digital, triggering increased attention to users of all ages around the world,

[*]Corresponding Author Márcio Mendonça, Email: mmendoncautfpr@gmail.com

https://dx.doi.org/10.25046/aj060425

while the extended popularization of the web helped in this direction. Electronic games are characterized by an enormous variety of levels in terms of their type and difficulty, targeting a wider age group, no longer being a niche market [1], as it used to be in the past.

This paper is an extended version of the previous work published in ILSA 2020 [2] approaching game-based learning. The proposal of active methodologies such as game-based learning aims to involve students by establishing a process that implies action-reflection-action and not just internalization of what has been exposed (as is the case in in-person classes) [3].

With Game-based learning, students learn while playing. Thus, making the learning process more enjoyable, causing a positive effect on cognitive development. Games are combined with traditional classes because the conventional learning process can be monotonous, and game-based learning can improve students' motivation to learn. It is not just about using games to review and reinforce concepts [4].

The games include many problem-solving features, adding elements of competition and opportunity. That is, the student player needs to deal with an unknown result, choose between several paths to an objective, construct a context of the problem, and collaborate with several players [4].

Among the benefits of games implementation in learning, according to [5], the following attributes should be mentioned:

- Games can easily attract the attention of individuals across various demographic boundaries (for example, age, sex, ethnicity and educational status).
- Games can assist young people in setting their goals as they can provide feedback and reinforcement or record changes in human behaviour.
- Games offer fun and excitement to the players. Hence, it is not difficult to attract and maintain a person's attention.
- Games also offer the chance to participants to explore their curiosity and new challenges, thus stimulating their motivation for learning.

On that basis, gamification is the practice of using elements of game design, mechanics and thinking in non-game activities to motivate participants. Gamification in the field of education exploits, among others, sets of games rules, players' experiences, and cultural roles in shaping students' behaviour [4]. Thus, it uses points, badges, ratings, and incentives to engage students in the learning process.

The benefits of gamification in education are [4]:

- Improved learning experience.
- Enhanced learning environment.
- Instant Feedback.
- Promotion of behavioral changes.
- Feasibility to integrate into different learning needs.

Nowadays, the utilization of novel learning technologies has overcome the conventional problems of distance, time, and cost in learning. However, the lack of student motivation is a problem that e-learning still faces. The application of gamification in e-learning is being used aiming at giving participation and increasing student motivation. Within the same learning content, the characteristics of different users and static gamification elements do not increase the expected motivation. To overcome this problem, gamification must be adapted to the characteristics and needs of the learners [6].

The development of computers, tablets, and smartphones with enhanced capabilities derived from improved processing power and low cost has boosted the popularity and wide use of digital games among young people. On this basis, researchers have recently focused on exploring the intrusion of games into the teaching and learning process. In this context, the authors, through the examined scenario, investigate how technological means could improve the teaching-learning process using digital games.

This approach is based on traditional teaching and learning methods, whereas it exploits diverse aspects of the learning process [7]. Additionally, it utilizes question games, where students are rewarded with marks when delivering the correct answers, thus creating active cooperation and healthy competition, which stimulates the learning process; this process is also known as gamification. This gamification scenario has been appropriately applied to exhibit the contribution of games in the learning processes in students' interests [8].

The successful integration of games in education to complement traditional learning is highly attributed to the valuable features of games, including the playful aspects, interactivity, feedback, problem-solving, experimentation, competition, and students' engagement in learning [7, 9]. To instantiate the methodology, the work [10, 11] presents, in addition to learning the proposed concepts, that the games favour students' cognitive and social development through the solution of problems and cooperation between them. An example that can be cited in the literature is the work with children playing with robots. Some elements are relevant in game development and were used in this article, shown as follows.

As a motivation for this research at university level, we can mention the work that uses game-based learning in conjunction with other methods as a teaching technique for children. In [12], a novel teaching structure assisted by a computer was presented for teaching maths in the 5th grade. Based on an award-winning curriculum program, this approach uses music and body gestures to develop certain connections between mathematical concepts and culturally inspired metaphors. Utilizing Virtual Augmented Reality (VAR) and sensors, the presented approach deals with the successful transformation of the class from traditional to digital.

Another factor that supports this research is the fact that the Brazilian Computer Society (SBC) considers the basic concepts of computing as important as those of mathematics, philosophy, physics, and other sciences for contemporary life. Thus, computer science, robotics and digital games have found new adherents with meaningful pedagogical experiences.

It is not the scope of this research to carry out any statistical analysis, only to present roughly the percentage of students who

suggested having abstracted the classical knowledge of robotics in each of the proposed experiments.

Digital games, since their introduction in 1974, have taken tremendous and rapid steps towards improving players' overall gaming experience and involvement. Actually, game developers intended to keep people engaged in playing their game uninterruptedly, targeting new goals every time, being determined to experience new challenges. Even though today's children have a different view towards the use of video and computer games than games were meant to offer, students still show an optimistic attitude characterized as interested, competitive, cooperative, results oriented. At the same time, they actively seek information, fundamentals (which is the purpose of this research, toward robotics) and even solutions [1].

As for the application of games based on Atari, a relevant consideration suggests a possible motivation of students in the application of games, even if it is worth mentioning that: the use of DeepMind Technologies Limited acquired by Google in 2014, has been breaking all records of Atari games and is capable of challenging any human to a match [13]. This motivated students about the importance of digital games in intelligent computational systems and for learning robotics. That said, the students answered questionnaires and developed state machines.

This article is divided into five sections. Section 1 presents the introduction and a brief review of the literature. Section 2 presents the theoretical aspects of games and learning. Section 3 presents the development of the research developed. The results are presented in Section 4, and, finally, in Section 5, the overall conclusions and discussions are presented.

2. Games and learning: theoretical aspects

This approach aims at improving learning, using the motivating effects of the elements and techniques of digital games. Student engagement is the criterion for integrating gamification into the learning process and therefore serves as an essential measure for its effectiveness. However, gamification should not be restricted in considering only games' main concepts such as points and leaderboards, which significantly reduce its educational value and the desired impact on students, who constitute the main target of this attempt [14].

Among the advantages of game-based learning that motivates the current study is the ability to provide learners with the understanding of concepts in a practical way, taking the student to a higher level of involvement with his learning in a more dynamic way. In this context, the work of [15] presents a model under study which illustrates the fact that an intense engagement in learning has a strong effect on human body as dopamine and serotonin do. Specifically, this neural model was based on the assumptions that dopaminergic activity increases as the expected reward increases and serotonergic activity increases as the expected cost of an action increases.

Following a brief history review regarding learning and technology, the view of Robert McClintock, Frank Moretti and Luyen Chou is devoted to the evolution of technology which goes side by side with the change and transformations in teaching and learning. Originally, education and training were a process of imitation and training - "picking up a stone and playing with the animal." If you are unable to do this the first time, practice several times until you succeed. "No, do it this way." The practice has become a way of playing to make this repetitive skill-based learning bearable and memorable. This type of "demonstration and practice" learning requires good coaches, usually in an individual relationship. This is how people learn to play sports, play musical instruments, and master other physical skills. In the most basic, not even language is necessary, e.g., athletes and musicians are often skillfully trained by people who do not or barely speak the same language [1].

According to [16], since the 1980s, some researchers carried out certain studies concerning the use of games in education and its benefits and approaches. In [17], the authors reported the popularization of gamification started only in 2010. The term gamification refers to the use of game elements, as aesthetics, game thinking and mechanics, in non-game-related contexts to involve people, motivate action, improve learning, and solve problems.

It also involves several concepts, such as rewarding and punishing the players. It has several connotations, such as the case of this research in which games are developed to assist in the learning process, in commercial games, or even in the commercial area. For example, in [6], the authors use a Systematic Literature Review (SLR) to explore adaptive gamification in terms of frameworks and methods proposed, as well as other research components. The first step is to define the research question (RQ) and then to search the literature published in popular scientific journal databases. Twenty-five selected articles were finally reviewed, in which the authors identified three elements that comprise the proposed framework. These are adaptive gamification engine, adaptive component, and gamification display. Additionally, eleven types of methods were implemented in adaptive gamification. Among them, Felder-Silverman Learning Style Model (FSLSM) is the most popular method. As for the components of adaptive gamification, four were mined, namely: player/learner profiles, learning style, behavior, and skill/knowledge.

Specifically, the following examples are related to this research. The authors in [18] conclude that in recent years, interest in the theme has been increased to such extend so that a theoretical game model was developed for educational purposes. In this case, not any particular game element could be used for gamification, except a subtly combination of elements that are used to contextualize learning. The initiative behind the game development was the integration of various types of activities into this game, as well as assisting teachers in working with modernized learning content.

Another study considers games as powerful experiences that exploit motivation and engagement [17]. In particular, the deployment of simplified elements reduces project complexity concerning badges, levels, points, and leaderboards, that fail to involve students and damage any existing interest for the learning process. A thorough consideration regarding game design must be given in gamification, apart from just implementing game components successfully. However, gamification is a broader

concept used in different areas of knowledge, for example, the management area [6].

In [19], it was observed that educators could increase the feedback mechanisms by exploiting specific game design elements by applying continuous feedback, visual cues, frequent question and answer activities as well a progress bar. Students are also prompted to interact with the content, experience real-life case studies, make decisions for specific tasks, and have realistic consequences for making wrong or unsatisfactory choices. In this way, they are getting involved within the game flow, which attracts students' attention. In conclusion, satisfactory learning outcomes come when facts are embedded adequately inside a game-based story rather than in a passive text format.

In the example of Sheldon's work [20], the professor of a higher education institution gamified his electronic games development class. Students' marks were starting from zero and increasing with their results. Moreover, there was an increase in the number of game-based activities against the traditional assessments. These activities had the form of missions to defeat enemies and were assigned to groups of students. The outcome of each mission was assessed and formed the student's final grade. Throughout this learning process, it was observed that the average grades of the students were notably improved at the end.

Gee also noted that students perform better when a level system is used in the game-based learning process. The players are more inspired when they apply what they have learned and get the corresponding feedback after completing each level. Each subsequent level requires, after all, skills acquired at previous levels [21]. The authors in [22] conclude that the game's dynamics are improved when levels or progress bars are involved. In addition, storytelling is another game design aspect that highly contributes to the successful integration of games into the learning process. Some examples of these types of games include SimCity, where players follow the story of building a city from scratch, as well as Monopoly, where the story behind the game is to become rich buying and selling property in every round, trying to avoid the risk of losing everything [16].

After an in-depth review of the available literature, certain characteristics were found mainly in [16,17,19-21] regarding the successful development of games integrated into learning environments. These include Freedom to Fail, Quick Feedback, Progression and Counting stories.

3. Development

There are several approaches that deal with the integration of digital games in learning environments. The simplest paradigm refers to the utilization of simple games used exclusively in the classroom whereas, more complex examples focus on the design of digital game concepts that are put into practice but are not necessarily used as learning simulators in virtual or real educational environments [7].

It is not the scope of this work, but it is possible to use board games, for example, as a learning method. Several research works can be found in the literature which study board games-based learning. More specifically, the authors in [7] propose cooperative learning as a means to design an educational course based on the board game. According to the curriculum, pre-service teachers cooperate in developing and delivering a board game prototype.

Because of the game's simplicity, the tutorial with the game instructions was presented during the beginning of the class in Brazilian and Portuguese (Brazil is the native language) and is presented in the metadata of this work.

In the first stage, a group discussion in an online platform took place where the participants were working on the fulfillment of specific activities according to the curriculum. Next, a questionnaire evaluated the performance of the pre-service teachers in the particular task. Based on the results produced, all participants' groups were able to successfully design games in line with the academic disciplines that are played sufficiently and without difficulties. Additionally, teachers' self-efficacy was improved through the process of cooperative learning.

Another study in this area that contributes to learning [23] presents a novel structure devoted to the general board game (GBG). GBG configures boardgame's standard interfaces, states and AI agents. This structure offers cooperation with various agents in different games and defines the parts of board game coding. GBG is particularly suitable for arbitrary 1, 2, 3 or multiplayer board games.

This work proposes three different games developed individually. The first one is devoted to designing a tanking battle game based on a classic Atari 2600. This game includes the definition of some concepts regarding robotics discipline fundamentals. These are the pose concept that complements the Cartesian coordinates (x, y) and the angle, which denotes the object's orientation. Two more games are also presented: one consisting of a manipulator arm and a ball, while the last one deals with car parking.

Goals: each experiment has its own distinct goals. In tanks, the objective is to destroy the opponent; in the robotic arm, the goal is to reach the ball and, finally, in the valet, a parking space must be found.

Rules: In the case of tank war, the opponent must be destroyed before the player is destroyed; in other games, the objectives must be accomplished in the shortest time possible.

The proposed digital game was built on the Scratch platform, which was developed by the Massachusetts Institute of Technology (MIT) team in 2007. Scratch is among the most accessible programming languages as it uses a graphical interface and does not require programming skills on any other programming language [24]. The scope of this game is to insert blocks, as shown in the example in Figure 1.

The methodology of the learning process was initially the presentation of various games, such as Pong, which was the first game of the era. Then an explanation followed on how the games work, and next, the commands and objectives were exhibited. In the next phase, the students were introduced to relevant areas of robotics and/or virtual agents related to the tank war game and continued with activities concerning the following two experiments presented.

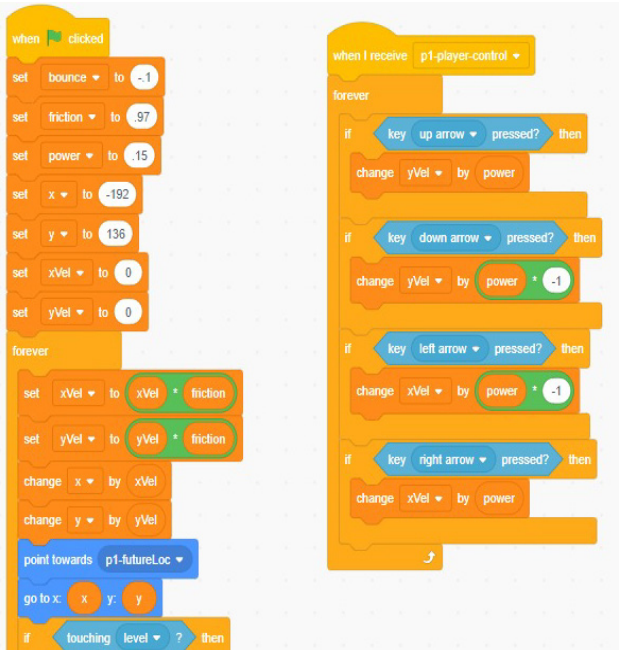


Figure 1: Screen example using Scratch.

4. Results

Three games built on Scratch were presented in different classes. The tanks war (experiment 1) and the autonomous valet (experiment 2) games were presented to the 2019/2 class, and various responses were collected. In sequence, the game with two robotic arms (experiment 3) was illustrated, in which students tried to simply catch a ball (target) by rotating the two arms joints (2 degrees of freedom, GDL) with the help of a keyboard. Concerning the first experiment, 27 students were involved. However, just a small number of results for all three experiments is presented for the purposes of this study due to the limited workspace.

4.1. Experiment 1

The purpose of this exercise was to assist students in understanding certain robotics concepts that concern the need for autonomy and hierarchy when actions are taken. Priority is given to some control actions in the scenario, such as obstacles avoidance and targeting the opposing tank. Additionally, new routines can be applied in the development of future works based on intelligent computer systems. For example, the opponent can become an autonomous entity by implementing intelligent computational methods based on Fuzzy logic. Fuzzy Cognitive Maps can help in this direction in the future, as it provides low computational complexity [25, 26].

In other words, this experiment added the abstraction of a robotic architecture inspired by fundamental concepts of Brooks' subsumption architecture [27]. The notion of priority when a robot, or in this game's case a virtual robot (bot), must prioritize its actions, such as staying alive, avoiding enemy attacks and obstacles before shooting the enemy.

In the initial phase of this game, as shown in Figure 2, both tanks are guided by players. The main objective of the first tank is to destroy the second by shooting at it while trying to avoid other fixed and mobile obstacles (opponent's shots). The next phase of the game design includes determining the game rules and then establishing the possible states of the tanks.

The game rules are defined according to possible damage suffered by the tanks. The game has a winning condition for the first player when he manages three hits to the opponent's tank. One point of damage is attributed either by a shot or physical contact with an obstacle. The distinct states of a tank are the following:

- Moving freely.
- Shooting.
- Avoiding a fixed obstacle.
- Avoiding a moving obstacle (opponent's shot).

According to the exercise proposal in the experiment developed in the seventh period of Control Engineering and Automation, the following tasks need to be accomplished by students:

- Develop tank's finite state machine.
- Identify the tank's position (coordinates) and pose (angle with the x-axis).

In this experiment, we had an abstraction of the mobile robotics concepts of approximately 80% for one class and 90% for another, despite having initial results with applications in only two classes. Objectivized concepts like hierarchy and pose in mobile robotics were abstracted, suggesting that this game is promising for introducing the basic concepts of programmed and mainly autonomous mobile robotics.

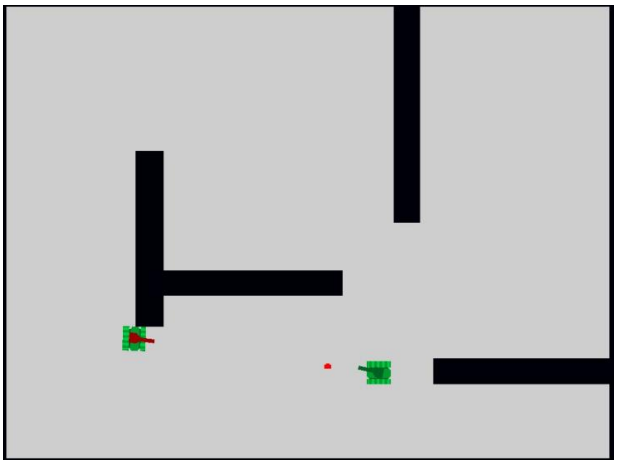


Figure 2: Representation of the first game

Based on the proposal, each student was asked to give a figure of the state machine and a description of the events in a doc or pdf file. In addition to the graphical representation, students were also asked to answer a group of questions to identify whether the nature of the activity made it possible for them to perceive aspects and concepts proposed. Such an activity can be considered as a Subjective Modeling when capturing the impression of the student in the short-term memory as listed below:

a) Is it possible to distinguish between the programmed and the autonomous control? Explain.

b) Is it possible to perceive the robot's hierarchy during the game, and in what way?

c) Is an exemplified angle during the battle necessary, in addition to the x, y coordinates?

d) Is It possible to perceive the different states, attacks and defences?

e) Develop a state machine that models the player's actions to defeat the opponent.

One exemplary and one complete answer provided by the students are listed below for each question:

Question a:

- Student 1: Yes, because the user-controlled tank responds to the given commands, whereas the autonomous only responds to the events following its logical construction.
- Student 2: It is possible. The opponent's tank moves without any human intervention. On the other hand, the player's tank is controlled at each cycle by the keyboard.

Question b:

- Student 1: The hierarchy is noticed as the robot tank is looking for its own defense, and then is trying to destroy the opponent with the intention to win. Our tank has no hierarchy since it is controlled by the user.
- Student 2: Hierarchy is perceived when activating opponents, avoiding obstacles and the opponent's attack for keeping its integrity (game's objective).

Question c:

- Student 1: The angle is used to increase the control precision. For example, position x = 0, y = 0 is where the tank is located in the environment. When X = 1, the tank will not necessarily move forward, but it may be rotated.
- Student 2: The question was not answered.

Question d: only two students answered the questions. However, the state machine abstraction was the main objective of this experiment.

Student 1: Student 1 provided a satisfactory interpretation of the robot's movement finite-state machine, shown in Figure 3. The possible states are: 1 - shoot avoidance; 2 - stopped; 3 - chasing the enemy; 4 - aiming at the enemy tank; 5 - shooting the enemy and 6 - obstacle avoidance.

Only two students were inserted due to a large amount of space for a more accurate analysis. However, students were chosen at random, and their responses can provide feedback on their learning from the most complex experiment, the tank battle. Other students have already managed to develop state machines for game strategy, which are shown below.

The response by the student is presented below.

a) Do not press anything.

b) Press arrow keys.

c) Mouse left-click.

d) Spin the mouse.

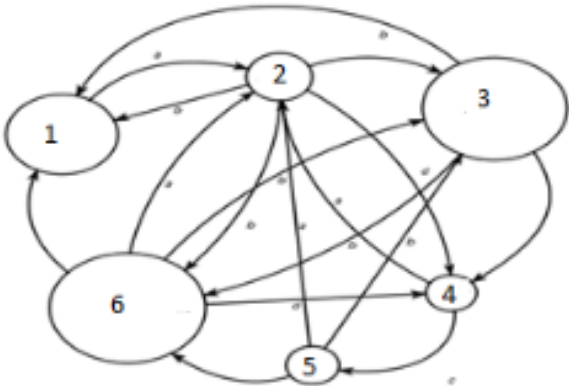


Figure 3: Finite-state machine by Student 1.

On the other hand, it is worth noticing that student 2 (Figure 4) provided a more complete interpretation than those submitted by the rest of the class. It seems to be more accurate and closer to the actual operation of the tanks in the game.

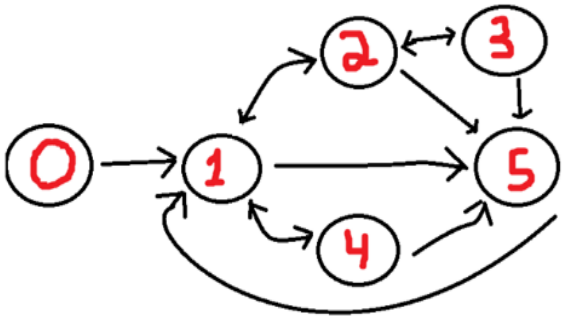


Figure 4: Finite-state machine by Student 2

- State 0: The game is stopped and has not yet started.
- State 1: The game started; tank not moving.
- State 2: The tank is moving.
- State 3: The tank moves, aiming and firing.
- State 4: The tank is stopped, aiming and firing.
- State 5: The tank is dead, awaiting re-entering to the game.
- State 0 → 1: Press the button to start the game.
- State 1 → 2: Press the directional buttons to move the tank.
- State 2 → 1: Release the movement buttons. Tank in motion.
- State 2 → 3: Press the shoot button to target. Tank in motion.
- State 3 → 2: Release the fire button. Tank in motion.
- State 1 → 4: Press the fire button while aiming at the target. Tank is stopped.
- States 1, 2, 3, 4 → 5: The tank got damage.
- State 5 → 1: The tank respawns after an x amount of time has passed.

Regarding student 3, he developed the state machine but left the vocabulary unfinished. Through this process, it emerges that the students have delivered various interpretations. Additionally, the students seemed to comprehend specific concepts in the field of robotics and their autonomous mode as well.

The graphical representation that student 4 provided (see Figure 5) illustrates the finite-state machine of the game combined with the respective vocabulary, offering a satisfactory interpretation of the activity concept. Although different interpretations emerged from students, these interpretations were similar concerning the game strategy. Overall, the number of different ways for learning assessment in each of the experiments needs to be highlighted.

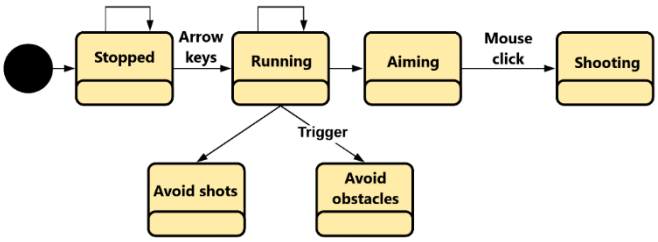


Figure 5: Finite-state machine by Student 4

In this tank warfare experiment, using the state machine was necessary due to different behaviours in combat, such as an attack, defence, avoiding obstacles, enemy shots, and chasing the enemy, among others. However, in the last experiment, four questions were enough to assess students' learning. In general, each experiment has a complexity level, whereas different evaluation methods were used.

State machines have been promising in terms of students' abstraction. Most of the students, about 40%, did not even know the state machines yet (because they are students of an Automation and Control course, whose robotics discipline is offered from the seventh period on). So, the results of this experiment showed promising abstraction of the game's strategy, especially the concepts of robotics involved, mentioned in the text.

New experiments took place in 2021 with distance learning due to the pandemic (University policy to mitigate Covid contagion). The difficulty of experimenting compared to previous ones before the pandemic was much greater [24]. It was necessary to do a live meeting and provide videos of the game's action to help students abstract the strategy and do their required tasks. Figure 6 shows details of the video provided to students.

UTFPR-CP, a federal university (in Portuguese Universidade Tecnológica Federal do Paraná Campus Cornélio Procópio), gave a subsidy to low-income students to purchase computers. In addition, it also promoted aid in cash to low-income students, which was also maintained during the pandemic. Only three examples (see Figures 7, 8, 9) will be presented, which took place in a distance learning class during the pandemic.

The ICT Panel COVID-19 (2020), which took place with young people aged 16 or over, showed that the cell phone was mentioned by 22% of users in social classes A and B, 43% of users in class C and 54% of users in classes D and E. The inequalities of student access to connected devices are striking. The same document also points out: three-quarters of Internet users aged 16 or over in classes D and E (74%) used the network exclusively by cell phone, a percentage that was 11% among users of classes A and B (no reference because it was written in Portuguese).

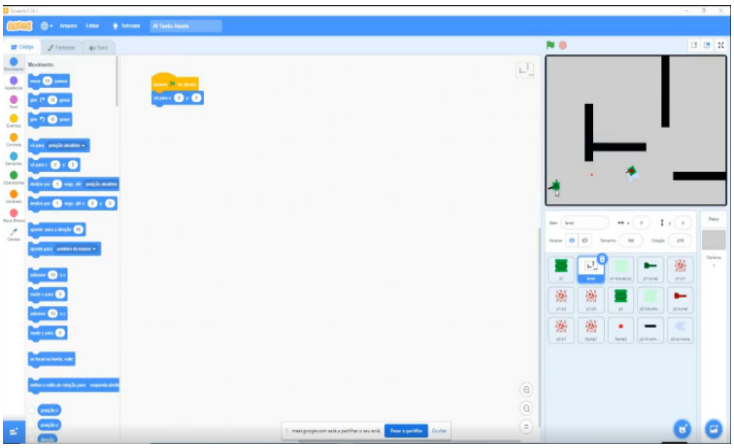


Figure 6: shows details of the video provided to students.

Student 1 (Distance Learning):

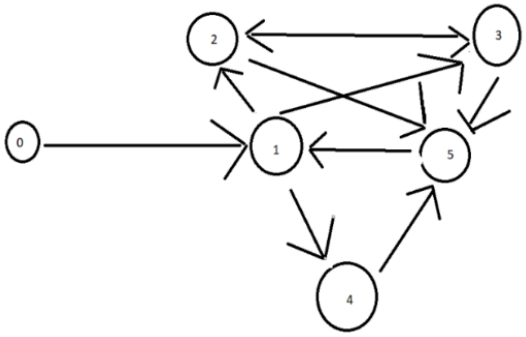


Figure 7: Finite-state machine by Student 1 (distance learning)

Vocabulary state machine.

- State 0 → 1: Start the game.
- State 1 → 2: The player moves the tank without shooting.
- State 1 → 3: The player moves the tank and shoots.
- State 1 → 4: The player stops the tank and shoots.
- State 4 → 5: The tank was stopped and died.
- State 2 → 5: the tank was moving and died.
- State 3 → 5: the tank was moving and shooting and died.
- State 5 → 1: The tank was dead and reappeared.
- State 3 → 2: Tank moving stopped shooting.
- State 2 → 3: Tank moving started shooting.

Student 2 (Distance Learning):

Vocabulary state machine.

- State 0 → 1: Starts the game, with the tank stopped.

- State 1 → 2: Tank is moving.
- State 2 → 3: Tank is moving and firing.
- State 3 → 5: The tank, while firing, dies and awaits reappearance.
- State 5 → 1: The tank was dead and reappeared.
- State 1 → 4: Stopped Tank aims and shoots.
- State 4 → 5: Tank was stopped and died.
- State 5 → 1: The tank was dead and reappeared.

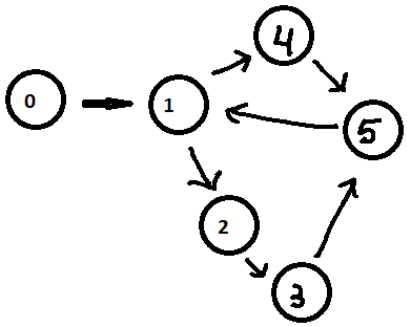


Figure 8: Finite-state machine by Student 2 (distance learning)

Student 3 (Distance Learning):

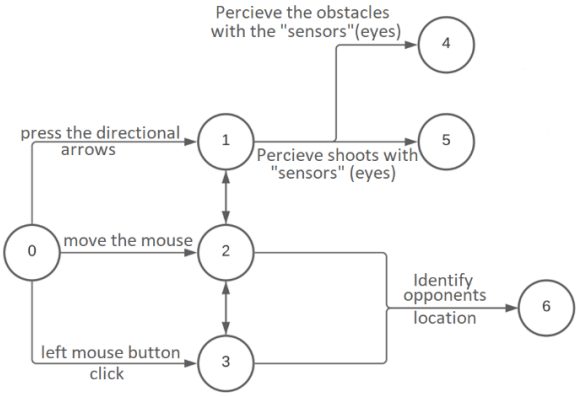


Figure 9: Finite-state machine by Student 3 (distance learning)

Vocabulary state machine.

- State 0: Game start.
- State 1: Player walking.
- State 2: Player aiming.
- State 3: Player shooting.
- State 4: Player dodging obstacles.
- State 5: Player dodging shots.
- State 6: Player hitting opponent.

The first two results were compatible with the experiments carried out in the classroom. The third result presented was one of the best in abstraction, if not the best of all the experiments presented in distance learning and on-site teaching. In addition to the states, the figure shows the events used to control the player's tank.

4.2. Experiment 2

Another simpler game was produced on Scratch, contributing to this work. In this game, the student must drive a car to find a parking space (autonomous valet), as shown in Figures 10 and 11. It was held in the same class (2019/2) following the tank war game.

It was also introduced in the robotics discipline, aiming to present students a brief notion of the maneuver difficulty, even if performed on the keyboard, by an autonomous valet using Fuzzy logic as an example.

The coexistence of several agents or virtual robots in the same scenario was emphasized. It is noteworthy that in the discipline of robotics, students have a posterior example of an autonomous valet. This game reinforces the concepts of tank warfare: autonomy and hierarchy of control actions. That is, it is necessary to avoid the cars reaching a parking space.

One student's answer: despite being a relatively simple activity for human beings, the game allowed us to observe the difficulties of developing an autonomous system, which would make these maneuvers without any human intervention. It helped us to understand a little more about autonomous robotics.

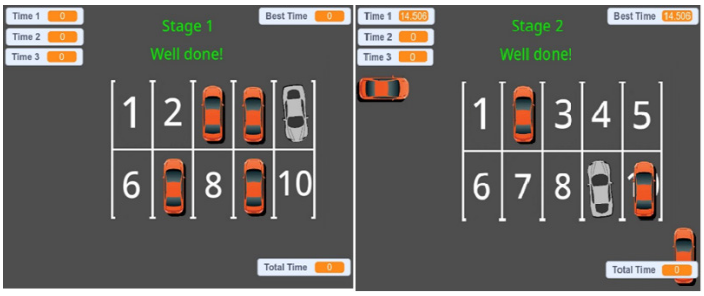


Figure 10: Student 1 answers; initial stages

Compared to Experiment 2, different results were obtained and aided in the autonomous vehicle abstraction. However, two of several examples will be cited in which can be seen that the second student was faster than the first one, as shown in Figures 10 to 13. Nevertheless, practically all students were successful in all three stages. The first objective is to help develop an autonomous valet through Fuzzy logic, for example, fuzzy cognitive maps [28].

The course has already clarified the difference between programmed and autonomous robots, a relevant concept in robotics [29].



Figure 11: Student 1 answers; final stage

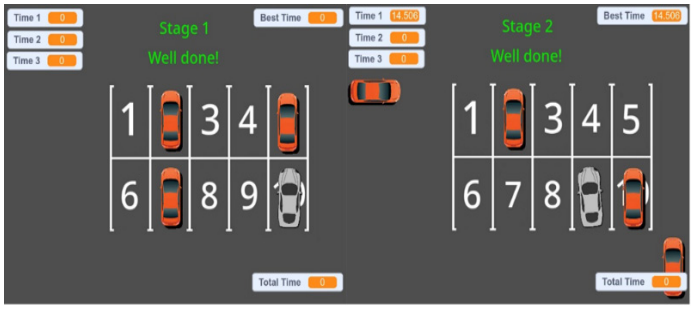


Figure 12: Student 2 answers; initial stages

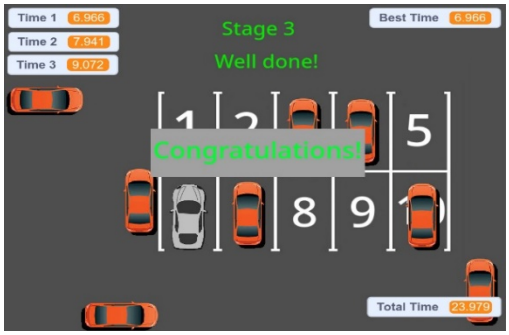


Figure 13: Student 2 answers; final stage

4.3. Experiment 3

The robotic arms game (experiment 3) was applied in the class of 2020/1, which has approximately 23 students. The game works as follows: a ball (target) appears on the scene at random as shown in Figures 14 and 15. The student's goal is to try and catch it using the cursors.

The game objective is to emphasize on direct and inverse cinematics concepts and the action radius (the robot inserts in a 2D environment). Future work could also implement a 3D game addressing a sphere or a search surface. In addition to the objectives, a classic problem regarding robotics manipulation became clear to students: the multiple solutions that can occur.

According to the figures presented, the game strategy is oriented to reaching three targets in sequence within the space area. However, the ball's movement cannot be achieved in certain situations, especially when the robotic arm is out of reach. It should be mentioned that the robotic arm has two degrees of freedom and also have rotational joints.

Due to workspace restrictions in this paper, the presentation of the answers to the provided questions will be limited, and only a few examples will be presented. It also needs to be highlighted that the results illustrated below were randomly selected to maintain the authenticity and truthfulness of the current research. Experiment 3 involves the following questions:

a) Is it possible to use more movements to catch the ball?

b) Would the movement become more natural if the joints move at the same time?

c) Does the game have direct or inverse kinematics?

d) Are there inaccessible positions of the arm?

In the meantime, the results in Figure 15 should be considered as they are the second successful attempt for student 2. The answers of five students are shown below for each of the four questions.

Question a:

- Student 1: Since the arm has two DOF, it is possible to have more than one solution.
- Student 2: It is possible for some ball positions, according to the image of the first line, in the same position, that there is more than one way to capture it. However, in the second line, there is only one possibility.
- Student 3: Yes, the movement is more harmonious when the joints move simultaneously, as this situation results in faster movements to reach the ball, more agile and more natural compared to a real human arm.
- Student 4: Yes, with the combination of the angles and the ball's position.
- Student 5: Yes, more than one solution is possible.

Question b:

- Student 1: Yes, this process would require greater computational power, but it would have a more harmonious movement closer to the human arm movement.
- Student 2: It would be, but it is easier to use keys to move both together simultaneously in the game.
- Student 3: Yes, it is possible to have different solutions to catch the ball, as each player has the freedom to choose the arm's movements by joining different angles for each joint but reaching the same point in the Cartesian plane.
- Student 4: Yes, it mimics the human body more.
- Student 5: Yes.

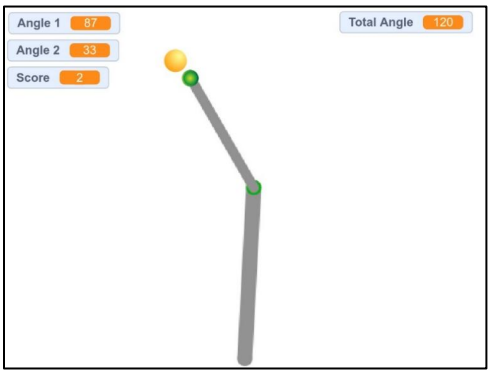


Figure 14: Student 1 result: first attempt

Question c:

- Student 1: Inverse, as the movement orientation of the joints occurs through the ball's position.
- Student 2: Inverse kinematics because the arm has a desired position and orientation (ball).

- Student 3: The game has inverse kinematics, as the player must define the joints' positions after a given ball orientation.
- Student 4: Inverse, the arm needs to follow the position of the given ball.
- Student 5: Direct, because the angles are already given, so the position of the arm can be found. In the game, we are going after the target "ball" visually.

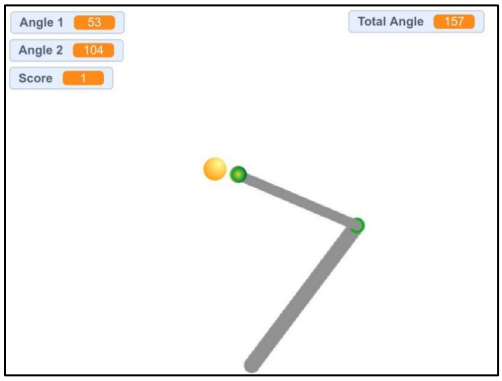


Figure 15: Student 1 result: second attempt

Question d:

- Student 1: Yes.
- Student 2: Yes, according to the figures, the ball travels practically the entire space of the screen and some points that are not unattainable.
- Student 3: The question was not answered.
- Student 4: Yes, so that a key is needed in the game to reset the ball.
- Student 5: Yes, then there is the ball moving.

The given answers emerge that they were similar since students could identify the problem just by playing the game. The correct answer to the fourth question is that the arm cannot reach the ball (target) because of its structure. Moreover, student 4 gave a more consistent answer to the second question, explaining his response assertively. Overall, the game seemed to have its objective fulfilled. In specific, considering a sample corresponding to approximately 20% of the class, it emerges that some answers were more competent than others. However, it is concluded that there was an abstraction of the contents and objective foundations.

Figures 14 and 15 illustrate in a more precise way the structure of the game's components and how the robotic arm changes its position to reach its target (ball), according to the game's random logic. The experiment was carried out in pairs, and the students had to repeat it three times and search for the target. The game attempted to leverage active learning [26] through objective concepts visualization. Additionally, the game demanded multiple solutions from students, as they had to find the target near or similar positions with different angles.

To provide a further understanding of the game, it is necessary to examine the arm's angles. Based on the representation inspired by the geometric model of MathWorks (see Figure 16), the angle θ1 is formed by the intersection of Arm 1 (L1) and the X-axis. The angle θ2 is formed by the continuation of Arm 1 (L1) and Arm 2 (L2), considering that (X, Y) is the desired position.

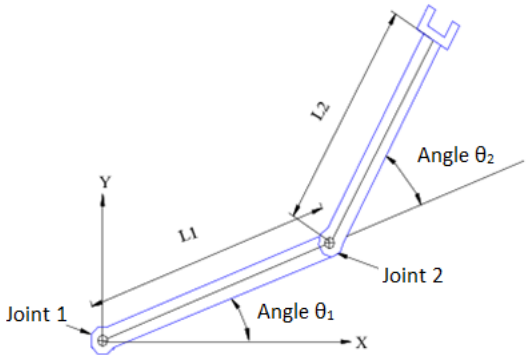


Figure 16: Representation of 2-DOF robotic manipulator.

5. Conclusion

The results, although initial, suggest the feasibility of the Scratch's proposals based on classic games of the 80s. The most important part of the experiments was the fact that the students experienced and became familiar with significant concepts in the domain of robotics through playing games. Among the produced outcomes was students' perception who considered the class "lighter" and liked the experiments. In short, the results were seemed convincing in terms of game-based learning since over 80% of the students had shown interest in these experiments, even if they were relatively simple. Therefore, the application of this methodology is considered quite promising.

In addition, it was possible to carry out some examples of experiments 1 and 3 in the distance learning modality, with a slight reduction in the percentage of students' interest, approximately 70%. Furthermore, experiment 3 had one of the most complex state machines and one of the best (if not the best) of the tank battle game, as mentioned above, which reinforces the promising results of the method.

The participation and motivation of the students in general in both classes was over 80%. According to the few results of the experiments, it can be observed that the objectives of introducing the basic concepts of robotics initially proposed were abstracted by the students with an even higher percentage than the motivational. That is approximately 85%.

Future works can emphasize the control of the opponent's tank and the appearance of obstacles like trees in the tank battle game. For the robotic arm, a 3D arm can be introduced to increase the difficulty of reaching the target significantly. Thus, the solution of the inverse kinematics would be more complex and the number of solutions to the problem. And finally, future work can be oriented to the development of realistic prototypes that would increase the difficulty of the game and introduce and utilize soft computing based on Fuzzy Cognitive Maps (FCMs). Considering that this intelligent technique has low computational complexity, it could be exploited to construct prototypes of low computational and financial cost, such as Arduino. Finally, as mentioned, the next generation of the games may also deploy other programming platforms, for instance, UNIT. Overall, considering the promising

results of distance learning, it would be interesting to conduct new experiments.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors gratefully acknowledge the Federal University of Technology – Paraná, campus Cornélio Procópio.

References

[1] M, Prensky, Digital game-based learning. McGraw-Hill, 2001.
[2] M. Mendonça et al., "Digital Game-Based Learning in a Robotics Course," in 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA), 1–7, 2020, doi: 10.1109/IISA50023.2020.9284366.
[3] R. Al-azawi, F. Al-faliti, M. Al-blushi, "Educational gamification vs. game based learning: Comparative study," in International Journal of Innovation, Management and Technology, **7**(4), 132–136, 2016, doi: 10.18178/ijimt.2016.7.4.659.
[4] M. Ebner, A. Holzinger, "Successful implementation of user-centered game based learning in higher education: an example from civil engineering," Computers & Education, **49**(3), 873–890, 2007, doi:10.1016/j.compedu.2005.11.026.
[5] F. Rozi, Y. Rosmansyah, B. Dabarsyah, "A Systematic Literature Review on Adaptive Gamification: Components, Methods, and Frameworks," in 2019 International Conference on Electrical Engineering and Informatics (ICEEI), 187–190, 2019, doi:10.1109/ICEEI47359.2019.8988857.
[6] S. Sena, et al., "Aprendizagem Baseada em Jogos Digitais: A Contribuição dos Jogos Epistêmicos na Geração de Novos Conhecimentos," in Revista Novas Tecnologias na Educação (RENOTE), **14**(1), 1–11. 2016.
[7] S. Freitas, "Are games effective learning tools? A review of educational games," Educational Technology & Society. **21**(2), 74–84, 2018.
[8] P. F. C. Santana, D. X Fortes, R. A. Porto, "Jogos digitais: a utilização no processo ensino aprendizagem," Revista Científica da FASETE, 1, 218–229, 2016.
[9] T.-H. Tsai, H.-C. Lin, K.-C. Huang, "Digital Game-Based Learning on Digital Archives: A Case Study of Taiwanese Classical Poems," in 2012 IEEE Fourth International Conference on Digital Game and Intelligent Toy Enhanced Learning. 132–134, 2012, doi:10.1109/DIGITEL.2012.36.
[10] N. P. Zea, J. L. G. Sanchez, F. L. Gutierrez, "Collaborative Learning by Means of Video Games: An Entertainment System in the Learning Processes," in 2009 Ninth IEEE International Conference on Advanced Learning Technologies, 215–217, 2009, doi:10.1109/ICALT.2009.95.
[11] A. Barmpoutis, et al., "Exploration of Kinesthetic Gaming for Enhancing Elementary Math Education Using Culturally Responsive Teaching Methodologies," in 2016 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR). 1–4, 2016, doi:10.1109/KELVAR.2016.7563674.
[12] E. Gibney, "DeepMind algorithm beats people at classic video games," Nature, **518**, 465–466, 2015, doi:10.1038/518465a.
[13] F. Ymran, O. Akeem, S. Yi, "Gamification Design in a History E-Learning Context," in 2017 International Conference on Information, Communication and Engineering (ICICE). 270–273, 2017, doi:10.1109/ICICE.2017.8479194.
[14] D. E. Asher, A. Zaldivar, J. L. Krichmar, "Effect of Neuromodulation on Performance in Game Playing: A Modeling Study," in 2010 IEEE 9th International Conference on Development and Learning, 155–160, 2010, doi:10.1109/DEVLRN.2010.5578851.
[15] E. Klopfer, Augmented learning: research and design of mobile educational games, Cambridge, 2008.
[16] S. Deterding, "Gamification: designing for motivation," Interactions, **19**(4), 14–17, 2012, doi:10.1145/2212877.2212883.
[17] E. Sanchez, V. Emin-Martinez, "Towards a Model of Play: an Empirical Study," in C. Busch (Ed.), Proceedings of the 8th European Conference on Games Based Learning, **2**, 503–512, 2014.
[18] K. M. Kapp, The gamification of learning and instruction: Game-based methods and strategies for training and education, Pfeiffer & Company, 2016.
[19] L. Sheldon, The multiplayer classroom: designing coursework as a game, Cengage Learning, 2011.
[20] J. Gee, The ecology of games: connecting youth, games, and learning, MIT Press, 2008.
[21] K. Hogan, M. Pressley, Scaffolding student learning: instructional approaches and issues, Brookline Books, 1997.
[22] W. Konen, "General Board Game Playing for Education and Research in Generic AI Game Learning," in 2019 IEEE Conference on Games (CoG), 1–8, 2019, doi:10.1109/CIG.2019.8848070.
[23] MIT Media Lab Scratch (2019, August 1), Scratch, Available at: scratch.mit.edu
[24] M. Mendonça, H. S. Kondo, L. B. de Souza, R. H. C. Palácios, J. P. L. S. de Almeida, "Semi-Unknown Environments Exploration Inspired by Swarm Robotics using Fuzzy Cognitive Maps," in 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 1–8, 2019, doi:10.1109/FUZZ-IEEE.2019.8858847.
[25] E. I. Papageorgiou, Fuzzy cognitive maps for applied sciences and engineering, Springer, 2014.
[26] R Brooks, "A robust layered control system for a mobile robot," IEEE Journal of Robotics and Automation, **2**(1), 14–23, 1986, doi:10.1109/JRA.1986.1087032.
[27] A. M. El-Sherbini, M. A. Aboul-Dahab, M. M. Fouad and M. F. Abdelkader, "Distance learning during Covid-19: Lessons learned and Case studies from Egypt," in 2021 IEEE Global Engineering Education Conference (EDUCON), 1743–1748, 2021, doi:10.1109/EDUCON46332.2021.9454051.
[28] A. Wilby, E. Lo, "Low-Cost, Open-Source Hovering Autonomous Underwater Vehicle (HAUV) for Marine Robotics Research based on the BlueROV2," in 2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV), 1–5, 2020, doi:10.1109/AUV50043.2020.9267913.
[29] C. Persello, L. Bruzzone, "Active and Semisupervised Learning for the Classification of Remote Sensing Images," IEEE Transactions on Geoscience and Remote Sensing, **52**(11), 6937–6956, 2014, doi:10.1109/TGRS.2014.2305805.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 230-241 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Sciences & Technology

ASTES Journal
ISSN: 2415-6698

A Case Study to Enhance Student Support Initiatives Through Forecasting Student Success in Higher-Education

Ndiatenda Ndou[*,1], Ritesh Ajoodha[1], Ashwini Jadhav[2]

[1]*School of Computer Science and Applied Mathematics, The University of the Witwatersrand, Johannesburg, 2050, South Africa*

[2]*Faculty of Science, The University of the Witwatersrand, Johannesburg, 2050, South Africa*

ARTICLE INFO

Article history:
Received: 07 November, 2020
Accepted: 19 December, 2020
Online: 15 January, 2021

Keywords:
Higher-Learning
Machine learning
Academic Performance
Student Success
Classification
Random Forests
Student Attrition

ABSTRACT

Enrolment figures have been expanding in South African institutions of higher-learning, however, the expansion has not been accompanied by a proportional increase in the percentage of enrolled learners completing their degrees. In a recent undergraduate-cohort-studies report, the DHET highlight the low percentage of students completing their degrees in the allotted time, having remained between 25.7% and 32.2% for the academic years 2000 to 2017, that is, every year since 2000, more than 67% of the learners enrolled did not complete their degrees in minimum time. In this paper, we set up two prediction tasks aimed at the early-identification of learners that may need academic assistance in order to complete their studies in the allocated time. In the first task we employed six classification models to deduce a learner's end-of-year outcome from the first year of registration until qualifying in a three-year degree. The classification task was a success, with Random Forests attaining top predictive accuracy at 95.45% classifying the "final outcome" variable. In the second task we attempt to predict the time it is most likely to take a student to complete their degree based on enrolment observations. We complete this task by employing six classifiers again to deduce the distribution over four risk profiles set up to represent the length of time taken to graduate. This phase of the study provided three main contributions to the current body of work: (1) an interactive program that can calculate the posterior probability over a student's risk profile, (2) a comparison of the classifiers accuracy in deducing a learner's risk profile, and (3) a ranking of the employed features according to their contribution in correctly classifying the risk profile variable. Random Forests attained the top accuracy in this phase of experiments as well, with an accuracy of 83%.

1 Introduction

The benefits that can be drawn from institutions of higher-learning extend beyond the degree holder. A study conducted on the relationships between the quality of life, human capital, and universities, revealed that valuable consumption amenities that enhance an areas quality of life are positively correlated with both the local level of human capital (measured by proportion of degree holders in an area) and the number of institutions of higher-learning in a region [1].

The higher-education system is one of great benefit to enrolled-individuals, the economy, and society, however, an inefficient system with high dropout rates and low throughput rates carries harsh costs and consequences for the individual student as well as the society financing the cost of service delivery [2]. The South African Human Science Research Council (HSRC) found that on average, 70% of the university drop-outs they surveyed came from families in the "low economic status civilian" category [3]. Students that belong in this category heavily depend on government study grants and subsidies to supplement the funding they receive from their parents or guardians. It is clear that student debt accumulation, or alternatively, costs to the government without a return on investment will be the outcome when these learners struggle and dropout. This was the case in 2005, where the national treasury reported R4.5 billion lost to student grants and subsidies that resulted in no graduates [3]. There is a student attrition problem in South Africa, as the expansion of enrolments has not come with a significant increase in the percentage of students completing their degrees [4].

Noting the value and possible severe-costs associated with in-

[*]Corresponding Author: Ndiatenda Ndou, The University of the Witwatersrand, Johannesburg, +27 71 168 8461 & ndiatenda.ndou@students.wits.ac.za

Table 1: This table introduces categories for the different features associated with student performance as discussed in Section 2. The features are divided into four groups, namely, "Socioeconomic Factors" (SEF), "Psycho-Social Factors" (PSF), "Pre & Intra-College Scores" (PICS), and "Individual Attributes" (IA). It is also important to take note of the colour coding scheme developed here for later reference to the four categories, [5](sic).

Socioeconomic Factors (SEF)	Psycho-Social Factors (PSF)	Pre & Intra-College Scores (PICS)	Individual Attributes (IA)
Family income	Academic self-efficacy	Mathematics	Age at first year
Parents education	Stress and time pressure	English	Work status
Head of house occupation	Class communication	Admission Point Score	Home language
Dwelling value	College activity participation	Accounting	Home province
Dwelling location (rural/urban)	Organization and attention to study	Economic studies	Home country
Financial support	Sense of loneliness	Statistics major	Interest in sports

stitutions of higher-learning, we see the clear need to explore the activities influencing student success or failure to solve the problem of student attrition and avoid the severe costs and consequences that it brings. Furthermore, we seek to develop advanced systems for the early-identification of vulnerable learners that may benefit from academic support systems.

In this research, we investigate the influence of biographical and enrolment observations on student success. This research was conducted through two published studies referred to as "phases of the current study" throughout this paper [5, 6]. In the first phase, we employ six machine learning models to predict three target variables that describe a learner's end-of-year outcome, namely, "first-year outcome", "second-year outcome", and "final outcome" [5]. The first phase of this research contributes to the current body of work by showing that various classification models can be used to predict a learner's end-of-year outcome from the first year of registration until qualifying in a three-year degree. We argue that if we can predict the academic trajectory of a student, early-assistance can be provided to students who may perform poor in the future, remediating their performance and promoting student success.

The second phase of this study involved the prediction of a "risk profile" variable, a variable that categorises the time taken by a student to graduate in a three-year degree by four values, namely: "no risk", where the student completes the degree in three years; "low risk", where the student completes the degree in more than three years; "medium risk", where the student fails/drops-out in less than three year; and "high risk", where the student takes more than three years to drop out [6]. We used six machine learning algorithms to predict the "risk profile" variable for a student based on biographical and enrolment observations. The second phase of this study contributes to the current body of related literature in three ways: (1) a comparison of six different classifiers in predicting the risk profile of a learner; (2) a ranking of the features employed according to their contribution when deducing the "risk profile"; and (3) an interactive program which uses Random forests classifier to deduce the distribution over a learner's risk profile. The contribution made by this research implies institutions of higher-learning can use machine learning techniques for the early-identification of learners that may benefit from academic assistance initiatives.

This paper continues with Section 2 which presents the background knowledge around the problem. We then introduce the procedure and system of methods applied in Section 3, followed by the results in Section 4. The work is concluded on Section 5 and we close this study with ideas of future work in Section 6.

2 Related Work

Predicting student performance is a multifaceted task that cannot be easily completed using attributes discovered in student enrolment records alone [4]. We therefore, set-out to explore the various factors influencing student performance, aiming to discover and develop an efficient methodology for solving the problem set up in this research. We begin this chapter with the discussion and grouping of the factors influencing student performance, followed by a brief presentation of the conceptual framework adopted for feature selection in this research, and we close the chapter with a comparison of the various methods of predicting student performance.

2.1 Factors affecting student performance

The South African Department of Higher Education and Training (DHET) released a report with results that portray an increasing number of first-time-entering university students, from 98095 students in the year 2000 to over 150000 students in the 2017 academic year [7]. This significant increase brings the idea that the various attributes that describe a university student must be more varied now than ever before, as more learners from different regions are now enrolling for degrees. To accurately early-identify a member of a given group of learners as successful or unsuccessful in their studies, we must explore a wide range attributes that describe these students, so that our decision is informed.

2.1.1 Socio-Economic Observations Determining Student Success

To determine an individual or their family's measure of social and economic position in the population, we explore the category of socio-economic attributes as determinants of student success. Studies have shown that a correlation exists between socio-economic status and academic achievement, furthermore, traditional measures of socio-economic status have been revealed to usually correlate strong enough with academic achievement to account for variations

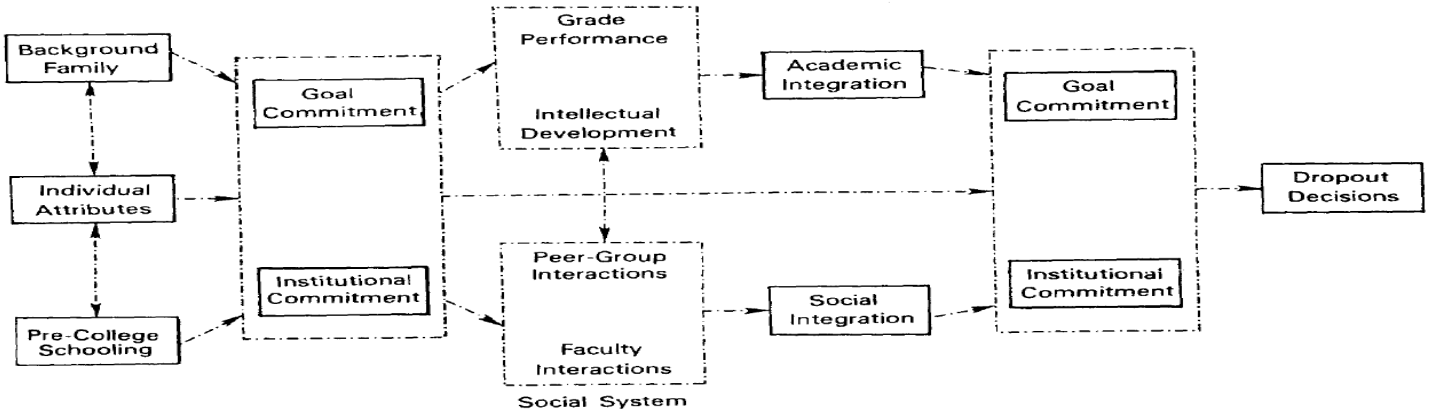


Figure 1: The graphical representation of the framework adopted to determine student success, [6](sic).

in a learner's performance [8]. A study of the determinants of student success also concludes that socio-economic factors contribute significantly when predicting a learner's performance [4]. The socio-economic attributes investigated in the literature we reviewed include; parents' education, dwelling value, family income, and parent/guardian's occupation. Family characteristics and financial support also form part of socio-economic factors revealed to explain student performance, specifically drop-out behaviour in university [2].

2.1.2 Psycho-Social Factors Affecting Student Performance

We seek to find more categories of factors that account for variations in a learner's academic performance. We therefore, explore the combined effects of a student's thoughts, social factors, and general behaviour at university on academic performance. Research conducted at a South African university found that to predict a learner's academic performance and adjustment to higher-education, psycho-social factor's such as; help-seeking attitude, workload, perceived stress, and self-esteem could be utilized as determinants [9].

Research on college students' performance utilized six psycho-social factors to predict first-year college student success [10]. The six psycho-social attributes used for prediction were; stress and time pressure levels, communication/participation in class, academic self-efficacy (a learner's belief in their ability to succeed academically), attention to study (measures time-management, planning, and scheduling behaviour), stress levels, and lastly, emotional satisfaction with academics. It was revealed that a strong correlation exists between the six psycho-social factors and a learner's GPA, these findings align with other related work [11]. A separate study also discovered that academic self-efficacy, student's optimism, commitment to schooling, and student health account for some of the variations in a student's performance, expectations and coping perceptions [12].

2.1.3 Factors Available For This Study

The review literature centred around factors affecting student performance revealed that a learner's mindset, social surrounding, and behaviour, explain significant variations in the performance of a university learner. In this research, student biographical and enrolment observations such as, majors enrolled for, age, pre-college scores, province and country of origin, are available to build machine learning models for the prediction of academic performance. We continue this chapter by introducing the framework we adopted for the rationale behind predicting academic performance from biographical and enrolment observations.

2.2 Conceptual Framework

We use the conceptual framework depicted in Figure 1 as a logical basis to predict the academic performance of a learner from biographical and enrolment observations [13]. The framework develops three categories of attributes contributing to student attrition (drop-out behaviour), namely, background attributes, individual attributes, and pre-college scores. The study reveals that these factors together influence a learner's goal and institutional commitment, which in turn contributes to the drop-out decision of a student via academic and social integration [13]. In this research, we partition background attributes further into socio-economic and psycho-social factors. Table 1 presents the grouping of features discovered to explain variations in academic performance during the review of related literature conducted in this section.

2.3 Methods For Predicting Student Performance

Various authors have already accurately predicted student performance utilizing machine learning. Table 2 compares the accuracies obtained in the various literature reviewed in this research. Acquiring the top accuracy position is a research based on the prediction of the success of a second-year student using an Artificial Neural Network (ANN) model [14]. The study uses individual attributes (IA), pre & intra-college scores, and socio-economic factors as predictors, these inputs are utilized for almost every result presented in the table [11, 17, 16].

3 Research Methodology

This research proposes an approach to the task of university-learner-performance prediction involving the prediction of a learner's end-

of-year outcome from the first year they register for a three-year degree until qualifying to graduate. We complete this task by employing six different machine learning algorithms to deduce the outcomes in the first, second, and final year of study in a South African university. This study extends further by attempting to predict a learner's risk-profile based on the time it will take the student to complete a three-year degree.

Table 2: The table compares the accuracy achieved by various authors using different machine-learning models to predict academic performance in each case. The table also illustrates for each case, the combination of features used as predictors, based on the feature categories developed in Table 1, [5] (sic).

Author(s)	Factors considered	Model used	Predictive Accuracy
Abu-Naser *et al.* (2015)[14]	IA, SEF, and PICS	Neural Networks	84.60%
Osmanbegovic and Suljic (2012)[11]	IA, SEF, PICS, and PSF	Naïve Bayes	84.30%
Osmanbegovic and Suljic (2012)[11]	IA, SEF, PICS, and PSF	Neural Networks	80.40%
Osmanbegovic and Suljic (2012)[11]	IA, SEF, PICS, and PSF	Decision Trees (C4.5)	79.60%
Mayilvaganan and Kalpanadevi (2014)[15]	IA and PICS	Decision Trees (C4.5)	74.70%
Ramesh *et al.* (2013)[16]	IA, SEF, and PICS	Multi-layer Perceptron	72.38%
Abed *et al.* (2020)[17]	IA, SEF, and PICS	Naïve Bayes	69.18%
Abed *et al.* (2020)[17]	IA, SEF, and PICS	SMO	68.56%
Ramesh *et al.* (2013)[16]	IA, SEF, and PICS	Decision Trees (C4.5)	64.88%

In this section, we introduce the procedure and system of methods implemented for the purpose of this study. The study incorporates two phases and thus, we begin by giving a description of the two phases, followed by subsection 3.2 giving a description of the data-sets. We present the feature selection technique in subsection 3.3 and follow this by a brief description of the machine learning models we used, closing the section with methods of evaluating and validating our results.

3.1 Phases of the Study

This study was conducted in two phases. The first phase, named, "Preliminary Phase", involved generating preliminary results on a synthetic dataset. In this phase we employed six different machine learning models, namely, Decision tree (C4.5), Logistic Model Trees (LMT), Multinomial Logistic Regression, naïve Bayes, Sequential Minimal Optimization (SMO), and Random forests. The purpose of the first phase was to reveal that machine learning models can be utilized for the early prediction of a learner's end-of-year outcome from the first year of registration until qualifying in a three-year degree, based on biographical and enrolment observations.

The second phase of this study, named, "Post-preliminary Phase", involved the prediction of the distribution over several risk profiles that describe the time it will take for a university student to complete a three-year degree. This phase is performed on a real dataset which the synthetic dataset in the first phase was modelled to resemble. In this phase we employed six different machine learning models, namely, Decision tree (C4.5), Linear Logistic Regression, Support Vector Machines (SVM), naïve Bayes, and Random forests. This phase provides an interactive program which calculates the posterior probability over a learner's "risk profile" as the main contribution o this study, and a ranking of features through Information Gain Ranking (IGR), to determine the features most contributing to student performance.

3.2 Data Description and Pre-Processing

Two sets of data were utilized in this study. We therefore, split the description of the datasets, starting with Subsection 3.2.1 which gives the synthetic dataset description, and Subsection 3.2.2 giving a description of the real dataset.

3.2.1 The Synthetic Dataset Description

The dataset used for the preliminary phase of this study is a synthetic dataset generated using Bayesian Network. The dataset was adopted from a recent prediction modelling study aimed at improving student placement at a South African university [17]. In this dataset, conditional independence assumptions were implemented to portray the relationships that exist between enrolment, socioeconomic, and individual attributes found in student records.

Three target variables are investigated in the preliminary phase, namely, "First Year Outcome" (FYO), "Second Year Outcome" (SYO), and "Final Outcome" (FO). The SYO and FYO variables contain two similar possible values: "proceed", and "failed", were proceed is the outcome for a student who met the requirements to proceed to the next year of study, and failed implies the student failed to meet the minimum requirements to proceed. The FO variable also has two possible values: "qualified", and "failed", were qualified implies the learner met the minimum requirements to graduate in a three-year degree, and failed implies the student failed to meet the requirements to graduate.

Data pre-processing is a crucial step when employing machine learning models. The pre-processing task incorporates, data preparation, data cleaning, data normalization, and data reduction tasks [18]. The synthetic dataset originally contained 50 000 instances. Three random samples (without feature replacement or bias to uniform class) containing 2000 instances were drawn from the raw dataset and several experiments were conducted to make the samples more suitable for our machine learning models. The first set of experiments focused on the detection and removal of outliers or anomalies. This involved the evaluation of classification results from different machine learning classifiers in an attempt to detect and remove instances that display significant deviations from the majority patterns. The second set of experiments conducted aimed to prevent over-fitting. This was done by the implementation of Synthetic Minority Oversampling Technique (SMOTE) which enforced an equal number of training instances for each value in the class variable. The third set of experiments conducted in the preliminary phase is feature selection, where 20 features were selected from each sample based on Information Gain Ranking (IGR) criterion. Subsection 3.3 presents the set of features selected from each sample and a discussion of how we arrived at this set.

3.2.2 The Real Dataset Description

The real data utilized for this study is from a research-intensive university in South Africa. The dataset composes of enrolment and biographical observations of learners enrolled in the faculty of science at the university, from the year 2008 to the year 2018.

The target variable for the post-preliminary phase of this study is "Risk Profile", a nominal variable that tells us how long it will take for an enrolled student to complete a three-year degree at a

South African university. The risk-profile variable has four possible values, namely; "no risk", where the student completes their degree in 3 years (the minimum allotted time); "low risk", where the student completes the three-year degree in more than 3 years; "medium risk", where the student fails to complete the three-year degree before the end of three years (student drops-out in less than 3 years' time); and "high risk", where the student fails to complete the three-year degree after exceeding the 3-years period (student drops-out after exceeding the allotted time).

Since our aim is to perform an accurate classification task on the dataset, pre-processing procedures had to be carried out before training the chosen classifiers. We performed the same experiments as those performed on the synthetic dataset to detect and remove outliers. We then employed a sampling procedure similar to the one we implemented in the preliminary phase of this study. Along with applying SMOTE, a random sample (with no replacement of features or bias to uniform class) of 200 instances was drawn from the original dataset. In Subsection 3.3 we present the features selected and the procedure followed to select them.

Table 3: A table providing the various features selected for input into the employed classifiers. The table groups the features according to whether they were used for the prediction of the learner's first year outcome, second year outcome, or final year outcome. The colour shading gives the category which the feature selected belongs to according to Table 1, [5] (sic).

#	Feature	1^{st} Year	2^{nd} Year	Final Year
1	English Home Language			
2	Plan Description			
3	Quintile			
4	Home Province			
5	Year Started			
6	Language			
7	Progress Outcome YOS1			
8	Home country			
9	Aggregate YOS2			
10	Rural or Urban			
11	Second Year Outcome			
12	Age at Third Year			
13	Mathematics Literacy			
14	NBTAL			
15	Age at First Year			
16	Computers			
17	NBTQL			
18	Age at Second Year			
19	Life Orientation			
20	NBTMA			
21	Plan Code			
22	English FAL			
23	Additional Mathematics			
24	Mathematics Major			

3.3 Feature Selection

In this subsection we present the methodology behind the features selected for both phases of the study. Subsection 3.2.1 provides the features we utilized to generate the preliminary results and Subsection 3.2.2 provides the features selected for the post-preliminary phase of the study.

3.3.1 Features for the Preliminary Phase

To select features for the purpose of predicting the three target variables investigated in this phase of the study we utilized Information Gain Ranking (IGR) criterion, which involves deducing the contribution of each feature when classifying an instance as a value of the class variable. Table 3 presents the features selected in all cases considered for the preliminary phase, based on the colour coding scheme developed in Table 1.

Feature selection was performed on each of the 3 samples drawn from the synthetic dataset. We investigated the contribution of 44 features using Information Gain (entropy). Through the entropy values and repeated experimentation with different sets of features, a total of 20 features were selected to predict the target variable in each of the three samples.

The features presented in Table 3 are not arranged according to IGR as there are significant differences in the entropy of most features across predicting the three target variables. The features selected align with our findings from the review of previous work, and more importantly, the conceptual framework we adopt in this research [13]. This is because for each target variable, there were features selected from each of the three investigated categories of features, namely, background (socio-economic) attributes, individual attributes, and pre-college scores.

3.3.2 Features for the second phase

To select features for the post-preliminary phase of this study, we continued utilizing IGR alongside the conceptual framework investigated in the related work section [13]. From the "individual attributes" category, the following features were selected; the National Benchmark Test (NBT) scores for academic literacy (NBTAL), quantitative literacy (NBTQL), and mathematical literacy (NBTMA). These scores give a measure of the individual learner's proficiency and ability to meet the demands of university-level work. To further determine an enrolled learner's professional career aspirations, we also considered the academic plan selected by the learner. The "plan code", "plan description", and "streamline" (Earth Science, Physical Science, Biological Science, or Mathematical Science) variables were selected for this task.

From the "background and family" category, the following features were selected; the home-country and province of the student, whether the school attended by the leaner is in the urban or rural areas, the quintile of the school attended, and the age of the learner at the first-year of registration. These variables combined give us a description of the learner's socio-economic status.

From the pre-college scores category, scores from the following subjects were considered; Mathematics major, Mathematics Literacy, Additional Mathematics, Physical Science, English Home Language, English First Additional Language, Computer studies, and Life Orientation. We note that in this phase of the study, we did not utilize college or university outcomes as inputs, these outcomes were only utilized in the prediction of target variables set up in the preliminary phase as indicated in Table 3. The features selected for the post-preliminary phase of this study are presented ranked according to entropy in Table 6.

3.4 Classification Models

This study composed of two phases, each involving a classification task. In total, we use nine off-the-shelf machine learning predictive models to perform the classification tasks in this research. The models used are: Random forests, naïve Bayes classifier, Decision tree (C4.5), Logistic Model Trees (LMT), Multinomial Logistic Regression, Linear Logistic Regression, Sequential Minimal Optimization (SMO), Support Vector Machines (SVMs), and the K-Star (K*) instance-based classifier. We continue this subsection by giving a brief description of the selected models.

Sequential Minimal Optimization: SMO is an algorithm derived for the improved (speed) training of Support Vector Machines. SVMs previously required that a large quadratic-programming problem be solved in their implementation. Traditional algorithms are slow when training SVMs, however, SMO completes the task much faster by breaking the large quadratic programming problem into a series of smaller problems which are then solved by analytical methods, avoiding the lengthy numerical optimization required. The SMO algorithm we use for the purpose of this study follows the original implementation [19].

Multinomial Logistic Regression: This model derives and extends from the binary logistic regression. It does so by allowing categories of the outcome variable to exceed two. This algorithm utilizes the Maximum Likelihood Estimation (MLE) to predict categorical placement or the probability of category membership on a dependent variable [20]. Multinomial Logistic Regression requires the careful detection and removal of outliers for accurate results, however, the model does not assume or require linearity, homoscedasticity, or normality [20]. An example of a four-category model of this nature, with one independent variable x_i can be given by: $\log(\frac{\pi_i^{(s)}}{\pi_i^{(0)}}) = \eta_0^{(s)} + \eta_1^{(s)} x_i$, s = 1,2,3,4. Where; $\eta_0^{(s)}$ and $\eta_1^{(s)}$ are the slope and intercept respectively, given the probability of category membership in "s" can be denoted by $\pi_i^{(s)}$, and the selected reference category by $\pi_i^{(0)}$. The Multinomial Logistic Regression implementation in this research follows the implementation by other authors before the current study [21, 22].

Naïve Bayes Classifier: The naïve Bayes model is a simplified example of Bayesian Networks. The model achieves learning with ease by assuming that features employed are all independent given the class variable.

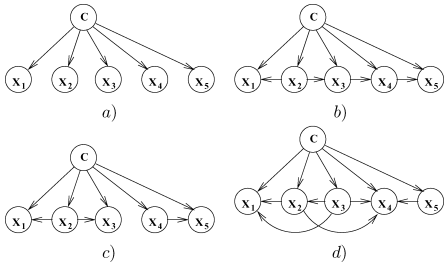


Figure 2: The diagrams (a) to (d) illustrate various examples of a Bayesian network model. The arrows travelling between nodes represent conditional dependencies among the features $x_1, x_2, ..., x_5, C$. Where C, represents the class variable and the models differ based on the existence of a statistical dependence between the predictors $(x_1, ..., x_5)$ [18].

The diagram (a) in Figure 2 is an illustration of the naïve Bayes model. This is because the features, $x_1, ..., x_5$ are conditionally independent given the class variable, C. The naïve Bayes independence assumption can be stated as the distribution: $P(C|X) = \prod_{i=1}^{n} P(x_i|C)$, where $X = (x_1, ..., x_n)$ is a feature vector and C is the class variable. In application, naïve Bayes often performs well when compared to more sophisticated classifiers, although it makes a generally poor assumption [23]. The naïve Bayes model implementation in this research follows other similar implementations in related studies explored; [24, 23].

K* Instance-Based Classifier: The K-Star algorithm classifies instances using training instances that are similar to them alongside a distance function that is based on entropy. The use of an entropy based function provides consistency in the classification of instances in our experiments that may be real-valued or symbolic.The K* implementation utilized for the purpose of this research followed the implementation by [25].

Support Vector Machines: The (SVMs) classification model separates classes of the training data with a hyper-plane. The test instances then get mapped on the same space with their prediction based on the side of the hyper-plane they belong after splitting. This task is performed by incorporating the training dataset into a binary linear classifier that is non-probabilistic. SVMs can be scaled through the one-versus-all partitioning, for various types of classification problems including high-dimensional and nonlinear classification tasks. The SVM model implementation in this research follows the implementation in other related work [26].

Linear Logistic Regression: The Linear Logistic Regression model utilizes additive logistic regression with simple regression functions as base learners of the algorithm [27].The implementation of this model followed in this research follows that of related work conducted in the past [28, 29].

Decision Tree: This model uses a decision support system to build a classification function that predicts the value of a dependent variable given the values of independent variables, through tree-like graph decisions and their possible after-effect, including costs of resources, chance results, and utility [30]. There are different algorithms for generating decision trees; C4.5, Random forest, and LMT are the tree-models selected for the purpose of this research.

The C4.5 algorithm uses information gain to build a decision tree, selecting features based on entropy and utilizing the ID3 algorithm recursively to build the tree. We follow the original structure and implementation of the C4.5 algorithm in this study [31].

LMT builds a single tree from a combination of logistic regression models and a tree structure. This model accomplishes the combination by using the Classification and Regression Tree (CART) algorithm to prune after building the regression functions through the LogiBoost algorithm. The LMT method used in this study follows from the original implementation [28].

Random Forests are a combination of decision tree predictors dependent on the value of a random vector, where the value also governs the growth of each tree in the generated forest. This algorithm involves utilizing the training data to generate an ensemble of decision trees and allowing them to decide on the most-popular class. Implementing this technique has several advantages, including that, the procedure abides by the Law of Large Numbers to prevent over-fitting, it is relatively robust to noise or outliers in

data, and the model achieves accuracy as good as similar techniques such as, Adaboost, Bagging, and Boosting, while still training faster than them [32]. The Random forest model implementation in this research is based on the original model [32].

3.5 Prediction and Evaluation

In this research, the dataset contains the dependent attribute and the values of the classes are known, we therefore have a classification task set up. This subsection provides the measures and techniques utilized to complete this task.

3.5.1 Evaluation and Validation

10-fold cross-validation procedure is applied to evaluate each model employed in this research. In this validation procedure, the training dataset is partitioned such that a portion (testing data) of it is not provided to the algorithm during training, but is used for the validation. The partition remaining for training is further split into 10 partitions (folds). Interchangeably, each of the 10-folds serve for validation while the remaining 9 are used for training until all 10-folds serve as the validation fold once.

3.5.2 Confusion Matrix

We use confusion matrices to illustrate classification outcomes. Table 4 provides an example of the format of a confusion matrix.

Table 4: This table depicts the structure of a confusion matrix. We denote "negative" and "positive" by -ve & +ve respectively. Where TP are the true positives (correctly classified positives), FP the false positives, FN are false negatives, and TN are the true negatives (correctly classified negatives).

	Predicted Class +ve	**Predicted Class -ve**
Actual Class +ve	TP	FP
Actual Class -ve	FN	TN

3.5.3 Accuracy

We extract precision and recall metrics from the confusion matrix in order to measure the accuracy of the employed machine learning models. "Precision" represents the correctly real-positive proportion of predicted positives, while "Recall" represents the correctly predicted proportion of the real positives. We calculate precision and recall as follows:

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

The accuracy follows directly, calculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

This representation of accuracy has been used in other related work [17], [33]. Other measures of accuracy explored include the Receiver Operating Characteristic (ROC) curve. This curve plots recall (the true positive rate) against the false positive rate (ratio between FP and the total number of negatives), and the area under the ROC reflects the probability that prediction is informed versus chance [34]. We desire the ROC-area to lie above 0.5, anything below 0.5 implies the prediction was guesswork and not informed. Another accuracy measure utilized is the F-beta measure (F1 score), which calculates a test accuracy as the weighted harmonic mean of precision and recall. The optimal value for the F-measure is 1 (indicating perfect precision and recall), and the worst value is 0.

4 Results and Discussion

This study was conducted in two phases. The first phase involved generating preliminary results on a synthetic data-set, while in the second phase, a similar set of experiments are performed on a real data-set, leading to conclusions and implications about the performance of the trained machine learning models in classifying the problem at hand, furthermore, the results drawn from the second phase provide a ranking of the employed features according to entropy, together with an interactive program which calculates the posterior probability over the students' risk profile so that support initiatives and programs can be focused on them.

4.1 Preliminary Results

This Subsection presents the results of six of the nine prediction models discussed in Section 3, namely; Decision tree (C4.5), Logistic Model Trees (LMT), naïve Bayes Classifier, Sequential Minimal Optimization (SMO), Multinomial Logistic Regression, and Random Forests. We present first the prediction outcomes through a table comparing predictive accuracy of the models as determined by Equation 3. This will be followed by an evaluation of the model's performance through F-measure and Receiver Operating Characteristic (ROC) curve.

4.1.1 Prediction Outcomes

Six different machine learning models were utilized to solve our classification problem. The predictive accuracy achieved by each model is recorded and presented in the Table 5.

Table 5: Predictive accuracy as calculated by Equation 3. After 10-fold cross-validation, all of the models utilized achieved an accuracy above 80%, with Random Forests achieving top accuracy in all three cases considered, [5](sic).

	Predictive Accuracy		
Model used	1^{st} Year Outcome	2^{nd} Year Outcome	Final Year Outcome
Random Forest	94.40%	93.70%	95.45%
LMT	91.90%	91.75%	93.15%
Decision Trees (J48)	87.55%	86.20%	91.45%
Multinomial Logistic	87.80%	86.20%	90.70%
SMO	87.25%	84.45%	89.20%
Naïve Bayes	83.95%	83.40%	84.40%

4.1.2 Model Performance Evaluation

A study based on evaluating classification results argued for the use of Precision, Recall, F-Measure, and the ROC curve, as accurate measures of a machine learning model performance [34]. We utilize these measures to evaluate results presented by Figure 3, 4, and 5.

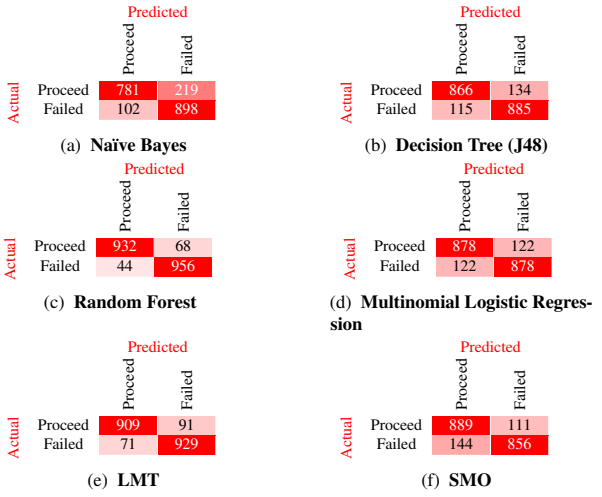


Figure 3: The set of confusion matrices resulting from the prediction of first-year outcome. Evaluation of the accuracy by class reveals that the weighted average of both precision and recall lies above 0.84 for all six models trained in our study. Further observations reveal that the f-measure of accuracy is more than 0.83 for all models, this value aligns with our accuracy as determined by Equation 3 in the Table 5. The test accuracy obtained in the table is further supported by the ROC curve obtained for all six models, as the weighted average of the ROC area for each model is more than 0.84 implying the models trained were making informed decisions in classifying the problem and not simply guessing.

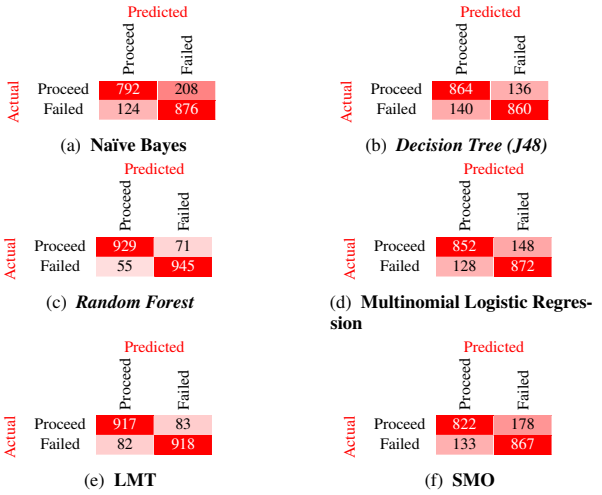


Figure 4: The confusion matrices obtained when classifying the second-year outcome variable. When we evaluate the detailed accuracy by class for each model, we find that the weighted average of both precision and recall measures is above 0.83, furthermore, the f-measure of test accuracy aligns with our findings in Table 5 with a value of more than 0.83 for all six models considered. The ROC curve obtained for all six models also aligns with our accuracy as determined by f-measure of test accuracy and Equation 3, as the weighted average of the area under the ROC curve lies above 0.85 implying the models trained are not attaining the great predictive accuracy through guess-work, but they are making informed predictions.

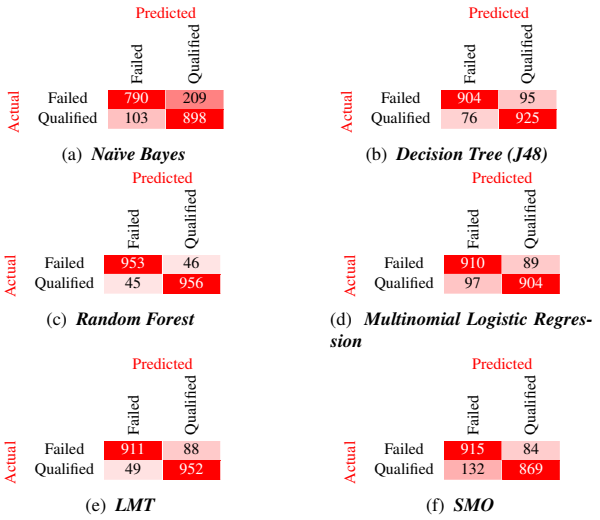


Figure 5: Confusion matrices obtained when classifying the third-year outcome variable with the six trained machine learning models. Evaluating the detailed accuracy by class results, we find that the weighted average of both precision and recall measures is more than 0.89 for all models except the naïve Bayes classifier which scored 0.85 for precision, and 0.84 for recall. The underperformance of the naïve Bayes model with respect to other trained classifiers can be explained by the naïve feature-independence assumption it makes. Results also show that the weighted average of the area under the ROC curve for each model is in alignment with our test accuracy measures, as it lies above 0.89, implying each model is making informed predications and not simply guessing the outcome.

4.2 Main Results (Post-preliminary)

The preliminary phase of this study revealed that we can utilize machine-learning models to accurately predict a learner's outcome from the first year of registration until qualifying in a three-year degree. The preliminary results confirmed our initial hypothesis but furthermore, revealed the kind of model we should utilize with such a wide variety of models available for use, but few fitted to the problem set up in this study. Evaluating the preliminary results, we note that Random Forests achieved top accuracy and performance as measured by all our model performance and accuracy evaluators, across all three test cases.

In this subsection, we present the results obtained when utilizing machine learning models to classify a learner into the four risk profiles ("No Risk", "Low Risk", "Medium Risk", and "High Risk") defined in Section 3, as the preliminary phase has confirmed this task can be completed. The classification problem set up in this phase is parallel to that in the prelim-phase in several ways, as the preliminary phase utilizes a synthetic data-set modelled to resemble the relationships that exist within the student enrolment data utilized for the second phase.

Subsection 4.2.1 presents the selection and ranking of features utilized, we follow this by a presentation of the classification outcomes and close the section with the presentation of an interactive program that can be utilized to calculate the posterior probability over a student's risk profile.

4.2.1 Selection and Ranking of features

We selected 20 features to predict the class variable. The features were selected using Information Gain Ranking (IGR) to deduce

the contribution of each feature in classifying the instances. The feature selection findings are illustrated through Table 6 with three columns below. The first column determines the rank of the features among the input set, the second column gives the amount of entropy attained by the feature, and the third column has the name of the feature associated with the ranking and entropy.

Table 6: A ranking through information gain (entropy) of the set of features selected to predict the student risk profile, ranked from the most contributing feature to the least contributing feature. The top seven features that are highlighted indicate an entropy greater than 0.1, [6] (sic).

Rank	Entropy	Feature Name
1	1.21960228	PlanCode
2	1.15086266	PlanDescription
3	0.59886383	Streamline
4	0.29582771	Year Started
5	0.20836689	AgeatFirstYear
6	0.18695721	SchoolQuintile
7	0.14234042	MathematicsMatricMajor
8	0.12166049	Homeprovince
9	0.06417526	isRuralorUrban
10	0.0568866	LifeOrientation
11	0.04978826	PhysicsChem
12	0.02780914	EnglishFirstLang
13	0.01253064	Homecountry
14	0.00550434	AdditionalMathematics
15	0.00000902	MathematicsMatricLit
16	< 0.00001	NBTAL
17	< 0.00001	NBTMA
18	< 0.00001	NBTQL
19	< 0.00001	ComputerStudies
20	< 0.00001	EnglishFirstAdditional

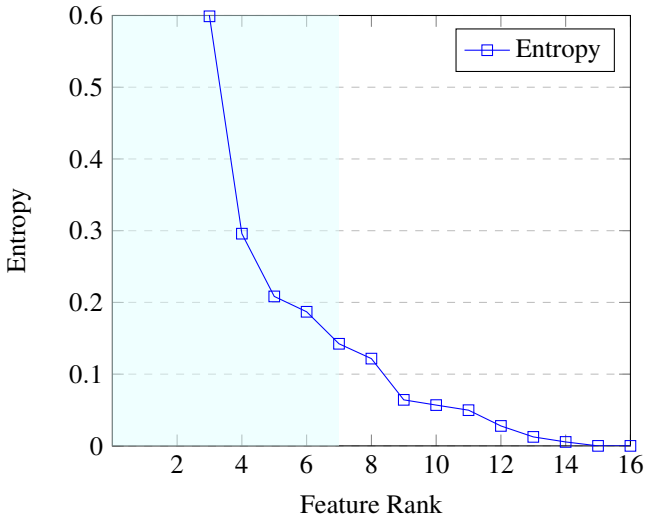


Figure 6: A graphical illustration of how the information gain (entropy) level varies across the chosen feature input set. The x-axis indicates the feature rank, and the y-axis indicates the information gain from utilizing the corresponding feature, [6] (sic).

The use of IGR also has implications on the contribution made by a feature relative to others within the chosen input set of features. We investigate the behavior of entropy as you move between subsequent features and present the findings through a graphical illustration in Figure 6, showing the monotonically decreasing behavior of the entropy function plotted versus rank. We see that the loss in entropy between each subsequent point (feature rank) is logarithmically decreasing, furthermore, we highlight the same top seven features from the IGR Table 6.

The highlighted features on Figure 6 illustrates their relative importance in forecasting the success of a learner. We note that the set of factors most contributing to a learner's success includes, "the plan-code and plan description" (combined, these two variables give a precise description of what the student is studying), "streamline" (mathematical, life, or physical science), "the year started", "school quantile", "age at first-year", and "the student's matric mathematics score".

4.2.2 *Classification Outcomes*

This section presents the results obtained from the classification algorithms trained to predict the class variable (risk profile). Six of the nine classification procedures discussed in Section 3 were employed for the post-preliminary phase of the study: Decision trees (C4.5), naïve Bayes Classifier, Linear Logistic Regression model, Support Vector Machines (SVMs), K*, and Random Forests.

Figure 7 (a) – (f) illustrates the results of each trained classifier after 10-fold cross-validation. Evaluating the performance of each model relative to other models employed, we note that the Random forests classifier attains the highest accuracy (83%) of all the models trained for the post-preliminary phase. This result aligns with our findings from the preliminary phase of the study as Random forests attained the highest accuracy among the selected models for both phases of the study.

We note further that the SVM classification model was revealed to be the least suited for the problem set up in this study. At 52%, SVMs attained the lowest predictive accuracy in this study, furthermore, SVMs took the longest time to train. When discussing training and testing times, it is also important to note that the K-star model took the least time to implement in this study.

Overall, the classification task was a success, with five of the six models employed attaining a predictive accuracy above 75%. Noting that Random forests was the most accurate in predicting the class variable for both phases of this study, this section continues by providing a web application utilizing the Random forests classifier to predict the risk-profile of a learner based on enrolment and academic factors. Severity of misclassified instances was also evaluated and taken into account to determine Random forests as the most suited model for the task, for example, the 27% of "No Risk" instances incorrectly classified by SVM as "Medium Risk" is far more severe and misleading than the misclassification of 5% "No Risk" instances as "Medium Risk" by Random forests classifier.

4.2.3 *Main Contribution of The Study*

In this subsection we provide an interactive program which can calculate the posterior probability over a student's risk profile utilizing the Random forests model employed in Section 4.2.2. This automated system makes predictions about a student's risk of fail-

Actual Risk \ Predicted Risk	No	Low	Med	High
No	171	20	9	0
Low	26	143	25	6
Med	17	22	145	16
High	2	7	16	175

(a) A confusion matrix resulting from the prediction of a student's "Risk Profile" utilizing the **C4.5** classification model. The **C4.5** algorithm achieved **79**% accuracy, furthermore, 634 instances were correctly classified and 166 instances were incorrectly classified by the model.

Actual Risk \ Predicted Risk	No	Low	Med	High
No	159	24	11	6
Low	25	145	14	16
Med	14	21	124	41
High	6	7	12	175

(b) A confusion matrix detailing the performance of the **K*** classifier when predicting a student's "Risk Profile" (class variable). The lazy K-Star algorithm achieved **75**% accuracy, furthermore, 603 instances were correctly classified and 197 instances were incorrectly classified by the model.

Actual Risk \ Predicted Risk	No	Low	Med	High
No	174	16	10	0
Low	31	142	22	5
Med	8	17	160	15
High	6	8	17	169

(c) A confusion matrix detailing the performance of the **naïve Bayes** classifier when predicting a student's "Risk Profile" (class variable). After 10-fold cross-validation the naïve Bayes classifier achieved **80**% accuracy, furthermore, 645 instances were correctly classified and 155 instances were incorrectly classified by the model.

Actual Risk \ Predicted Risk	No	Low	Med	High
No	108	23	55	14
Low	38	80	65	17
Med	29	34	125	12
High	37	26	34	103

(d) A confusion matrix resulting from the prediction of a student's "Risk Profile" utilizing the **SVM** classification algorithm. The SVM algorithm achieved **52**% accuracy, furthermore, 416 instances were correctly classified and 384 instances were incorrectly classified by the model.

Actual Risk \ Predicted Risk	No	Low	Med	High
No	174	17	9	0
Low	23	150	21	6
Med	6	15	161	18
High	5	4	14	177

(e) A confusion matrix resulting from the prediction of a student's "Risk Profile" utilizing **Random Forests** classifier. After 10-fold cross-validation the random forests model achieved **83**% accuracy, furthermore, 662 instances were correctly classified and 138 instances were incorrectly classified by the model.

Actual Risk \ Predicted Risk	No	Low	Med	High
No	168	21	11	0
Low	37	138	19	6
Med	6	22	149	23
High	4	4	18	174

(f) A confusion Matrix detailing the performance of the **linear logistic regression** classifier when applied to a dataset of biographical and enrolment observations. The linear logistic regression model achieved **78**% accuracy, furthermore, 629 instances were correctly classified and 171 instances were incorrectly classified by the model.

Figure 7: A set of confusion matrices obtained when classifying the "Risk Profile" variable. We provide the accuracy of each classification model as determined by Equation (*****) along with a count of the correctly and incorrectly classified instances by each model, [6] (sic).

ure based on the conceptual framework developed in a "drop-out from higher education" study [13]. The conceptual framework connects dropout decision to categories of input features, namely, background (family) attributes, individual attributes, and pre-college scores. This framework is better depicted in the Figure 1 provided under the related work section.

The web application depicted in Figure 8 provides a practical tool which university student support programs can utilize for early detection of learners in need of academic support. We argue that the early detection and assistance of students at risk of failure is likely to lead to improved academic performance and eventually higher pass-rates, which translate to increased throughput rates.

The example depicted in Figure 8 illustrates the calculation of the risk profile of a learner based on biographical and enrolment observation. The program predicts the posterior distribution over the four "Risk Profiles", namely, "No Risk", "Medium Risk", "High Risk", and "Low Risk" using the Random forests classifier. The learner in the example is from an urban quantile 3 school in Kwazulu-Natal South Africa, furthermore, individual attributes are provided as follows; scores of 48%, 50%, and 55% for the quantitative, academic, and mathematical literacy National-Benchmark-Tests (NBT) respectively. The learner also completed pre-college courses with scores of; 50% for both core Mathematics and Life Orientation, and 60% for English Home Language. The output presented at the bottom of Figure 8 illustrates that hypothetically the student is 10% likely to complete their degree in 3 years (No Risk), 50% likely to complete their degree in greater than 3 years (Low Risk), 35% likely to drop-out before the end of 3 years (Medium Risk), and 5% likely to drop-out in greater than 3 year (High Risk). With the output obtained, students support-program-coordinators can then decide what assistance will prove most beneficial to the student in the example as the learner posses a high chance of struggling to complete their degree in the allocated time (50% Low Risk).

5 Implications and Conclusions

The expansion of enrolments in South African universities has not been accompanied by a proportional increase in the percentage of learners graduating. In this study, we took on the task of early prediction of a learner's academic trajectory, aiming at identifying those who may struggle in universities, so that proactive learner

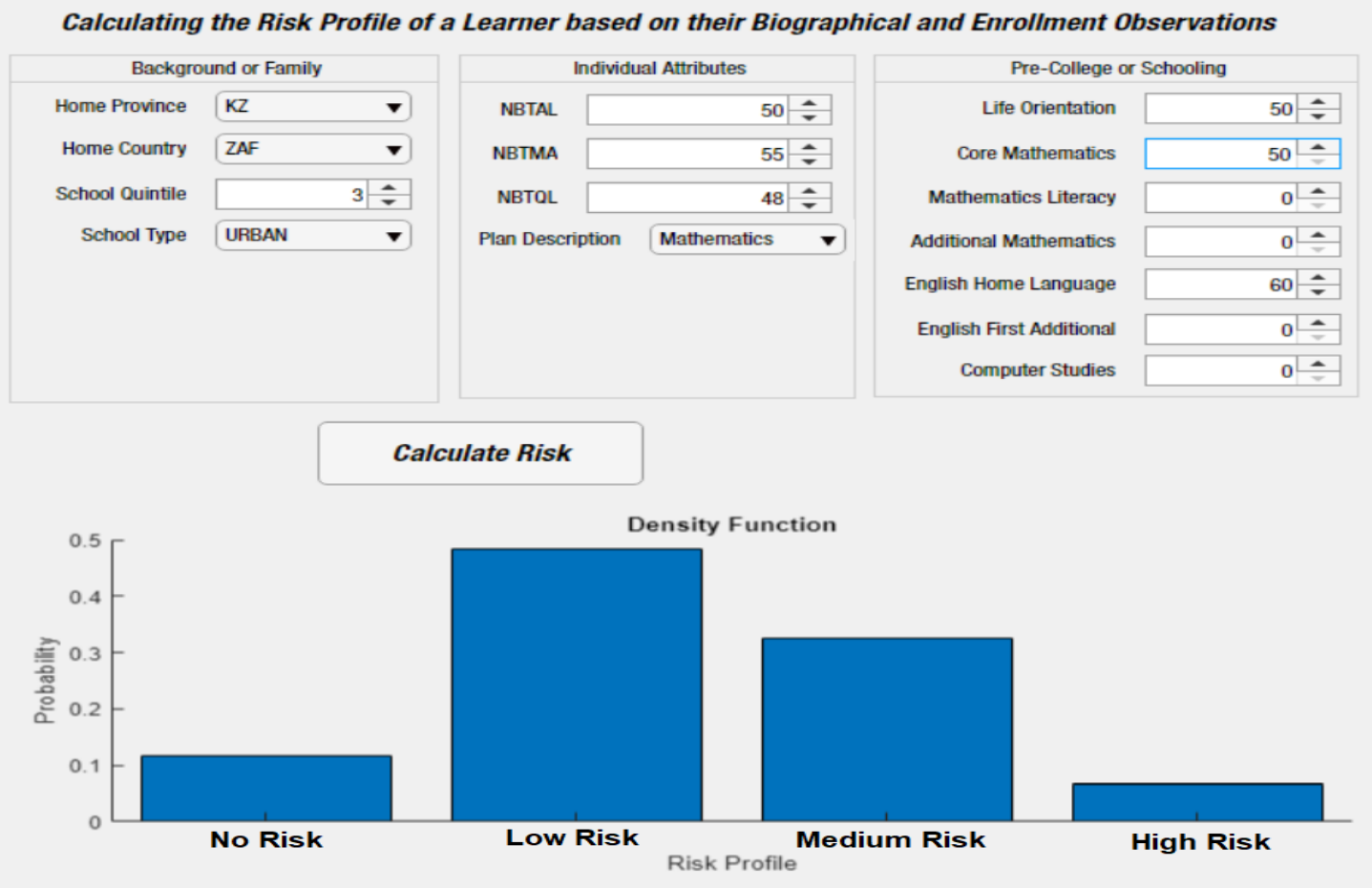


Figure 8: The graphical user interface for the at-risk program, [6](sic).

remediation to promote success may be provided to them.

This paper contributes to the current body of knowledge firstly by introducing an approach involving the prediction of a learner's outcome from the first year of registration until qualifying in a three-year degree. We argue for the early prediction of a learner's entire academic trajectory with the aim to detect those who are likely to benefit from student academic support initiatives. We trained six machine learning models to predict first, second, and final year outcomes from a synthetic data-set. After 10-fold cross-validation this task was completed with great success as all six models attained an accuracy above 83%. Furthermore, an evaluation of the F-measure of accuracy and ROC-curve reveal that these models are making informed-accurate decisions and not simply guessing, therefore, leading to our second contribution involving a real data-set from a research-intensive university in South Africa.

The second contribution of this study is a ranking (through entropy) of features according to their contribution in correctly predicting a learner's "risk profile" (class variable). The ranking of features according to entropy reveals which features are stronger determinants of student success relative to others employed and, in this study, we highlight the seven top-ranked features, namely, "plan code", "plan description", "year started", "age at first year", "streamline", "school quantile", and "Matric Mathematics major".

The third and main contribution made by this study is an interactive program which can predict the distribution over a learner's risk profile utilizing biographical and enrolment observations. The interactive program proposed in this paper can be utilized for early identification of university learners who are most likely to benefit from student support initiatives aimed at improving academic performance. The implication of this study is that university learners can be assisted early in their academic journey increasing their chances of success. Furthermore, the early detection and assistance of learners in need of academic support will result in an improved and enriched learning experience beyond what the student would have experienced if support initiatives were implemented after failure has been detected.

6 Future work

To continue with the work done in this paper, future work may involve: (a) incorporating into our models features from categories not considered such as the "psycho-social attributes category", (b) exploring what courses offered in university possess high failure rates and how good the set of features we employed predict success in these courses, or (c) approaching the problem from a different perspective by attempting to predict which courses are students likely to struggle completing so that support initiatives may be focused on the specific courses.

Conflict of Interest The authors declare no conflict of interest.

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

References

[1] J. V. Winters, "Human capital, higher education institutions, and quality of life," Regional Science and Urban Economics, **41**(5), 446–454, 2011, doi:10.1016/j.regsciurbeco.2011.03.001.

[2] G. Lassibille, L. N. Gómez, "Why do higher education students drop out? Evidence from Spain," Education Economics, **16**(1), 89–105, 2008, doi:10.1080/09645290701523267.

[3] M. Letseka, S. Maile, High university drop-out rates: A threat to South Africa's future, Human Sciences Research Council Pretoria, 2008.

[4] K. G. Bokana, D. D. Tewari, "Determinants of Student Success at a South African University: An Econometric Analysis," The Anthropologist, **17**(1), 259–277, 2014, doi:10.1080/09720073.2014.11891436.

[5] N. Ndou, R. Ajoodha, A. Jadhav, "A Case Study: Educational Data-mining to Determine Student Success at Higher Education Institutions," in The International Multidisciplinary Information Technology and Engineering Conference, 2020.

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

[7] D. H. E. T. Department of Higher Education & Training Republic of South Africa, "2000 to 2017 first time entering undergraduate cohort studies for public higher education institutions," Annual report, 16–28, 2020.

[8] K. R. White, "The relation between socioeconomic status and academic achievement." Psychological Bulletin, **91**(3), 461–481, 1982, doi:10.1037/0033-2909.91.3.461.

[9] M. Sommer, K. Dumont, "Psychosocial Factors Predicting Academic Performance of Students at a Historically Disadvantaged University," South African Journal of Psychology, **41**(3), 386–395, 2011, doi:10.1177/008124631104100312.

[10] E. J. Krumrei-Mancuso, F. B. Newton, E. Kim, D. Wilcox, "Psychosocial Factors Predicting First-Year College Student Success," Journal of College Student Development, **54**(3), 247–266, 2013, doi:10.1353/csd.2013.0034.

[11] E. Osmanbegovic, M. Suljic, "Data mining approach for predicting student performance," Economic Review: Journal of Economics and Business, **10**(1), 3–12, 2012.

[12] M. M. Chemers, L. Hu, B. F. Garcia, "Academic self-efficacy and first year college student performance and adjustment." Journal of Educational Psychology, **93**(1), 55–64, 2001, doi:10.1037/0022-0663.93.1.55.

[13] V. Tinto, "Drop-Outs From Higher Education: A Theoretical Synthesis of Recent Research," Review of Educational Research, **45**, 89–125, 1975, doi:10.2307/1170024.

[14] S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah, E. Alajrami, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology," International Journal of Hybrid Information Technology, **8**(2), 221–228, 2015, doi:10.14257/ijhit.2015.8.2.20.

[15] M. Mayilvaganan, D. Kalpanadevi, "Comparison of classification techniques for predicting the performance of students academic environment," in 2014 International Conference on Communication and Network Technologies, IEEE, 2014, doi:10.1109/cnt.2014.7062736.

[16] V. Ramesh, P. Parkavi, K. Ramar, "Predicting Student Performance: A Statistical and Data Mining Approach," International Journal of Computer Applications, **63**(8), 35–39, 2013, doi:10.5120/10489-5242.

[17] T. Abed, R. Ajoodha, A. Jadhav, "A Prediction Model to Improve Student Placement at a South African Higher Education Institution," in 2020 International SAUPEC/RobMech/PRASA Conference, IEEE, 2020, doi:10.1109/saupec/robmech/prasa48453.2020.9041147.

[18] S. García, J. Luengo, F. Herrera, Data preprocessing in data mining, Springer, 2015.

[19] J. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," Technical Report MSR-TR-98-14, 1998.

[20] D. Böhning, "Multinomial logistic regression algorithm," Annals of the Institute of Statistical Mathematics, **44**(1), 197–200, 1992, doi:10.1007/bf00048682.

[21] B. Krishnapuram, L. Carin, M. Figueiredo, A. Hartemink, "Sparse multinomial logistic regression: fast algorithms and generalization bounds," IEEE Transactions on Pattern Analysis and Machine Intelligence, **27**(6), 957–968, 2005, doi:10.1109/tpami.2005.127.

[22] Y. Wang, "A multinomial logistic regression modeling approach for anomaly intrusion detection," Computers & Security, **24**(8), 662–674, 2005, doi:10.1016/j.cose.2005.05.003.

[23] I. Rish, et al., "An empirical study of the naive Bayes classifier," in IJCAI 2001 workshop on empirical methods in artificial intelligence, volume 3, 41–46, 2001.

[24] R. Ajoodha, Influence Modelling and Learning Between Dynamic Bayesian Networks Using Score-based Structure Learning, University of the Witwatersrand, Faculty of Science, School of Computer Science and Applied Mathematics, 2018.

[25] J. G. Cleary, L. E. Trigg, "K*: An Instance-based Learner Using an Entropic Distance Measure," in 12th International Conference on Machine Learning, 108–114, 1995, doi:10.1016/b978-1-55860-377-6.50022-0.

[26] C.-C. Chang, C.-J. Lin, "LIBSVM," ACM Transactions on Intelligent Systems and Technology, **2**(3), 1–27, 2011, doi:10.1145/1961189.1961199.

[27] J. Friedman, T. Hastie, R. Tibshirani, "Additive Logistic Regression: A Statistical View of Boosting," The Annals of Statistics, **28**, 337–407, 2000, doi:10.1214/aos/1016218223.

[28] N. Landwehr, M. Hall, E. Frank, "Logistic Model Trees," Machine Learning, **59**, 161–205, 2005, doi:10.1007/s10994-005-0466-3.

[29] M. Sumner, E. Frank, M. Hall, "Speeding up Logistic Model Tree Induction," in 9th European Conference on Principles and Practice of Knowledge Discovery in Databases, 675–683, Springer Berlin Heidelberg, 2005, doi:10.1007/11564126_72.

[30] M. J. Zaki, W. Meira, Jr, Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, 2014, doi:10.1017/CBO9780511810114.

[31] J. Quinlan, C4.5: Programs for Machine Learning, Ebrary online, Elsevier Science, 2014.

[32] Y. L. Pavlov, Random Forests, De Gruyter, 2000, doi:10.1515/9783110941975.

[33] S. Sahu, B. M. Mehtre, "Network intrusion detection system using J48 Decision Tree," in 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2015, doi:10.1109/icacci.2015.7275914.

[34] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 237-248 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Leveraging Energy Efficiency Investments: An Innovative Web-based Benchmarking Tool

Filippos Dimitrios Mexis*, Aikaterini Papapostolou, Charikleia Karakosta, Elissaios Sarmas, Diamantis Koutsandreas, Haris Doukas

School of Electrical and Computer Engineering, Decision Support Systems Laboratory, Energy Policy Unit National Technical University of Athens, Athens, 104 31, Greece

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 21 August, 2021
Online: 11 October, 2021

Keywords:
Energy Efficiency Investments
Sustainable Financing
Integrated benchmarking

ABSTRACT

Energy Efficiency (EE) plays a key role in decreasing energy consumption at a European level, while it is considered as one of the most cost-efficient means to achieve carbon reduction and reinforce energy sufficiency and security. EE financing is imperative to implement measures that will lead to achieving the desired carbon neutrality and, thus, avert climate change. The majority of EE investments ideas are abandoned during the first stages of investment generation as there is not enough interest by the involved actors to support the maturing of the idea. The present paper aims to boost EE investments by developing a web-based Tool that evaluates project ideas, connecting them with real financing proposals. All the above are being realised through standardised procedures, establishing a concrete typology of five (5) EE sectors, a well-structured risk assessment methodology of five (5) risk categories and (9) risk factors, and a benchmarking procedure that takes into account four (4) broadly used economic criteria and eleven (11) verified sustainability indicators. All the parameters are calculated using the candidate project data and EU official statistics, formulated into four (4) main criteria that are fed into a MultiCriteria Decision Analysis that performs the project's benchmarking. The presented methodology is being practically tested through the development of three (3) innovative Tools (Assess, Agree, Assign) and a stakeholder consultation process with around 200 participants. The Tools filter and benchmark candidate project ideas, based on the standardised benchmarking and the EU Taxonomy sustainability principles, while connecting the most promising project ideas with state-of-the-art financing methods, such as the Green Loans, the Green Bonds and the Energy Efficiency Auctions. By this token, the developed Tools provenly provide added value to the respective stakeholders, offering standardisation in EE project benchmarking and financing, building trust between investors and projects developers.

1. Introduction

This paper is an extension of work originally presented in 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA 2020) [1]. The conference was realised virtually in 15-17 July 2021, originally organized in Piraeus, by the University of Piraeus and the University of Thessaly. The conference aims to create a forum for researchers and professionals in all areas of Information, Intelligence, Systems and Applications, while its proceeding are published by the Institute of Electrical and Electronics Engineers (IEEE) and uploaded on the its online platform, IEEEXPLORE.

The attainment of the energy and environmental targets set by the European Union (EU) in view of moving to a carbon-free economy by 2050 requires that more intense policies and measures be put into effect, ensuring that the target trajectory is on track [2]. Out of the available policy measures, the improvement of energy efficiency (EE) can play a key role in this transition, leading to the decrease of energy demand and emissions, as well as bringing about economic and social benefits [3] [4]. For this reason, EE has been recognised as the "first fuel" [5], indicating its essential role and the large-scale effects that it could result in.

Mainstreaming EE is the most feasible solution to maintain increased levels of economic and social activities while being a more sustainable and less energy demanding society [6]. The increase of EE implementation goes together with the upscale of EE investments since a prerequisite for realising EE projects is that the capital needed is secured at their development phase. However,

*Corresponding Author: Filippos Dimitrios Mexis, E-mails: pmexis@epu.ntua.gr

according to the current state of play, investments in EE are well below the required level to meet the established long-term targets [7].

Interestingly enough, despite the existence of many EE investments with great prospects at the development phase, only a few of them achieve to get financed; something that has been called as "efficiency paradox" or "energy efficiency gap" [8]. This phenomenon can be ascribed to the transparency gap that EE investments involve [9]. First, there is a shortage of evidence about successfully implemented EE projects, such as about their financial performance and the stakeholders involved within their context. The main factors that halt the financing of EE are the lack of evidence on the performance of such projects, standardised procedures and standards for energy efficiency investments, and the underestimation of its multiple benefits [10]. Moreover, postulating the financial performance of EE investments is a stochastic task at its nature since it is related to several uncertainties (e.g., energy prices), while their evaluation presents high complexity [11].

To the best of our knowledge, a standardised procedure for this scope does not exist up to date [12]. As a result, project developers do not have the expertise or resources to make a convincing financing case for investors, and, on the other hand, investors face difficulties in identifying the EE projects that merit attention [12]. Likewise, financial institutions, given the high transaction costs for evaluating EE investments with respect to their entailed technical aspects, focus only on the creditworthiness of the borrower for deciding whether they provide a loan or not [13]. Therefore, they may exclude from their portfolios many EE projects of great potential. In order to bridge the transparency gap that exists at the initial phases of EE investments, it is of paramount importance that innovative decision support schemes and standardisation Tools will become available towards providing the required information to the involved actors and key players in EE financing [14], also enabling their connection with each other.

The energy efficiency gap is a great field of interest for energy policy scientists, policymakers and research initiatives. Several efforts have been made to mitigate the energy efficiency gap and stimulate the energy efficiency project implementation in Europe by making use of a variety of different financing and funding methods. In [15], the author have developed a benchmarking framework to encourage EE Investments in South Europe under the financing method of Energy Performance Contracting, while the Energy Efficiency Financing Institutions Group (EEFIG) has evolved the EEFIG Underwriting Toolkit to assist financial institutions in scaling up the deployment of capital into energy efficiency [11]. A different approach is followed in [16] by introducing public-private partnerships to finance EE projects.

In addition, the European Union has funded projects and initiatives so as to research and develop benchmarking and standardisation methods for EE. Some examples are the LAUNCH project [17], which aims to accelerate deal closure and pipeline growth for Sustainable Energy Assets through standardised material and the RenonBill project [18] that provides tools to address the residential sector's energy renovation financing demand and to assess bundle investments. Also, the EEnvest project [19] aims to secure investors´ trust in EE actions for existing buildings by developing a combined technical-financial risk evaluation framework focused on the renovation of commercial buildings and the E2DRIVER [20] project is developing a training platform that will boost the collective intelligence of the automotive industry on EE. In the scope of certification, X-tendo [21] and its toolbox introduce features of energy performance certificates to provide public authorities with improved compliance, reliability, usability and convergence of next-generation energy performance assessment and certification.

In this context, this paper proposes a methodological framework that supports, on the one hand, capital providers, i.e., financiers and investors, in finding an attractive project idea that merits attention, giving them access to a pool of project ideas whilst providing a detailed assessment of them. On the other hand, it provides a better understanding of the EE projects' framework to those looking for capital (e.g., project developers), enabling them to acquire the evidence towards performing complete proposals to investors [22].

The presented approach is implemented across three steps, each of which is materialised via a respective standardised Tool. At the first step ("*Assess Tool*"), the initial filtering and identification of attractive project ideas take place, while also the entailed uncertainty is postulated through the calculation of the total risk of failure of the EE investments under examination. Following, the benchmarking of the EE projects that qualified from the first step is realised (*"Agree Tool"*), taking into consideration multiple criteria of diverge nature (e.g., economic, energy), thus assessing them from a holistic point of view. In this effort, a multi-criteria approach is employed, resulting in the categorisation of investments. Finally, at the last step *("Assign Tool"*), investors are connected to EE project ideas with a detailed description of their characteristics, while the most suitable financing instruments are proposed (e.g., Green Bonds), according to their needs. The pool of available investments is composed of the ones that have passed the first two steps successfully.

Apart from this introductory section, the rest of the paper is organised into five sections. Section 2 analyses the methodological framework of the proposed methodology and Tools' development, while Section 3 describes the materialisation of this methodology along with the methods and data that have been used. Section 4 presents the standardised Tools developed, and, finally, Section 5 concludes the paper and suggests the key points for future research.

2. Methodological Framework

The methodology has been structured into five steps, as depicted in the figure below:

Step 1: Identification of the eligible projects.

The proposed methodology aims to benchmark project ideas, which can be considered potential investments at an early elaboration stage. The project ideas should consist of EE measures and interventions that will provenly lead to a decrease in the energy demand of the site implemented. The eligible sectors of activity and type of EE measures should be adequately defined.

Step 2: Review/ Selection of Problem Parameters:

When assessing an EE project, the parameters that should be taken into account are diverse and multivariant, similar to the

nature of EE measures. The indicators and parameters had to be carefully reviewed and selected so as to include and quantify all aspects of EE measures properly, but also reflect in a standardised way the expected financial performance of these potential investments at the early stage of the project's conceptualisation. State-of-the-art typologies methods have been used in order to produce an integrated and standardised benchmarking.

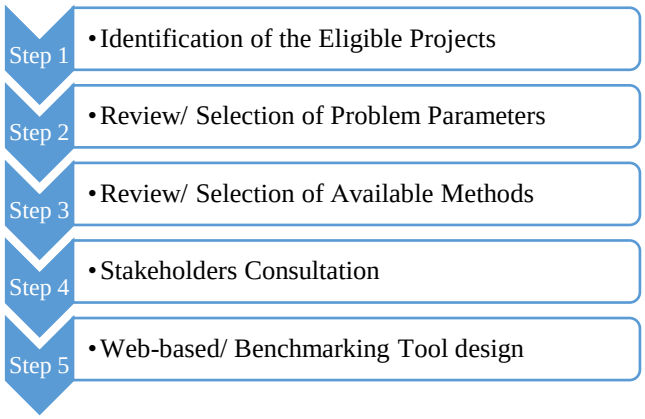


Figure 1: Main methodological steps

Step 3: Review/ Selection of Methods

In this step, a thorough review has taken place in order to run across the most suitable method that can incorporate and weight the selected criteria and benchmark candidate EE projects into predefined classes. Considering all the diverse and distinct factors that EE projects consist of while maintaining versatility, the review was primarily focused on Multi-Criteria Decision Analysis methods (MCDA). Methods such as the PROMETHEE UTADIS and the descendants of the ELECTRE-Tri method, such as the ELECTRE-Tri-B, ELECTRE-Tri-C, ELECTRE-Tri-nC, and ELECTRE-Tri-nB, have been taken into consideration [23].

Step 4: Stakeholders Consultation

The proposed methodology and the Tools developed are oriented towards EE companies and investors interested in investing in these kinds of projects. In order to properly assess and cover their needs and develop Tools that will provide added value to the EE sector, rigorous stakeholder consultation has been realised. The consultation activities included bilateral contacts, discussions, structured interviews, webinars, and workshops.

The consultation process has been organised into two distinct phases. The first phase took place before developing the respective methodology and Tools and is focused on identifying the design parameters. The second part was released after the initial development of the Tools in order to receive feedback and fine-tune the benchmarking process and methods. The stakeholder consultation steps are summarised in the following diagram.

The second part of the consultation included demonstrations and testings, which have been conducted in bilateral meetings with EE companies, project developers, and financers interested in sustainable financing. In addition, feedback on the methodology and the developed Tools has been received from 198 stakeholders by the dissemination of dedicated questionnaires.

Step 5: Web-based/ Benchmarking Tool design

The final step consists of the design and the technical development of the benchmarking Tools, which will assist energy efficiency stakeholders, such as investors seeking to finance sustainable projects and developers that design and implement such projects. The Tools will perform a complete series of actions corresponding to the value-chain of energy efficiency financing. First, the risks and maturity of the inserted investment ideas together with their EU Taxonomy compliance are being evaluated, and then the projects that pass the pre-evaluation are classified into predetermined categories to identify the Triple-A ones. Finally, the last step of Tools matches the investments with state-of-the-art green financing schemes.

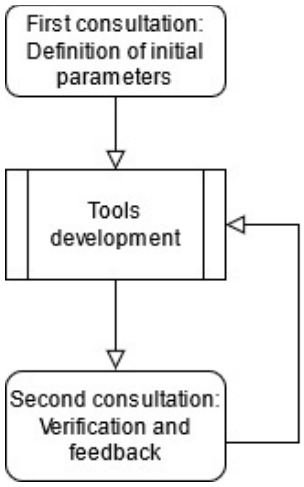


Figure 2: Stakeholder consolation steps

The Tools have been developed in Python 3.0 programming language, taking into advantage the versatility and power framework that it provides. In addition, the Python Django framework has been used to combine the individual scripts for the risk assessment, MCDA method, and criteria calculations into an integrated web-based environment with easy to use front-end and back-end functionalities. The back-end service utilises the SQLite database, while the app is hosted on the National Technical University of Athens' servers.

The Tools' design, functionalities, input-output data and the respective procedures, along with flowcharts, are presented and analysed in depth in Section 4: Web-based Benchmarking Toolbox.

3. Methods and Data

The proposed methodology is based on the actual needs of the EE financing sector, while the design parameters and requirements of the problem have been identified and verified through stakeholder consultation. All the respective processes, such as the identification, review, stakeholder consultation and development, have been realised within H2020 Triple-A project activities and are of vital importance for the assessment and mainstreaming of EE investments. For example, the ELECTRE Tri multi-criteria analysis method has been utilised for the necessary classification of the investments, the project risk elicitation methodology has been crucial for the identification of risky projects which could be immediately rejected from the Triple-A Assess Tool, and the Key Performance Indicators (KPIs) serve as the most suitable criteria

concerning the financial prosperity of the investments. Therefore, before presenting the Triple-A Tools step-by-step, it is essential to deliver the underlying methodological basis, referring to the tools' components.

3.1. Identification of the eligible projects.

The potential investments in EE measures covered by the presented methodology should be pertaining to the following five (5) sectors (i) Buildings, (ii) Manufacturing, (iii) Transportation, (iv) District energy networks, and (v) Outdoor lighting.

The project sectors have been identified in order to form a consistent typology that will include the majority of EE measures in such a way as to simplify the user experience of the developed Tools and batch them in a manner that will benefit the calculations of the respective benchmarking. The eight identified sectors are analysed in the following paragraphs:

- **Building sector**: Building sector projects ideas include any renovation compliant with the energy performance standards set in the applicable building regulations for major renovations transposing the Energy Performance of Buildings Directive (EPBD) or renovations that achieve energy savings of at least 30% in comparison to the baseline performance of the building before the renovation. The baseline performance and predicted improvement shall be based on a specialised building survey and be validated by an accredited energy auditor. Individual renovation measures and installation of renewables on-site that contribute to climate change mitigation by reducing GHG emissions and professional, scientific and technical activities necessary to support and validate buildings' renovation are also included in the Building sector [24]. Also, the construction of new buildings can be included if they meet at least the national requirements for nZEB and have a level of energy performance equivalent to the EPC rating of B (or above).
- **Manufacturing sector:** Under the Manufacturing sector are included projects of EE measures in industries in order to manufacture products with high EE, compared to relevant products already in the market, of manufacture products, key components and machinery that are essential for implementing projects that achieve high EE. More generally, any intervention in industries that aim to reduce energy losses and to improve the energy performance of equipment, machinery or install renewable energy systems can be included. Equally, improvement of existing industrial production lines and processes to achieve significant improvement of their EE and reduction of GHG emissions per unit of product are covered.
- **Transportation sector**: The Transportation sector includes measures in public transport vehicles, which can achieve substantial GHG emission reduction in terms of CO2e emissions per passenger- kilometre (gCO2e/km). The potential project should be referred to zero direct emissions land transport solutions, such as light rail transit, metro, tram, trolleybus, bus and rail, or fleets with direct emissions below 50 gCO_2e/pkm until 2025 (non-eligible thereafter). Additionally, projects to finance the replacement of corporate fleet with new, more efficient passenger cars are included. The vehicles can be light commercial vehicles and category L vehicles (2- and 3-wheel vehicles and quadricycles).
- **District Energy Networks**: As District Energy Networks are considered the construction projects, retrofit, expansion and operation of pipelines and associated infrastructure for distributing heating and cooling. For a project to be eligible, the system should meet the definition of efficient district heat/cool systems in the EU Energy Efficiency Directive and EIB energy lending policy regarding phasing out fossil fuels ("Supporting the energy transformation"). In detail, "efficient district heating and cooling" is defined as a district heating or cooling system using at least 50% renewable energy, 50% waste heat, 75% cogenerated heat or 50% of a combination of such energy and heat.
- **Outdoor lighting:** The Outdoor lightning sector includes any project focused on the renovation or renovation and expansion of existing outdoor lighting installation (e.g. street lighting, lighting of public squares) that lead to a high EE performance. All projects should foresee an appropriate control and measurement system.

3.2. Review/ Selection of Problem Parameters

The problem parameters are the key elements that form the criteria used in the methodology. The criteria, in turn, are defined and quantified by relevant indicators that can reflect EE project characteristics. In the presented methodology, the indicators for the benchmarking include data regarding the EU Taxonomy compliance of projects, risk estimations per country, sector, economic and financial KPIs based on project's characteristics, and the multi-benefits of projects as depicted by the potential contribution of EE measured by the Sustainable Development Goals. The typology and formation of the criteria have been inspired by the Final Standardised Tools Deliverable [25] published within the context of the Triple-A H2020 project. In the following paragraphs, each indicator type will be presented and explained.

- **EU Taxonomy compliance:** The methodology has pointed out the EU Taxonomy Technical screening criteria for each the identified project's sector. The criteria include thresholds regarding the technical characteristics of the proposed EE projects, which vary from one sector and type of activity to another. The technical screening criteria of the EU Taxonomy have been used to all the sectors, except the Outdoor Lighting sector, for which a dedicated set of criteria has been used, based on the outcomes of the PREMIUM LIGHT PRO project [26].
- **Risk Assessment:** The risk assessment of the potential investments in EE focuses on estimating in a qualitative and quantitative approach the total investment risk caused by a variety of factors. These factors are linked to the nature and principles of EE or country and sector individual characteristics [27]. The factors have been adequately identified and grouped into five (5) distinct risk categories. The categorisation aims to group the identified risk factors based on their conceptual characteristics and to facilitate the quantification procedure of the investment's risk. The aggregation of the risk parameters into a total investment risk

has been realised by applying a weighted average formula, as analysed in Section 3. The risk assessment methodology has been inspired by the Final Report on Risks of Energy Efficiency Financing and Mitigation Strategies Typology published within the context of the Triple-A H2020 project [28].

- **Economic KPIs:** The economic parameters could not be missing from a benchmarking methodology of potential investments. A selection of classic financial indicators, broadly used by investors, has been carried out, while the calculation methods have been parameterised to fit the cashflows and particularities of EE measures.
- The indicator's calculation has been based on EU Directives and Regulations on Cost-Benefit Analysis of Investment Projects, while data from EU official statistics have been used to provide a standardised, unbiased result. The indicators included in the benchmarking procedure are the Net Present Value (NPV), the Discounted Payback Period, the Internal Rate of Return (IRR), and the Cost-Effectiveness (or Avoidance Cost) [25].
- **Sustainable Development Goals:** An aggregated indicator has been included as a criterion in the benchmarking methodology, reflecting the necessity to improve EE in the case study countries and the identified sectors. The indicators used are official Eurostat statistics linked with the progress of the UN's Sustainable Development Goals. These indicators reflect the current situation of EE per country and sector, the situation regarding energy poverty, environmental pollution and greenhouse gas emissions, and energy security - sufficiency. Overall, a total of 11 sustainability indicators have been included, linked with the multi-benefits of EE and aggregated into a single criterion.

3.3. Review/ Selection of available methods

The review and selection of the most appropriate method is a crucial element enabling the benchmarking of the EE investment ideas to be performed in an affluent and efficient way. The problem parameters and requirements, such as the benchmarking of EE project ideas, in predefined classes based on several criteria, emphasised the need for a multi-criteria approach.

ELECTRE Tri was considered the most appropriate methodology to address the problem, as it accumulates key advantages that match the needs of the problem in question. The determinant factor that led to select the ELECTRE-Tri Multiple Criteria Decision Analysis as the benchmarking method is that it allows the classification of the candidate projects into predefined categories (alternatives) by constructing outranking relations and comparing each alternative to a limit [29].

Also, one of the strong points is that ELECTRE-TRI handles both qualitative and quantitative data, meaning that it can deal with the imperfect nature of knowledge [30]. It is a multi-criteria decision-making method which is mainly utilised for classification problems, providing labelling of the alternatives based on a set of conflicting and non-overlapping criteria.

Not to mention the discrimination thresholds that the selected method incorporates, which eliminate the serious drawbacks of imprecision, ill-determination, and uncertainty that can undergo with the data available and used. The theoretical background, steps and processes performed in the ELECTRE Tri method are thoroughly presented in the relevant bibliography in [31], [32].

3.4. Calculation of the EE investments risk

In the context of the proposed methodology, an important pillar is calculating the investment's total risk, which was determined by combining both a qualitative and a quantitative approach. The major factors that affect the security of the implementation of investment that mainly consist of EE measures have been classified into the following five categories. Each risk category consists of risk factors that are estimated based on formulas or linguistic values, according to their nature.

- Financial risk is related to the creditworthiness of the applicant for the loan/financing. It consists of the risk factor "creditworthiness of the borrower". This factor indicates the financial capacity of the borrower to pay off his debt. It is considered a critical factor from the perspective of a financing institution or bank when considering giving a loan [33,34]. The risk factor is qualitative, and it is estimated by assigning linguistic values out of the scale ("Low", "Medium", "High", "Unknown") for evaluating the creditworthiness of the applicant for the loan.
- Behavioural risk is related to the rebound effect that may exist in the context of the inspected EE investment. It consists of the risk factor "rebound effect", which describes a specific behavioural bias in sites that have been implemented EE measures. This bias affects the end-user, and it usually emerges when the implementation of an EE project leads to lower costs for energy services combined with an increase in the demand for such services. It is a qualitative risk factor, and it is being quantified by assigning linguistic values out of the scale ("Insignificant ", "Low ", "Medium", "High", "Very High") to each type of project based on literature review and questions replied by the applicant of the project idea, as inspired by the Industrial Energy Accelerator project by UNIDO [35] .
- Energy Market & Regulatory risk consists of factors related to the country's energy prices and energy taxes volatility (and in which the investment will be implemented) and the requests for issuing work permits that may exist in the context of the inspected project. Energy prices and taxes volatility is associated with the price risk in EE investments, and it is a quantitative factor calculated based on the coefficient of variation (CV) of monthly values of Consumer Price Index (CPI) for the last fifteen years (01/2005-12/2019), as reported by OECD. The risk factor "request for issuing project permits" signifies the legislative complexity for the completion of a project and consists of a qualitative analysis considering the amount and type of project permits needed for the implementation of the project.
- Economic risk is differentiated from financial risk, and it is related to the country's economic environment that the investment takes place. The risk factor that facilitates the quantification of this project category is the "weak economic environment", which is connected to, among others, the country's interest rates, inflation, availability of finance, etc. [36]. It is a qualitative factor, quantified by the credit ratings

provided by Standard & Poor's (S&Ps) for the case-study countries.

- Technological, Planning and Operational risk is related to the technical complexity, the initial savings assessment, the implemented equipment, the project design, and the Operation & Maintenance of the inspected project. In more detail, the risk factors that are included in this Risk Category, are qualitative and they are quantified by assigning linguistic values out of the scale ("Insignificant", "Medium", "Very high") to each type of project based on questions answered by the applicant of the project. Specifically, the "Low quality of initial savings assessment" risk factor considers how the energy savings assessment and the baseline definition have been conducted. The "Implementation of low-quality equipment or poor project design" considers the existence of proof about the quality of the equipment to be installed, the experience of the team for planning and implementing the project and who conducts the technical implementation of the project. The "Inadequate Operation & Maintainance" considers the experience of the end-user in using and operating the proposed equipment, the existence of a maintenance plan, product warranties and Measurement & Verification protocols and standards.

Following the determination of the risk factors' values, the risk categories' values are then calculated by averaging the values of the risk factors of which each category is composed. The risk is being calculated by deploying the method of the weighted average

The total project risk value is the weighted arithmetic mean of the risk categories' values and is calculated as follows:

$$\boldsymbol{Project\ risk\ value} = \sum_{i=1}^{n} \boldsymbol{w_i} \times \boldsymbol{R_i} \quad (1)$$

w_i : i = 1, …, 5 are each risk category's weight

R_i : i = 1, …, 5 are each risk category's value

The weights used in the equation emerge from the stakeholder consultation process that has been realised within the Triple-A project's activities. The weights values have been calculated through the respective answers in linguistic variable questions regarding the investors' preferences and their perception of energy efficiency projects' risk, and project developers' risk prioritisation according to the estimated impact on the EE projects implementation.

3.5. Final formation of the KPIs serving as classification criteria

The application of the ELECTRE Tri method, which was selected to follow through with the problem in question, is based on the elicitation of a series of criteria. These criteria emerge from the identified project parameters, as presented in Paragraph 3.2, and should follow the principles and restrictions implied by the ELECTRE-Tri methodological background. The final formation should produce non-overlapping criteria and which must affect directly the evaluation of the examined alternatives (EE investments). The selected criteria reflect the economic prosperity, the risk and the sustainability of the potential investment, based on specific KPIs.

The first two criteria consist of economic KPIs broadly used in the classic benchmarking of potential investments in the financing sector. The KPIs are calculated taking into account the expected economic savings of the potential EE investments, while the KPIs parameters are based on the EU Directives and Regulations on Cost-Benefit Analysis of Investment Projects or reflected directly from EU official statistics to provide standardised unbiased results. The origin of the applied KPIs aspires to build confidence among investors and facilitate financing bodies and EE funds to detect and finance projects that meet the necessary criteria rapidly.

The **first criterion** is the economic indicator of Cost-Effectiveness. Cost-Effectiveness refers to the ratio of dis-counted benefits over the discounted costs. In its simplest form is a measure of whether an investment's benefits exceed its costs [37]. In the proposed methodology, the Cost-Effectiveness is calculated based on the project cost per kWh saved

The **second criterion** is also an economic indicator selected by the respective stakeholder of each potential investment. The selection can vary between the KPIs listed below:

- Net Present Value (NPV), reflecting the risk and cashflows discount by quantising it through the discount rate, the profitability of the investment by involving in the calculations the yearly income, the operational costs and the initial investment [38].
- Discount Payback Period, which quantifies the number of years necessary to recover the investment's budget while accounting for the time value of money [39].
- Internal Rate of Return (IRR) is a rate of return used in capital budgeting to measure and compare the profitability of investments.
- Cost-Effectiveness. Which refers to the ratio of dis-counted benefits over the discounted costs. In its simplest form is a measure of whether an investment's benefits exceed its costs [37]. In the proposed methodology, Cost-Effectiveness is calculated based on the project cost per kWh saved

The **third criterion** comprises the aggregated total investment risk, reflecting the likelihood of the unwanted case that the potential investment will not achieve the proposed savings in energy and thus fail to deliver the expected economic benefits. The aggregated total risk is calculated through the procedure described in detail in Paragraph 3.4.

The **fourth criterion** refers to the sustainability aspects of the proposed investment. Expect the financial and risk criteria, sustainability criteria are considered essential and have been incorporated into the methodology, as EE investments provide multi-benefits in social and environmental aspects, except energy savings. For example, they contribute to environmental and social factors while reducing energy consumption, leading to lower GHG emissions, improving comfort in living spaces, and industry productivity.

The sustainability criterion for the proposed methodology has been based on the United Nations' Sustainable Development Goals indicators have been selected based on quantitative analysis, as recorded by Eurostat's statistical indicators. These indicators reflect the current situation of EE, energy poverty, and environmental pollution. The identified indicators are directly linked with the United Nations' Sustainable Development Goals (SDG) Agenda [40], mainly to those related to the energy sector and environmental protection. The criterion emerges from the aggregation of statistical data per country or per country and sector, depending on the dataset's nature.

The utilised SDG factors that were aggregated in order to construct the SDG criterion for the ELECTRE Tri method are presented in Table 1 Selected SDG indicators.

Table 1: Selected SDG indicators.

Name	Criterion
Arrears on utility bills	It reflects the share of (sub)population (%) having arrears on utility bills, based on the question "In the last twelve months, has the household been in arrears, i.e. has been unable to pay on time due to financial difficulties for utility bills (heating, electricity, gas, water, etc.) for the main dwelling?"
Total population living in a dwelling with a leaking roof, damp walls, floors or foundation, or rot in window frames or floor	It indicates the share (%) of the population experiencing at least one of the following basic deficits in their housing condition: a leaking roof, damp walls, floors or foundation, or rot in window frames or floor.
Population unable to keep home adequately warm by poverty status	It indicates the share (%) of the population, who are unable to keep home adequately warm. Data for this indicator are being collected as part of the EU Statistics on Income and Living Conditions (EU-SILC) to monitor the development of poverty and social inclusion in the EU.
Primary energy consumption	It quantifies the Gross Inland Consumption in toe, excluding all non-energy use of energy carriers (e.g. natural gas used not for combustion but for producing chemicals).
Energy import dependency	The criterion shows the share (%) of the total energy needs of a country met by imports from other countries. It is calculated as net imports divided by the gross available energy
Final energy consumption in the industry sector	It includes all the energy supplied to the industry sector in toe; excluding deliveries to the energy transformation sector and the energy industries themselves).
Final energy consumption in the transportation sector	It measures the energy consumption of the transportation sector in toe, excluding deliveries to the energy transformation sector and the energy industries themselves.
Final energy consumption in other sectors or commercial and public services	It indicates the energy supplied to non-categorised sectors, commercial and public services in toe.
Final energy consumption in households per capita	The indicator measures how much electricity and heat every citizen consumes at home (kgoe/capita), excluding energy used for transportation. Since the indicator refers to final energy consumption, only energy used by end consumers is considered.
GHG emissions from energy consumption	The data are based on measures of the European Environmental Energy Agency and represent the GHG emissions caused by the energy sector in ktn $CO_{2\text{-eq}}$.
GHG emissions from the industrial sector	Similar to C_{10}, the C_{11} criterion contains the GHG emissions (in ktn $CO_{2\text{-eq}}$) caused by the industrial sector, as reported by the European Environmental Energy Agency.

The four criteria mentioned above are calculated automatically by the Triple-A Web-based Benchmarking Toolbox presented in Section 4. The calculation is being performed in the back-end of the respective Tool, while the user only needs to provide the data required for the calculations and reply to the requested questions. As described thoroughly in Paragraph 3.7, the four criteria feed a MultiCriteria Decision Analysis Method, which in turn performs the final benchmarking of the Web-based Benhcmaking Toolbox.

4. Web-based Benchmarking Toolbox

The methodology, typology, and expected results analysed in the previous paragraphs have come into existence by the implementation of web-based Tools [1]. The Tools form part of an integrated Toolbox, namely the Triple-A Toolbox, which has been implemented in the Python 3.0 programming language.

The Toolbox aims to perform an initial filtering of the EE project ideas, benchmark them into predefined classes, and connect them with investors via concrete financing methods. The examined projects, along with their attributes (project country, project sector, aggregated project risk etc.), are saved into a database in order to be identifiable and open for financing requests.

The Toolbox consists of three distinct Tools, namely the Assess Tool, the Agree Tool and the Assign Tool. The project architecture is designed in a way that each Tool can be used separately, thus offering interoperability among the different modules. The steps for each module of the Triple-A Toolbox are presented in Figure 3.

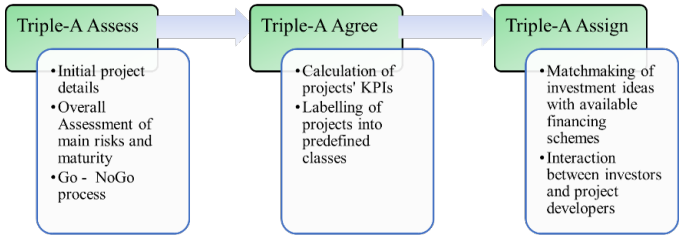


Figure 3: Web-based Tools functionalities

4.1. Assess Tool

The Assess Tool aims to perform initial filtering and screening of the EE project ideas. The main idea of the Assess Tool is the overall assessment of the main risks and the maturity, the candidate EE investments, and their EU Taxonomy compliance. To evaluate the risk of each EE investment, its main characteristics are considered, such as the country of implementation and the technology implemented within its context. On the one hand, the considered risks are related to the region where the project is implemented, the proposed technologies that are used, and the project management. On the other hand, the maturity of the investments is related to their readiness for implementation. In this respect, several key factors that could possibly affect the project's success and its access to sustainable financing have been considered (e.g., risk level, size of the investment, type of EE projects, EE measures eligible, EU Taxonomy compliance etc.), serving as eligibility criteria, aiming to assure the security and trust of the potential investments, which is a necessary action in order to proceed to further analysis.

The flowchart of the Assess Tool methodology is presented in Figure 4:

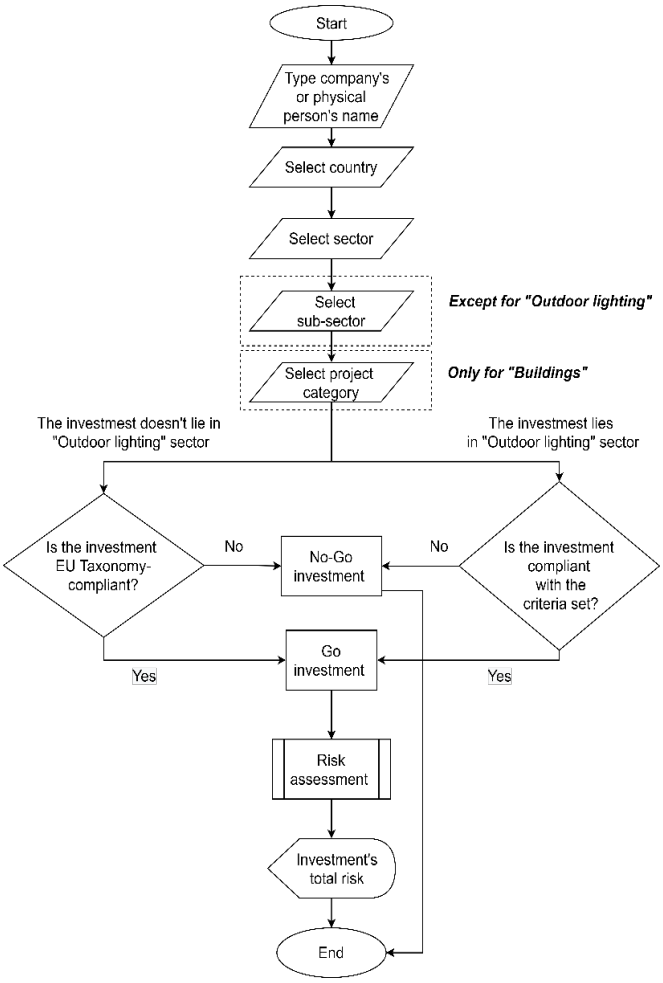


Figure 4: Assess Tool Flowchart

From the user's perspective, this Tool includes the following steps:

Step 1: Inserting the basic information of the candidate EE project (Company's or Physical person's name, country, sector).

Step 2: Selecting the project's subsector (except for the "Outdoor Lighting" sector). The project sectors and subsectors are presented in **Table 2**.

Step 3: Selecting the project category (Only for the building's sector; Table 4). In case that the investment concerns a building, select the project category, as presented in Table 3.

Table 2: Assess Tool Sectors

A/A	Sectors	Sub-sectors
S_1	Buildings	Residential
		Non-Residential
S_2	Manufacturing	Hydrogen
		Iron and Steel
		Aluminum
		Cement
		Low carbon technologies
		Fertilisers and Nitrogen
		Other organic basic chemicals
		Other inorganic basic chemicals
S_3	Transportation	Public Transport
		Passenger cars and light commercial vehicles
S_4	District Energy Networks	District Heating / Cooling Distribution
		Installation and operation of electric heat pumps
		Cogeneration of Heating / Cooling and Power
		Production of Heating / Cooling
S_5	Outdoor Lighting	-

Table 3: Assess Tool project categories.

Sectors	A/A	Project categories
Buildings	P_1	Building envelope retrofits
	P_2	HVAC&R retrofits
	P_3	Lighting appliances' retrofits
	P_4	Automatic control retrofits
	P_5	RES installations
	P_6	Construction of new buildings
Manufacturing	P_7	Manufacturing-specific retrofits
Transportation	P_8	Purchase of new vehicles
District Energy Networks	P_9	District Energy Networks retrofits/ expansion
Outdoor Lighting	P_{10}	Outdoor lighting retrofits

Step 4: After having imported all the project-specific information, select whether the investment is EU taxonomy eligible. Depending on the project sector, as well as the additional information is given (Project Type, Project Category, etc.), an EU taxonomy list appears, including the thresholds that need to be met by the investment and some explanations on the requirements, such as the metrics of each threshold. Then, the user should carefully check if the investment fulfils all the thresholds. In case that all the requirements are met, then the Yes checkbox should be selected, stating that the investment is taxonomy compliant. Finally, the user must fill in a checklist with taxonomy criteria, according to whether the investment is compliant with the criteria set in each case or not.

Step 5: Message notification to inform the user whether the project is Go or No-Go. If the answer to the EU Taxonomy compliance questions sets a project as No-Go, the user is notified immediately, without being able to complete the remaining steps of the Tool.

Step 6: Answer a series of project-specific questions, which are taken into consideration for the aggregated investment risk calculation.

Step 7: Calculation of the project's aggregated risk, according to the methodology described in **Paragraph 3.4**.

Step 8: Analytic presentation of the project's risk calculations, including the calculation for each distinct risk category, as described in Paragraph 3.4 and the total calculated risk of the investment.

4.2. Agree Tool

The Agree Tool implements the second pillar of the proposed benchmarking scheme for the EE project identification and labelling. The Agree Tool can only be applied to EE project ideas

that have successfully passed the Assess Tool and have been identified as "Go" projects. The Tool supports and facilitates the identification of the so-called Triple-A investments [12], e.g. the investments that have a high performance in the selected KPIs, as described in Paragraph 3.5. In other words, Triple-A EE investments are defined as the ones with extremely strong potential to meet their energy-saving targets, already from their conceptual phase. The candidate projects that are inserted into the Agree Tool are classified into one of the following categories: "Triple-A", "Reserved", or "Rejected", according to their performance on the evaluation criteria.

The flowchart of the Triple-A Agree Tool methodology is presented in Figure 5.

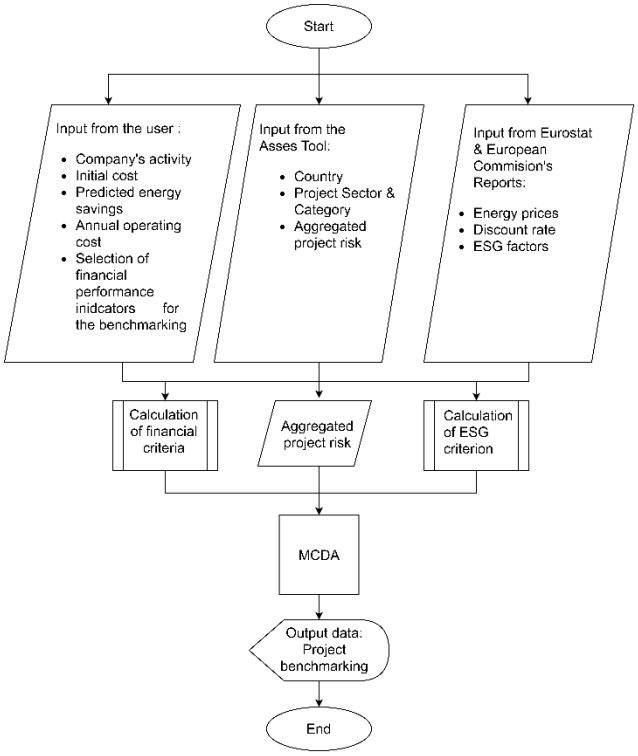


Figure 5: Agree Tool flow chart.

The steps for the Agree Tool from the user's (investor, banker, financer) point of view are described in detail as follows:

Step 1: Definition of project country and project sector in order to set the ELECTRE Tri algorithm's thresholds to the customised values.

Step 2: Select the criteria which will be used for the evaluation of the project and also insert the required input for the calculation of the criteria. The ELECTRE Tri algorithm is applied on the basis of 4 evaluation criteria, as described in Paragraph 3.5. There are two financial criteria (K1 & K2) combined with one aggregated risk criterion (K3) and one SDG criterion (K4). The criteria of ELECTRE Tri must comply with the restriction of consistency. Therefore, two financial KPIs should be applied to the algorithm. The aggregated risk criterion is fed to the model from the Assess Tool's risk calculation process. Finally, the SDG criterion stems from the methodology discussed in Paragraph 3.5.

Step 3: Weight adjustment of the ELECTRE Tri criteria according to the importance of each factor and the user's preferences. The established weights reflect the effect of each criterion on the final benchmarking. Extreme weight values (0% or 100%) are rarely assigned to a criterion. At the same time, the sum of all criterion weights should be 100%. The Agree Tool supports default values for weights which can be easily changed by the user from the Tool's interface. The criteria thresholds of the benchmarking classes have been set according to outcomes of similar web-based Tools and databases that accumulate data and knowledge from implemented EE projects, such as the DEEP platform [41], as well as from input provided by experts via the stakeholder consultation. Until March 2021, a number of almost 100 users (project partners, Triple-A project Advisory Board Members and external stakeholders) have been signed up and tested the developed Tools, providing valuable feedback. In addition, the Triple-A Tools have been tested with real projects that have been collected in an EE project pipeline within the Triple-A project activities.

Step 4: After the execution of the multi-criteria algorithm analysis, the candidate EE project is classified into one of the predefined categories: Triple-A, Reserved, and Rejected. The three categories are explained in the paragraphs below:

Triple-A Projects: This category includes projects which could be considered as the best investment opportunities and, thus, merit attention by the funding organisations. These are investments that demonstrate extremely strong potential to meet their energy-saving targets, from their primitive phase, when they are still considered as project fiches from the funding institutes. The above characterisation comes in accordance with the definition of the respective Triple-A investment grade, which refers to investments with a strong capacity to fulfil their financial commitments by achieving the expected performance targets.

Reserved Projects: This category includes a series of projects which are profitable, but they do not guarantee that they will fulfil the predefined performance targets. Such projects have achieved a good but not outstanding performance according to the evaluation criteria. Even though they have not managed to achieve the top score in the evaluation methods, they are capable of repaying the initial capital invested and of contributing significantly to the energy savings.

Rejected Projects: This category includes projects which have a poor total performance according to the evaluation criteria. Their risk is higher than the predefined threshold, and they seem incapable of recovering the total investment. Also, projects which were labelled as "No-Go" during the Triple-A Assess phase are also characterised as Rejected.

4.3. Assign Tool

The Assign Tool implements the final step of the integrated methodological framework. The main functionality of this Tool is the matchmaking of investment ideas, which have successfully passed the Assess and Agree tools, with a list of potential financing schemes. Thus, it aspires to support all involved parties to achieve project delivery, such as contracts, underwriting procedures, etc.

The Assign Tool includes numerous user interfaces in order to facilitate all beneficiaries and EE financing instruments. The platform's main target groups are, on the one hand, financing bodies and, on the other hand, companies and project developers. From the aspect of financing bodies, commercial or green investment banks are included, as well as investment funds, developers and managers of financial products, Unit Investment Trust and several types of financers that target financing Triple-A

opportunities to include in their portfolio. From the aspect of project developers, the Tool is oriented towards EE companies, credit professionals, ESCOs, construction companies and management investment companies that try to attract financing in order to implement their projects.

The flowchart of the Assign Tool methodology is presented in Figure 6.

The Assign Tool provides three different types of financing schemes, aiming to be as flexible as possible and to provide several options for both financing bodies and project developers. The three supported types of financing are (a) *Green Loans and Mortgages* and (b) *Green Bonds*, and (c) Energy Efficiency Auctions.

Loans are one of the most traditional financing mechanisms, although their characteristics may vary significantly for different financing cases. During the last decades, financing institutes have prioritised environmental issues and objectives, and Green financing products have emerged as a new trend. [42]. The developed Tool has classified the financing cases to Green Loans and Green Mortgages. The need for such classification has emerged because of the popularity of such financial products, even though mortgages are a form of a loan.

The Assign Tool supports the procedure of financing proposals between investors and project developers. Investors are enabled to send financing proposals for projects, having the flexibility to set the preferred loan parameters. Any type of financing proposal is acceptable, e.g., investors are not obliged to cover the total cost of a project. They could partially finance the project and select the preferred percentage of leverage. Project developers can assess the proposal and accept or reject it by using their personalised interface of the Tool.

The second type of financing, which is supported by the Triple-A Assign Tool is Green Bonds. A bond is a type of investment that represents a loan between a borrower and a lender, being considered a debt investment. Green Bonds are any type of bond instrument where the proceeds will be exclusively applied to finance or re-finance, partly or fully, new, or existing eligible Green Projects [43]. The Assign Tool introduces a new methodological approach, enabling an investor to come across companies and project developers to aggregate their projects into a portfolio. EE projects are often small and fragmented in nature in comparison with large energy infrastructure projects such as power plants [44]. Project aggregation expedites EE financing, as bundling small and medium-size projects into a larger build-up reduce the overall risk and is more appealing to investors.

The Assign Tool creates the ideal circumstances to create links between primary entities that are involved in the procedure of bond issuance. Based on the "Proposal for an EU Green Bond Standard" [45] published in June 2020 by the EU Technical Expert Group on Sustainable Finance, the Assign Tool supports Green Bond issuers to be compatible with the EU Green Bonds Standard by providing all necessary functionalities and information.

The third and final financing method covered by the Assign Tool is the EE auctions. The EE auctions is a policy instrument aiming to achieve energy savings at highly attractive prices and avoiding deadweight effects associated with financial support [46]. The procedure resembles the renewable energy auctions already implemented in a number of countries. Lessons learned and previous experience of EE bidding and auctions in Switzerland have proved that tendering could possibly help to save 26% of public funds as compared to more traditional subsidise programmes [47].

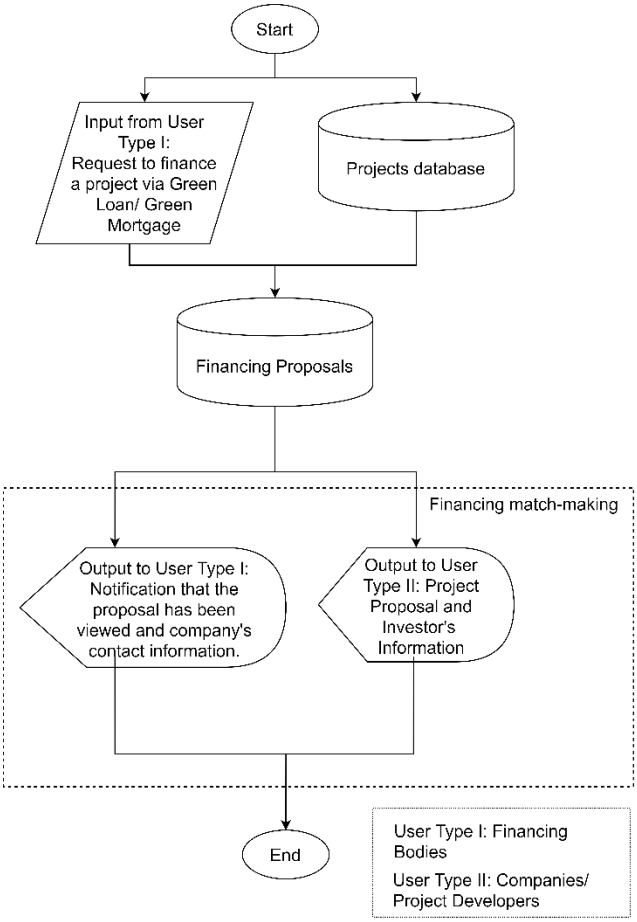


Figure 6: Assign Tool flowchart

The Assign Tool supports the publication of EE auctions in the web-based Tool, enabling interested key actors to get informed about the available auctions. In addition, with the support of the database of projects, followed by the respective information about each project characteristic, stakeholders can filter projects and identify the ones that perfectly match the parameters of the auction. As well, the creation of a printable portfolio of projects is obtainable, and the contact details of each company or project developer are accessible so as to proceed to further actions.

The Triple-A Tools are accessible through the Triple-A Standardised Toolbox platform (https://aaa-h2020.eu/tools), which was implemented in a user-friendly environment, rendering the whole benchmarking procedure flexible and easy to follow, while also offering a communication canal between project developers and financing bodies through flexible options such as green loans and green bonds and energy efficiency auctions.

5. Conclusions

It is widely accepted that many investments in EE are well below the required level to meet the established long-term targets, both in economic and energy savings levels. The present manuscript introduces both an integrated **methodology** to facilitate the financing of energy efficiency projects, as well as a fully-

featured **Web-based Tool**, namely the **Triple-A Standardised Tools,** that incarnates that methodology.

The Triple-A Standardised Tools aspire to set the way for **identifying** and **financing** Triple-A investments, based on a solid methodological procedure that takes into consideration many different aspects of the examined investment. The Triple-A Tools are designed with an integrated process, constituting them orientated towards different types of stakeholders in EE financing.

In other words, the developed Tools facilitate project developers to compare and label their projects in a standardised way through the Assess & Agree Tools, while also providing easy access to financiers, bankers, and investors to finance bankable green projects through the Assign Tool. The Triple-A Standardised Tools have been practically presented to project developers and investors, receiving positive feedback and raising interest in their functionalities through targeted stakeholder consultation activities.

The **methodology** described in the present paper facilitates EE financing in a practical way, differentiating from existing methodologies, by providing integrated standardized procedures. Also, the methodology:

- Offers a concrete typology of sectors and project types, risk categories, KPIs, benchmarking, and financing Tools widely accepted and known by stakeholders and key actors that do business in the EE sector.
- Incorporate state-of-the-art elements of sustainable financing, such as the EU Taxonomy and the UN's Sustainable Development Goals,
- Ensures interoperability with similar databases and Tools, providing an integrated experience and standardisation to the users.
- Supports innovative financing methods, which are still in the rollout stage but are expected to change the scenery of EE financing, such as the EE auctions.

The **Tools** developed based on the described methodology provide an integrated approach in a variety of aspects of EE financing:

- the **Assess Tool** addresses the need for evaluation of the risks and maturity of the examined EE investment ideas along with their EU Taxonomy compliance.
- the **Agree Tool** provides a solid classification mechanism for the projects that successfully pass the Assess Tool, identifying the remarkable project as the Triple-A ones. Finally,
- the **Assign Tool** supports the matchmaking of the potential investments with state-of-the-art green financing schemes and creates a community of EE stakeholders (Companies, Project Developers, Investors, Financing Bodies).

The Triple-A Tools provide a one-stop-shop approach to connect project developers with investors, by assessing the maturity and the EU Taxonomy compliance of potential EE projects, while benchmarking them in a standardised way. The respective tools have been already used and tested by stakeholders within the consultation activities of the H2020 Triple-A project [48]. Valuable feedback has been provided and considered and updates have been realised in the Web-based Benchmarking Tool. In addition, in order to confirm that the developed Tools provide added value to energy efficiency professionals, respective questions have been added to the stakeholder consultation questionnaires and the 53% have replied that they find the Triple-A Web-based benchmarking Tool useful to their profession.

Furthermore, engineering and energy efficiency companies have demonstrated interest in learning more about the Tools and their functionalities, while the Tools' developing team was asked to make a live presentation in the Corporate Social Responsibility School 2021 [49] in July 2021. The Tools will also be demonstrated in the Regional Workshop organised by the Triple-A project. In the events mentioned above participate green and sustainable financing investors, engineering companies, project developers that develop energy efficiency projects and policymakers.

Until August 2021, 627 users have visited the Triple-A Tools online platform, while almost 120 stakeholders have been signed up. In addition, 113 projects have been inserted into the Tools in order to be benchmarked and connected with financing methods. Furthermore, the Triple-A Tools have reached 336 participants in the Triple-A events (workshops and Capacity Building Webinars), 72 participants in 59 bilateral meetings and 17 meetings with the Triple-A Advisory Board Members and 198 replies in the respective questionnaires.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

The current paper was primarily based on the research conducted within the framework of the EU Horizon 2020 project "Triple-A: Enhancing at an Early Stage the Investment Value Chain of Energy Efficiency Projects" (grant 846569, https://aaa-h2020.eu) aiming to assist financial institutions and project developers increase their deployment of capital in EE, making investments more transparent predictable and attractive. The contents of the paper are the sole responsibility of its authors and do not necessarily reflect the views of the EC.

References

[1] A. Papapostolou, F.D. Mexis, E. Sarmas, C. Karakosta, J. Psarras, "Web-based Application for Screening Energy Efficiency Investments: A MCDA Approach," 11th International Conference on Information, Intelligence, Systems and Applications, IISA 2020, 2020, doi:10.1109/IISA50023.2020.9284403.
[2] International Energy Agency (IEA), Net Zero by 2050, Paris, 2021.
[3] Council of Europe Development Bank, Energy Poverty in Europe: How energy efficiency and Renewables can help, 2019.
[4] OECD/IEA, Energy Efficiency 2018: Analysis and outlooks to 2040, 2018.
[5] M. Andrei, P. Thollander, I. Pierre, B. Gindroz, P. Rohdin, "Decarbonization of industry: Guidelines towards a harmonized energy efficiency policy program impact evaluation methodology," Energy Reports, **7**, 1385–1395, 2021, doi:10.1016/J.EGYR.2021.02.067.
[6] A. Papapostolou, C. Karakosta, K.-A. Kourti, H. Doukas, J. Psarras, "Supporting Europe's Energy Policy Towards a Decarbonised Energy System: A Comparative Assessment," Sustainability, **11**(15), 4010, 2019, doi:10.3390/su11154010.
[7] Efficiency Financial Institutions Group (EEFIG), Energy efficiency – the first fuel for the EU economy, 2014.
[8] A.B. Jaffe, R.N. Stavins, "The energy-efficiency gap," Energy Policy1, **22**(10), 804–810, 944.
[9] Horizon H2020 Triple-A project, Updated Web-Based Database on Energy Efficiency Financing and Supporting, , Deliverable 3.5, 2021.
[10] C. Karakosta, A. Papapostolou, Transformation Pathways towards a Clean,

Secure and Efficient European Energy System: A MCDA Approach, Nova Science Publishers, Inc., 2019.
[11] Financing Energy Efficiency | EEFIG Underwriting Toolkit, May 2021.
[12] H. Doukas, "On the appraisal of 'Triple-A' energy efficiency investments," Energy Sources, Part B: Economics, Planning, and Policy, **13**(7), 320–327, 2018, doi:10.1080/15567249.2018.1494763.
[13] T. Fawcett, G. Killip, "Re-thinking energy efficiency in European policy: Practitioners' use of 'multiple benefits' arguments," Journal of Cleaner Production, **210**, 1171–1179, 2019, doi:10.1016/J.JCLEPRO.2018.11.026.
[14] C. Karakosta, K. Papapostolou, G. Vasileiou, J. Psarras, Financial Schemes for Energy Efficiency Projects: Lessons Learnt from In-Country Demonstrations, Academic Press, 2020.
[15] T. Tsoutsos, S. Tournaki, E. Farmaki, P. Sonvilla, P. Lensing, J. Bartnicki, A. Cobos, M. Biscan, "Benchmarking Framework to Encourage Energy Efficiency Investments in South Europe. The Trust EPC South Approach," Procedia Environmental Sciences, **38**, 413–419, 2017, doi:10.1016/j.proenv.2017.03.125.
[16] N. Carbonara, R. Pellegrino, "Public-private partnerships for energy efficiency projects: A win-win model to choose the energy performance contracting structure," Journal of Cleaner Production, **170**, 1064–1075, 2018, doi:10.1016/j.jclepro.2017.09.151.
[17] LAUNCH H2020 | Sustainable Energy Assets in Europe, May 2021.
[18] RenOnBill | About, May 2021.
[19] The Project - EEnvest, May 2021.
[20] E2DRIVER, May 2021.
[21] X-tendo Toolbox, May 2021.
[22] C. Karakosta, A. Papapostolou, Linking Stakeholder Engagement to Multiple Future Policies in the European Energy Sector, 383–392, 2020, doi:10.4018/978-1-7998-1196-1.ch022.
[23] A. Ishizaka, P. Nemery, Multi-Criteria Decision Analysis: Methods and Software, Wiley, West Sussex, 2013.
[24] C. Karakosta, A. Papapostolou, J. Psarras, Concrete Steps towards the Promotion of Renewable Energy Deployment in an Effort to Tackle Climate Change, Nova Science Pub Inc, 2016.
[25] F.D. Mexis, K. Touloumis, A. Papapostolou, C. Karakosta, Deliverable 4.2, Final Triple-A Standardised Tools, Triple-A Project: Enhancing at an Early Stage the Investment Value Chain of Energy Efficiency Projects funded by European Union's Horizon 2020 Grant Agreement No 846569, Athens, 2021.
[26] Austrian Energy Agency, LED Street Lighting Procurement & Design Guidelines, developed for "Premium Light Pro" Project, funded by European Union's Horizon 2020, Grant Agreement No 695931; Road lighting - Part 5: Energy performance indicators, CSN EN 13201-5, 2015, 2017.
[27] T. Loureiro, M. Gil, R. Desmaris, A. Andaloro, C. Karakosta, S. Plesser, "De-Risking Energy Efficiency Investments through Innovation," Proceedings, **65**(1), 3, 2020, doi:10.3390/proceedings2020065003.
[28] Triple-A Final Report on Risks of Energy Efficiency Financing and Mitigation Strategies Typology, Athens, 2020.
[29] F. José, S. Greco, M. Ehrgott, Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Nature, 2006.
[30] B. Roy, J.R. Figueira, J. Almeida-Dias, "Discriminating thresholds as a tool to cope with imperfect knowledge in multiple criteria decision aiding: Theoretical results and practical issues," Omega, **43**, 9–20, 2014, doi:10.1016/j.omega.2013.05.003.
[31] W. Yu, Aide multicritère à la décision dans le cadre de la problématique du tri: concepts, méthodes et applications, Diss. Université Paris IX-Dauphine, 1992.
[32] V. Mousseau, R. Slowinski, P. Zielniewicz, ELECTRE TRI 2.0 a Methodological Guide and User's Manual, Laboratoire d'Analyse et de Modélisation de Systèmes pour l'Aide à la Décision (LAMSADE), Paris, 1999.
[33] E. Mills, "Risk transfer via energy-savings insurance," Energy Policy, **31**(3), 273–281, 2003, doi:10.1016/S0301-4215(02)00040-X.
[34] L. Wang, J. juan Peng, J. qiang Wang, "A multi-criteria decision-making framework for risk ranking of energy performance contracting project under picture fuzzy environment," Journal of Cleaner Production, **191**, 105–118, 2018, doi:10.1016/j.jclepro.2018.04.169.
[35] Industrial Energy Accelerator (2020): Energy Efficiency Risk Management - A Report for Financial Intermediaries in Brazil. Report prepared by adelphi (Germany) and SITAWI (Brazil) under the UNIDO lead IEA initiative. January 2020.
[36] T. Pettinger, Factors affecting investment, Economics Help, 2019.
[37] U.S. Environmental Protection Agency, "Understanding Cost-Effectiveness of Energy Efficiency Programs," 2008.
[38] EBF, Sustainable Finance.
[39] Bankers Almanac, Standard & Poor's Definitions, 2018.
[40] United Nations, SDG Indicators, Jul. 2021.
[41] Energy Efficiency Financial Institutions Groups (EFFIG), De-risking Energy Efficiency Platform, Apr. 2020.
[42] D. Kahlenborn, C. Annica, I. Georgiev, F. Eisinger, D. Hogg, Defining 'green' in the context of green finance, Brussels, Belgium, 2017.
[43] Internatioanl Capital Market Association (ICMA), Green Bond Principles, May 2020.
[44] Aggregation and bundling of loans for Energy Efficiency – the holy grail for the European Investment Bank: An opportunity for Housing Federations? | Build Up, Dec. 2020.
[45] Technical Expert Group on Sustainable Finance, "Teg Report Proposal for an Eu Green Bond Standard," (June), 1–79, 2019.
[46] P. Radgen, K. Bisang, I. Koenig, "Competitive tenders for energy efficiency – lessons learnt in Switzerland," in ECEEE INDUSTRIAL SUMMER STUDY PROCEEDINGS, 81–89, 2016.
[47] Swiss Office for energy, Energy efficiency bidding and auctions: Ten years of experience in Switzerland, 2019.
[48] H2020 Triple-A Project, Aug. 2021.
[49] CSR School - Corporate Sustainability & Responsibility, Aug. 2021.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 249-259 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Task-based Paradigm for Promoting an Alternative Thinking Style in Teaching Mathematics

Mikhail Rodionov[1], Zhanna Dedovets[*,2], Irina Akimova[2]

[1]*Department of Informatics and Methods of Teaching Informatics and Mathematics, Penza State University, Penza, 440026, Russia*

[2]*Faculty of Humanities and Education, The University of the West Indies, St. Augustine Campus, Trinidad and Tobago*

ARTICLE INFO

Article history:
Received: 12 July, 2021
Accepted: 01 October, 2021
Online: 11 October, 2021

Keywords:
Alternative style of thinking,
Teaching of Mathematics,
Heuristic Procedure,
Mathematical Task,
Secondary School

ABSTRACT

The article identifies an alternative style of thinking as one of the important components of human intellectual development. It is shown that it can be effectively implemented in mathematics lessons at school. The purpose of this study is to develop and substantiate a strategy for the formation of an alternative style of thinking among students in mathematics lessons. The innovation in this research lies in the development of a new approach to the formation of an alternative style of thinking of students, involving the progression of schoolchildren up the "ladder of levels", purposefully correlating task structures with their "alternative analogs". There are essential research findings. The levels of formation of the alternative style of thinking of schoolchildren are defined and their multiple characteristics are given. It is shown that as the main means of actualizing an alternative style of thinking, it is advisable to set tasks that provide alternative options for analyzing the elements of their content area. The stage-by-stage work of students as they move up the "ladder of levels" is presented. Methodological recommendations for teachers of mathematics have been created and partially tested. These were presented at several seminars / training sessions and were successfully applied in practice throughout the year. Statistical processing using Pearson's criterion χ2 was applied at the end of the year to the results of the performance of special tasks by 52 students of grades 3-4 of one of the schools in Penza. Some students applied the traditional method (27 students), and the remainder applied the methodology proposed by the authors (25 students). Analysis of the results revealed a higher level of mastery of the alternative style of thinking among students in the experimental group.

1. Introduction

Being ready to perceive novel and unusual feelings, contradicting ideas initially estimated as the only correct ones, is an indicator of human intellectual and personal development. According to Vladimir Zinchenko, a person develops when he "rises above the available types of activity, chooses some, or builds a new one" [1].

The indicator under consideration, being an alternative thinking style, presupposes openness to all possible aspects of experience, including an unusual one; freedom from personal bias, stereotypes, prejudices and other external constraints. It admits a solution to problems opposite to generally accepted ones, based on the willingness of a person to "contradict himself", being the basis for the formation of tolerance of any dissenting opinion and action in the future [2–12].

This readiness is one of the factors in overcoming stereotypes. Provided an individual considers a particular situation from different angles, takes into account various points of view on the essential relationships between concepts established in a situation, he will be able to at least partially decentralize his thinking, and overcome his initially unilateral perspectives.

Despite numerous studies somehow affecting the role and place of an alternative style of thinking in the intellectual development of a person [2, 3, 7, 8, 10, 13, 14], the patterns and possibilities of the formation of such a style in the educational space at school has not been considered as an object of special research. In particular, issues of the main conditions and possibilities for the formation of the considered phenomenon in school education, the stages and levels of such formation, the diagnostic means to achieve these levels, the nature of educational goals in various academic subjects targeting the actualization of

*Corresponding Author: Zhanna Dedovets +18687337811
dedzhanna333@gmail.com

https://dx.doi.org/10.25046/aj060527

alternative forms of educational and research activity, have not been considered by the scientists [14, 15].

As a result, purposeful formation of an alternative thinking style remains outside the attention of teachers and methodologists, leading to passive thinking in school students. In particular, they are not ready to transit from standard methods for performing educational and research activities to alternative methods. In other words, these standard methods for performing activities are fixed in their minds as rigid algorithms that do not allow any structural transformation.

Thus, the above conditions determine the relevance of the proposed research.

2. Problem Statement

An alternative style of thinking can be effectively realized within the framework of mathematical content, since the presence of deep and often hidden connections between different branches of mathematics, different "mathematical languages" and general methods provide a possibility to correlate different perspectives on various mathematical concepts, facts and patterns. These possibilities are described in the works of many well-known mathematicians and methodologists [1, 4, 8, 15–20, etc.].

In particular, the actualization of the alternative style of thinking was very vividly illustrated in [15]. The great scientist correlated numbers and algebraic quantities: "Calculations are made either using numbers, as in ordinary arithmetic, or using species as in algebra. Both methods are based on the same principles and lead to one goal, but arithmetic using a specific and particular way, and algebra applying an indefinite and general way" [15]. However, "in arithmetic, when solving problems, one only passes from the given quantities to the sought ones; in algebra one works backwards, passing from the sought quantities, being considered as data, to the given values, being considered as the sought ones, in order to come to the conclusion or the equation, from which it would be possible to find the sought quantity" [15].

Newton implemented these considerations in finding a binominal series expansion of the function $(1+x)^{\propto}(\forall\propto),$, and in calculating the area bounded by the curves$(1-x^2)^{\frac{n}{2}}$ for various natural numbers n [21]. Newton was well aware that, according to the logic of his construction, the obtained expansion was hypothetical, and not a proven fact. Therefore, he proposes other ways to confirm the correctness of the obtained expansion, which are of interest in the context under consideration.

In particular, Newton applied Viète's method for square root extraction for numbers to polynomials [21]. Meanwhile, he formulated the rule for "square root extraction" from a polynomial, and he explained this rule with examples, in particular, an example of obtaining a binomial series for [21]

$$(1+x)^{\frac{1}{2}} = \sqrt{1+x}$$

$$\sqrt{1+x} = 1.$$

Despite the fact that there are tasks for actualizing an alternative style of thinking in modern textbooks in mathematics, preparation materials for the final certification of school students, and materials for various Olympiads, this orientation has not been fully implemented in mass practice of math education.

According to our observations and numerous publications, many school students realize the possibility of an alternative consideration of a particular problem situation only as reasoning that must be remembered and applied when solving problems of a specific type, rather than as a universal approach to its analysis [9–11, 22–24]. To prevent this, an adequate pedagogical format, considering the purposeful actualization of an alternative style of thinking as one of the main tasks of school mathematics education, is urgent.

Considering the above, the purpose of this study is to define theoretical foundations for the formation of an alternative thinking style for school students in the educational process, and to develop an appropriate methodological structure to ensure such thinking goes beyond the range of stereotyped patterns and ideas.

3. Literature Review

Various aspects of an alternative thinking style, being an adequate indicator of personal intellectual development, have been considered by many scientists who associated it with the categories of decentralized, lateral, tolerant, or divergent thinking. The absence of this style is evident in the unilateral and tendentious thinking of a person using mainly stereotyped patterns and algorithms [2, 3, 8, 10, 17]. Therein, his ability for creative activity is poorly developed, which, in turn, negatively affects his ability to solve educational problems.

In this context, a special role is played by the mathematical content, which is characterized by a high degree of generalization, universality, and alternativeness of research and proof methods used. For example, one can mention the duality principle, which plays an important role in projective geometry. According to this, the projective space can be isomorphically mapped onto itself with the "transformation" of points in the plane, and vice versa. Obviously, the above regularity significantly helps to formulate statements of the mentioned theory, since any proof can automatically be applied to two different meaningful alternatives.

Accordingly, the targeted process of mathematics education creates the greatest opportunity for teaching generalized methods of mental activity to students, providing an opportunity for a multilateral analysis of the problem situations under consideration. It is convenient to organize such an analysis by changing the semantic status of the components of these situations. In this case, the initially key element of the task can be considered as an auxiliary element, and the initially auxiliary element as being a key one.

This approach appears in many psychological and pedagogical studies in various modifications. For example, the founder of Gestalt therapy, introduced thought experiments that contribute to the personal growth of a person ("feeling of opposing forces", "attention and concentration", etc.) [25]. The same procedures can be used in teaching mathematics. In particular, when solving geometrical problems, school students often consistently analyze the corresponding geometric configurations, highlighting the most promising one for further exploration, realizing the feeling of "background and figure" [8, 11, 22]. Having reconstructed the

problem situation, students get the opportunity to actualize the "feeling of opposing forces" [1,11, 22, 26, 27].

The role of the considered characteristic of thinking in the development of scientific ideas was emphasized by many famous scientists [1, 18]. In particular, it is noted that the concepts and facts under consideration should be not investigated in the "normal", correct, and usual form, but "in a critical situation, in fever and passion..." [28]. The authors also suggested paying special attention to characteristics that do not correspond to "the usual intuitive ideas, and allow us to look at well-known facts from a new angle, to see new and unexpected content therein" when conducting research [18]. Similar considerations in favor of the implementation of a "pathological" approach to the research subject are expressed by the famous mathematician in [1].

In a didactic context, updating the considered approach can be viewed as a kind of heuristic method used at the stage of problem solving [2, 3, 11, 22, 29]. In particular, this technique is the base of many non-standard math tasks. Mastering it, along with solving purely didactic problems, allows for the forming of a multilateral and "multidimensional" approach to the subject of study, "decentralizing" the viewpoint on the considered mathematical content and, thus, contributing to the formation of a sufficiently flexible and "variable" style of thinking of students.

From the foregoing, it follows that it is expedient to develop a special methodological structure that ensures the development of an alternative thinking style when teaching mathematics at school. The appropriate orientation should be a constant component of the educational process, one of the methods of cognition that allows for effectively organizing an educational heuristic search in the development of new subject content.

Based on the analysis of a number of studies in the context of the problem under consideration [2–4, 8, 11, 22, 25, 28, 30–33], and our own pedagogical experience, the following methodological provisions can be taken as the basis for such work:

- It is advisable to actualize and form an alternative thinking style for school students on the basis of special tasks implying an alternative consideration of a certain mathematical model and its components. These tasks should be interrelated with the general context of mathematical material, ensuring the effective achievement of the educational goals planned by the teacher.
- In accordance with active learning strategies, the implementation of the previous condition in the learning process presupposes an organization of adequate independent activity of students, during which certain connections and relations are established between the considered alternative positions, and a qualitatively new integrative knowledge arises. The highest developmental effect is present when the idea of applying some heuristic procedure comes directly from students in the form of conjecture and suppositions. George Pólya writes, "A conjecturing student associates himself therewith; his prestige and self-esteem depend on the outcome of the case to some extent, and he is eager to find out whether his guess turns out to be correct or not; he will be actively interested in his task and the work of class, he will not fall asleep, and will not be distracted by other tasks" [8].
- Actualization of various semantic positions, being complementary and mutually exclusive, presupposes interaction and cooperation of students' efforts. According to Rada Granovskaya and Yuliya Krizhanovskaya, "during joint activities, the initial attitudes of some participants on the inability to solve the problem are removed. Interstimulation of partners ... reduces self-criticism and prevents the displacement of original ideas into the subconscious" [3]. In particular, the so-called "effect of a teasing interlocutor" [33], which defends the position opposite to that of the author is very effective in the considered perspective. This approach makes it possible to impart a subjective-personal character to the educational exploratory process, which ensures the effective mastering of the corresponding heuristic techniques.
- The role of a teasing interlocutor can be effectively performed by a modern virtual educational environment, which allows for quickly moving and changing elements of a particular mathematical model when expedient. During the process it becomes possible to actualize alternative key elements for a relatively short period of time, and determine further prospects for their use when considering this specific task situation, or, to select the optimal alternatives [33].

The use of virtual educational and developmental environments can be implemented locally when organizing a heuristic search for a solution to a given specific problem, or it can become a means for full-fledged creative reflection on the meaningful links between its elements, due to their optimal dynamization and visualization. In the course of such work, it is given an initially boring (task) situation, after "having started to formulate and think it over" students see that "it acquires components, details, and new storylines, sometimes completely unexpected from the viewpoint of its initial perception" [33]. This approach is mentioned in a number of our publications [6, 30, 31, 34, 35].

4. Materials and Methods

The development of this research concept was carried out in line with the main trends in the system of mathematical education, in particular, problem-based and student-centered approaches, taking into account the specifics of the mathematical content [3, 4, 8, 9, 11, 22, 28, 34, 36–38]. The main instrument to implement these approaches is a system of special math tasks, leading to actualization of an alternative thinking style in school students regulated by a teacher. The study involved three clusters of materials.

Firstly, we have analyzed normative and program documents, educational standards of general secondary education, programs of mathematical disciplines, scientific and methodical works of leading domestic and foreign experts in the field of developmental psychology, didactics, theory and methods of mathematical education, materials on the history of mathematics, current textbooks, teaching aids, methodical recommendations, and educational software [1–3, 18, 21, 28, 33, 39]. The study of these materials indicates that the problem of the formation of alternative thinking in school students currently has not been singled out as a research subject for specialists in the field of mathematical education.

Secondly, we have conducted a longitudinal study of teaching mathematics in a number of educational institutions in Penza (Russia). The background competence of school students was revealed in terms of their readiness to manifest an alternative thinking style when solving exploratory math tasks. It was shown that about 80% of the school students surveyed in a number of Penza schools were not ready for consideration of variable alternatives to a given mathematical models, limiting themselves to analyzing the only possible alternative. Insufficient attention of teachers to the described component of educational and exploratory mathematical activity seems to be the main reason for such a problem.

Finally, the third cluster of materials comprised the results of an exploratory pedagogical experiment to determine the possibilities of introducing the developed mathematical content and make an assessment of its effectiveness in terms of creating the readiness of school students to apply an alternative thinking style when solving exploratory math tasks. The exploratory experiment was carried out in the course of a two-year mathematical training of school students at Lyceum No. 3 for gifted students of the Penza region, and a number of educational institutions in Penza. In total, the experiment included about 80 school students at the basic and senior levels of education.

The content of the materials being tested is partially presented in a number of our textbooks and articles [7, 11, 22, 30, 31].

The data collection technique was implemented using special educational and exploratory tasks. The results of their solving by school students were subjected to qualitative and quantitative analysis on the basis of a special quantitative scale at the background and productive stages. The following indicators were considered as the main characteristics of the alternative style of thinking: readiness to actualize unusual associative links; the ability to perceive the problem in an unusual way; readiness to build initially unobvious alternative constructions (an openness to an extraordinary combination of elements (“game-research”); the ability to alter actions; the ability to re-construct the acquired information in accordance with the structure of the problem domain [25]).

The data collection technique was implemented using special diagnostic approaches based on the use of two-level tasks. Each diagnostic task included two parts. In the first part, it was proposed to solve a task that involves finding a solution based on the actualization of an alternative task structure with a similar solution method. The second part included questions for which the students did not have clear guidelines. The purpose of posing these questions was to give students a certain "hint" about the possibility of developing the initial task. Attempts to find such answers testifies to the presence of a certain orientation of the students towards the search for alternatives. At the same time, students were required to set goals independently, (in particular, when they formulated tasks); focus on finding the most general way of action; effectively "translate" the information contained in the original problem into an alternative mathematical language; engage in full-fledged reflection on the revealed meaningful interrelationships.

The data collection technique was implemented using a special diagnostic intervention based on the use of two-level educational search tasks. Each diagnostic task included two parts. In the first part, it was proposed to solve a task that involves finding a solution based on the actualization of an alternative task structure with a similar solution method. The second part included questions for which the students did not have clear guidelines. The very purpose of posing these questions was to give schoolchildren a certain "hint" about the possibility of developing the initial task situation.

Already the presence of attempts to find such answers testifies to the presence of a certain orientation of the senior pupils towards the search for alternatives. At the same time, the ability to engage in relatively independent goal-setting was required (in particular, when drawing up tasks); focus on finding the most general approach; effective "translation" of the information contained in the original problem into an alternative mathematical language; full-fledged reflection on the revealed meaningful interrelationships (for example, when completing the definition of the task situation). The following are examples of such tasks for consideration.

Task 1:

- Let D be the point of the AC side in the ABC triangle; r, r1 and r2 the radiuses of the inscribed circles of the triangles ABC, ABD and BDC, accordingly. Prove that $r < r1 + r2$.
- Is it possible to extend the proved inequality to a larger number of inscribed circles? (You can, for example, divide the AC side into more parts, or take point D inside the ABC triangle).

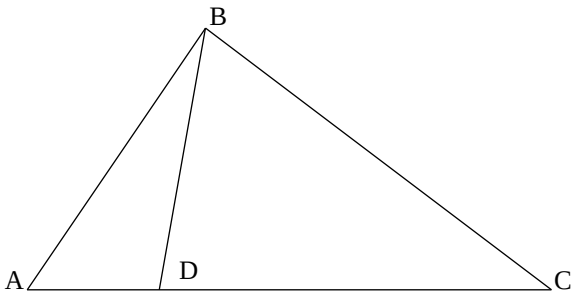


Figure 1: Illustration of the task

Task 2:

- Prove that the bisector of the exterior angle at the vertex of an isosceles triangle is parallel to the base.
- How many ways of proof can you suggest? (The number of proposed and successfully implemented methods is considered).

To establish an ordinal measurement scale, a preliminary assessment of possible answers was carried out and a weighted assessment in points was made. Completion of only the first part of the assignment was assessed by one point, a purposeful attempt to answer the question of the second assignment added one more point, while the performance of the second part of the assignment was assessed by two points. If, when completing the second part of the tasks, the student offered at least three answer options, then one more point was added to the total number of points. Thus, the student could receive a maximum of four or five points for one task. The total grade that could be obtained for completing the entire work, including four tasks, did not exceed 18 points. Three categories of students were identified. A student who received from 0 to 6 points corresponds to the initial level of the formation

of an alternative style of thinking; students who received from 6 to 12 points were assigned to the intermediate level, and from 13 to 18 - to the high level.

Comparison of the results of the tasks performed by the students of the experimental and control groups showed that the purposeful use of the methodology for devising meaningful alternatives in the study of mathematical material significantly influenced the increase in the readiness of students to successfully solve mathematical search problems.

5. Research Results

5.1. Levels of formation of an alternative thinking style

With regard to mathematics, the characteristics defined above can be actualized in solving math tasks, the solution of which involves implementation of a certain set of intellectual procedures. These procedures are reflected in the exploratory activity of school students, corresponding to the levels of its development [16]. However, their status varies at different levels (Table 1).

Table 1: Levels of Formation

Levels of formation of an alternative thinking style	Object of the analysis	Dominant way for actualizing an alternative style of thinking	Selection criteria for an alternative search	Nature of the intellectual procedures involved
Initial	A particular task	External (based on strict external requirement)	Desire to "function" strictly within the prescribed range of choice	Random chaotic or systematic search for alternatives
Medium	General task situation	Mixed (result-oriented)	The choice is determined by different model representations of the task situation	Selection of a technique by analogy with the solution of "related" tasks, accompanied by its experimental verification
High	System of related tasks	Internal (search-oriented)	Independent and reasonable variation of the conditions and requirements of the task	Selection of a heuristic procedure generated by purposeful modification of the components of the task domain

This scale of levels allows us to outline a strategy for the consistent formation of an alternative style of thinking, based on the use of systems of special tasks which are presented below.

5.2. Substantive and Methodological Tools

5.2.1. Enumeration of Alternatives

We now consider the main methodological features of the formation of an alternative thinking style of school students during their moving up the "ladder of levels".

At the initial level, alternativeness is directly related to the selection of possible task options, and its transformation. It is important that the expediency of such a choice should acquire a subjective personal meaning for a school student. According to Zinchenko, it means that the subject has a special type of "awareness of ignorance of some knowledge", or "an emotional experience of ignorance of some knowledge", as stated by Konstantin Pigrov. Johann Amos Comenius qualifies such awareness as the source of the thirst for knowledge.

Let us consider an example of actualizing an alternative thinking style of students by enumerating alternatives presented in a particular task situation.

There are three criteria for congruent triangles with different combinations of angles and sides of a triangle studied in school course of planimetry. If we enumerate possible combinations, we can formulate some more statements (with two angles and a side opposite to one of the angles; with two sides and an angle opposite to one of the sides; and with three angles).

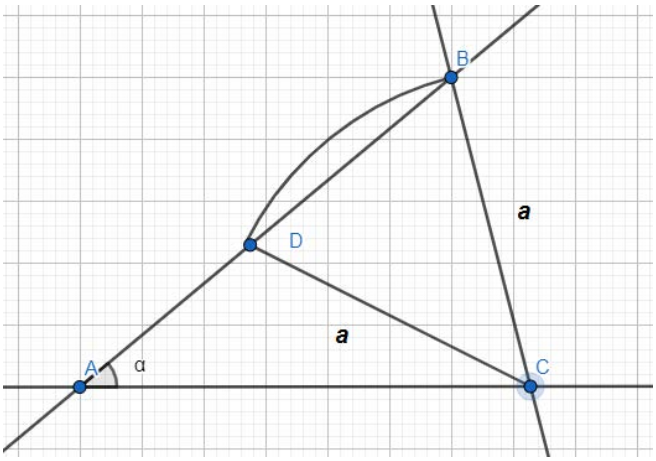


Figure 2: Illustration of the task

The first criterion formulated by students can serve as a starting point for studying the sum of angles of a triangle.

In the second case, we consider the fourth criterion for congruent triangles with two possible solutions for constructing a triangle from the corresponding elements based thereon (Figure 2). The unambiguity of the solution can be provided by the additional condition of uniformity of the angles at the vertex B of one triangle, and the corresponding vertex D of another triangle.

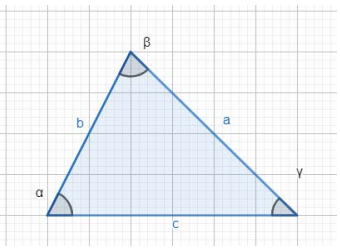
Finally, the obvious inequity of a criterion for congruent triangles by three angles naturally raises the question of congruence of triangles.

The described alternatives matrix can be clearly presented for school students on the basis of a directive that includes all the selected cases (Table 2).

This example shows the possibility of actualizing an alternative thinking style of a student by enumerating various specific models of a common task, aimed at identifying all combinations with a minimum number of triangle elements, unambiguously determining its construction when considering standard mathematical material.

To conduct further exploratory activity, we should move away from traditional stereotypes and apply well-known heuristic procedures [2, 3, 8, 9, 11, 22, 26].

Table 2: An Example of a Directive for School Students

Criteria for congruent triangles	Constructing triangles given three elements	Solving triangles
1. Two triangles are congruent if two corresponding angles and the side in between are equal (a – side; α and γ – angles).	 (Construction is possible and unambiguous at $\gamma + \beta < 180^0$)	Angle α is found according to triangle angle sum theorem; side c is calculated by the law of sines; side b is found according to the law of cosines.

For example, we consider the replacement of structural components of the task domain. With regard to school mathematical content, the mentioned procedure can be actualized with the help of a specific teaching technique, namely, "he (she, it) does not know what he (she, it) really is."

We illustrate this statement with the following artificially constructed example. Consider the usual quadratic equation $x^2 - 13x + 36 = 0$. This equation obviously involves numbers and letters. Now we make an assumption that "numbers do not know that they are numbers, and letters do not know that they are letters." In other words, "the number 6 does not know that it is a known number, the letter *x* does not know that it is an unknown number." Let us find the number 6:

$$6^2 - \left(\frac{13x}{6}\right) \times 6 + x^2 = 0$$

$$D = \frac{169x^2}{36} - 4x = \frac{25x^2}{36}$$

$$6 = \frac{\frac{13x}{6} - \frac{5x}{6}}{2} = \frac{2x}{3} => 2x = 18 \quad x = 9$$

$$6 = \frac{\frac{13x}{6} + \frac{5x}{6}}{2} = \frac{3x}{2} => 3x = 12 \quad x = 4$$

This example illustrates the potential for an alternative search for thought, being opposite to the standard algorithm. According to Zinchenko, it is "at the junction of opposite processes of value comprehension and signification of meanings that new images are born, conveying certain meanings and making the meaning visible" [1].

The specificity of this approach is constant correlation of the original design for task solution with its alternatives forming the basis of this solution in a new semantic paradigm. The logic of an exploratory process using the adopted approach is shown in Figure 3.

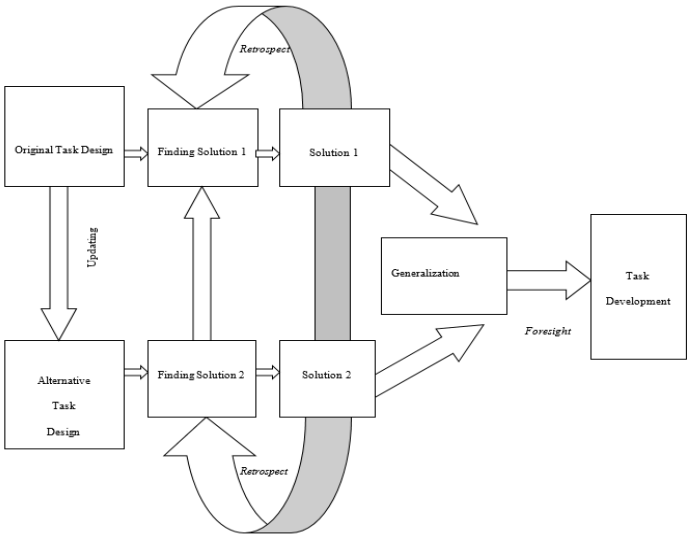


Figure 3: Schematic for an exploratory process

This schematic demonstrates updating of possible alternative problem-solving in the process of sequential reformulation of the task situation, ensuring effectiveness of the exploratory process being implemented.

Here are some examples of such updating when solving non-standard tasks:

"The number does not know that it is a letter"

Calculate: $\frac{254\times399-145}{254+399\times253}$

Solution: Let a=254, b=399. Then $\frac{ab-(b-a)}{a+b(a-1)} = 1$

(Soros Olympiad, 1994/95, Round 1, Grade 10)

Calculate: $(\frac{1-2(17.3)^3}{1+(17.3)^3})^3 + (\frac{17.3\cdot(2-(17.3)^3}{1+(17.3)^3})^3 + (17.3)^3$

Solution: Let $(17.3)^3$=a. Then $\frac{(1-2a)^3+a(2-a)^3}{(1+a)^3} + a = \cdots = 1$

(Republican Mathematics Olympiad, Kiev, Ukraine)

Which number is greater?

$\frac{10^{1967}+1}{10^{1968}+1}$ or $\frac{10^{1968}+1}{10^{1969}+1}$;

Solution: $\frac{a+1}{10a+1}$ ∨ $\frac{10a+1}{100a+1}$;

$$100a^2 + 101a + 1 \ \vee \ 100a^2 + 20a + 1$$
$$101 > 20$$

Prove that $1993\times1995^3 - 1994\times1992^3$ is an integer cube.

Solution: $1994 = a$;

$(a\text{-}1)\times(a+1)^3 - a(a-2)^3 = \cdots = (2a-1)^3$

"Algebraic or numerical expressions do not know that they are letters"

(Intellectual Marathon, 1992)

Find the sum: $\frac{1}{1+\frac{1}{1+\frac{1}{3+\frac{1}{4+\frac{1}{5+\cdots+\frac{1}{n}}}}}} + \frac{1}{2+\frac{1}{3+\frac{1}{4+\frac{1}{5+\cdots+\frac{1}{n}}}}}$.

Solution: $\frac{1}{1+\frac{1}{A-1}} + \frac{1}{A} = \frac{A-1}{A} + \frac{1}{A} = 1$.

Solve the equation $\frac{x^2+x+2}{3x^2+5x-14} = \frac{x^2+x+6}{3x^2+5x-10}$.

Solution: $\frac{a}{b} = \frac{a+4}{b+4} => \frac{a}{b} = 1, 3x^2 + 5x - 14 = x^2 + x + 2 \ldots$

Answer: x=2, x= - 4.

It is clear that the described principle of transition to meaningful alternatives can also be used when considering geometric configurations: "The angle does not know that it is exterior"; "The circle does not know that it is escribed," etc.

Let us consider an example of the exploratory process to analyze one of configurations.

Example. Prove that the sum of angles DBE and CAM does not depend on the choice of the line l (Figure 4).

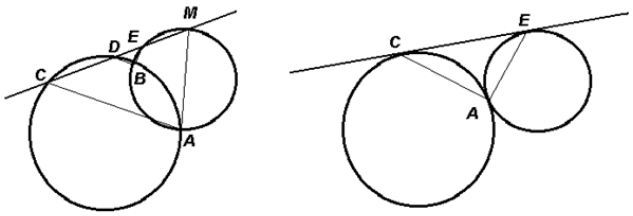


Figure 4: Illustration of the task

It is proposed to consider the line *CM* as a dynamic image that can move relative to circles, and turn into a tangent to both of them ("*a secant does not know that it is a tangent*"). Similarly, it's possible to move from an intersection to an external contact of circles ("*intersecting circles do not know that they are just tangent to each other*").

We obtain a chain of alternative configurations, which are degenerate cases of the initial one (Figure 3). The latter naturally leads to a derivative task situation.

Example. Find the angle CAE, where point A is the point of external contact of two given circles, and CE is their common tangent.

Solving the task, it is possible to assume that the angle CAE, considered as the sum of two angles, obtained at the intersection of the line *CE* and common tangent to two circles, is equal to 90°. As for the original geometric configuration, we consider the secant lines *CM* and *AB* as analogs of common tangents in a degenerate alternative construction. As in the degenerate case, the solution consists of expressing the angles forming the angle *CAM*, via the angles of the triangle *BDE*.

$\angle CAB = 1/2 \cup CD\mathrm{B} = ½(360° - \cup \mathrm{CAB}) = ½(360° - 2\angle CD\mathrm{B}) = 180° - \angle \mathrm{CDB} = \angle BD\mathrm{E}$.

Likewise, $\angle \mathrm{BAM} = \angle \mathrm{BE}D$. Then, $\angle \mathrm{CAM} = \angle \mathrm{CAB} + \angle \mathrm{BAM} = \angle BD\mathrm{E} + \angle \mathrm{BE}D = 180° - \angle D\mathrm{BE}$.

Finally, we get: $\angle \mathrm{CAM} + \angle D\mathrm{BE} = 180°$.

Thus, we have shown that by highlighting the key link (determining the successful search) based on a meaningful analysis of the alternative configuration, one can successfully update the relevant considerations when examining the original one.

5.2.3. Structuring the Task

Since the use of an alternative style of reasoning is regarded by the student mainly as a local intellectual act in understanding the task situation at previous levels of mathematical development, then it is associated with the possibility of free and informed choice from a combination of implicitly presented meaningful alternatives at a higher level. It is quite obvious that in this case the heuristic potential for choosing alternatives increases significantly. Thenceforward, the larger set of heuristic procedures the student actualizes, and the more systemic they are, the more efficient a person's intellectual development is.

Successful procedure for an informed alternative choice becomes one of the main instruments for maintaining an active exploratory activity in schoolchildren. One of the most effective methods for creating an opportunity for such a choice is a step-by-step transformation of components of a task situation as applied to mathematics education, leading to the construction of a sequence of new task structures, somehow related to the original one, which is accompanied by their comparative analysis. The advantage of such work stems from the fact that the initial intellectual activity of a student does not stop when the solution of the original problem is completed. Moreover, it receives an additional heuristic impulse during its transfer to new substantive content, ensuring full realization of his individual trajectory of development.

Transformation of task structures can be potentially implemented in many alternative formulations, determined by the choice of the involved techniques, or their expedient combinations. As a result of specially organized joint work of subjects of the educational process, one can obtain a whole cycle of alternative derivative tasks based on some initial mindset. The main methodological problem in this case is to ensure effective anticipation of the most promising and available areas of exploratory activity for implementation at the particular stage of training.

We consider one of such complexes, obtained by successive generalization of the following original situation.

Original Task. Construct a circle that is tangent to the lines *m* and *n* (*m*//*n*), and to the circle (*O*, *r*) lying between these lines.

Having changed the location of this circle ("*the circle lying between two given lines does not know that in fact it does not lie between them*"), we get the task:

Task 2. Construct a circle that is tangent to two different parallel lines, and a given circle.

If in Task 2 we consider that the lines coincide ("*the lines do not know that they coincide*"), then we get the following task:

Task 3. Construct a circle that is tangent to a given line at a given point *M* thereon, and a given circle that is tangent to this line.

Having ignored the condition of tangency of a given circle and line ("*the circle that is tangent to the line does not know that it actually is not tangent to it*"), we get the following task:

Task 4. Construct a circle that is tangent to a given line at a given point *M* thereon, and a given circle.

Having ignored the condition of tangency at a given point ("*the circle does not know at what point it is tangent to the line*"), we get the following task:

Task 5. Construct a circle that is tangent to a given line and a given circle.

Having replaced a line with a circle in Task 5 ("the circle does not know that it is the line"), we get the following task:

Task 6. Construct a circle that is tangent to two given circles, and to one of them at a given point.

Having replaced a circle with lines in Task 6 ("the circle does not know that it is two lines"), we get the following task:

Task 7. Construct a circle that is tangent to the lines *m* and *n* (*m*//*n*), and a line *c* intersecting the lines *m* and *n* at points *K* and *L*, $K \in m$, $L \in$n.

Using alternative task structures in the educational and exploratory work, ideologically related to the original task situation, a teacher gets the opportunity to stimulate an exploratory activity of school students, while enriching and improving their intellectual and personal potential.

It is important to note that similar procedures can be visualized quite effectively using a modern software environment. By selecting, swapping, and moving the initially actualized and background elements of considered mathematical structures with a manipulator, one can obtain alternative constructions that allow "seeing" a potentially successful area of an exploratory activity [6, 31, 32, 35].

6. Approbation

A unified, generally recognized methodological concept of the formation of an alternative style of thinking by students in mathematics lessons presumably is not provided in real school practice. The test of this hypothesis was carried out within the framework of the first stage of the experiment, the main task of which was: to study the nature of the spontaneous influence of the process of "natural" development of the content of school mathematics courses and the personality traits of the teacher on the effective actualization and formation of an alternative style of thinking in schoolchildren.

To confirm this fact, the results of the tests performed by 80 students of grades 3-4 of a number of schools in Penza (Russia) were analyzed. Control work included tasks that imply elements of the study of task structures and the feasibility of enumerating alternative options for the search process. Evaluation of their implementation in points took into account the characteristics of the levels of development of the alternative style of thinking presented above. The data obtained indicate that more than half of the surveyed schoolchildren can be assigned to the initial level of such development (52%), 39% to the average level, and only 9% to the high level. These data indicate that teaching mathematics in itself in accordance with the current curriculum does not yet provide a high level of development of the alternative style of thinking of schoolchildren.

Approbation of the proposed methodological solutions was carried out in two segments, differing from each other in the number of students involved and the nature of the experimental work.

1. A "free" experiment was carried out in several schools, gymnasiums and lyceums in Penza (№№ 3, 6, 11, 12, 42, 47), covering about 200 students. It was preceded by a lot of preparatory work. In the course of its implementation, we read a series of lectures for teachers in the field of advanced training, as well as mathematicians of the final year of Penza State University (as part of elective courses), in which the listeners were familiarized with the main theoretical provisions of the developed concept and a number of teaching materials.

Additional group and individual consultations were held with teachers who expressed their willingness to take part in the experiment. A special role in the implementation of this strategy was given to the initial and current diagnostics of relatively stable indicators of the development of an alternative style of thinking in order to clarify the individual educational trajectory of a particular student.

The experimental work was carried out for two years in accordance with the above recommendations, taking into account the achieved level of development of objective motivation. Throughout this period, regular observations of the course of the educational process were carried out, open lessons were organized, laboratory experiments and conversations with students were conducted.

Without purporting to be an exhaustive categorical conclusion, we consider it possible to note certain positive shifts in the development of the alternative style of schoolchildren's thinking, which were found in a significant number of the schoolchildren who participated in the experiment. These schoolchildren have significantly increased their interest in solving tasks of a creative and research nature. Moreover, when performing such tasks, they in most cases independently transformed their conditions on the basis of implementing alternative search strategies possible in a given situation. In the course of studying the theory, the students involved in the experiment tried to "get ahead" of the teacher's plan, "to see the goal from afar." At the same time, a certain independence of judgment became more characteristic of them, expressed in a more critical and reflexive attitude to the decision made, the desire to supplement and improve it in relation to both the analyzed situation and a wider class of mathematical concepts.

2. The "local" experiment was carried out in one of the lyceums of the city of Penza, where one of the authors of the article worked for 10 years as a methodologist and teacher of mathematics. Purposeful experimental work in line with the studied issues was carried out at the Lyceum for three years from 2018 to 2020. At the beginning of this work, the teachers of mathematics of the experimental classes were given explanations of its goals and content, discussed the possibilities of implementation, and proposed educational materials that included systems of tasks of a special type corresponding to different levels of formation of an alternative style of thinking. Ongoing consultations were offered to teachers in the course of attending lessons, analyzing the results of control and cross-sectional works.

To partially confirm the effectiveness of the proposed tools, the results of 52 students in grades 3-4 were analyzed separately, some of whom worked according to the traditional method (27 students), and the other part according to the methodology proposed by the authors (25 students). At the first stage of the study, a number of introductory tests were carried out, consisting of tasks that imply elements of the study of task structures and the feasibility of enumerating alternative options for the search process. The distribution according to the levels of formation of alternative thinking of schoolchildren was carried out in accordance with the criteria reflected in the section "Materials and Methods".

Table 3: The results before the experiment

Levels of formation of ATS	Before the experiment			
	Experimental group		Control group	
	Quantity	%	Quantity	%
Initial	12	48	12	44,4
Medium	8	32	10	37
High	5	20	5	18,6

Table 3 and Figure 5 show that the level of formation of the alternative style of thinking in both groups is rather low. About half of all subjects are at the elementary level, and only about 20% of students have reached a high level.

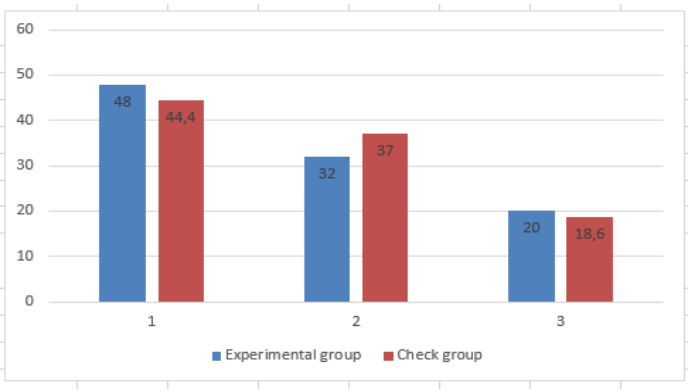


Figure 5: The results before the experiment

The data obtained were statistically processed using Pearson's chi square test (goodness of fit). The processing results are shown in Table 4.

Table 4: The results before the experiment

Levels off ATS	Q_{emp}		Q_1-Q_2	$(Q_1-Q_2)^2$	Q_1+Q_2	$\frac{(Q_1-Q_2)^2}{Q_1+Q_2}$
	Experimental group	Control group				
Initial	12	12	0	0	24	0,00
Medium	8	10	-2	4	18	0,22
High	5	5	0	0	10	0,00
Σ	25	27				0,22

The data in the table indicate the absence of significant differences in the level of development of the alternative style of thinking among students in both groups (empirical value of Pearson's criterion $X^2 = 0{,}22$). In other words, the nature of the distribution of schoolchildren according to the levels of development of this style was initially approximately the same.

Experimental work within the framework of this study was carried out for two years. In the course of teaching, the pupils of the experimental classes were offered the tasks described above, corresponding to their level of formation of an alternative style of thinking. For the initial level, emphasis was placed on an alternative enumeration of mathematical structures, to which tasks were added to replace the structural components of these structures. At the second and third levels, schoolchildren purposefully worked on a relatively independent compilation and analysis of task systems in accordance with the concept presented in the article.

As a result of the experimental work control measures were again carried out.

Comparison of the dynamics of changes in motivational characteristics in the selected groups of students (Table 5) showed that over the same time interval in the experimental classes, about 40% of students moved to a higher level of development of an alternative style of thinking, while the number and distribution of pupils at each of the levels in the second group changed insignificantly. A few students in this group even showed a slight decrease in interest in mathematics, explained by their future professional choice (not related to mathematical activity).

Table 5: The results after the experiment

Levels of formation of ATS	After the experiment			
	Experimental group		Control group	
	Quantity	%	Quantity	%
Initial	3	12	10	37
Medium	9	36	13	48,2
High	13	52	4	14,8

The data obtained were again subjected to statistical processing using Pearson's chi square test (goodness of fit). The processed results are shown on Figure 6 and in Table 5.

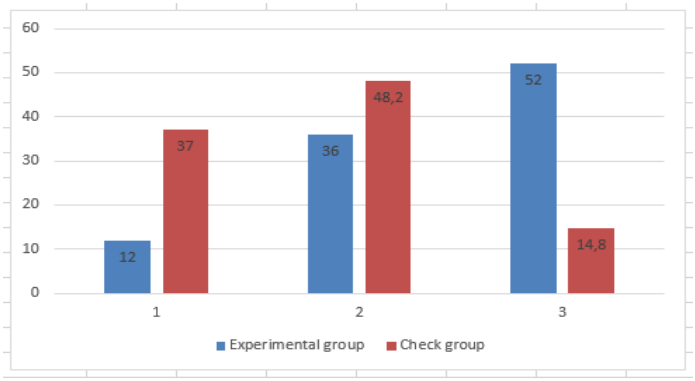


Figure 6 The results after the experiment

The results of statistical processing of the data obtained indicate the appearance of significant differences in the analyzed indicator between the schoolchildren of the control and experimental groups (Pearson's criterion $X^2 = 9{,}26$)

Table 6: The results after the experiment

Levels off ATS	Qemp		Q_1-Q_2	$(Q_1-Q_2)^2$	Q_1+Q_2	$\frac{(Q_1-Q_2)^2}{Q_1+Q_2}$
	Experimental group	Control group				
Initial	3	10	-7	49	13	3,77
Medium	9	13	-4	16	22	0,73
High	13	4	9	81	17	4,76
Σ	25	27				9,26

In other words, we found a significant difference between the distribution of characteristics of the formation of the alternative style of thinking of schoolchildren in the studied groups of schoolchildren. This difference indicates that the developed methodological solutions are more effective than the traditional ones in terms of the formation of the specified style.

In general, it can be stated that, as evidenced by the results of approbation, the described level-by-level strategy for the formation of an alternative style of thinking of schoolchildren in the process of teaching mathematics proved to be quite effective both in terms of the gradual formation of students' readiness to overcome their initially one-sided attitudes, and in terms of the development of their cognitive activity generally.

7. Conclusions and Suggestions

In the course of the study, the author's concept of the level-by-level formation of an alternative style of thinking of schoolchildren in mathematics lessons was developed. As the main means of such formation, it is advisable to consider special tasks that provide for the presence of alternative options for analyzing the elements of its subject area. Within the framework of this concept, the following scientific and methodological results were obtained:

- The important role of the alternative style of thinking is revealed, which underlies the realization of human intellectual activity and which ensures effective interaction with the environment and the ability to tolerantly respond to its diversity.
- It is shown that in the framework of mass school education, insufficient attention is paid to the formation of such a style. As a result, as many publications testify, many schoolchildren master the methods and techniques of educational and search activity as rigid algorithmic structures that hardly allow variation of semantic positions with respect to the concepts under consideration and the relations between them.
- A mathematics course has great potential in terms of the formation of an alternative style of thinking, as it is characterized by the presence of a variety of problem material, the mastering of which is possible from various positions and using various methods. For the full realization of this potential, an adequate pedagogical format is required, within which an alternative style of thinking can be actualized. As such a format, it is advisable to consider the joint activity of a teacher and students in solving special tasks that provide for the availability of alternative options for analyzing the elements of their subject area.
- When planning work on the development of an alternative style of thinking for schoolchildren in the process of teaching mathematics, it is advisable to single out a number of levels of formation of this style, the multifaceted characteristics of which are presented in the text of the article. This characteristic formed the basis for the development of diagnostic tools that ensure the assignment of a student to the appropriate level.
- The content of the work revealed the transition of schoolchildren up the "ladder of levels". At the initial stage, such work is directly related to a basic enumeration of possible options for the state of a particular task structure and its possible transformation. At the next stage, a purposeful search for opportunities to connect the problem situation under consideration with a more familiar alternative problem structure by reformulating its condition is carried out, as well as identifying how promising this structure is in terms of choosing a possible search direction. Finally, at the third stage of the work, the alternative style of thinking is actualized as a result of a conscious free choice from a number of implicitly presented meaningful alternatives. Such a choice can be provided on the basis of a step-by-step transformation of the components of the task situation, leading to the construction of a sequence of new alternative task structures, which is accompanied by their comparative analysis.
- On the basis of the selected theoretical positions, teaching materials were developed, which were successfully tested in a number of schools in Penza (Russia) over a number of years. As evidenced by the results of approbation, subjected to statistical processing, the level-by-level strategy, as described by the author, for the formation of an alternative style of thinking for schoolchildren in the process of learning mathematics proved to be, on the whole, quite effective both in terms of forming students' readiness to overcome their initially one-sided attitudes, and in terms of developing their cognitive activity on the whole.

At the next stage of the research, we plan to organize a wider pedagogical experiment to determine the optimal means of implementation of the proposed recommendations into a meaningful compilation of various programs in mathematics for different countries and regions.

References

[1] M. Uzunoglu, M.S. Alam, “Dynamic modeling, design and simulation of a P 1. V.P. Zinchenko, Psychological Foundations of Pedagogy (Psychological and Pedagogical Bases for the Construction of Elkonin-Davydov System of Developmental Education), Moscow, Gardariki, 2002.

[2] E. de Bono, Lateral Thinking: A Textbook of Creativity, London, Penguin Books, 1977.

[3] R.M. Granovskaya, Yu.S. Krizhanovskaya, Creativity and Overcoming Stereotypes, St. Petersburg, OMS Publ., 1994.

[4] M.A. Kholodnaya, Cognitive Psychology. Cognitive Styles, Moscow, Yurait Publ., 2018.

[5] V.G. Thomas, B.A. Parsons, “Culturally responsive evaluation meets systems-oriented evaluation,” American Journal of Evaluation, **38**(1), 7–28, 2016, doi: 10.1177/1098214016644069.

[6] N.N. Khramova, M.M. Rodionov, “The development of thinking variability of schoolstudents in mathematics using 1C: Math designer,” Computer Science and Education, **7**(256), 15–21, 2014.

[7] W. Lake, M. Wallin, G. Woolcott, W. Boyd, A. Foster, C. Markopoulos, W. Boyd, “Applying an alternative mathematics pedagogy for students with weak mathematics: meta-analysis of alternative pedagogies,” International Journal of Mathematical Education in Science and Technology, **48**(2), 215–228, 2017, doi:10.1080/0020739X.2016.1245876.

[8] D. Pólya, Mathematical Discovery: On Understanding, Learning and Teaching Problem Solving, Moscow, Nauka, 1976.

[9] M. Rodionov, Z. Dedovets, “The development of students’ intellectual tolerance in the process of teaching mathematics at secondary level,” London Journal of Research in Humanities and Social Sciences, **18**(2), 1–11, 2018, https://journalspress.com/LJRHSS_Volume18/397_The-Development-of-Students-Intellectual-Tolerance-in-the-Process-of-Teaching-Mathematics-at-Secondary-Level.pdf.

[10] M. Rodionov, Z. Dedovets, “The forming of students’ intellectual tolerance: theoretical study and implementation,” International Journal for Cross-Disciplinary Subjects in Education, **4**(1), 775–781, 2014, doi: 10.20533 / ijcdse.2042.6364.2014.0259.

[11] M.A. Rodionov, E.V. Marina, Formation of Variable Thinking of School Students in Solving Construction Problems: A Manual, Penza, PSPU Publ., 2006.

[12] A.G. Stepanov, “The style of thinking in the formation system mythologemes paintings of socio-historical reality,” Bulletin of Chelyabinsk State University, **8**(430), 29–34, 2019, doi: 10.24411/1994-2796-2019-10805.

[13] M.A. Gadzhimuradov, Z.D. Gadzhieva, “Features of mathematical thinking

and its development when learning geometry," The World of Science, Culture and Education, **2**(69), 189–191, 2018.
[14] K.G. Kozhabaev, R.S. Gabdullin, "About pupils thinking development in the process of mathematics training," Modern High Technologies, **5**(3), 554–559, 2016, http://top-technologies.ru/ru/article/view?id=35952.
[15] E.Kh. Galyamova, "Methodological peculiarities of the process of geometry teaching considering the cognitive styles," Prepodavatel XXI Vek, 1(1), 153–158, 2019.
[16] S.N. Lipatova, O.A. Milinis, "Psychological and pedagogical support of the development of students's giftedness in the system of additional education: innovative model," Otechestvennaya i Zarubezhnaya Pedagogika, **2**(49), 200–210, 2018.
[17] C. Plumpton, Generalization and structure. In Process of Learning Mathematics, L.R. Chapman, Ed., Oxford, Pergamon Press, 87–121, 1972.
[18] H. Poincaré, The Foundations of Science (Science and Hypothesis. The Value of Science. Science and Method), Moscow, Nauka, 1990.
[19] T.L. Killpack, L.C. Melón, "Toward inclusive STEM classrooms: What personal role do faculty play?" CBE Life Sciences Education, **15**(3), 2016, doi:10.1187/cbe.16-01-0020.
[20] E.Yu Ilaltdinova, V.V. Kisova, "Teaching talent and teaching intellect: theoretical and applied approaches design," Vestnik of Minin University, **4**(25), 2018, doi:10.26795/2307-1281-2018-6-4-9.
[21]. I. Newton, Mathematical Works. Transl. by D.D. Mordecai-Boltovsky, Moscow, Leningrad, ONTI Publ., 1937.
[22] M.A. Rodionov, Formation of Search Motivation in the Process of Teaching Mathematics: A Manual for Students and Teachers, Penza, PSPU Publ., 2001.
[23] M.K. Watfa, D. Audi, "Innovative virtual and collaborative teaching methodologies," Behaviour & Information Technology, **36**(7), 663–673, 2017, doi: 10.1080/0144929x.2016.1275806.
[24] M.M. Thomson, D. DiFrancesca, S. Carrier, C. Lee, "Teaching efficacy: exploring relationships between mathematics and science self-efficacy beliefs, PCK and domain knowledge among preservice teachers from the United States," Teacher Development, **21**(1), 1–20, 2017, doi:10.1080/13664530.2016.1204355.
[25] F. Perls, R. Hefferline, P. Goodman, Experiences in the Psychology of Self-Knowledge: A Workshop on Gestalt Therapy, Moscow, Gil-Estel Publ., 1993.
[26] M. Rodionov, Z. Dedovets, E. Pavlova, N. Sharapova, I. Akimova, "Design and implementation of adaptive technology for teaching mathematics to school students based on integrated diagnostic approach to subject preparation and competence development," Revista Amazonia Investiga, **9**(26), 458–472, 2020, doi:10.34069/AI/2020.26.02.53.
[27] M.A. Rodionov, E.V. Marina, Developmental Potential of Mathematical Tasks, and Possibilities for its Actualization in the Educational Process: A Manual, Penza, PSPU Publ., 2010.
[28] I. Lakatos, Proofs and Refutations: The Logic of Mathematical Discovery, Moscow, Nauka, 1967.
[29] J. Engelbrecht, A. Harding, "Interventions to improve teaching and learning in first year mathematics courses," International Journal of Mathematical Education in Science and Technology, **46**(7), 1046–1060, 2015, doi:10.1080/0020739X.2015.1070441.
[30] T.A. Chernetskaya, M.A. Rodionov, "Interactive creative environments as a means for forming elements of mathematical research activity in schoolstudents," Computer Science and Education, 3, 36–41, 2014.
[31] V. Mkrttchian, G. Vishnevskaya, M. Rodionov, Avatar-Based Learning and Teaching in Modern Educational Environments: Emerging Research and Opportunities, Hershey, PA, IGI Global, Information Science Reference, 2018.
[32] P.C. Abrami, R.M. Bernard, E. Borokhovski, D.I. Waddington, C.A. Wade, T. Persson, "Strategies for teaching students to think critically: a meta-analysis," Review of Educational Research, **85**(2), 275–314, 2015, doi:10.3102/0034654314551063.
[33] H. Heckhausen, Motivation and Activity, Vol. 1, Moscow, Pedagogica, 1986.
[34] M.A. Rodionov, I.V. Akimova, T.A. Chernetskaya, "Creation of the constructive creative environment on the basis of realization of interactive dynamic models," In the World of Scientific Discoveries, **9**(57), 21–34, 2014, doi:10.12731/wsd-2014-9-2.
[35] M. Rodionov, Z. Dedovets, E. Pavlova, N. Sharapova, I. Akimova, "Design and implementation of adaptive technology for teaching mathematics to school students based on integrated diagnostic approach to subject preparation and competence development," Revista Amazonia Investiga, 9(26), 458–472, 2020, doi:10.34069/AI/2020.26.02.53.
[36] V.A. Mazilov, Yu.N. Slepko, "Pedagogical giftedness as a key prerequisite for efficient modern educational system," Integration of Education, **23**(1), 37–49, 2019, doi:10.15507/1991-9468.094.023.201901.037-049.
[37] S. Abramovich, A.Z. Grinshpan, D.L. Milligan, "Teaching mathematics through concept motivation and action learning," Education Research International, 7, 1–13, 2019, doi:10.1155/2019/3745406.
[38] G.T. Klinkov, M.A. Rodionov, O.E. Kozlova, E.V. Veszetiu, E.V. Vovk, "Implementation of discussion technologies in a personality-centered professional education," Revista Amazonia Investiga, **9**(26), 82–87, 2020, doi: 10.34069/AI/2020.26.02.9.
[39] K.W. Remijan, "Project-based learning and design-focused projects to motivate secondary mathematics students," Interdisciplinary Journal of Problem-Based Learning, **11**(1), 2017, doi:10.7771/1541-5015.1520.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 259-277 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Software Development Lifecycle for Survivable Mobile Telecommunication Systems

Mykoniati Maria*, Lambrinoudakis Costas

Department of Digital Systems, University of Piraeus, Piraeus, 18532, Greece

ARTICLE INFO

Article history:
Received: 21 March, 2021
Accepted: 19 July, 2021
Online: 03 August, 2021

Keywords:
Telecommunication Systems
Software Development Lifecycle
Survivability

ABSTRACT

Survivability of systems is a very important system property and consists major concern for organizations and companies. Survivable systems should maintain their critical services functional in a timely manner. There are several approaches, proposed in the literature, on how to develop survivable telecommunication systems, but the majority is based on node outages or path failures, missing the main scope of survivability which is service failure. The contribution of this paper is that it presents a SDLC (Software Development Life Cycle) for developing survivable mobile telecommunication systems. Additionally, the main characteristic of a mobile telecommunication system is that it consists of different types of nodes (ex. MME, SGSN, etc.) that are connected to systems (ex. 5G, 4G, 3G, 2G etc.) and thus form an intersystem that provides services to end users. This interconnection and interoperability of network nodes is of high complexity constituting a threat to system survivability. Thus, another contribution of the current research work is that it provides a systematic approach for handling this complexity.

1. Introduction

Availability and continuity of critical IT infrastructures is a matter of concern in many of scientific fields like security, robustness, fault tolerance etc. In fact, the unavailability and failure of such infrastructures causes severe financial losses to many organizations.

Survival of IT infrastructures, like information systems or network systems is a matter of concern for any company that develops and maintains network systems. That means that such systems should continue to support the critical services even during attacks, failures or accidents. A definition of survivability is: "survivability is the capability of a system to fulfil its mission, in a timely manner, in the presence of threats such as attacks or large-scale natural disasters" [1], with security, robustness, fault-tolerance and recovery of systems to be among survivability's main disciplines.

It is important to highlight that survivability focusses on the survival of the mission of the system and not of the system itself. This is the core principle of survivability.

There is much research on survivability measures and approaches that should be adopted by a system to be survivable. But how can we be sure that a system is survivable? What are those capabilities that should be tested in order for a system to be characterized as survivable and against which threats? Additionally, what are the interconnections and interoperability threats that should be considered when survivability of large complex system of systems, like mobile telecommunication systems, is examined and how could these be analysed at everyday work when building such systems?

Through literature review, a detailed research on survivability approaches is presented highlighting that most of them address survivability of telecommunication networks by handling node or path outages. However, survivability should be based on service failure and not on system failure. In fact, even if the entire network is performing as expected, there could be failures in services for many other reasons. For example, a software bug could result in a specific service failure, or delays caused by excessive load in specific network nodes could result in random service failures. Another reason could be that robustness requirements are not considered during system design. A very representative example is the handling of collision scenarios, where two messages requesting a service arrive at the same node simultaneously. Robust system design could resolve this conflict.

To conclude, the contributions of the current paper are:

- The solution proposed by the current paper is Survivability by Design, meaning that survivability should be part of the software development lifecycle (SDLC) of the telecommunication system. The idea comes [2] which is a

*Corresponding Author: Mykoniati Maria, mmykoniati@gmail.com

paper titled as "Life-Cycle Models for Survivable Systems", that proposes survivability to be part of the SDLC phases and describes how this could be achieved. This is the theory that the current research is based on to describe how survivable telecommunication systems shall be developed.

- Another contribution of the current paper is that it addresses the risk of service failures arising from the increased complexity of interconnection and interoperability of mobile telecommunication network nodes. This is a major concern since most of the times development teams tend to focus only on the node under development, when new features are to be developed, without taking into consideration requirements or threats coming from connectivity with the other nodes. More specifically, even if the entire system is tested end-to-end, when a mobile telecommunication network node is operating in the provider's environment, it may be connected to nodes developed by other companies. The behaviour of that node is unpredictable, and this should be considered during SDLC phases, by setting appropriate survivability requirements and design practices, and by testing without ignoring specific failure scenarios.

During the next chapter, survivability as a term is examined in order to present the main principles and requirements of survivability. Following this literature review, the general framework in the form of a software development lifecycle (SDLC) is presented. Finally, the paper closes with overall conclusions.

2. Literature Review

2.1. Survivability as a term

As described in [1], survivability is the ability of a system to maintain its critical services that serve system's mission in a timely manner in case of attacks, failures or disasters. As a result, survivability itself is a system property that the system should emerge and should be considered as a requirement during the design phase and not as an ad-on characteristic [2]. Additionally, since the focus is on critical services and system mission, survivability should be considered as a different set of characteristics for each system, based on system's scope. For example, for a telecommunication network, survivability as a requirement may include, define and implement mechanisms that would allow the system to feature robustness, fault-tolerance, interoperability, restorability, security, safety, resilience, dependability etc, for its critical services in order to provide uninterrupted communication to end users. For an e-shop, usability or secure transactions would also be key principles for the survival of the mission of the system. There is much research on gathering these characteristics to a general set for systems' design, with the most representative one being the research described in [2]. They argue that for any system survivability is succeeded if it has the ability to provide Resistance, Recognition and Recovery (3Rs) from attacks or failures. In extend the system should provide Adaptation and Evolution by improving system survivability and increasing its resistance by knowledge gained from previous attacks or failures.

Threat for the survivability of a system, according to [3], is anything that may prevent the system from providing its essential services under the "minimum acceptable level of service", or affecting the provision of its essential services for more time than the one predefined as acceptable. As a result, the threat against a system's survivability is unknown and not always predictable through a risk analysis. Therefore, it is critical for survivability to gather, analyse and deal with the impact threat incident may cause, rather than focussing on predicting all possible threats. For instance, from the "survivability point of view", it is more important to focus on how a network node would behave under a Denial of Service attack and how it could recover rather than identifying measures that would prevent this attack.

2.2. Survivable Systems

Having defined survivability, a brief description of different approaches that have been adopted for designing and implementing a system that satisfies the survivability requirements follows.

Starting with the Survivability Analysis Framework (SAF) [4], survivability is considered as a set of peoples' capabilities, a set of actions and of technology working together to achieve operational effectiveness. The focus is on interoperability of organizational components and how to cope with complexity arising from this interoperability in order to analyse potential failure conditions, likelihood of error conditions, impact of occurrences, or recovery strategies. This analysis yields requirements for the design and implementation of the system.

The second approach considers survivability as part of the system's development life cycle. It is described by research [2] and claims that "survivability goals and methods must be addressed for each action of the life-cycle", as survivability should be integrated into the primary development phase of system and not treated as an add-on property of an already implemented system. Starting with requirements specification, the system should be able to monitor itself in order to recognise attacks or failures, resist and recover from attacks and failures and reconfigure to adapt to attacks and failures. Additionally, after mission definition, essential services of system should be depicted, and the system should be designed in such a way so that to maintain these services when it is under attack or failure. Continuing with requirements, intrusion requirements should be defined, in order for the performance of the system under attack or failure to be defined, in order to ensure that acceptable levels of quality of service are always reached. What is important here is that intrusion scenarios are considered as usage scenarios to be handled. The testing of these requirements should include three attack phases, the penetration phase, where the intruder attempts to gain access to the system, the exploration phase, where the intruder has gained access and is exploring the integral system organization and capabilities to find possible exploitation targets, and the exploitation phase where the intruder performs attacks against system facilities. According to these phases, survivability strategies for resistance, recognition, recovery, adaptation and evolution must be enforced. By considering these requirements, the system may be designed and implemented as survivable.

The third approach is presented in [5], and it is based on analysing the different states of quality of service, that the system may fall into during a failure, and on estimating the probability of

the essential services being available during the failure. After changes to the environment or attacks to the system, the system may degrade to the next quality of service level. When failure is restored, the system may return to the higher QoS level. Acceptable QoS levels for the system and transitions between them, may be modelled with the use of a transition matrix.

Another approach for providing survivability is the one proposed by [6]. Contrary to the security approaches that try to prevent an attacker to gain access, the assumption here is that the attacker has gained access and the objective is to try to find ways to prevent him from interfering with systems' critical services. Methods of prevention are based on frustrating the attacker to believe that he or she has gained access to essential services.

A fifth approach is presented in [7] known as the WILLOW architecture. It is a proposal that focuses on proactive and reactive reconfiguration of a system in order to achieve survivability for its services. During proactive reconfiguration, it is possible to add, remove and replace components and interconnections of the system, as well as to adjust their mode of operation. This is called posturing and is used to minimize the system's vulnerabilities that can be exploited by various threats. For instance, such a reconfiguration may be to turn-off non-essential services and networking links as well as to strengthen the cryptographic keys if a virus has infected the system. The reactive configuration does the same actions, aiming to restore a system from damage or intrusions, in specific time intervals. In fact, as proposed, the most appropriate approach for reacting is fault tolerance. An example, of reactive reconfiguration against an attack or damage is the activation of applications' copies.

A similar approach of reconfiguring the system and switching to different level of quality of service is also provided in [8], where the authors claim that QoS and survivability are firmly connected. As a result, if QoS is to be measured, reconfiguration approaches may be triggered under certain measurements to provide survivability for the system. Firstly, as "survivable system", may be characterized, any system that may repair itself or degrade in such a way that will provide as much functionality as possible. This may be done if the system is able to switch between alternatives of acceptable predefined levels of functionality. Secondly, a survivable system is a system that may adapt threats in its environment and environmental changes and reallocate essential processing to most robust resources. All these may be achieved through dynamic reconfiguration. Such reconfiguration may be "process/host restart, migration of objects to alternate hosts, replication, transparent rebinding of clients and servers, use of service alternatives, and approximate services". [8] These reconfigurations may be based on several metrics like "available battery power, varying communication bandwidth, available memory or faults in software components" [8] and must be done in predetermined time and based on QoS service levels. Then a survivable system must provide a minimum level of QoS under changing environments. For that purpose, the best-suited elements are to be chosen at each time, based on these QoS factors.

2.3. Evaluation of System Survivability

According to related literature, evaluation of systems' survivability, is mainly based on defining different acceptance levels of system performance and on evaluating the impact by measuring the key properties like number of outages, time needed for system recovery etc. Though, these evaluation models are mostly based on node failures or link failures, but they are not giving the whole idea about the quality of service the system provides to end users. As a result, they seem to be based on system availability and continuity and not on critical services or system's mission availability. Of course, system's availability is of vital importance for supporting system's mission and providing end to end functionality. So, system availability should be part of any survivability analysis and evaluation plan. Thus, the purpose of this paper is to provide an entire evaluation framework of all survivability aspects and not only providing system - centric evaluation methods. As a result, many of these evaluation models could be very useful to pinpoint any possible network failures and include these in a test suite that would test if the system could recover from them or if it could function as expected while the system is suffering from these failures. But it is very important to provide guidance for testing or evaluating systems' survivability from the requirements specification step of a SDLC, up to the release of new product.

Starting with [9], the authors use a Markov model to map the possibility of a failure. They base survivability measurements on the frequency of failure events, on the duration of outages and on the impact of failure. Since the research is conducted through a case study with wireless networks, as a failure is considered node failure, power faults and link failures. A similar approach is proposed by [10], where the authors are using a semi-Markov survivability evaluation model for intrusion tolerant database systems. As key attributes for quantification of a database's survivability, integrity and availability are proposed. Much focus is paid on system's functionality under failure and how system performs against these attributes.

To continue with quantification of system's survivability, the author in [11], proposed network condition metrics which are density (based on topology and its changes), mobility (speed of node, predictability etc.), channel (bit error rate, capacity distribution etc.), node resources (memory, computing power etc.), network traffic (QoS, packet size, distribution etc.), derived properties (degree of connectivity, queueing delay, propagation delay etc.). In addition to those metrics, service requirements are also defined. Again, every adverse event, transits system's performance to another state which is quantified by these measurements (based on network and service performance) in order to be marked as acceptable or not. Another approach based again on a Markov model is being presented in [12]. It is focused on call losses of a telecommunication switching system because of various system failures like hardware/software faults, human errors, impairment damage from adverse environments etc. As key survivability metrics, system performance, availability and performability are used and the measurements proposed are measurements that can be used to describe system survivability such as the number of functional units, the number of connected nodes, the maximum traffic capacity, blocking probability, throughput/goodput, and the service restoration time.

To continue with evaluation methods, in [13], authors propose a testing survivability framework, focusing again on the recovery part of the survivability attributes. They firstly present the idea of 5-step phases of survivability of a system under failure, normal

phase, resistance phase, destroyed phase, recovery phase and adaptation phase. Then they propose a scheme for representing the different stages of system performance against time during these phases. For quantification of network performance, two factors are proposed to be used, the Node Connectivity Factor (NCF) and the Link Connectivity Factor (LCF). Practically though, they try to focus on the availability of an end-to-end activity for the end user which is what really matters. This is why their research focuses on source-destination pairs "SD-pairs", to describe connectivity and service quality "SD-quality" and test these factors by applying different failures in order to calculate SD Recovery time for each pair. Finally, NRD metric is calculated to give an overall idea about the entire system's survivability.

Another very important research on evaluation of survivability has been conducted by authors in [14]. The framework proposed, is based on developing a general measurement model, which may be specified based on specific domain requirements, a network survivability testing model, which is based on testing network performance against survivability metrics during different steps of system performance (resistance, destroy, recovery), and the network survivability evaluation, which includes measurement of the entire system's survivability based on different metrics, evaluation models or algorithms. The method concludes to a mechanism which if applied to the system under test, may provide all possible combinations of test schemes to test failures of a network and to measure them in order to extract conclusions on the overall system's survivability.

In [15] the authors propose measuring survivability through four attributes, Process-Weighted Average Availability (PWAA), Process-Weighted Average Controllability (PWAC), Process-Weighted Average Robustness (PWAR), Process -Weighted Average Adaptability (PWAD). These depict the state of the system through survivability life cycle, which is normal state, resistance state, destroy state, recovery state and adaptation phase.

Finally, another important approach for quantifying survivability is coming from authors in [16], who propose to base quantification, on system's reaction to specific attacks and vulnerabilities modelled by attack graph. The attack graph represents the nodes that the attacker may exploit, while the way chosen to transverse these nodes in order to cover all possible system functionality states is forward-search, breadth-first and depth-limited.

To conclude, what may be observed is that most approaches on quantifying survivability are based on measuring availability and robustness characteristics of the system. Though, survivability is a more complex attribute that the system as a whole should emerge and should be based on the ability of the system to continue serving critical services. As a result, the approach proposed in this paper for evaluating survivability, is a testing framework focussing on testing services available against systems failures, attacks or accidents.

2.4. Survivability and Telecommunication Systems

Before concentrating on the proposed SDLC for mobile telecommunication systems, we conclude the current literature review with a brief presentation of a few representative approaches for designing and implementing a survivable telecommunication system. It becomes clear that all these approaches are focussing on outages and path failures and not on service failures as survivability preserves.

In [17], the authors investigate the impact of possible failure scenarios and possible survivability strategies to contend with spatial and temporal network behaviour in mobile cellular networks. The failures for this paper are restricted to loss of BS, BSC-MSC or VLR. In [18], the authors analyse architectural principles for achieving minimization of services loss and service restoration through certain disaster recovery plans. The failure scenarios that are considered are central office switch fires, earthquakes, flooding, large-scale power outages, signalling network outages, fiber cuts, and terrorism. The result of these scenarios are outages to network devices for which the paper introduces a four-phase methodology to handle such cases. Another approach for providing survivability to Universal Mobile Telecommunication Systems (UMTS) networks is based on Markov chains, semi-Markov process, reliability block diagrams and Markov reward models [19].

What we may observe from these approaches is that the designs proposed are based on fault tolerance techniques and on how to mitigate the failure of network nodes. There are many other approaches in literature that indicate various techniques to handle the impact of the failure of a node or a link. Though, survivability is far more than that. Survivability should be part of every step of the SDLC. The current research focuses on providing survivability requirements for mobile telecommunication systems that should be taken into consideration during the requirements elicitation phase of the SDLC, and on how to validate the satisfaction of these requirements during the testing or development phases.

To sum up this literature review on survivability as a term and on approaches for providing survivability to a system the following requirements should be adopted:

- Survivability is a mission driven attribute which means that the mission of the system is what should survive at the end, and not the system itself. Additionally, the majority of approaches discriminate and mark system services at essential and non-essential services with the essential services being the ones that should survive, and perform at an acceptable level of QoS, when a system is under attack or failure.
- Threat against survivability is any failure that may affect its critical services. So, the system should be able to react to any failure even if the root cause is unknown.
- A system must be designed as survivable and for this to be succeeded, survivability requirements, based each time on system's nature, must be defined during requirements specification of every development life cycle. These requirements may be organized to 3Rs (recognition, resistance, recovery and adaptation methodology) Additionally, survivability requirements should be considered during all stages of system's development lifecycle and as part of the everyday work.
- For a system to be compliant with survivability requirements specification, a monitoring system that monitors and evaluates system's survivability is of vital importance. Additionally, if a monitoring system is available, the state of the system may be

known each time and preventive or corrective actions, like re-configuration or other system's self-healing processes, may be applied for providing survivability to the system, even when unplanned threats are realised.

- Finally, testing and evaluation of system's survivability should contain investigation of intrusion scenarios and failure incidents in order survivability requirements to be raised. This could be very useful if test driven development methodologies are used.

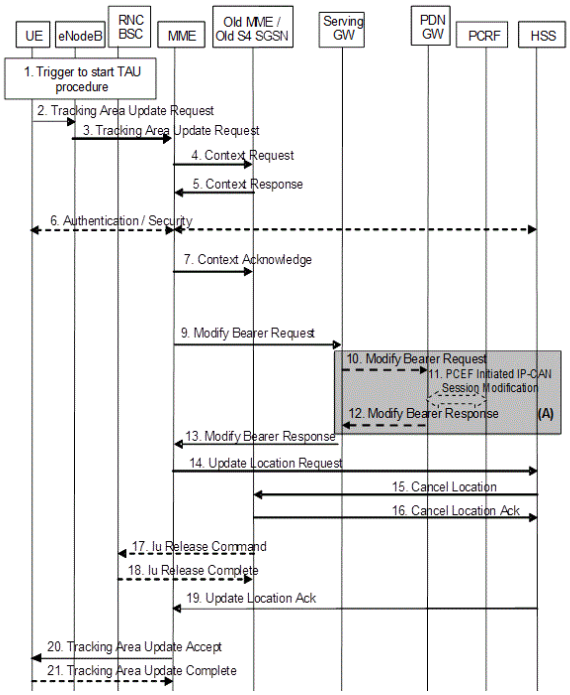
2.5. Mobile Telecommunication Systems

Before closing literature review, we will present some information on mobile telecommunication networks. Nowadays mobile telecommunication networks consist of a combination of 2G, 3G 4G and 5G mobile networks. Each network consists of the radio access network and the core network, which is finally connected to various networks like internet, IP Multimedia Subsystem (IMS) etc, to serve system's main mission which is to facilitate voice and data communications. Among network nodes, the communication in control-plane layer and user-plane layer is being established through specific interfaces.

Each of these systems has several nodes connected to each other. The particularity of mobile systems compared to other systems, like the internet, is that all services need an exchange of messages between a set of nodes to be established and performed. This significantly increases the risk of failure since problems may occur at any time during the exchange of the aforementioned messages. An example of such a message flow and possible failures that may occur, can be found in [20] or in the 3rd Generation Partnership Project (3GPP) standards. To continue with this logic, the network nodes that are connected to realize a service may be part of the same or a different network. For example, in 3G to 4G intersystem Tracking Area Update service, the nodes that may participate are from 4G, network nodes eNodeB, Mobility Management Entity (MME), Packet Gateway (P-GW), Serving Gateway (S-GW), Home Subscriber Server (HSS) and radio network controller (RNC), and Serving GPRS Support Node (SGSN) network nodes from the 3G network. This scenario is depicted in figure (1) bellow. Additionally, nodes may be manufactured from different organizations, a fact that increases the risk of interoperability failures. As a result, with various nodes interconnected, new networks are formed adding new system and survivability requirements that must be considered through the development of any new feature. The whole picture of a mobile network is shown in figure (2) bellow. This figure depicts the interconnection between 2G, 3G and 4G mobile networks through relevant interfaces. Though, the 5G network and the way it is connected with the rest of the mobile networks is missing. For this purpose, we utilize another picture from [21] that depicts the connection of the 5G network with the 4G network. This is figure 3 below.

The view of such interconnected systems adopted by the current work for all stages of the software development lifecycle is a multi-layered logic with the following levels:

- **Node level**: Any node of a mobile telecommunication network for which a new functionality or feature is to be developed. For example, MME should be considered to perform in node level.
- **System Level**: 2G, 3G, 4G and 5G, or any other that follow, are considered as systems. Nodes forming a system could be part of different PLMN operators. Any development task for a service that includes network nodes from the same system should be considered in system level.
- **Intersystem Level**: The entire telecommunication system may be considered as an intersystem. Nodes forming a network for serving an inter-system scenario may be considered as an intersystem. For example, in the scenario below, an Intersystem Tracking Area Update is depicted. The scenario includes nodes from 3G and 4G systems.



].

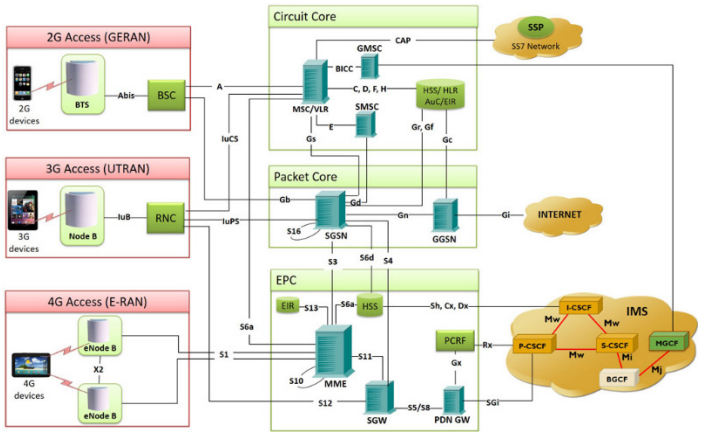


Figure 2: Common telecommunication network – The whole image (http://www.gl.com/telecom-test-solutions/communications-networking-2G-3G-4G-lab.html)

What is also important is that nodes supporting system or intersystem scenarios could even be part of different public switched telephone network (PLMN) operators. This means that when developing a new feature, the behaviour of nodes should not be considered as "known". Any possibility of receiving an unexpected message should be taken into consideration and the system should be able to resist to such a threat and recover from failure.

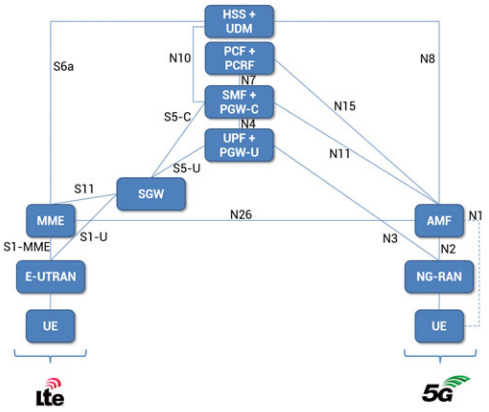


Figure 3: Common telecommunication network – 5G system added to the whole image (https://www.rfglobalnet.com/doc/g-core-network-architecture-network-functions-and-interworking-0001)

3. SDLC of Survivable Telecommunication Systems

Nowadays, systems development is mostly based on iterative models, or spiral models, in order to support continuous delivery of new functionality with certain predefined criteria. At the end of all iterations, an updated system, or a new release, is tested against its overall functionality in order to be delivered to the telecommunication operators.

Current research aims to improve this process by considering the survivability of critical services as the main requirement of the system under development. The main idea is to consider the whole (inter)system as a deliverable of any new release, instead of just focussing on a small part of the network. In this way, all survivability requirements at all system levels are considered and tested. The contribution of the current research is that it provides a complete proposal on how to handle survivability requirements and quality assurance of developed telecommunication system based on these requirements. The requirements are categorized to those related to service and those related to network since without it the system will not be available to perform any service. Additionally, the methodology proposed takes into consideration any arising requirement from the complicated interconnections of the telecommunication subsystems. All these requirements are gathered and grouped into 3Rs categories as described in literature; recognition, resistance, recovery and adaptation. In other words, requirements are enriched to include the whole network's survivability requirements. The result of not taking into consideration system and node inter-operability is a very important increase on the number of defects. Additionally, the testing methodology proposed by the current paper, considers all possible service failure scenarios and possible impact of any new functionality to the legacy code for critical services already developed.

The inputs to the aforementioned methodology are new features that will be developed or/and possible defects. When a new feature or a defect is planned to be developed, a new SDLC starts.

According to related literature, any methodology for designing survivable systems should start by defining the system's mission and the critical services that serve that mission. These should be documented and dealt with as requirements to any new functionality.

For mobile systems, critical are all services related to voice or data transmission from user perspective, and charging services from operator's perspective. This is also depicted in table (1) below, with service level requirements. So, for example, a voice bearer may be considered as critical service. A handover to such a bearer is critical also.

After definition of the mission and critical services that should survive, the general software development lifecycle (SDLC), is modified and used, with respect to special characteristics of the developed system, in such a way that at the end of the cycle the delivered (inter)system to emerge survivability. The SDLC that is proposed is depicted in figure (4).

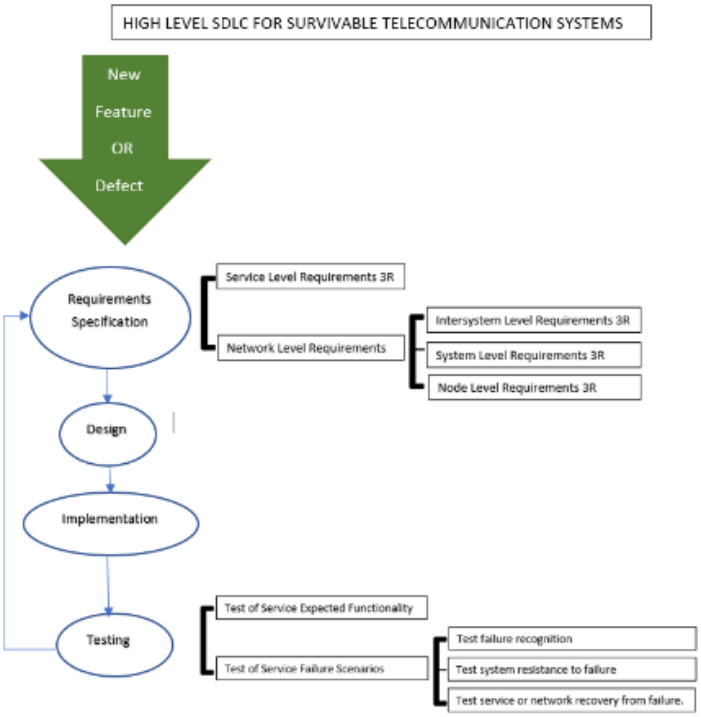


Figure 4: SDLC of Survivable Telecommunication Systems.

3.1. Requirement's specification

Requirements for extending the system's functionality are predefined and described in 3GPP documents. Survivability requirements should be based on a risk analysis study and detailed examination of the potential threats. As already explained, threats against survivability of the system are those that can directly affect the critical services of the system. This is the most effective way to protect critical services as such a service should survive even if the root cause of the failure is unknown. Thus, requirements are grouped to service level requirements that are related to services and network level requirements that are related to network

availability in order to support the operation of the services. For each group, requirements related to 3Rs (recognition, resistance, recovery, adaptation) methodology are presented.

In the tables below, high-level requirements related to survivability and defined by 3GPP are depicted. All these requirements are related to survivability and should be considered additionally to any requirement related to a new functionality or to any maintenance task. Additionally, any requirement that is an outcome of our research may also be depicted in service level survivability requirements table under columns titled "Our contribution". These requirements are related to **failure recognition** and **resistance** and are presented to previous papers [20], [22] related with survivability on telecommunication systems. Furthermore, the error handling requirements proposed from the current paper may be summarized to the following ones:

1. The system should be able to resist to failures related with **loss of messages**.

2. The system should be able to react to messages **arriving later or earlier** than expected. This should not have any impact to the service or to any other following services.

3. The system should be able to resist to failures related with **duplicate messages** sent to the nodes.

4. Any new functionality should be considered as a **threat to the critical services** already developed and any possible failure should be handled.

5. "**Hanging processes**" should also be considered as possible causes of failure.

Table 1: Service Level Requirements

3GPP Title	**3GPP Doc Num**	**3GPP Service Survivability Requirements related to failure Recognition**
UMTS Terrestrial Radio Access (UTRA) system	2101-301	*"*Set of attributes to describe UMTS bearer service (**delay variation tolerance, maximum transfer delay, maximum bit error rate**) information transfer rate attributes **(peak bit rate, mean bit rate, occupancy)**."* [23] *"*Performance: inherent transmission delay and level of traffic blocking"* [23]
Performance Management (PM)	32401	*"Data gathered through telecommunication management system are gathered to support performance evaluation on:* *- **Quality of Service** (e.g. delays during call set-up, packet throughput, etc) QoS can indicate the network performance expected to be experienced by the user."* [24]
Found across multiple 3GPP documents		
Error Causes **Please refer to certain interface 3GPP document for more details**		Specific error causes may be returned to the request message each time indicating a certain failure. For example, in GPRS Tunnelling Protocol (GTP) messages error cause "Mandatory IE incorrect" may be returned. From this the root cause of failure may be depicted and corrected by development team in case it can be corrected. Otherwise, there may be causes like "network failure" with root cause some failure to the network where all connections of the node with the node that returned this value, should be deleted.
Our contribution		
"Self-Diagnosis Framework for Mobile Network Services" [20]		Using the management reference mode of 32.101 we have proposed a self-diagnosis framework that may recognize and report different kinds of failure of service flow between nodes. Using this framework, the root cause of failure may also be depicted. Failures that have been analyzed are any possible failures that may occur when a message of a flow leaves a node to reach the neighboring node. The contribution of the paper is that focuses on diagnosis of service failure and not of system failure opposed to other proposals and to telecommunication management standard.
3GPP Title	**3GPP Doc Num**	**3GPP Service Survivability Requirements related to failure Resistance**
UMTS	2101-301	*"Handover should be transparent. In case of speech call loss of information may be tolerated but handover should be quick to **avoid connection break.** In case of data service temporary break is tolerable but not loss of information. Handover between terrestrial environments should be seamless within the same network"* [23]
		"Handovers should not increase the load on the fixed network significantly" [23]

		"The level of security should not be affected by handovers" [23]
		"Bearer services cannot be handed over between two environments if they are not supported in both. However, handover to an alternative bearer ***offering reduced capabilities*** *should be possible where this is supported by the service in use. The radio interface should have the capability to provide for handover and roaming between networks run by different operators"* [23]
Services and System Aspects;	22 101	*"Any handover required to maintain an active service while a user is mobile within the coverage area of a given network, shall be seamless from the user's perspective."* [25]
		"The 3GPP system shall be able to provide ***continuity between CS voice services*** *and the full duplex speech component of IMS multimedia telephony service with no negative impact upon the user's experience of the voice service. The same should be true for IMS Services."* [25]
		"The system shall support either *- transparent relay of the IP signaling and traffic;* *- service aware interconnection"* [25]
3G security; Security threats and requirements	21 133	*"Service Integrity: "It shall be possible to protect against unauthorized modification of user traffic"* *Service availability: It shall be possible to prevent intruders from restricting the availability of services by logical means"* [26]
Security Objectives and Principles	33 120	*"Security Objectives:* *1. to ensure that the security features standardized are compatible with world-wide availability* *2. to ensure that the security features are adequately standardized to ensure world-wide interoperability and* *roaming between different serving networks;"* [27]
Security architecture	33 401	The standard presents: *"- user identities confidentiality: MSIN, the IMEI, and the IMEISV should be confidentiality protected* *- user data signaling confidentiality: All S1 and X2 messages carried between RN and eNB shall be confidentiality-protected. Synchronization of the input parameters for integrity protection shall be ensured for the protocols involved in the integrity protection.* *- Integrity protection, and replay protection, shall be provided to NAS and RRC-signaling.* *- authentication and key agreement procedure between the mobile device and the core network,* *- security interworking of mobile networks (EUTRAN-UTRAN-GERAN)"* [28]
Technical Specification Group Services and System Aspects;	23 401	*"Authentication: NAS security mode control procedure is to take an EPS security context into use, and initialize and start NAS signaling security between the UE and the MME with the corresponding EPS NAS keys and EPS security algorithms"* [21]
5G; Security architecture and procedures for 5G System	33.501	The standard presents: *"- network access security: enable a UE to authenticate and access services via the network securely, including the 3GPP access and on-3GPP access, and in particularly, to protect against attacks on the (radio) interfaces* *-network domain security secure exchange of signaling and user plane data between networks.* *- User domain security: user access to mobile equipment.* *- Application domain security: enable applications in the user domain and in the provider domain to exchange messages securely"* [29]
		As it is presented to the current standard part of network life-cycle includes: *"the PLMN network is being adjusted to meet the long-term requirements of the network operator and the customer, e.g. with regard to performance, capacity and customer satisfaction through the enhancement of the network or equipment up-grade"* [29]
Found across multiple 3GPP documents		

Error Handiling		Some error causes indicate failures that can be handled in order to avoid dropping the service. Sometimes these handlings may be found across 3GPP documents or there may be implementation specific approaches that each organization implements during development of the device. To the example above "Mandatory IE incorrect" if we assume that the mandatory IE that is not correct is bearer ID. And the message causing this error is an answer to a previous message, then we may conclude which is the correct bearer id and ignore the error instead of dropping the service. The same may happen with network errors if we use relocation through selection functions to relocate the service that may be dropped in case it is critical (voice bearer for example)
Collision Handling		Collision is the case where two messages requesting a service arrive at a network and at the same time or one request arrives before the whole process of messages of the previous one has been completed. Then a handing of these requests should take place. This handling may be for example to serve both requests by a priority sequence, or to drop one of the two. For example, in case a request arrives for a UE that is already in process of a handover there is no meaning in processing it since the UE will leave from current Tracking area. Though there are cases that the service should continue to the Tracking area the UE will move to.
Our conrtibution		
"Fault Prediction Model for Node Selection Function of Mobile Networks" [22]		Our proposal regarding service resistance to failure is the fault prediction model proposed. This model takes into consideration DPMO (Defects per million opportunities) value which is a value that may be used to evaluate the operational performance of a node against 6sigma value. Then this value is used as a parameter in selection algorithm of mobile systems. This function is used to select a node which will be used to successfully complete a service flow.
Error Handiling		Apart from error causes defined by 3GPP documents and robust measurements that should be developed in order such cases to be handled, here we introduce some other error handline requirements: 1. The system should be able to resist to failures related to loss of messages. The failure should be ignored if this is possible. For example, if an acknowledgement message has not arrived, the service could be considered as established to avoid dropping it. If it could not be ignored, then the system should consider if there is a failure of neighboring node. In this case, the node should inform network management system and release any connection associated with this node. 2. The system should be able to react to messages arriving later or earlier than expected. This should not have any impact to the service or to any other following services. 3. The system should be able to resist to failures related with duplicate messages sent to the nodes. 4. Any new functionality should be considered as a threat to the **critical services** already developed and any possible failure should be handled.
Hanging Processes		As "hanging processes" we mean a service that fails, and leaves resources reserved causing failure to future services. For example, if a PDN Connection fails to be released and it is found as "already established" when a new PDN Connection is requested. This PDN Connection may be a critical service like voice bearer.
3GPP Title	**3GPP Doc Num**	**3GPP Service Survivability Requirements related to service Recovery from failure and adaptation.**
Restoration procedures	23 007	*"The data stored in location registers are automatically updated in normal operation; the main information stored in a location register defines the location of each mobile station and the subscriber data required to handle traffic for each mobile subscriber. The loss or corruption of these data will seriously degrade the service offered to mobile subscribers; it is therefore necessary to define procedures to limit the effects of failure of a location register, and to restore the location register data automatically"* [30]

Services and Systems Aspects;	22 101	*"The voice call continuity user's experience shall be such that, to the greatest degree possible, a consistency of service is provided regardless of the underlying communication infrastructure and technology"* [25]
UMTS	2101-301	*"Flexibility: Negotiation of bearer service attributes (bearer type, bit rate, delay, BER, up/down link symmetry, protection including none or unequal protection), parallel bearer services (service mix), real-time / non-real-time communication modes, adaptation of bearer service bit rate"* [23]
		"UTRA should adapt flexibly into changes and should have the capability to serve a variety of traffic densities (up to very high densities) and a variety of traffic mixes in an economical way." [23]
		"Flexibility and dynamic reconfiguration: minimum set of bearer capabilities, operating modes and features to ensure that inter-operability is always possible; continuity of operation during dynamic updating of terminal capabilities." [23]
Self-Organizing Networks (SON); Self-healing concepts and requirements	32541	*"In the case of software faults, the recovery actions may be :* *a) system initializations (at different levels),* *b) reload of a backup of software,* *c) activation of a fallback software load,* *d) download of a software unit,* *e) reconfiguration, etc.* *In the case of hardware faults, the recovery actions depend on the existence and type of redundant (i.e. back-up) resources."* [31] *'[If the faulty resource has no redundancy, the recovery actions may be:* *a) Isolate and remove the faulty resource from service so that it does not disturb other working resources;* *b) Remove the physical and functional resources (if any) from the service, which are dependent on the faulty one. This prevents the propagation of the fault effects to other fault-free resources;* *c) State management related activities for the faulty resource and other affected/dependent resources;* *d) Reset the faulty resource;"* [31] *e) Other reconfiguration actions, etc.* *"If the faulty resource has redundancy, the recovery action shall be changeover, which includes the action a), c) and d) above and a specific recovery sequence. The detail of the specific recovery sequence is out of the scope of the present document"* [31]

Table 2: Network Level Requirements

3GPP Title	**3GPP Doc Num**	**3GPP Network Survivability Requirements related to failure Recognition**		
		Node Level	**System Level**	**Intersystem Level**
Telecommunication management; Principles and high-level requirements	32.101	*"Telecommunication management system consists of an architectural framework or management reference model, that is used to collect measurements for management functions. Some of which are related to survivability like performance management, fraud management, fault management, security management, etc. With the use of performance measurements, configuration of system due to load needs may be executed. Additionally, for fault management, alarms or events may also imply a needed re-configuration for avoiding failures. Failure may be detected; isolated and root cause may be depicted."* [32]		

Performance Management (PM)	32401	*"Data sent at node level are gathered through telecommunication management system to support performance evaluation on:* *- traffic levels within the network, including the level of both the user traffic and the signaling traffic* *- verification of the network configuration: evaluation of effectiveness of changes of network plan related to traffic levels.* *- resource access measurements* *- resource availability (e.g. the recording of begin and end times of service unavailability)*" [24]	*"Network Operators are informed of PM - related events through alarms and may act accordingly."* [24]	
Fault Management;	32.111-1	"*If the faulty resource has no redundancy, the recovery actions shall be:* *- Generate and forward appropriate notifications to inform the **OS** about all the changes performed*." [33]		
3GPP Title	**3GPP Doc Num**	**3GPP Network Survivability Requirements related to system Recovery from failure and adaptation**		
		Node Level	**System Level**	**Intersystem Level**
Restoration procedures	23 007	*"The data stored in location registers are automatically updated in normal operation; the main information stored in a location register defines the location of each mobile station and the subscriber data required to handle traffic for each mobile subscriber. The loss or corruption of these data will seriously degrade the service offered to mobile subscribers; it is therefore necessary to define procedures to limit the effects of failure of a location register, and to restore the location register data automatically. The document describes data restoration procedures for VLR, HLR, HSS, GGSN, SGSN, MME. Triggering point is receiving a request for unknown IMSI in cases when the failing node has not detected the failure or receiving a message with restoration indicator set to not confirmed. These indicators show data corruption and procedure for restoring of these data through message exchange follows."* [30]		
		"Node restart. If a node restarts it sends a reset indicator to the neighboring nodes. Upon receiving such an indicator, the neighboring node shall inform its neighbors about the failure and release and re-initiate any PDN connection associated with failing node." [30]		

Fault Management	**32.111-1**	*"After a fault has been detected and the replaceable faulty units have been identified, some management functions are necessary in order to perform system recovery and/or restoration, either automatically by the NE and/or the EM, or manually by the operator. If the faulty resource has no redundancy, the recovery actions shall be:* *a) Isolate and remove from service the faulty resource so that it cannot disturb other working resources;* *b) Remove from service the physical and functional resources (if any) which are dependent on the faulty one. This prevents the propagation of the fault effects to other fault-free resources;* *c) State management related activities for the faulty resource and other affected/dependent resources."* [33]		
Self-Organizing Networks (SON); Self-healing concepts and requirements	32541	*"In the case of **software faults**, the recovery actions may be :* *a) system initializations (at different levels),* *b) reload of a backup of software,* *c) activation of a fallback software load,* *d) download of a software unit,* *e) reconfiguration, etc.* *In the case of **hardware faults**, the same as line of fault management above plus this:* *a) **Reset** the faulty resource;* *b) Other reconfiguration actions**, etc.* *If the faulty resource has redundancy, the recovery action shall be changeover.* ***Here we see that reconfiguration is something proposed by 3GPP but not a "must have" attribute."* [31]		

3GPP Title	**3GPP Doc Num**	**3GPP Network Survivability Requirements related to failure Resistance**		
		Node Level	**System Level**	**Intersystem Level**
(UMTS); protocol description and error handling	25.921	*"The error handling shall be specified in the protocol for the cases when the requirement for presence or absence of an IE indicated by the condition is not followed."* [34]		
Technical Specification Group Services and System Aspects;	23401	*"SGW-MME / SGW-PGW* ***GTP-C Load Control*** *feature is an optional feature which allows a GTP control plane node to send its Load Control Information to a peer GTP control plane node which the receiving GTP control plane peer node uses to augment existing GW selection procedure"* [21]	***"APN level load control*** *may be supported and activated in the network. If this feature is activated, the PDN GW may convey the Load Control Information at APN level (reflecting the operating status of the resources at the APN level), besides at node level."* [21]	
		"SGW-MME / SGW-PGW ***GTP-C Overload Contro****l feature is an optional feature. Nodes using GTP control plane signaling may support communication of Overload Control Information in order to mitigate overload situation for the overloaded node through actions taken by the peer node(s)"* [21]	***"NAS Level Congestion control****: The MME may detect the NAS signaling congestion associated with the APN and start and stop performing the APN based congestion control based on criteria: (max number of EPS bearers and EPS bearer activation per APN, one or multiple PDN GWs of an APN are not reachable or indicated congestion to the MME, Maximum rate of MM signaling requests associated with the devices with a particular subscribed APN, Setting in network management)"* [21]	
		"MME-Enb The ***MME Load Balancing*** *functionality permits UEs that are entering into an MME Pool Area to be directed to an appropriate MME in a manner that achieves load balancing between MMEs".* [21]	*"PDN* ***GW control of overload*** *by rejection of PDN connection requests from UE."* [21]	
		"*MME-Enb The* ***MME Load Re-balancing*** *functionality permits UEs that are registered on an MME (within an MME Pool Area) to be moved to another MME*" [21]		

		"MME *The MME shall contain mechanisms for avoiding and handling overload situations"* [21]		
		"SGW-MME ***Throttling of Downlink Data Notification*** *Requests. MME may restrict the signaling load that its SGWs are generating on it, if configured to do so."* [21]		
		"MME-UE *UE Level* ***NAS congestion****: The MME may detect the NAS signaling congestion associated with the UEs belonging to a particular group. The MME may start and stop performing the group specific NAS level congestion control based on criteria (maximum rate of MM and SM signaling requests associated with the devices of a particular group, Setting in network management)"* [21]		
Configuration Management (CM);	32.600	*"Configuration Management (CM), in general, provides the operator with the ability to assure correct and effective operation of the PLMN network as it evolves. CM actions have the objective to control and monitor the actual configuration on the Network Elements (NEs) and network resources, and they may be initiated by the operator or by functions in the Operations Systems (OSs) or NEs. CM actions may be requested as part of an implementation program (e.g. additions and deletions), as part of an optimization program (e.g. modifications), and to* ***maintain the overall Quality of Service (QoS)****. The CM actions are initiated either as single actions on single NEs of the PLMN network, or as part of a complex procedure involving actions on many resources/objects in one or several NEs."* [35]		

3.2. Design and Implementation

After requirements specification, design and implementation phases follow which are not worth analysing further since they are organization specific. **Robust and secure code design** techniques should be part of this phase. Additionally, **risks related to survivability** should be part of risk assessment which is usually conducted through the design phase.

3.3. Testing or Evaluation of System's Survivability

To continue, the testing phase of the proposed SDLC is presented. Testing is the way to evaluate a system's survivability. Testing phase should also follow the same model and test cases should be designed for node, system and intersystem level. In this way the whole system will be tested each time. Additionally, test cases should include tests against services' correct functionality, and they should be extended to also test any resistance, recognition or recovery survivability requirement to all testing levels (node, system, intersystem). For this to be achieved test-driven development is the most appropriate approach. Modern SDLC approaches are test-driven which is what is also proposed for the current SDLC.

Test-driven means that the tests are designed according to the requirements and are constructed even before the development of new features or maintenance tasks like bug fixing. Additionally, through this work we propose another approach that is related to test-driven development and has to do with failure impact evaluation. In other words, testing may be also used to evaluate the impact of any failure to critical services, and having this information available, new tasks may be extracted for the next iteration cycle regarding failure recognition, resistance or recovery. So, in this case tests are indeed driving the development and are a tool to discover many issues that may occur from any combination of services. So, any time a new service is to be

developed or updated, testing any possible combination of it with critical services will reveal any threats to critical services from the newly inserted code.

Impact analysis could be applied in any iteration of SDLC providing new requirements related to survivability requirements. Tests related to impact analysis may be:

1. Executing critical services before and after newly developed or modified service.

2. Executing critical services after failure of newly developed or modified service.

3. Executing critical services in collision with newly developed or modified service.

Additionally, another proposal is to test all survivability requirements for each new or modified functionality. So apart from just testing failure scenarios, recognition of failure and recovery from failure or resistance to failure should be also tested in order testing procedure to be considered complete.

All tests related to survivability evaluation and corresponding test approaches that could be used, are depicted in the following table (4) below. Test scenarios are also related to corresponding threat to survivability and impact of realization of this theat. Finally, any test case should be added to regression testing in order to ensure that future changes will not affect the existing functionality.

Table 3: Evaluation of Survivability Requirements through testing.

Survivability Threats		**Root Cause of Failure**	**Failure Impact in Node Level**	**Test Scenarios**	**Testing Methods**	**Examples from 4G network**
3GPP Fault Management; 32.111-1 Categories of faults for which an NE (network element) may	**Hardware failures**, i.e. the malfunction of some physical resource within a NE.	Device damage	Messages sent from one node to neighboring node may not be answered.	Testing of scenario where device is forced out - no information of the event to management system. The NE should be able to track the issue and report to management system. The impact from this failure to service under development and the restoration time should be defined.	Functional testing	Unplug of the device.
				Testing of scenario where device fails and sends alarm to management system. Service under development should be released or served by alternative resources after system re-configuration.	Functional testing	Enforcement of the NE to send a failure alarm to the management system
		CPU / Memory Overload		Testing of scenarios that		

raise alarms are:		System missconfiguration	Messages sent from one node to neighboring node may not be answered or answered with delay.	the message is not answered from neighboring NE in all phases of service establishment and test requirements related to handing of this situation.	Functional testing Unit or Module Testing Static Analysis,	Test scenarios where message is not answered.
			Faulty messages may arrive to NEs.	Testing of scenarios that the message arrives with wrong configuration information.	Functional testing Unit or Module testing Static Analysis Fuzzy-testing Fault-injection testing	Test message with wrong information about MMEs capability of supporting IOT devices.
	Software problems, e.g. software bugs, database inconsistencies	Any S/W bug that results in wrong functionality of service or non-compliance with standards	Service rejection or faulty service establishment.	Test-driven development with tests that are designed due to 3GPP standards requirements.	Functional testing Unit or Module Testing Static Analysis Fault-injection testing	Test all scenarios that reflect 3GPP requirements.
		S/W Bug lead to hanging processes	Future service requests may be rejected.	Enforce processes to be hanged and see if system reacts according to requirements. Test critical services impact if attempted.	Functional testing Unit or Module Testing Static Analysis,	Test if after deletion of a voice bearer it can re-established.
3GPP Fault Management; 32.111-1		Missing of robustness measurements like handling collision scenarios or handling of wrong Information Elements in messages	Service rejection or faulty service establishment.	Testing of all possible collision combination, especially with critical services, and test scenarios during which messages have wrong IEs that could be	Functional testing Unit or Module Testing Static Analysis,	PDN connection consists of a series of messages. A test case could include the modification of bearer id to a wrong one and see

				handled by robustness measurements.		if system is robust enough to handle this error.
		S/W bug that may lead to unanswered messages	Service rejection.	Testing of scenarios where messages of process under development are not answered.	Functional testing Unit or Module Testing Static Analysis,	A test case could be the PDN establishment and testing if service is properly rejected.
				Test the impact to critical services. Test cases with critical services already established and the above scenario following should be tested. The opposite is also valid scenario and should be tested. In this case failures from hanging processes will also be tested.	Functional testing Unit or Module Testing Static Analysis,	Testing of the above scenario after and before voice bearer handover.
		S/W bug that may lead to message sent twice	Service may be re-established if there is no mechanism for ignoring repeated messages	Testing of scenarios where messages of service under development are sent twice.	Functional testing Unit or Module Testing Static Analysis,	A test case could be sending PDN request twice for the same bearer.
	Functional faults, i.e. a failure of some functional resource in a NE and no hardware component can be found responsible for the problem.	Any other failure that may lead to service unavailability	Service Failure	Any related test	Functional testing Unit or Module Testing Static Analysis,	Any related test

	Loss of some or all of the NE's specified capability due to overload situations.	System Overload of requests	Messages sent from one node to neighboring node may not be answered or answered with delay.	Testing of service impact after increasing system load. Testing service impact after increasing load of service.	Stress testing Load testing Stability testing	Try to establish a voice bearer in a loaded system and an overloaded system. And try to see the impact to the system and voice bearer when system is loaded by voice bearer requests.
	Communication failures between two NEs, or between NE and OS, or between two OSs.	S/W failures, H/W failure, Overload situations, Path / Link failures, Network timing issues.	Messages sent from one node to neighboring node may not be answered or answered with delay.	Testing of scenarios of scenarios that the message is not answered from neighboring NE in all phases of service establishment and test requirements related to handing of this situation.	Functional testing Unit or Module Testing Static Analysis,	Testing of scenarios of scenarios that the message is not answered from neighboring NE.
Security Testing	Any security threat should be considered and tested. Details on security testing will not be provided to current document.					
Failure Recognition	In all possible errors, network management system should be tested. Network management system should be informed about any kind failure and should be able to trigger system resistance or recovery mechanisms. So, any NE that is under development should be tested against this functionality also.					
System Recovery	In all possible failure scenarios, recovery mechanisms following should also be tested.					

4. Conclusions and Future Work

To sum up, during the current paper, a development framework of a survivable mobile telecommunication system, based on system's mission and critical services, has been presented and proposed. This framework was based on the available survivability approaches through literature review with its main contribution to be that it provides a solution that is more focussed on interconnection and interoperation of systems forming larger intersystem. By this any survivability requirement from any level of service is considered through everyday development work and the focus is not only based on correct system functionality. Additionally, by this any interoperability and interconnection requirements and threats related to survivability may be examined through development life cycle.

Contrary to other approaches for evaluation of survivability, the one proposed is a more practical guide for testing the critical services of systems and evaluating measurements correlated to survivability of (inter)system, end to end from the requirements specification phase of the system and it does not only focus on node or link failure as most of proposals of literature review. This approach has been adopted because survivability is a built-on and not an add-on characteristic.

To sum up, the major outcomes of the current research are:

- The current research improves the traditional SDLC process, by enriching requirements analysis and testing phases with approaches related to survivability. The resulting proposed methodology is the Survivability Software Development Lifecycle presented in Chapter 3 that may be applied to telecommunication systems.
- The current research provides a systematic approach for handling the complexity arising from the interconnection of different network nodes of a telecommunication system.

Finally, as future work, we are planning to apply the proposed methodology in order gather and analyse metrics related to overall system survivability.

Conflict of Interest

The authors declare no conflict of interest

Acknowledgment

This work has been partly supported by the University of Piraeus Research Center.

References

[1] R.J. Ellison, D.A. Fisher, R.C. Linger, H.F. Lipson, T.A. Longstaff, N.R. Mead, "Survivable network systems: An emerging discipline", Carnegie Mellon University Digital Library, 1997, doi:10.21236/ada341963

[2] R.C. Linger, et al., "Life-Cycle models for survivable systems", Carnegie Mellon University Digital Library, 2002, doi: 10.1184/R1/6575138.v1

[3] V. R. Westmark, "A definition for information system survivability", in 37th Annual Hawaii International Conference on System Sciences, 10-19, 2004 doi: 10.1109/HICSS.2004.1265710.

[4] R.J. Ellison, "Survivability analysis framework", Carnegie Mellon University Digital Library. 2010, doi: 10.1184/R1/6584474.v1

[5] J. C. Knight, E. A. Strunk and K. J. Sullivan, "Towards a rigorous definition of information system survivability" in DARPA Information Survivability Conference and Exposition, 78-89, 2003, doi: 10.1109/DISCEX.2003.1194874.

[6] P. Pal, "Survival by defense – enabling" , in 2001 workshop on New security paradigms, 71–78, 2001, doi: 10.1145/508171.508183

[7] J. Knight, Dennis Heimbigner , Alexander Wolf, Antoinio Carzaniga, Jonathan Hill, Premkumar Devanbu, Michael Gertz, "The WILLOW survivability architecture", in Fourth Information Survivability Workshop, 2001, doi: 10.1.1.96.5316

[8] W. Li, L. Shu and Y. Feng, "A Dynamic Survivability reconfiguration framework based on QoS", in 2009 International Conference on Advanced Computer Control, 103-106, 2009, doi: 10.1109/ICACC.2009.107.

[9] D. Chen et al., "Network survivability performance evaluation: a quantitative approach with applications in wireless ad-hoc networks", in 5th International Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems, 61-68, 2002, doi: 10.1145/570758.570769.

[10] A. H. Wang, S. Yan and P. Liu, "A semi-Markov survivability evaluation model for intrusion tolerant database systems," in 2010 International Conference on Availability, Reliability and Security, 104-111, 2010, doi: 10.1109/ARES.2010.90..

[11] A.J. Mohammad, "Towards quantifying metrics for resilient and survivable networks", in 14th IEEE International Conference on Network Protocols (ICNP 2006), 2006, doi: 10.1.1.1.6900

[12] L.Y. Trivedi, "A general framework for network survivability quantification.", in 12th GI/ITG Conference on Measuring, Modeling, and Evaluation of Computer and Communication Systems, 369-378, 2004, doi: 10.1.1.94.3404

[13] M. Liang et al., "A novel method for survivability test based on end nodes in large scale network", KSII Transactions On Internet Ans Information Systems 9(2), 620-636, 2015, doi:10.3837/tiis.2015.02.008.

[14] C. Wang et al., "A general framework for network survivability testing and evaluation", Journal of Networks **6**(6), 831-841, 2004, doi: 10.4304/jnw.6.6.831-841

[15] M. Liang et al., "Research on survivability metrics based on survivable process of network system", in 4th international conference on security of information and networks, 247-250, 2011, doi: 10.1145/2070425.2070470 .

[16] L. Zhang, W. Wang, L. Guo, W. Yang and Y. Yang, "A survivability quantitative analysis model for network system based on attack graph", in 2007 International Conference on Machine Learning and Cybernetics, 3211-3216, 2007, doi: 10.1109/ICMLC.2007.4370701.

[17] D.W. Tipper et al., "Survivability analysis for mobile cellular networks", in Communication Networks and Distributed Systems Modeling and Simulation Conference, 2731-2738, 2002, doi: 10.1.1.361.411

[18] M. C. Baker, C. A. Witschorik, J. C. Tuch, W. Hagey-Espie and V. B. Mendiratta, "Architectures and disaster recovery strategies for survivable telecommunications services", Bell Labs Technical Journal, **9**(2), 125-145, 2004, doi: 10.1002/bltj.20030.

[19] S. Dharmaraja, V. Jindal and U. Varshney, "Reliability and survivability analysis for UMTS networks: an analytical approach", IEEE Transactions on Network and Service Management, **5**(3), 132-142, 2008, doi: 10.1109/TNSM.2009.031101. 8

[20] M. Mykoniati et al., "Self-Diagnosis framework for mobile network services", JACN, **7**(2), 2019, doi: 10.18178/JACN.2019.7.2.268

[21] "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", Version 16, Technical specification 3GPP org The Mobile Broadband Standard 23401, 2021.

[22] M. Mykoniati et al., Lambrinoudakis, "Fault prediction model for node selection function of mobile networks", in 9th International Conference on Information Communication and Management ICICM 2019, 153-159, 2019, doi: doi.org/10.1145/3357419.3357452.

[23] "Requirements for the UMTS Terrestrial Radio Access (UTRA) system", Version 3.0.1, Technical specification 3GPP org The Mobile Broadband Standard, 21.01U, 1997

[24] "Telecommunication management;Performance Management (PM);Concept and requirements", Version 16.0.0., Technical specification 3GPP org The Mobile Broadband Standard 32401, 2020

[25] "Service aspects; Service principle", Version 18.1.1, Technical specification 3GPP org The Mobile Broadband Standard 22101, 2021

[26] "3G security; Security threats and requirements", Version 4.1.0, Technical specification 3GPP org The Mobile Broadband Standard 21133, 2002

[27] "Security objectives and principles", Version 4.0.0, Technical specification 3GPP org The Mobile Broadband Standard 33120, 2001

[28] "System Architecture Evolution (SAE); Security architecture", Version 16.3.0, Technical specification 3GPP org The Mobile Broadband Standard 33.401, 2020

[29] "Security architecture and procedures for 5G System", Version 17.2.0, Technical specification 3GPP org The Mobile Broadband Standard 33.501, 2021

[30] "Restoration procedures", Version 17.1.0, Technical specification 3GPP org The Mobile Broadband Standard 23.007, 2021

[31] "Telecommunication management; Self-Organizing Networks (SON); Self-healing concepts and requirements", Version 16.0.0, Technical specification 3GPP org The Mobile Broadband Standard 23.541, 2020

[32] "Telecommunication management; Principles and high level requirements", Version 16.0.0, Technical specification 3GPP org The Mobile Broadband Standard 32101, 2020

[33] "Telecommunication management; Fault Management; Part 1: 3G fault management requirements", Version 16.0.0, Technical specification 3GPP org The Mobile Broadband Standard 32111-1, 2020

[34] "Guidelines and principles for protocol description and error handling", Version 7.0.0, Technical specification 3GPP org The Mobile Broadband Standard 25.921, 2007

[35] "Telecommunication management; Configuration Management (CM); Concept and high-level requirements", Version 16.0.0, Technical specification 3GPP org The Mobile Broadband Standard 32600, 2020

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 267-272 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Low-Power Primary Cell with Water-Based Electrolyte for Powering of Wireless Sensors

Dmitry Petrov*, Ulrich Hilleringmann

Sensor Technology Group, Paderborn University, Paderborn, 33098, Germany

ARTICLE INFO

Article history:
Received: 31 July, 2021
Accepted: 18 September, 2021
Online: 14 October, 2021

Keywords:
Power Supply
Wireless Sensors
Primary Cell
Water Detection
Water as Electrolyte
Voltage Converter
Maximal Power Point

ABSTRACT

In this work we discuss a special technique for powering of sensor systems, based on a low-power primary cell consisting of two electrodes, made from different metals, with water (lake, rain or tap water) used as an electrolyte. Once placed into an aqueous solution, the primary cell generates a small electric current, which may be utilized for powering of sensor systems. The generated electrical energy is fed into an energy storage (capacitor). After transformation of the voltage by a step-up converter, it is used for supplying the electrical sensor circuit. The expected output power of the developed circuit is 10-15 mA by 2 V output voltage during 0.2-0.5 second. The improved voltage converter topology with implemented maximal power point techniques allows significant reduction of the energy storage's size in the second revision of the circuit and thus reduction of the resulting size of the board. The implemented sensor board with discussed powering technique, assembled in Paderborn University was already tested in different practical scenarios.

1. Introduction

This paper is an extension of work originally presented in 2020 IEEE Sensors [1]. Energy harvesting technologies are often used for powering of wireless sensors, especially for industrial monitoring applications [2]. Powering techniques like photovoltaic and thermoelectric generators are widely presented on the market, but there is always a need for new ideas for energy powering techniques [3]. The Sensor Technology Group at Paderborn University is directly developing the power supply system for wireless sensors based on a galvanic cell used for generation of small electrical currents with a water-based electrolyte (lake, rain or tap water) [1, 4].

The wireless sensors powered by the developed electrochemical cell can be used e.g., for construction monitoring and detection of water leakages in roofs or wall structures of industrial and private buildings. Such sensors integrated in building structures can get activated by contact with water, caused by a water leakage. This water serves as an electrolyte for the galvanic cell, which provides the electrical energy for the sensor electronics. The main advantage of this approach is a long sensor lifetime in the not activated state, longer than most batteries' durations that are currently available on the market (usually up to 10 years).

2. Implementation of a galvanic cell

The implemented galvanic cell with two metal electrodes mounted on the non-conductive substrate (composite epoxy material) is depicted in Figure 1, left. To perform the measurements, the developed cell is placed inside the vessel with tap water serving as an electrolyte (Figure 1, right).

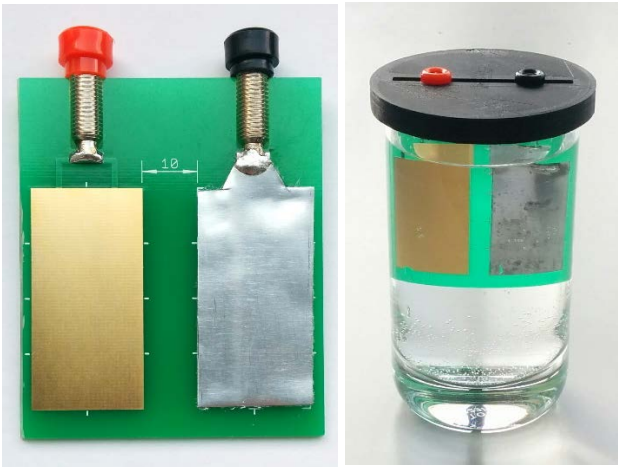


Figure 1: The implemented setup of a galvanic cell with two metal electrodes (left), and galvanic cell placed in water, serving as electrolyte (right)

*Corresponding Author: Dmitry Petrov, Email: petrov@sensorik.uni-paderborn.de

The difference of electric potential between two electrodes results from the standard potential of the chosen electrode metals [5-7].

During the experimental measurements, multiple combinations of electrode materials have been compared and tested. The practically established parameters of the galvanic cells (short-circuit current and open-circuit voltage) are depicted in Figure 2.

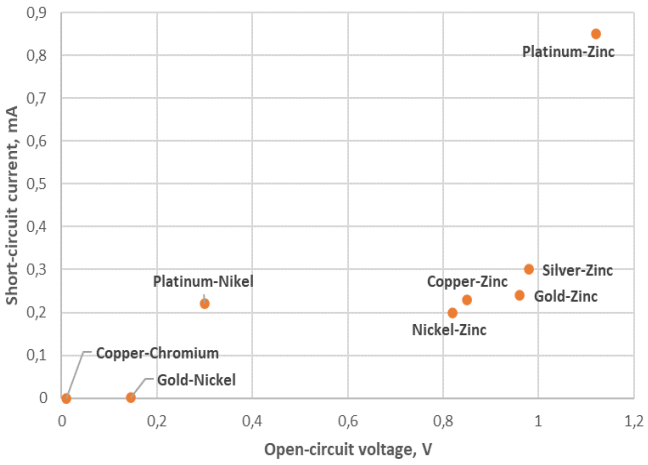


Figure 2: Established parameters for different metals, used for galvanic cell electrodes with electrode areas of 8 cm2 (4 x 2 cm), in tap water [1]

Table 1 shows the standard electrochemical potentials of commonly used electrode materials [5].

Table 1: Standard electrochemical potential of metals commonly used as electrode materials

Electrode material	Reaction pair	Standard potential, V
Gold	$Au^{3+} + 3e^-$	1.5
Gold	$Au^{2+} + 2e^-$	1.69
Platinum	$Pt^{2+} + 2e^-$	1.2
Silver	$Ag^+ + e^-$	0.80
Copper	$Cu^+ + e^-$	0.52
Zinc	$Zn^{2+} + 2e^-$	-0.76

Although the pair platinum and zinc results in the highest value of output power, a combination of zinc and gold metals has been selected for the developed galvanic cell [8], because gold finish is a standard coating technique in the production of printed circuit boards (PCB). A gold layer of 100-200 nm is typically used to protect the contact pads of the PCB against corrosion and simplifies the soldering process [9]. An adequate gold thickness allows the stable operation of the implemented cell, although, any defect in the gold layer (like fracture or porosity) reduces the generated power by reducing the open-circuit voltage.

Figure 3 shows the established performance of the implemented cell with gold and zinc electrodes of 8 cm^2 area (4 x 2 cm), loaded to a capacitance of 50 mF. The distance between the electrodes was 10 mm. The calculated energy saved in the capacitor achieved 9,8 mJ after 5 minutes and about 11 mJ in 10 minutes.

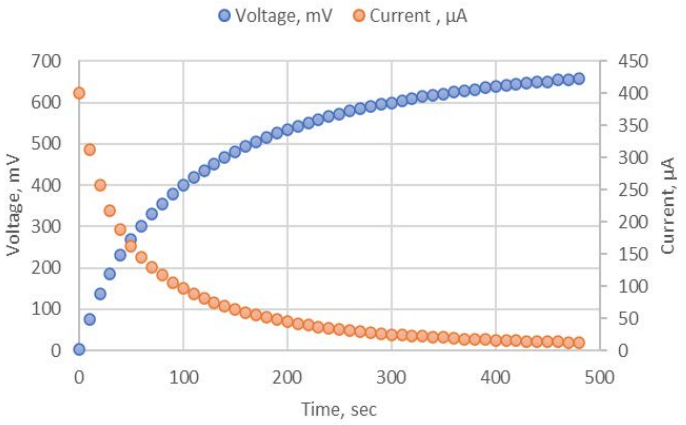


Figure 3: Current and voltage curves charging a capacitance of 50 mF [1]

Figure 4 depicts the load curve of the developed galvanic cell. The established open-circuit voltage is around 700 mV, the corresponding maximum power point (MPP) is about 250 mV. The calculated generated power achieves up to 44 µW, if the galvanic cell operates close to the MPP voltage.

The performance of the tested galvanic cell working on the capacitive load (shown in Figure 3) could be improved by connecting a voltage converter working with the established maximum power point between the output of the galvanic cell and charged capacitance. This design of the voltage converter will be discussed in detail in chapter 4.

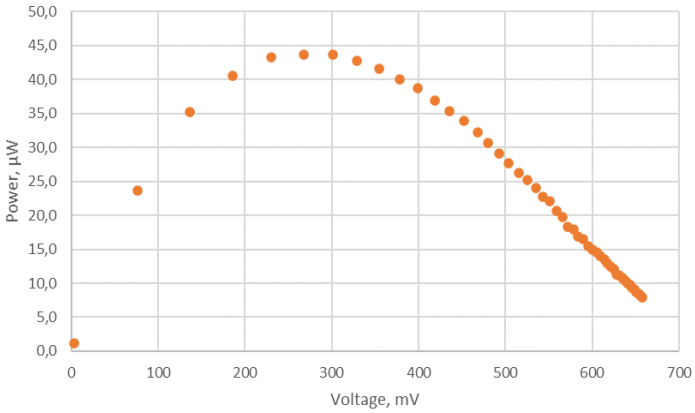


Figure 4: The load curve of the galvanic cell consisting of Au and Zn.

3. Sensor board design

The selected combination of gold and zinc electrodes connected by water ensures an open circuit voltage of about 700 mV. This voltage is lower than the operating voltage of typical electronic circuits (from 1.8 V to 2.0 V for common low-voltage microchips). The maximal output current of the developed galvanic cell is between 100 and 200 µA (short-circuit current), this is much lower than the typical current consumption of wireless transmitter ICs (10 to 15 mA). Therefore, a step-up voltage converter circuit was implemented to boost the output cell voltage. The harvested energy has to be integrated and must be stored temporarily in an energy storage circuit, e.g., a capacitor or super capacitor, until the energy amount is high enough to power the sensor circuit for the required time (0.2 to 0.5 seconds).

The first revision of the developed sensor board uses a super capacitor as an energy storage component. It is connected between the output of the galvanic cell and the input of the voltage boost circuit. The block diagram of this first revision is shown in Figure 5. After charging the super capacitor up to 700 mV (threshold voltage), the chosen step-up converter is starting up. It boosts up performs the voltage to 2.0 V to supply the electrical circuit of the sensor board. After complete discharging of the super capacitor, the charging cycle will be repeated. The low-voltage synchronous boost converter, TPS61201 from Texas Instruments [10], was implemented in the first revision of the developed sensor board.

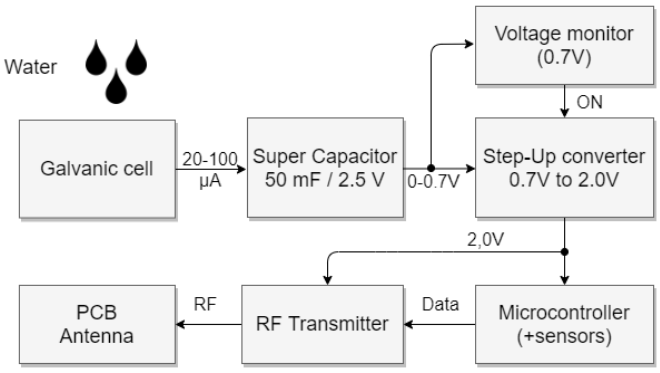


Figure 5: Block diagram of the developed circuit board with energy storage connected to the input of the voltage converter

The required capacitance of the super capacitor is between 20 and 50 mF, which is enough to deliver the estimated current of 10-15 mA over a time interval of 0.2 to 0.5 seconds. For the super capacitor, a model produced by AVX company with a capacity of 50 mF was used in the voltage converter circuit [11]. To supervise the charging process of the super capacitor, a discrete low-voltage detector by RICOH with 0.7 V threshold voltage is used [12].

The first revision of the sensor PCB (with a size of 27 mm x 100 mm) is depict in Figure 6. Two metal electrodes, each at a size of 8 cm^2, are located on the bottom side of the board. The electronic components are placed on the upper side of the PCB. The electrical contact between the gold coated positive electrode and the electrical circuit is performed by a through-hole connection (the black point at the top right corner of the left electrode). The negative (zinc) electrode is cut out of a zinc plate, the mechanical and electrical connections of the zinc electrode to the upper side of the PCB is realized by two flaps, inserted into board openings and connected to the contact pads on the top side. The openings are sealed with silicone to prevent the electrical contact of the sensor circuit with the electrolyte.

On the top side of the developed sensor board the supercapacitor is located on the left, the voltage converter and transmitter circuit in the middle of the board, and the RF antenna on the right side. A complete sealing of the electrical circuit (the complete top side of the board) could also be implemented for practical applications to protect the circuit from the moisture and dust.

The total component cost for the first revision of the developed sensor PCB is about 9.0 € (calculated for 1000 pieces).

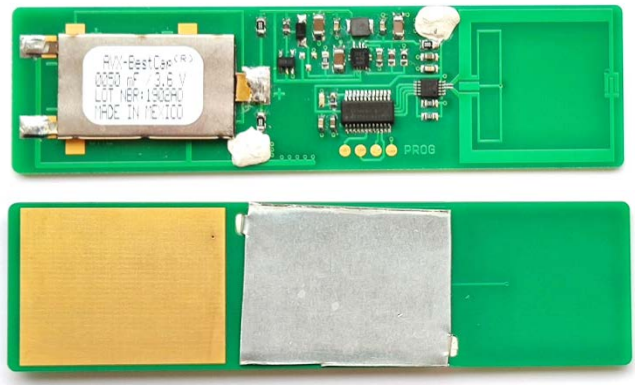


Figure 6: Design of the first revision of the sensor board (top and bottom side)

4. Optimized sensor board design

The minimal size of the first revision of the developed board is defined by the size of the super capacitor (energy storage), which occupies about 30% of the board area. The new improved topology of the power supply circuit which is currently implemented, uses the maximum power point (MPP) tracking technique. It allows an operation of the galvanic cell close to the point of maximal efficiency. In this topology the energy storage is placed between the output of the step-up converter and the sensor electronics. Here the storage capacitor can be charged to a higher voltage level (3.3 V instead of 0.7 V) giving the possibility to save considerably more electrical energy, thus reducing the required capacitance. This allows less expensive alternatives to the supercapacitor, like tantalum capacitors. By using the new topology, the sensor PCB area may be reduced by about 30%. The block diagram of the second revision of the sensor board with implemented MPP tracking is depict in Figure 7.

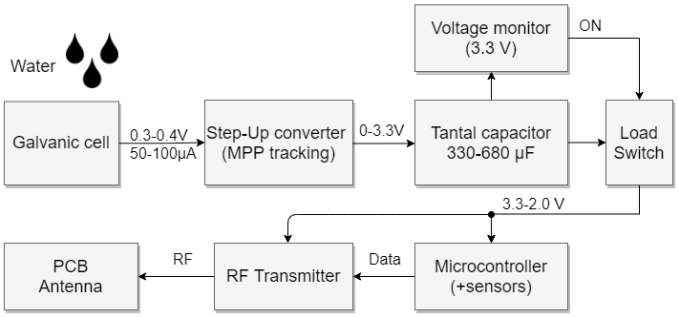


Figure 7: Block diagram of the sensor board with energy storage at the output of the voltage converter and implemented MPP tracking

Using the selected ADP5091 voltage converter with implemented MPP tracking by Analog Devices [13] increases the output power of the galvanic cell and significantly reduces the charging time (from 8-10 minutes to 3-4 minutes). Further optimisation of the electrode sizes and relation between gold and zinc area help to reduce the total size of the sensor board from 100 mm x 27 mm to 75 mm x 27 mm (reduction by 25%). The component costs in this version were also reduced by about 30%. The total costs for the second revision of the sensor PCB are about 6.5 € (from 1000 pieces).

The second revision of the sensor board is shown in Figure 8. The tantalum capacitors for storing the generated electrical energy (two orange blocks on the top side of the PCB) occupy only a small

part of the sensor board area. The voltage converter is located on the left and the RF antenna with the transmitter circuit on the right side of the sensor PCB.

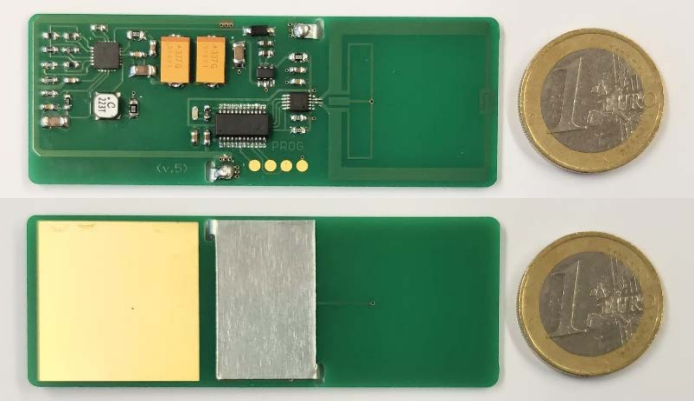


Figure 8: Design of the second revision of the sensor board (top and bottom side)

Both, gold and zinc electrodes are placed on the bottom side of the board. The optimized relation between gold and zinc areas (up to 1:2) allows to significantly reduce the electrode area. The size of the gold electrode was reduced from 25 x 32 mm to 25 x 25 mm (reduction by 22%) and the size of the zinc electrode was reduced from 25 x 32 mm to 25 x 14 mm (reduction by 56%).

5. Wireless data transmission

The amount of collected electrical energy should be high enough to perform the initialization of the EFM8 low-power MCU by Silicon Labs [14], the measurement tasks, and additionally the wireless transmission of the data packages via Si4012 RF transmitter [15]. The power consumption of the RF transmitter circuit is most significant compared to other parts of the circuit. Additionally, the current consumption changes very quickly, depending of the data transmission cycle. Knowing the current profile of the RF transmitter helps to design the power supply circuit and to calculate the required amount of electrical energy for the sensor operation. The measured real-time current consumption of the transmitter circuit is shown in Figure 9.

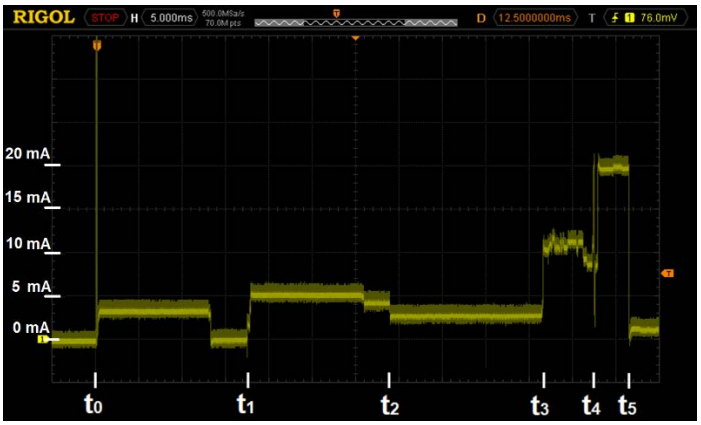


Figure 9: Current consumption of the RF transmitter with the time scale according to the Table 2

This measurement was realized by measuring the voltage drop over a shunt resistor, placed in the supply line of the transmitter circuit.

The important time intervals of the data transmission are shown in Table 2. The calculated energy value is 0.7 mJ for start-up of the transmitter module and the first transmitted data packet, and 0.3 mJ for following data packets. Between transmission of data packets, the transmitter IC goes in a power down mode, reducing the current consumption down to few microamperes.

Table 2: Measured Current consumption of RF transmitter

Time interval	Task	Current, mA	Power, mW
t_0 - t_1	Power-on	3.5	7.0
t_1 - t_2	Initialize	5.0	10.0
t_2 – t_3	Compose data package	4.0	8.0
t_3 – t_4	Initialize transmission	10.0	20.0
t_4 – t_5	Data transmission	19.0	38.0

Sending of multiple sequential packages reduces the probability of a data loss (e.g., by interfering of a single packet with electromagnetic noise spikes). However, the multiple data packets increase the energy budget of the sensor system. In the first revision of the sensor board, the number of repetitions counts was 4 to 6 data packages, depending on the charging state of the super capacitor.

The implemented randomized time intervals between the individual packages, shown in Figure 10, should additionally increase the safety of the data transmission.

In the second revision of the sensor board a number of 2 to 3 transmitted packets was implemented and successfully tested, which helps to reduce the required energy budget of the sensor circuit almost by 50% (considering the power consumption of the MCU, not affected by reducing the number of transmitted packets).

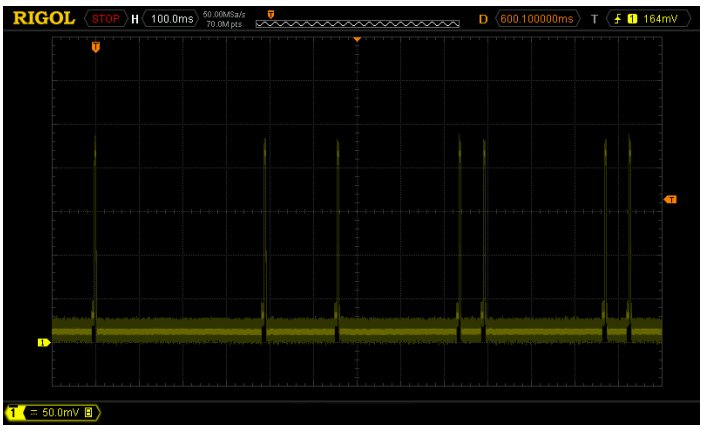


Figure 10: Multiple data packets with randomized time intervals

To keep the time of a single data transmission as short as possible, the length of the data package was also decreased to a minimum: the package consists out of two bytes preamble, four bytes of payload and additional two bytes for the control sum (CRC-16), leading to a total length of the data package of only 8 bytes (Table 3). Since the implemented sensors operate as simple water detectors with transmitting only their unique ID number, the payload bytes are currently holding the four bytes of the sensor ID. For other sensor implementations for measuring and transmitting of additional parameters the packet length needs to be adjust correspondingly.

Table 3: Transmitted data packet

Byte 1	Byte 2	Byte 3	Byte 4	Byte 5	Byte 6	Byte 7	Byte 8
Preamble	Preamble	ID-1	ID-2	ID-3	ID-4	CRC-16	CRC-16

6. Receiver board

To receive and evaluate the sensor data, a receiver board based on the Si4355 866 MHz RF receiver module by Silicon Labs [16] was developed. The block diagram of the receiver board is shown in Figure 11.

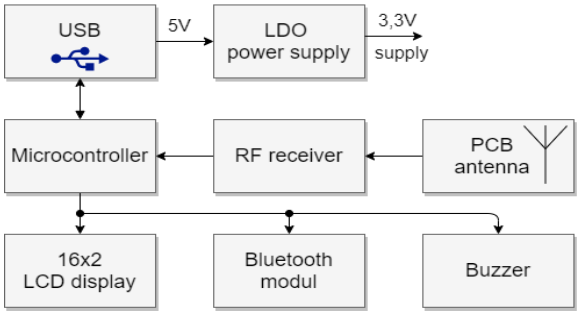


Figure 11: Block diagram of the receiver board [1]

The received data packets are decoded by the RF receiver module, the payload data are processed by the MCU and send via USB interface to the external system. The received data packets are also shown on the built-in LCD display, additionally the buzzer tone signalizes the receipt of a new packet. The received packets can also be transmitted wirelessly via Bluetooth interface to a tablet or a mobile phone. The developed receiver board is powered via USB interface from the host device.

The photo of the developed receiver board is depicted in Figure 12. The USB interface is located on the left side of the board, the receiver IC and the RF antenna are placed on the right side of the main PCB, separated from the main part of the electrical circuit and having their own ground area. The pin-header on the top side of the PCB is used to connect the external LCD display, in the middle the FT232RL USB to UART interface by FTDI Chip [17] and the power supply circuit are located.

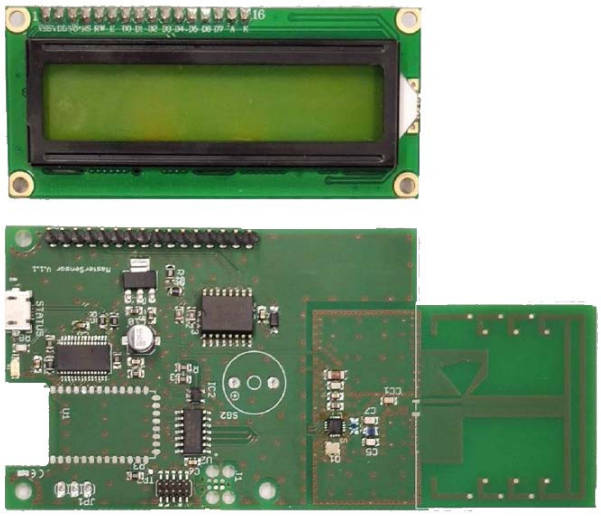


Figure 12: Receiver board with external LCD display module

The receiver board is controlled by an ATtiny1614 microcontroller by Microchip [18]. To address all the implemented interfaces an additional 16-bit port expander IC is used [19]. The Bluetooth module [20] is placed on the bottom left side of the PCB (not assembled in the current revision of the board).

The connected and powered receiver board is shown in Figure 13. The LCD display is mounted above the receiver circuit, covering most electronic components. The connected USB cable also powers the receiver board from the host PC. The LCD display shows the programmed firmware revision in the first line and the last received data packet (sensor ID number) in the second line.



Figure 13: Assembled and powered receiver board showing the last received data packet on the LCD display

The measured transmission range was up to 150 meters outdoors and 30 to 40 meters indoors. According to the orientation of transmitter and receiver antennas, the transmission range may deviate by 20-30%.

A bidirectional communication between sensors and receiver board has not been implemented in the current revision of the sensor system, caused by the small energy budget and limited by the size of the sensor board.

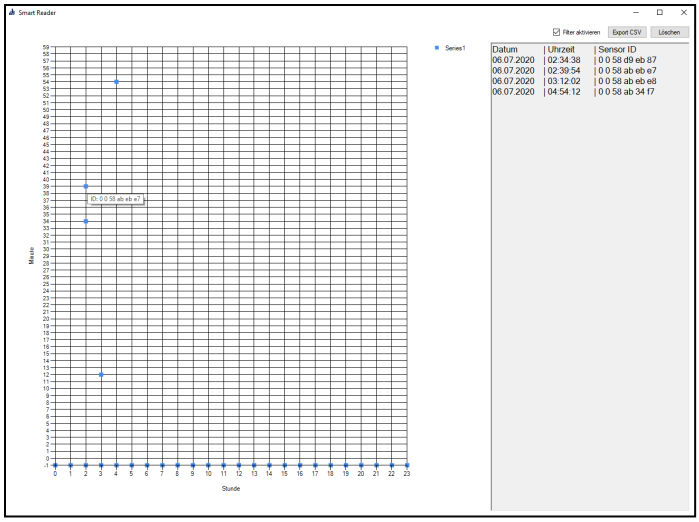


Figure 14: Developed sensor mapping software for time-based representation of activated sensors

7. Developed PC software

For some practical applications (e.g., construction monitoring and detection of water leakages), a graphical depiction of sensor mapping with a time-based representation of activated sensors may help to establish the starting point for water leakages inside the

monitored building. An example of the developed software with graphical representation of activated sensors is shown in Figure 14.

On the left side of the display the last activated sensors are shown with the blue points. The list of activated sensors including the sensor ID with a corresponding time stamp is shown on the right side of the window. The sensor mapping feature with graphical representation of the physical sensor placement inside the monitored area is not implemented in the current revision of the software; it will be developed in the next revision.

8. Conclusions

The developed system for powering of wireless sensor systems uses a build-in galvanic cell. It may be used in different scenarios, where the use of conventional batteries is not desired or not possible. Another advantage compared to battery powered systems is a significantly longer storage time of the galvanic cell.

After initiating the galvanic cell, that means both cell electrodes get in contact with water, the lifetime of the cell will be limited by corrosion of the negative (zinc) electrode. The confirmed lifetime can take some days to some weeks, depending on the application and the type of used electrolyte. The confirmed lifetime makes possible the implementation of different scenarios, like water detectors, one-way sensors for environment parameter control etc.

The first revision of the sensor PCB with a size of 27 x 100 mm assembled at Paderborn University was already tested in different practical scenarios.

The second revision of the energy supply circuit with implemented MPP tracking has a higher efficiency in combination with a smaller PCB size of 27 x 75 mm. It also allows a reduction of the resulting board dimension by 26%. The total components cost in the second revision were reduced by about 30%.

The energy supply circuit in combination with the galvanic cell generates enough energy for sending data packages every 3-5 minutes. By increasing the size of the electrodes, more complex sensor systems with higher energy budgets can be implemented.

The developed receiver board receives and evaluates the transmitted sensor data. The received data packages are also shown on the built-in LCD display and forwarded via USB interface to the host-PC. They can be transmitted wirelessly via Bluetooth interface to a tablet or a mobile phone.

The corresponding PC software with the time-based representation of activated sensors in a graphical form including the list of activated sensors with corresponding time stamp was also developed.

9. Outlook

Further optimisation of the sensor board could focus on the transmitter circuit including the transmitter antenna, which currently occupies about 30% of the board size. Another antenna topology with a smaller size should be investigated and implemented.

Furthermore, the operation characteristics of the developed cell under varying environment conditions (humidity and temperature) should be investigated.

Additionally, methods to increase the generated output power could be explored, e. g. by coating the electrodes with material influencing the conductivity of the water and therefore increasing the output power (e.g., natrium chloride, or citron acid).

The developed PC software can be improved by some additional features like a sensor mapping tool with a graphical representation of the physical sensor placement inside the monitored area. The option of combining the registered sensors to different sensor groups (e. g. for monitoring of different objects) is also possible.

References

[1] D. Petrov, U. Hilleringmann, “Water-based primary cell for powering of wireless sensors”, IEEE Sensors 2020

[2] G. Tuna, V.C. Gungor, „Energy harvesting and battery technologies for powering wireless sensor networks”, Industrial Wireless Sensor Networks, 2016

[3] F. Mazunga, A Nechibvute, “Ultra-low power techniques in energy harvesting wireless sensor networks: Recent advances and issues”, Scientific African, Volume 11, 2021

[4] H. F. Rezaei, A. Kruger and C. Just, “An energy harvesting scheme for underwater sensor applications”, Electro/Information Technology (EIT), 2012 IEEE International Conference on. IEEE, 2012

[5] R. A. Huggins, Advanced batteries: materials science aspects, Springer, 2009

[6] J. Marple, Primary Battery Design, Encyclopedia of Applied Electrochemistry. Springer, New York, 2014

[7] D. Linden, “Handbook of batteries”, Fuel and energy abstracts, 1995

[8] K. B. Lee, L.Lin, “Electrolyte-based on-demand and disposable microbattery”, Journal of Microelectromechanical Systems, **12**, 1-8, 2003.

[9] Multi-CB, Printed circuit board surface finish, [Online]. Available: https://www.multi-circuit-boards.eu/en/pcb-design-aid/surface/surfaces.html. Accessed on: Apr. 25, 2021.

[10] Texas Instruments, TPS6120x Low Input Voltage Synchronous Boost Converter, [Online]. Available: http://www.ti.com/lit/ds/symlink/tps61200.pdf. Accessed on: Apr. 25, 2021.

[11] AVX, SuperCapacitors, BZ Series, [Online]. Available: http://www.avx.com/products/supercapacitors/bestcapreg-bz-series. Accessed on: Apr. 25, 2021.

[12] RICOH, Low Voltage Detector, [Online]. Available: https://www.mouser.de/datasheet/2/792/r3116-ea-1770467.pdf. Accessed on: Apr. 25, 2021.

[13] Analog Devices, Ultralow Power Energy Harvester PMU with MPPT and Charge Management, [Online]. Available: https://www.analog.com/en/products/adp5091.html. Accessed on: Apr. 25, 2021.

[14] Silicon Labs, Ultra Low Power Microcontrollers, [Online]. Available: https://www.silabs.com/mcu/8-bit/efm8-sleepy-bee/device.efm8sb10f8g-qfn20. Accessed on: Apr. 25, 2021.

[15] Silicon Labs, Crystal-less RF Transmitter, [Online]. Available: https://www.silabs.com/documents/public/data-sheets/Si4012.pdf. Accessed on: Apr. 25, 2021.

[16] Silicon Labs, FSK Sub-GHz receiver, [Online]. Available: https://www.silabs.com/documents/public/data-sheets/Si4355.pdf. Accessed on: Apr. 25, 2021.

[17] FTDI Chip, USB Full Speed to Serial UART IC, [Online]. Available: https://ftdichip.com/products/ft232rl. Accessed on: Apr. 25, 2021

[18] Microchip, 8-bit AVR microcontroller, [Online]. Available: https://www.microchip.com/wwwproducts/en/ATTINY1614. Accessed on: Apr. 25, 2021

[19] NXP Semiconductors, Remote 8-bit I/O expander [Online]. Available: https://www.nxp.com/part/PCF8574T#/. Accessed on: Apr. 25, 2021

[20] DSD TECH Bluetooth 4.0 module, [Online]. Available: http://www.dsdtech-global.com/2017/08/hm-10.html. Accessed on: Apr. 25, 2021

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 273-278 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Discover DaVinci: Blockchain, Art and New Ways of Digital Learning

Marko Suvajdzic*,[1], Dragana Stojanovic[2]

[1]*Digital Worlds Institute, University of Florida,* Gainesville, *FL, 32601, USA*

[2]*Faculty of Media and Communication, Singidunum University, Belgrade, 101801, Serbia*

ARTICLE INFO

Article history:
Received: 16 July, 2021
Accepted: 20 September, 2021
Online: 14 October, 2021

Keywords:
Blockchain
Art
Gamification

ABSTRACT

Discover DaVinci is a novel augmented reality system that incorporates blockchain technology with experiential learning to engage participants in an interactive discovery of Leonardo da Vinci's oeuvre. The software was created by Marko Suvajdzic, first author of this paper, and it was produced at the University of Florida Digital Worlds Institute. In the true spirit of this "Renaissance man", Discover DaVinci explores new ideas and technologies "ahead of their time", opening up questions about usage of blockchain system in the domain of art and technology. This paper discusses some of these questions, such as relation of art and technology, usefulness of blockchain system for digital art, and new materiality of art in digital and informational age. Proposed work of this manuscript is to present the field of digital learning through a general review and more specifically through a prism of Discover DaVinci project created by Digital worlds Insititute at the University of Florida.

1. Introduction: Art and Technology, the Unseparable Pair

Speaking about art in contemporary times, together with a still present idea about its alleged autonomy [1], we may tend to understand it as framed within a particular institutional, academic or actual scene [2]. If we concentrate on its manifestations on the stage, in galleries or museums, where we experience it as a certain kind of a screen given to the audience, it is easy to completely miss its deep, continual and historical relation to technology and experimentation. In the other words, as traditional art audience, we often regard art as an object to be watched, listened to, assessed, valued or sold/bought. In such a constellation, the art audience is left in a quite passive state, although art, as a communicational model of encoding and decoding (note the obvious relation to technology and programming techniques!), offers its audience much more [3].

One of the dimensions of art production and placement that is explored in the beginning of 21st century is certainly connected to contemporary communicational, digital and web technologies. It may seem as a new field to be questioned and explored, but the relation of art and technology goes a lot more back into history. It is enough to recall a philosophical term technē, which was used to refer to a making or a doing, but it was also used to designate a kind of an episteme, and it was a name for knowledge in the widest sense [4], [5]. What is more important here is that it was used for any kind of artistic practice that involved making-doing, as well as gaining knowledge through technological experimentation. It was particularly used for crafts and mechanical art, but Larry Shiner argues that it might not have been strictly about art as such, but much more about human ability to make and perform, or, more directly, to use technology to improve the possibilities of making-doing, of understanding art, and of presenting it to others in a concrete historical context [6].

It seems that the current context in which matters of art and art production are just being developed and innovated belongs to digital reality, or, even more forward, to a technological and artificial intelligence (AI) environment. In these digital circumstances, the focus is transferred from an art-object to an art-information, and from human to machine agency, but it does not mean excluding human factor from the art worlds, institutions and economies – on the contrary. We may say that by introducing machine intelligence and transactional systems, their human users are getting chance to consider these changes as motivational and developmental factor not only in the field of technology, but also in the fields of social sciences and art.

One of the technological platforms, or, better to say, a technological system that is found right at the crosspoint of mathematics, sociology, economy, law and politics is blockchain

*Corresponding Author: Marko Suvajdzic, Email: marko@digitalworlds.ufl.edu

system. As this paper will propose and as we will show in further discussion using Discover DaVinci project as an example, the usefulness of blockchain system for the world of art is, although contested, quite prominent and more and more explored – not only within new forms of art production, presentation, protection, marketing and trading, but also within new forms of learning about art in contemporary times. Researches in this field might prove important not only for technology-oriented studies, but also for the education and learning studies.

2. Blockchain and Art: Is It All About the Money?

Looking back to beginnings of blockchain system application, it is obvious that it was created as an "electronic payment system based on cryptographic proof instead of trust…" [7]. This means that blockchain was first and foremost meant to be a more secured and complex system for monetary transactions. This system is based on cryptographically carved currencies of which the first one was Bitcoin, after which it was followed by many more. Blockchain system is thus a distributed database whose security, paradoxically enough, lies in an increase of parties included. So, instead of a trust in one mediator as a representative of a desired transaction, there are hundreds and thousands of nodes, mutually connected, that carry the information about the transaction. The security factor is ensured by peer-to-peer system (P2P), where the server is shared between all the computers and included parties – not only the parties involved directly in transaction, but also all the parties that embody this complex system of humans, codes and computational nodes. What is interesting here is that by increasing the number of involved participants we decrease the risk of a problematic transaction. This is a very specific cooperation of machines and humans, working together on the task of transferring information and ensuring that it is not copied or doubled illegally. The information is condensed into a block, and blocks are packed into chains (thus the name blockchain). This is done through a hash function, which is a form of cryptography that functions as a suture between blocks, and it is placed in such a way that it guarantees that the information within a block cannot be changed internally or externally without the complete change being detectable in all the chain nodes. The only, though theoretical way of manipulating blockchain system lies in an option that at least 51% of elements in the chain decide to manipulate the chain in the same way, and in an exact same time. From here we conclude that the more peers in the blockchain network, the less possibility of manipulating it in any kind of way. On the other side, some downsides of blockchain system include work efficiency problems and high resource consumption due to the process of mining, and that is currently in the focus of research of all those who work on blockchain system inclusion in any of the fields it may find its usage [8].

But is the blockchain system really all about money, and why is its decentralized mechanics so important for the world of art, and, more so, for the process of learning about art? First of all, blockchain system is a system – which means, a specifically operating platform of communication. In this way, any value we put in it, money, art or otherwise, gets included into a specific, mechanical, human-machine built world, a network-like system in itself. Although some can argue that blockchain system is well outside of traditional institutions, and those arguments are given in a very strong way [9], it seems that blockchain system in itself forms a specific set of communicational and transactional rules that inscribes itself in the digital entities it works with. In the sense of previous understanding of art institutions and art worlds [10], it definitely skips the logic of material art exhibition and valuation, but it creates a new world for art existence and presentation, or, more precisely, it adds a new, digital, immaterial layer to previously known logics of art creation, trading and consumption. Thus we cannot really speak about the new, potentially foreign system to the world of art, but of a new technological option for art development and market.

Speaking of that, blockchain system is not only about the art market, although it is a big part of innovation it brings. Of all the other benefits it may bring, besides better possibilities to protect digital artworks and authors' rights, blockchain system offers many options for solving the issues of copying digital works [11]. It also helps promoting yet unestablished authors, and it can resolve the issues of returning material artworks to the countries where they originated from, or to the individuals that have rights to its ownership – in the same time not disturbing work of museums and galleries that currently exhibit these works. Of course, and on the top of it all, blockchain system offers new possibilities of learning about art, especially through gamification strategies. But, before we delve deeper into a concrete example of this possibility that will be described through the analysis of Discover DaVinci project of the University of Florida Digital Worlds Institute, there is still enough space to think through another perspective that the blockchain system brings to the world of art, and that is the issue of (a new) materiality of art in the domain of digital art.

3. Digital Age and New Materiality of Art

In her text "Collections Management on the Blockchain: A return to the Principles of the Museum" Helen Kaplinsky mentions posthuman museums in the context of human-machine, or human-algorithm dialogical or network interaction [12]. Mentioning posthuman museums (and maybe we can extend this into proposing a thought of posthuman classrooms?), she poses the question which can be condensed into the following: do we still need bodies in the age of digital and blockchain platforms prevalence? And, how it will influence the world of art? Or, more broadly said, do we still need materiality of the artworld? As Martin Zeilinger says, "Digital art can thus help us rethink questions related both to the materiality of the digital and to the immateriality of the art object as an abstract container of value" [13]. But how is digital art linked to a material, and an art object to immaterial in this quotation? It seems that a digital contemporaneity, and, especially blockchain technology is contributing further to the complexity of this issue, becoming a relevant and interesting topic not only for art itself, but also for the theoretical discourses around art.

Looking briefly into the nature of digital art, we may easily conclude that digital art is a perfect example of an ephemeral, immaterial art practice – so much that it is actually hard to pinpoint the storing "place", or even a "source" of a digital artwork – be it because of its many copies and shares, or due to its (non)physical, digital qualities. It is important to note that digital art is not the first attempt to immaterialize the artwork and

art practices itself – that attempt can be traced back to conceptual art, as an institutional critique towards capitalist ideologies, as a way of disintegration of traditional art norms, and also as an experimentational practice and exploration of art in a dematerialized, "post-aesthetic" condition [14]. It was in 1973 when Lucy Lippard hoped that "'conceptual art' would be able to avoid the general commercialization… of modernism" [15]. Since then, it seems, we can trace a specific hope in the act of dematerializing art, in the means that in that way it is thought to be able to escape commodification logic of the market, and avoid becoming a high-efficiency container for commercial value". In the age of digital art, the first example of this sort is certainly net art, and for a couple of decades digital art, as a continual line drawn from net art subverting practices, was regarded as an art that is able to contradict the logic of institutionalized and monetized art world. Furthermore, it was regarded as "temporary autonomous zone with no tolerance for institutional dogma and the ideologically suspect economic value systems propagated by institutional bureaucracies of the art world". Perhaps that is why there is still so much resistance towards blockchain technology within the scope of digital art practices, as well as among digital artists and digital art theoreticians. These resistance voices usually point to the dangers of hypercommodification of art; also, they are worried about the possibility of creating a technologically dependent society, or of strengthening the capitalist logic in the domain of art [16]-[20].

To be able to understand relations between digital art and blockchain technology, we have to go back to possible definitions of digital art. According to Christiane Paul, digital art is "art that is digital-born, computable art that is created, stored and distributed via digital technologies and uses the features of these technologies as a medium" [21]. Similarly enough, Hui defines digital objects as appearing to human users "as colourful and visible beings. At the level of programming they are text files; further down the operating system they are binary codes; finally, at the level of circuit boards they are nothing but signals generated by the values of voltage and the operation of logic gates" [22]. Both the authors stress the informational element of the digital object, or digital art object that needs digital technology to be able to exist, or to present itself to the audience. This certainly opens up the path for thinking about contemporary digital technologies, together with blockchain technologies, as relevant factors of enhancing the digital media. In the end, in order to serve and inspire digital art and artists, blockchain system should be both medium and subject of artistic practices, both conventional and critical/subversive. In order for this to happen, artists should have a voice in how this technology might be applied to the field they act in, but to be able to gain this voice, they would have to say *yes* to technology, and to research the ways it could work for their ideas, and not against them. All the different artworks of emerging blockchain artists already show this, and they successfully overcome duality between human and machine factor, as well as between materiality and immateriality of the artwork (see Figure 1 as an example, also [23]).

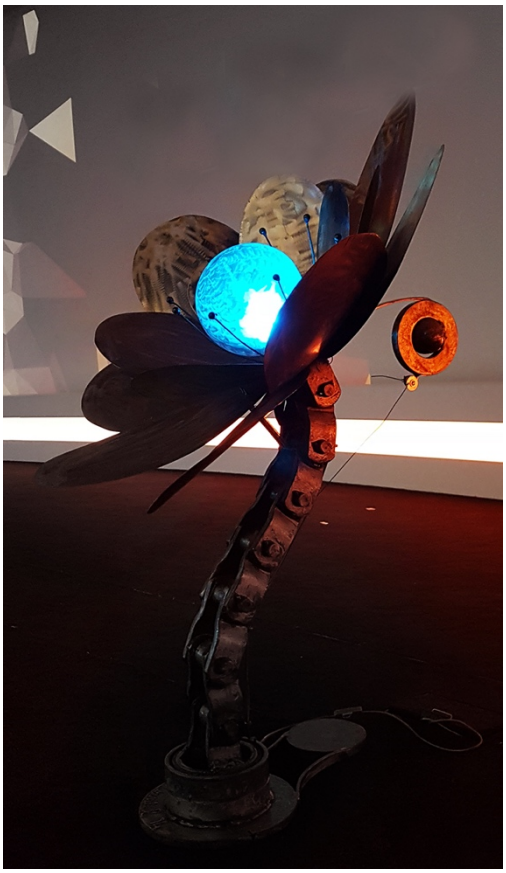


Figure 1: Primavera De Filippi, Plantoid

As innovations in art are always linked to innovations in technology and vice versa, and as making of art, in its quality of *technē* is always in the same time a knowledge creation, the authors of this text will propose a model of technology research in art, or, differently said, of art research in technology. One of the ways to do it is to work on the intersections of technology, art and learning, and to create gamification strategies of learning about art through blockchain system. One of such examples is *Discover DaVinci* project, an interactive system produced at the University of Florida Digital Worlds Institute.

4. Introduction to "Discover DaVinci" Project

In order to illustrate the emerging potential at the intersection of art and blockchain, we present a case study of a new interactive system produced at the University of Florida Digital Worlds Institute. This software was created by Marko Suvajdžić, associate professor and associate director of Digital Worlds Institute. Suvajdžić's experience includes various digital startups and educational projects involving artificial Intelligence-intensive video game titles for major corporate clients. Suvajdžić also co-founded 5 of his own startups.

The technologies of mobile computing, augmented reality (AR), and blockchain are starting to merge, creating new opportunities and scenarios to interact with our environment. In AR we can look at virtual objects superimposed within a real environment and resize them, rotate them, explore and interact with them on multiple levels. With the combination of AR and blockchain, we can create a system capable of keeping track of

digital assets located virtually in 3D space (i.e., spatial computing). The global scale of blockchain and related technologies heightens the potential for trade and digital distribution with a fully automated and trusted way to keep track of their creations without a "middle-man".

Discover DaVinci is a novel educational tool that teaches concepts of blockchain technology through an augmented reality experiential learning game.

5. Discover DaVinci Project Partnerships

This project was developed in collaboration with several units from the University of Florida and industry partners:

- Digital Arts & Sciences Faculty (Computer Science and Digital Worlds Institute)
- Digital Worlds Studios' Artists and Programmers
- Gator Blockchain Club (gatorblockchainclub.com) – Student-run blockchain club at the University of Florida
- Center for Innovation and Entrepreneurship (College of Business)
- Creative Campus Committee at the University of Florida

Industry Partners:

- DLUX, decentralized content network (dlux.io)
- Steem (steem.com), and Steemit (steemit.com)
- A-Frame, web VR platform (aframe.io).

6. Research Study: "Discover DaVinci"

Discover DaVinci (Fig.2) utilizes the format of a digital, collectible trading & drafting card game with AR elements on the STEEM blockchain. Although each player "owns" their cards, all transactions are public. This is enabled by the blockchain system itself, that defines private and public key for every user. The private key always stays private and cannot be recovered, and the public key is visible to the others. In this way blockchain system guarantees all the participants both their privacy, and in the same time public visibility of every change in the chain, being enacted by purchases and transfers in ownership.



Figure 2: Discover DaVinci game interface, University of Florida (2019)

Discover DaVinci collectible cards are all unique tokens, owned by the player - a digital asset registered to the player's account. The aim is to draw new question cards daily, answer the questions about Leonardo DaVinci, collect the special AR invention cards, and ultimately submit the accumulated card collection into a drawing for prizes. The app was developed to honor the 500th anniversary of Leonardo DaVinci by promoting new and innovative technologies. In the same time, this project is an homage to Leonardo DaVinci's fascination with art and technology, which is an *oeuvre* of the artist of the Renaissance period, but it may also be a quality of artists *maneuvering* their own artwork pathways in the age of blockchain technology development.

By using a novel combination blockchain, and spatial computing/AR technologies, the *Discover DaVinci* mobile app allows users to explore holographic images of DaVinci's devices. By combining AR with blockchain technology, we have created a system capable of assigning unique values and user comments to these digital assets, thus providing a personal sense of ownership, the potential for trade, and a fully automated and trusted way to keep track of these digital artifacts (e.g. collected 3D objects in AR space). Finding the artifacts themselves is an act of discovery, further enhanced by subsequent investigation, collection, and sharing of the artifacts across a diverse social media ecosystem.

The sense of ownership, and not only accumulating of goods is provided by the technology of non-fungible tokens (NFTs), which is especially valuable in the field of art. In the same time, using NFT in tokenization of an artwork or a collectible makes it possible to distinguish the art token from any other cryptovalue, or a cryptocurrency, which are defined as fungible tokens. In a way, this gives us an answer to a previously posed question – is blockchain based art only about the money? Following the logic of non-fungible tokens, it is not. NFTs are not interchangeable, nor convertible in any other value. They possess their own unique value – just as it is the case with material, traditional art. In this way we can see that blockchain system can be regarded as the next step of efficiency and inspiration material within the already existing art forms, art market and art institutions. Of course, as there is always a room for critique and resistance in traditional art world, it is the same in technological world. That is why it is important for the artists and art/art theory teachers to delve deeper into blockchain technology research, and to tailor new media to their own expressions and needs – just as it has always been, as it is shown in the case of Leonardo DaVinci.

Back to *Discover DaVinci* project, the blockchain system was implemented using the STEEM API, which handled the writing of new user accounts and the reward transactions to the Blockchain (Fig. 3 bottom). The front end was developed as a web app that rendered the various 2D and 3D/AR components of the game (Fig. 3 top right). Finally, the authentication and account creation was handled using a third-party service, called STEEM Ninja (Fig. 3 top left).

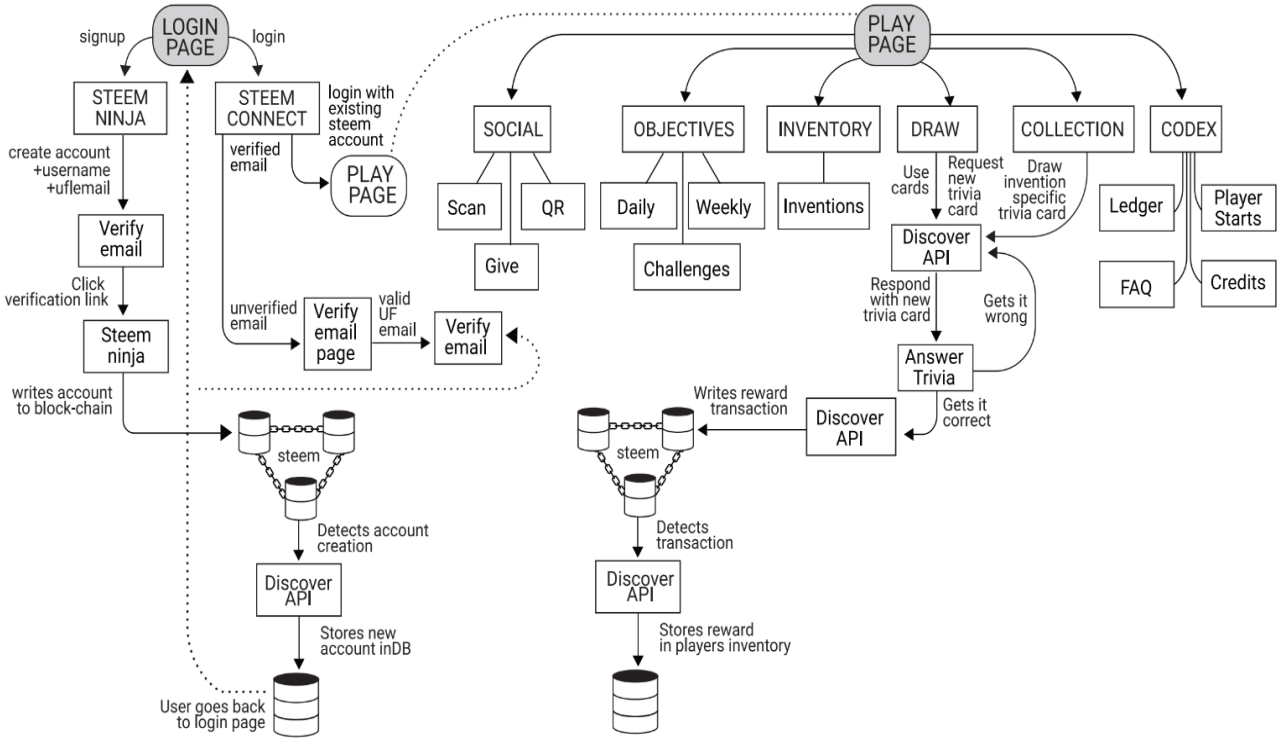


Figure 3: Discover DaVinci-System Architecture illustration, University of Florida (2019)

7. Discussion

We launched Discover DaVinci app to the initial audience of University of Florida students and faculty in November of 2019. Over the period of 60 days, we engaged ~150 participants who played the game, and left their feedback on the project's discussion board.

The player feedback has been analyzed, and the subsequent versions in the "Discover" series are currently being evaluated. The project is intended to continue as a platform for gamified education that promotes participant interaction, and supports focused learning on a topic of choice.

In this review paper with the main example of Discover DaVinci App, the project conducted by the team at the University of Florida, Digital Worlds Institute, the authors wanted to refer to Discover DaVinci application and its potential within the fields of gamification and learning concepts. Certainly, there is still a space for further tests to be done in the future, so we could effectively reveal the impact of such an application on learning methods and contemporary educational approaches involving blockchain technology. However, at this point, the authors would like to propose a formal conclusion, stating that these kinds of applications offer effective possibilities for understanding teaching and learning processes within generation that adopt knowledge at the present moment. What has been noticed is certainly an ease of a new technology adoption in the entertaining educational setting, and that is the point from which we can continue to think about the intersection of technology and education.

Discover DaVinci is an innovative app that uses blockchain technology as a tool that utilizes gamification of education, specifically focused on presenting life and works of Leonardo DaVinci. The app has been custom created for this purpose by the team at the Digital Worlds Institute at the University of Florida, and as such is not directly comparable with general educational tools on the market. However, it is used to illustrate how blockchain technology can be introduced to new users through gamified content.

We believe that this novel confluence of art, technology, and digital communications represents a tangible example of how artists and cultural researchers can leverage the potential of diverse digital resources to support both creative development and cultural inquiry. The Discover DaVinci app allows a diverse player community to utilize their existing mobile devices and participate in acknowledging DaVinci's life work while exploring a new paradigm of digital convergence. The outcomes included an enhanced cross-campus appreciation of personal creativity, social and cultural engagement, and a working framework that exemplifies how emerging technologies can serve esthetic ends while also ensuring fair use and exchange of intellectual property in the digital age.

Acknowledgements

The authors would like to thank the students and staff who contributed to the development of the app: Liam Wood, Paul Burgermeister, Steven Ettinger, Mark Giles, Jan Zamojski, Kelvin Chang, Ryan Trimble, Nick Heitzman, Seung Hyuk Jang, Anna

Walters, Kristen Eldridge, Anthony Monteagudo, Aaron Karlson, Chelsea Cantrell, Gabriel Coleman, Jared Crane, Justin Marlin, Tim Difato, and members of the UF Bachelor of Arts in Digital Arts and Sciences Society. We would also like to express our appreciation to the University of Florida Creative Campus Committee for providing funding for this project through the catalyst fund program.

References

[1] R. Cronk, "The Pseudo-Autonomy of Art," Art on the Rebound: A Collection of Essays on Art and Culture, https://www.westland.net/venice/art/cronk/autonomy.htm, 1996, Retrieved 23 March 2021.
[2] G. Dickie, Aesthetics, An Introduction, Pegasus, 1971.
[3] R. Rojcewicz, The Gods and Technology: A Reading of Heidegger, State University of New York Press, 2006.
[4] D. Carpi, Bioethics and Biolaw through Literature, Walter de Gruyter, 2011.
[5] L. Shiner, The Invention of Art, University of Chicago Press, 2001.
[6] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," www.bitcoin.org, Retrieved 22 August 2020.
[7] N. Chalaemwongwan, W. Kurutach, "State of the Art and Challenges Facing Consensus Protocols and Blockchain," ICOIN 2018, 957-962, 2018. doi: 10.1109/ICOIN.2018.8343266.
[8] S. Ren, "The Application of Blockchain and its Advantage to the Field of Art," 2nd International Symposium on Big Data and Applied Statistics, Journal of Physics: Conference Series, IOP Publishing, 2019.
[9] R. Myers, "Blockchain Poetics," Catlow R., Garret M., Jones N, Skinner S. (Eds.), Artists Re:Thinking the Blockchain, Torque Editions & Furtherfield, 239-250, 2017.
[10] A. Danto, "Symposium: The work of Art, the Artworld," The Journal of Philosophy 61(19), 571-584, 1964. doi: 10.2307/2022937.
[11] R. O'Dwyer, "Limited Edition: Producing artificial scarcity for digital at on the blockchain and its implications for the cultural industries," Convergence: The International Journal of Research into New Media Technologies, Sage, 1-21, 2018, doi: 10.1177/1354856518795097.
[12] H. Kaplinsky, "Collections Management on the Blockchain: A return to the Principles of the Museum," Catlow R., Garret M., Jones N, Skinner S. (Eds.), Artists Re:Thinking the Blockchain. Torque Editions & Furtherfield, 261-274, 2017.
[13] M. Zeilinger, "Everything you've always wanted to know about the Blockchain*(*but you were afraid to ask Mel Ramsden)," Catlow R., Garret M., Jones N, Skinner S. (Eds.), Artists Re:Thinking the Blockchain. Torque Editions & Furtherfield, 287-295, 2017.
[14] N. Jones, S. Skinner, "A Quasi Proto Preface," Catlow R., Garret M., Jones N, Skinner S. (Eds.), Artists Re:Thinking the Blockchain, Torque Editions & Furtherfield, 9-19, 2017.
[15] A. Alberro, B. Stimson, Conceptual Art: A Critical Anthology, MIT Press, 1999.
[16] M. Zeilinger, "Digital Art as 'Monetized Graphics': enforcing intellectual property on the blockchain," Philos.Technol, 2016, doi: 10.1007/sl3347-016-0243-1.
[17] L. Lessig, Free Culture, Penguin, 2004.
[18] A. Moore, K. Himma, "Intellectual Property," Stanford Encyclopedia of Philosophy, http://plato-stanford.edu/entries/intellectual-property/, Retrieved 15 November 2016.
[19] J. Boyle, The Public Domain: Enclosing the Commons of the mind, Yale University Press, 2008.
[20] R. J. Coombe, The Cultural Life of Intellectual Properties: Autorship, Appropriation, and the Law, Duke University Press, 1998.
[21] C. Paul, A Companion to Digital Art, Wiley Blackwell Publishers, 2016.
[22] Y. Hui, "What is a Digital Object?," Metaphilosophy 43(4), 380-395, 2012. doi: 10.1111/j.1467-9973.2012.01761.x.
[23] M. Suvajdzic, D. Stojanovic, J. Appelbaum, "Blockchain Art and Blockchain Facilitated Art Economy: Two Ways in Which Art and Blockchain Collide,"2019 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), 2019, doi: 10.1109/TIMES-iCON47539.2019.9024403.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 278-291 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

An Alternative Approach for Thai Automatic Speech Recognition Based on the CNN-based Keyword Spotting with Real-World Application

Kanjanapan Sukvichai*, Chaitat Utintu

Department of Electrical Engineering, Faculty of Engineering, Kasetsart University, 10900, Thailand

ARTICLE INFO

Article history:
Received: 02 April, 2021
Accepted: 26 July, 2021
Online: 03 August, 2021

Keywords:
Thai ASR
MFCC
KWS
CNNs

ABSTRACT

An automatic speech recognition (ASR) is a key technology for preventing an ongoing global coronavirus epidemic. Due to the limited corpus database and the morphological diversity of the Thai language, Thai speech recognition is still difficult. In this research, the automatic speech recognition model was built differently from the traditional Thai NLP systems by using an alternative approach based on the keyword spotting (KWS) method using the Mel-frequency cepstral coefficient (MFCC) and convolutional neural network (CNN). MFCC was used in the speech feature extraction process which could convert the voice input signals into the voice feature images. Keywords on these images could then be treated as ordinary objects in the object detection domain. The YOLOv3, which is the popular CNN object detector, was proposed to localize and classify Thai keywords. The keyword spotting method was applied to categorize the Thai spontaneous spoken sentence based on the detected keywords. In order to find out the proposed technique's performance, real-world tests were carried out with three connected airport tasks. The Tiny-YOLOv3 showed the comparative results with the standard YOLOv3, thus our method could be implemented on the low-resource platform with low latency and a small memory footprint.

1. Introduction

Automatic speech recognition (ASR) is still an intriguing and demanding subject among researchers around the world. It is the mechanism that enables human-machine interaction through speech orders. In the present of the Coronavirus outbreak, this technology could decrease the propagation of the infection because no physical interaction such as touchscreen is required. Moreover, the automatic speech recognition system could also provide the hand-free experience that produces the huge advantage among the handicapped and visually impaired people. This manuscript is an extension of work that was initially presented in the 2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP) [1]. In Thailand, according to the computers and the Thai language research [2], the first development of Thai ASR system concerned about isolated word recognition, which could be used for short commands [3]. The other improved method, which might be employed for continuous voice in little jobs, such as e-mail access or telephone banking, was also offered. More researches were publishing on huge vocabulary continuous speech recognition, or LVCSR [4], [5]. Inadequate speech databases and Thai-language complexity such as letter-to-sound mapping, phoneme set selection, segmentation, and tonality, were the key problems when developed a Thai automatic speech recognition system as mentioned in [6]. Nevertheless, many Thai organizations put effort in developing Thai speech corpus such as the LOTUS corpus by NECTEC [7]. For the standard automatic speech recognition system in natural language processing (NLP), it consisted of three basic models: Acoustic model, Lexicon, and Language model. For Thai language, the limitation of available large speech database and the complexity of Thai language could contribute to delayed speech recognition improvement [2].

In an effort to solve the problem, the automatic speech recognition was constructed using the different approach. The proposed method used the keyword spotting methodology to recognize the Thai sentence linked to the categorized keywords instead of using the traditional ASR system. This approach could ignore the presence of unpredictable words in the utterances such as unprecedented words and non-speech sounds. In such a case,

*Corresponding Author: Kanjanapan Sukvichai, Faculty of Engineering, Kasetsart University, 10900, Thailand, : +662-797-0999, fengkpsc@ku.ac.th

https://dx.doi.org/10.25046/aj060431

the typical method must include the large vocabularies, as well as grammars.

More details about the keyword spotting were clearly mentioned in the following section. For investigating more corresponding Thai research, the use of acoustic modeling based on Hidden Markov Models (HMMs) method [8], [9]. The models consisted of filler models, syllable models, and keyword models, which used for handling out-of-vocabulary sound elements in the speech recognition using the keyword spotting algorithm. Another paper proposed the nondestructive determination of maturity of the Monthong Durian based on Mel-Frequency Ceptral Coefficients (MFCCs) and Neural Network. Then we have further investigated more excellent works that related to ours like SpeechYOLO paper [10], this paper used first version of YOLO with keyword spotting algorithm for localizing boundaries of utterances within input signal by considering audio as objects and use Short-Time Fourier Transform (STFT) technique to extract the audio feature. In our research we have identified a similar approach in order to solve the limitations in Thai speech recognition field, but we use the improved version of YOLO which was YOLOv3 and we also use Mel-Frequency Cepstral Coefficients (MFCC) technique which could artificially implement the behavior of the human auditory perception due to the fact that it is capable of using the non-linear Mel scale which relies on the non-linear frequency scale of the real human.

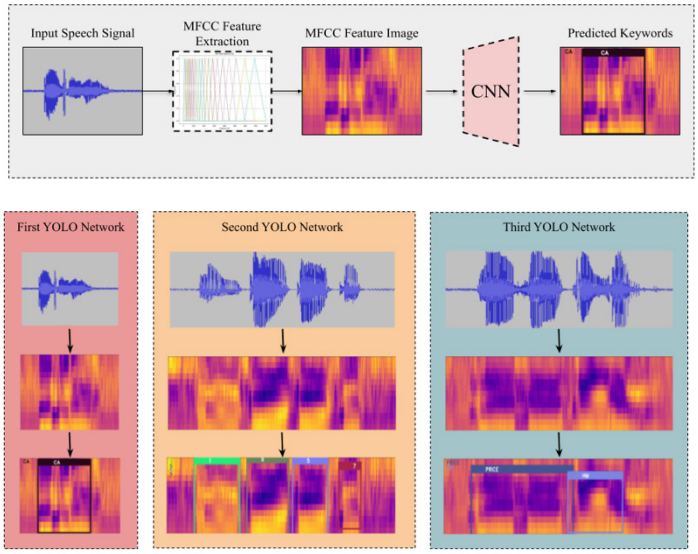


Figure 1: **Top:** the core methodology pipeline composed of the Mel-Frequency Cepstral Coefficient (MFCC) as the feature extraction method and the YOLOv3 as the popular convolutional neural network (CNN) based object detector which was then applied to localize and classify keywords. **Bottom:** our proposed method for the real-world application, which consisted of three separated YOLO networks including airline names, numbers (0-9), and frequently asked questions (FAQs). Both full and tiny version of YOLOv3 were proposed.

This manuscript started with explaining how Keyword Spotting (KWS) to clarify the speaker propose and also investigated the KWS prior works. Then, the system pipeline will be explained which started from the speech feature analysis extraction method. This process was crucial since it could represent the voice signal information in the form of a feature image. Consequently, in the next section, the convolutional neural network was explained including the history of object detection techniques, the novel CNN-based object detectors, and the major improvement in YOLO family. After that, the voice dataset collection is described. Then, the highlights of our core methodology are explained. This manuscript also explained about the YOLO annotation and network training. Finally, the proposed method consisted of three networks was experimented. These three networks were responsible for three real-world tasks that would then introduce in the following section. In the last section of this manuscript, the comparison between the regular and the light-weight version of YOLOv3 performances was discussed for the additional usage in the low-resource platform. The overall design of our work was displayed in Figure 1.

2. Keyword Spotting for the Spoken Sentence

In general, most people tend to struggle to understand the spoken sentences especially with the native speakers. Unlike a written sentence, the spoken sentence is more difficult and variety in terms of speed, accent, style, or tone depending on the morphological richness of each language such as Thai language as previously mentioned. The most reasonable answer why the spoken sentence is difficult is that we don't experience enough vocabularies. Many scientists also say that listening requires more vocabulary knowledge than reading. In addition, it is difficult to tell whether the sentence contains unknown terms or is quickly spoken. Eventually, the spoken words are not always finished or carefully delivered with the native speakers like syllable stressing or skipping some words. Just like the human, the machine must transcribe the entire sentence with a huge model of vocabulary. To address this issue, the keywords are employed to involve the main idea rather than to use every single word in the sentence. This idea was motivated by the scanning strategies used in the reading comprehension test. Many skillful readers used this technique to reduce time to search for some specific information by just looking for the related words instead of reading the whole paragraph line by line. This scanning technique also seemed to help the system to come across the complexity and the variety of the sentence structures.

For the ASR system, there were certain technologies known as wake word, the particular word or phrase which would enable some additional functions. These voice-related systems, such as "Hey Siri" by Apple's Siri or "OK Google" by Google Now, exploited the keyword spotting (KWS) method which would listen until the specific keywords could be recognized [11]. The prior work for KWS introduced the Keyword/Filler Hidden Markov Model (HMM) [12]. Each keyword was trained with the HMM model, then the non-keyword segments, so-called fillers, was trained separately with a filler model HMM. Due to the decoding need of Viterbi, this strategy ended with considerable computational complexity. Then, the following enhanced approaches concerned the use of the discriminative model on the keyword spotting task, namely, large-margin formulation [13] and recurrent neural networks [14]. Although it could produce many significant improvements over the traditional HMM method, the high computational time was still occurred due to the entire utterance processing in order to locate an optimal area of keywords. These limitations were still challenging until Google proposed the novel deep learning keyword spotting approach base on the deep neural network, also called DeepKWS. Since this method could provide high detection performance with smaller

memory consumption and shorter runtime computation, then it was appropriate for the mobile device usage. Unlike the HMM approach, this process did not require the excessive sequence search algorithm. After that, Google explored a small-footprint keyword spotting (KWS) system using the Convolutional Neural Networks (CNNs) [15]. The CNN approach could outperform the DeepKWS in many ways. For instance, it provided more robustness over different speaking styles, reduced model size and more practical for spectral representation input. Thus, in this research, we chose the CNN-based object detection for localizing and classify Thai keywords from the spoken sentence.

3. Speech Feature Analysis and Extraction

Automatic Speech Recognition (ASR) system is influenced by feature analysis and extraction, since the high-quality features provide an acceptable and dependable results in the localization and classification process. The main purpose of feature extraction is to reveal acoustic information in terms of a sequence of feature vectors, which can successfully characterize a given speech data. Several feature extraction techniques are available to extract the parametric representation of the audio signal, such as perceptual linear prediction (PLP), linear prediction coding (LPC) and Mel-frequency Cepstral Coefficients (MFCC). MFCC has been proven to be the most prevalent and powerful technique [16], [17].

Mel-Frequency Cepstral Coefficients (MFCC) is the transformation from the speech waveform to frequency domain mathematical features, which are considered to be much more accurate than time domain features. This technique artificially implements the behavior of the human auditory perception due to the fact that it uses the non-linear Mel scale which relies on the non-linear frequency scale of the real human, so it results in the parametrically resemblances between the extracted vectors and human sense of hearing. The output of the MFCC is a short-term power spectrum coefficient of a windowed signal produced from the original signal Fast Fourier Transform (FFT). This transformation is done using the logarithmic power spectrum through a linear transformation, Discrete Cosine Transform (DCT). MFCC for speech can reveal more compact spectrum since Mel-scale coefficients are countable.

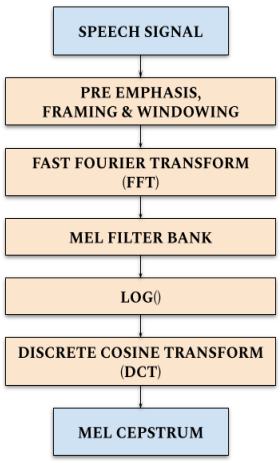


Figure 2: The overall MFCC derivation

In detail, the overall MFCC feature extraction technique is shown in Figure 2. It consists of windowing the speech signal, applying the Fast Fourier Transform (FFT), taking the log of the magnitude, warping the frequencies on a Mel scale, applying the Discrete Cosine Transform (DCT). In order to remove the MFCC from the speech signal, pre-emphasis begins. Compensate filtering for the high frequency area disappearing during the mechanism of voice creation are considered in the pre-emphasis. In addition, the significant of the high-frequency component is also strengthened. This step is therefore followed by the high-pass filter as explained in (1).

$$S_y(n) = S(n) - \alpha S(n-1) \tag{1}$$

where,

$S(n)$ is the input signal.

$S_y(n)$ is the output signal.

α is a control slope of the filter ranged from 0.9 to 1.0

The z-transform of the filter is defined as (2).

$$H(z) = 1 - \alpha z^{-1} \tag{2}$$

After the pre-emphasis process, the spectrum of the signal is balanced and some glottal effects from the vocal tract parameters are removed. Then the speech signal needs to be analyzed over a short period of time by capturing into a discrete frame, with some overlapping between frames. The purpose of the overlapping analysis is to approximately center each speech signal at some frame and avoid significant information loss. On each frame, the Hamming window is multiplied individually to maintain the continuity of the first and the last points in a frame. If the signal in a frame is denoted by $S(n)$ then the signal after windowing is applied will be $S(n)*w(n,\alpha)$ where $w(n,\alpha)$ represented the Hamming windowing defined as (3)

$$w(n,\alpha) = (1-\alpha) - \alpha\cos\left(\frac{2\pi n}{N-1}\right) \tag{3}$$

where,

$w(n,\alpha)$ is the Hamming window.

N is the total number of samples in a frame.

n is a sample number.

Each short-time frame was then transformed into the spectral features by applying the Fast Fourier Transform (FFT), which was the accelerated version of the Discrete Fourier Transform (DFT). This technique converted each frame of N samples from the time domain into frequency domain and was shown in (4).

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{(2\pi nk)j}{N}} \tag{4}$$

where,

N is the total number of points in the FFT computation.

Next, the Fourier transformed output was passed through a set of band-pass filters, so called Logarithmic Mel-filter bank, in order to cover from a real frequency, estimated in the Hertz unit,

f m, by the equation (5). The Mel scale was about a linear frequency spacing of less than 1 kHz and a logarithmic spacing of more than 1 kHz, respectively.

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{5}$$

The frequency of Mel-scale is proportional to the linear frequency logarithm. Due to the fact that the behavior of human's aural perception is non-linear, this concept can be implemented using Mel-scale filter bank, which is commonly the combination of the K triangular filters. The example of a filter bank with K=20 is illustrated in Figure 3. The higher frequency filters contain more bandwidth than the lower frequency filters, but they share similar temporal constraints. Each triangular filter is centered with a maximum amplitude of 1 and is reduced linearly to zero until it approaches the central frequency of the two neighboring filters, where zero response is present and can be derived as (6).

$$F_m(k) = \begin{cases} \dfrac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) < k < f(m) \\ \dfrac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) < k < f(m+1) \\ 0, & \text{Otherwise} \end{cases} \tag{6}$$

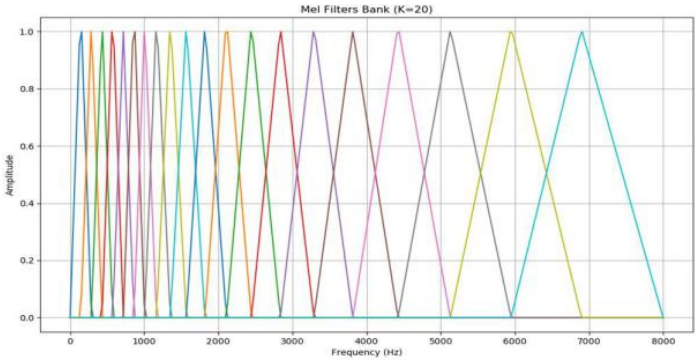


Figure 3: Mel filters bank when K=20

where,

$F_m(k)$ is the filter number m at k frequency.

m is the filter number in Mel filters bank.

Therefore, the Mel spectrum of the frequency spectrum $X(k)$ was calculated by multiplying the spectrum by each of the triangular Mel weighting filters as shown in (7).

$$S(m) = \sum_{k=0}^{N-1}\left(\left|X(k)\right|^2 F_m(k)\right) \tag{7}$$

where,

M is the total number of triangular Mel weighting filters.

The Discrete Cosine Transformation (DCT) was then applied on the transformed Mel frequency coefficients in order to produce a set of twelve cepstral coefficients. Since this algorithm, which was described in (8), resulted in a signal with a quefrency peak in the time-like domain, so called cepstral domain. Thus, Mel-Frequency Cepstral Coefficients, or MFCC, was nominated from the final features which were similar to cepstrum. The zeroth coefficient was excluded because it contained the unreliable information.

$$C(n) = \sum_{m=0}^{M-1} \log_{10}\left(S(m)\right) \cos\left(\frac{\pi n (m - 0.5)}{M}\right) \tag{8}$$

where,

$S(m)$ is the frequency coefficient.

$C(n)$ is the cepstral coefficient.

L is the number of MFCCs.

M is the number of triangular bandpass filters.

$n = 0, 1, \ldots, L$

Following this phase, a speech signal in the form of an MFCC might be employed to perform machine learning technology or treated as an ordinary image. The MFCC for each speech frame was stacked and converted into an RGB image with Plasma color map. Figure 4 shows examples of MFCC's Thai keyword images which include check-in, price and what time is it respectively.

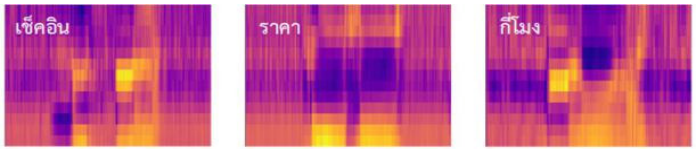


Figure 4: The examples of Thai keyword MFCCs

4. Convolutional Neural Networks for Voice Image

The sequence of the cepstral vectors or MFCC characteristics was received from the recorded voice as an input as previously discussed in the last section. The next step is to derive the corresponding keywords. In this section, computer vision technique for the Natural Language Processing (NLP) task is detailed. In order to resolve this challenge, we opted to employ a state-of-the-art object detection algorithm because it can be used to estimate the location (object localization) and category (object classification) of each object in a given image. Therefore, this technique should be able to extract the important sort of information for better semantic understanding of images [18]. From the preceding progress, as a result of considering MFCC features as the normal image, then keywords can be comparative to objects. Thus, the object detection is well-suited and also helpful for solving this situation as shown in Figure 5. The object detection technique that we chose in our work is the novel deep learning approach based on Convolutional neural network (CNN) [18][19]. CNN is a specialized type of neural network model modified especially for dealing with two-dimensional input data.

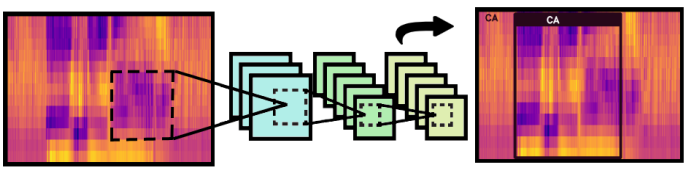


Figure 5: Object detection on MFCCs.

It is typically composed of three fundamental types of layer consisting of convolutional, pooling and fully connected layers. Convolutional and pooling layers contain filters, whose depth increases from left to right while width and height have the inverse proportion, have an interactive role in extracting features. The Convolutional layer is denominated from the important linear mathematical operation in the layer, so-called convolution, or filtering. The convolution function is described in (9).

$$S(i,j) = \sum_m \sum_n I(m,n)F(i-m, j-n) \tag{9}$$

where,

I is the two-dimensional array such as Image.

F is filter or kernel.

$S(i,j)$ is the output or feature map.

This operation involves the dot product between the input array and the filter or kernel, by convolving across its width and height, then extended throughout its depth. Next, the result from convolution is passed through an activation unit called Rectified Linear Unit (ReLU) calculated by (10). It is a piecewise function that will return the input directly if the input value is positive, otherwise, it will return zero.

$$R(z) = \max(0, z) \tag{10}$$

Subsequently, it is down sampled or normalized by the pooling operation, e.g. max pooling, in the pooling layer. This operation summarizes the initial activation feature map to become more robust. After that, fully connected layers then map them into output via activation function, which normally uses the Softmax activation defined in (11). It provides the probability distribution for each neuron, in the same way as a traditional neural network.

$$\sigma(y)_i = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}} \tag{11}$$

where,

y_i are the element of the input vector.

n is the total number of classes.

y is the input vector.

The learning parameters can be optimized via the Stochastic gradient descent (SGD) method and the back propagation, called the training, in each layer across the models. The algorithms aim to minimize the difference between prediction and ground truth. Therefore, in the training process, the dataset images are fed forward with initial kernels and weights, so-called forward propagation. Then a model's accuracy is calculated by a loss function. The error value is used for updating kernels and weights later in the back propagation using gradient descent optimization. The overview of convolutional neural network architecture and the parameter optimization process are illustrated in Figure 6.

The convolutional neural network has been shown to be outperformed to previous methodologies in many ways. Firstly, Spatial Hierarchical feature representation can be learned automatically and the multiple nonlinear mappings can reveal the hidden patterns. Then deeper architecture, which significantly increases the model competency, allows related task optimization. Finally, due to the uprising of CNN learning performance, some challenging computer vision problems might be solved from a new perspective. CNN method has also been popular among various research topics such as facial recognition [22] and pedestrian detection [23]. CNN-based object detection technique aims to identify the location and the class of every object existing in the input image. Nowadays, its frameworks can be divided into two main categories, which are two-stage and one-stage approaches. A two-stage model is a technique that is constructed based on the standard object detection pipeline. It proposes two separate consecutive methods which are region proposal and object classification. In detail, region proposals are generated first, then each proposal will be classified according to the object classes. The second methodology, which is a one-stage model, considers transforming the traditional object detection problem to the regression problem by merging all separated tasks, consisting of localization and classification, into a single unified model.

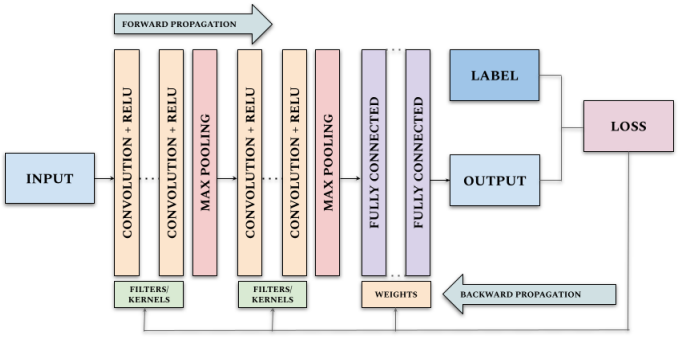


Figure 6: CNN architecture and the training process.

For the two-stage method, in [22], the author introduced the proposal of the regions with CNN features (R-CNN), which took a giant step towards object detection. It extracts a set of object candidate boxes or proposals by using a selective search method, which leads to major improvement in detection accuracy. However, this technique experiences many drawbacks. The enormous quantities of overlapped repetition boxes greatly delay the overall process and some feature losses can also be due to proposals in fixed candidate region. To solve the problems, in [23], the author proposed Spatial Pyramid Pooling Networks (SPPNet). The Spatial Pyramid Pooling (SPP) layer in SPPNet allows a CNN to construct a representation without considering the input image size. Although SPPNet utilizes less processing time than R-CNN since it requires only the single computing for the convolutional layer, the downside of this approach is the extended training period and the high use of disk space. The several advantages from R-CNN and SPPNet were investigated and integrated in the following work, Fast RCNN, but the detection speed is still bounded by the selective search [24]. Therefore, in [25], the author proposed Faster RCNN, which was known as the first ever end-to-end object detector that could almost yield real-time performance. It can break through the speed bottleneck of past efforts by demonstrating the use of the region

proposal networks (RPN) instead of the traditional selective search method. In addition, it uses a separate network to predict the region proposals only. The region proposal network is trained together with the model which results in predicting more accurate proposed regions. Later, there were some further improvement efforts on Faster RCNN, such as Feature Pyramid Networks (FPN) introduced in [26]. With the small object constraints, a basic image pyramid can be used to scale input images into multiple sizes before sending to the network. The FPN has enhanced the performance of multi-scale object detection and has become a model for numerous novel approaches.

On the one hand, the two-stage detector comes up with high localization and classification accuracy. On the other hand, its inference speed is impractical for the real-time applications, especially for the low-resource computational platforms like embedded systems, and the complex procedures of its core methodology could possibly diminish the opportunities of further optimization and improvement over components. The one-stage models have been built and researched in order to overcome these restrictions. It recognized objects based on two distinct neural networks, which lead to high computational time as indicated before in Faster RCNN section. Accordingly, the one-stage method combines feature extraction, region proposal and object classification altogether into a single network. The object classification is also performed regarding the predefined size and number of anchor boxes specified in the model configuration, then the regression concept is applied for the object localization process. Consequently, this novel breakthrough could implement the object detection as a single regression problem by predicting bounding box coordination and class probability simultaneously. The one-stage object detection pipeline used in our work is shown in Figure 7.

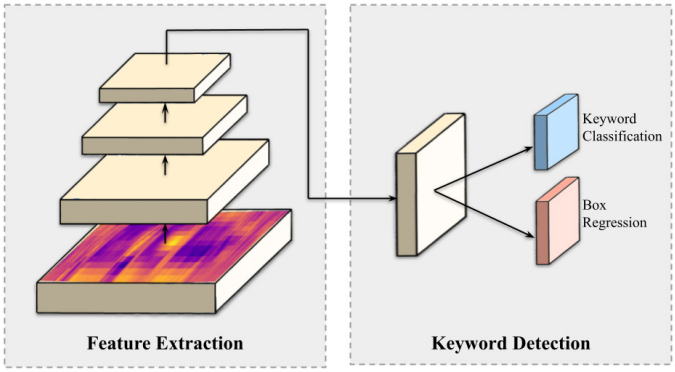


Figure 7: The basic architecture of one-stage object detector using MFCC voice image as the model input.

The one-stage object detectors were demonstrated to attract a great deal of attention amongst researchers not only because they can overcome the problem of a real-time bottleneck in the two-stage approaches, but because it could generate similar detection performance, it the detection speed and accuracy trade-off is well-optimized. Since our application is based on the Automatic Speech Recognition (ASR), both processing time and detection performance must be concerned. Therefore, we have chosen the state-of-the-art one-stage approach called YOLO, which is You Only Look Once. Based on real-time criteria, the computational time must be considered more than 10 frames per second (FPS) or less than 100 milliseconds. The performance evaluation results published in the YOLO article meet our requirements and stated that the third version of YOLO, or YOLOv3, outperforms other popular methods such as Single Shot Multibox Detector (SSD) [27], which had the original implementation on Caffe [28].

You Only Look Once (YOLO) is the prominent one-stage CNN-based object detector which we have chosen as a Thai keyword localizer and classifier in our research due to the high calculation speed and acceptable detection accuracy on the real-time task. YOLO frames object detection as a regression problem, thus it can simultaneously construct the object bounding boxes and predict class probabilities, while it needs only one forward propagation from the input image. The model name, You Only Look Once, might therefore be designated from such properties.

YOLO divides the input image into $S \times S$ grid. Then, each grid cell takes responsibility for detecting the object whenever its center approaches that particular grid cell boundary. Each grid cell will predict B bounding boxes and their predicted confidence scores. The bounding boxes can be represented by five parameters including x, y, w, h and the confidence scores. The parameters (x, y) indicate the center coordination of the box relative to the grid cell area. The next two parameters (w, h) describe the width and height of the box relative to overall image dimensions. The last parameter, the confidence scores can be derived from (12).

$$Pr(object) * IOU_{Pred}^{Truth} \tag{12}$$

where,

$Pr(object)$ tells how likely the object existence is.

IOU_{Pred}^{Truth} tells how accurate the predicted box is.

At that moment, each grid cell also predicts C conditional class probabilities, which is the conditional probabilities of the grid cell detecting an object, regardless of the number of boxes. The prediction is finally encoded as a $S \times S \times (\mathrm{B} \times 5 + \mathrm{C})$ tensor and the final confident scores can be described by Equation (13).

$$Pr(object) * IOU_{Pred}^{Truth} * Pr\left(class_i \middle| object\right) = Pr(class_i) * IOU_{Pred}^{Truth} \tag{13}$$

where,

$Pr\left(class_i \middle| object\right)$ is the conditional class probability.

During the model training step, the loss function is optimized and could be described in (14). From the following equation, we can notice that the loss function only penalizes the classification errors when an object exists in that corresponding grid cell, and likewise, the errors of the bounding box coordinate are penalized when a prediction is literally responsible for the ground truth.

$$Loss_{total} = Loss_{bonding\,box} + Loss_{confidence} + Loss_{classification} \tag{14}$$

where,

$$Loss_{bonding\ box} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left\{ \left(x_i - \hat{x}_i\right)^2 + \left(y_i - \hat{y}_i\right)^2 \right\} + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left\{ \left(\sqrt{w_i} - \sqrt{\hat{w}_i}\right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i}\right)^2 \right\}$$

$$Loss_{confidence} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left(C_i - \hat{C}_i\right)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left(C_i - \hat{C}_i\right)^2$$

$$Loss_{classification} = \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in classes} \left(p_i(c) - \hat{p}_i(c)\right)^2$$

In the first version of YOLO [28], the network architecture was designed based on the GoogLeNet[29] architecture for image classification. YOLOv1 network consists of 24 convolutional layers and 2 fully connected layers. It replaced GoogLeNet's inception modules with 1 x 1 reduction layers and 3 x 3 convolutional layers respectively. The first 20 convolutional layers of the network were pretrained on the classic ImageNet 1000-class competition dataset with the input size of 244 x 244 to have 88% top-5 accuracy for a week. The PASCAL Visual Object Classes (VOC) in 2007 and 2012 were then used for training and validation. On this dataset, 98 bounding boxes per images are predicted. The weak predictions are then filtered out by the Non-maximum Suppression (NMS) technique. The overall training pipeline was originally implemented on the Darknet framework. YOLOv1 could achieve 63.4 mAP (mean average precision) and 45 FPS (frames per second). This achievement implicated that YOLO could perform real-time performance while the detection accuracy was as comparable as Faster R-CNN. To improve YOLO's detection speed, Fast YOLO was introduced [30]. It optimized the original YOLO architecture by decreasing the convolutional layers from 24 to 9 and using fewer filters in those layers. Therefore, it could reach up to 155 FPS, but the accuracy dropped to 52.7% mAP.

For considering the generalizability, YOLOv1 performed well on PASCAL VOC 2007 and it seemed to be the best method when evaluated on the artwork dataset which contained the challenging difficulty on a pixel level. Although YOLOv1 could achieve significant achievement over other methods, it also struggled with some limitations. Since the classification and localization network of this version of YOLO could detect only one object, any grid cell could detect one object too. This constraint resulted in the limit maximum number of objects, which was a total of 49 objects per one detection for 7x7 grid cells. As a result, it caused relatively high localization error especially the small objects that close to each other, such as groups of pedestrians. Moreover, this network architecture found difficulty in object generalization when evaluating on the other input dimensions instead of 244 x 244. To overcome these YOLOv1 constraints and achieve better performance, the second version of YOLO (YOLOv2) was then developed [31]. YOLOv2 concerned mainly about maintaining the prior classification performance while it also worked on improving recall and localization. Thus, it applied a variety of significant modifications to increase mean average precision (mAP). By adding batch normalization on the convolution layers, the model showed drastic impact on the convergence. In detail, the activations in the hidden layer were shifting and scaling which leaded to improving model stability and also reducing the overfitting issue. This technique could raise YOLO's mAP up to 2% and has been widely used for the model regularization propose. Consequently, YOLOv2 also proposed the high-resolution classifier from 224x244 in the first version to 448x448. Initially, the model was trained on the 224x224 input images, then the classification network was fine-tuned with 448x448 resolution on the ImageNet dataset for 10 epochs before detection training. This process gave a room for kernel adjustment on higher resolution images, which therefore resulted in better detection performance and reached slightly below 4% mAP improvement. Moreover, YOLOv2's researcher has introduced the remarkable progress, motivated by Faster R-CNN, which was the usage of anchor box. Instead of using fully-connected layers for bounding box prediction on the feature map, the convolutional layers were modified to predict locations of anchor boxes which leaded to recall increasing, but slightly degraded the overall mAP. The bounding box prediction could then relative to these anchor boxes, which was a small area consisted of width and height, unlike in YOLOv1 that relative to full image dimension. As mentioned in Faster R-CNN, the initial sizes of anchor boxes were chosen by hand-picking method. In contrast, YOLOv2 developer designed the suitable anchor box dimensions using k-mean clustering. The distance metric was calculated based on IoU scores, which is shown in (15).

$$\left\| x, c_i \right\| = 1 - IoU(x, c_i) \tag{15}$$

where,

x is a ground truth candidate box.

c_i is one of the centroids.

The YOLOv2 bounding box prediction spatially related to the offsets of the anchors. It predicted 5 parameters including tx, ty, tw, th and to, then applied the sigma function to limit the offset range of an output, which was calculated by (16).

$$\begin{aligned} b_x &= \sigma(t_x) + c_x \\ b_y &= \sigma(t_y) + c_y \\ b_w &= p_x e^{t_w} \\ b_h &= p_h e^{t_h} \end{aligned}$$

$$Pr(object) * IoU(b, object) = \sigma(t_0) \tag{16}$$

where,

t_x, t_y, t_w, t_h are the predicted offset and scale.

c_x, c_y are the top left corner of the anchor grid cell.

p_w, p_h are the size of the anchor box.

b_x, b_y, b_w, b_h are the predicted box center and size.

$\sigma(t_0)$ is the box confidence score.

These anchor boxes generated by clustering technique could provide higher average IoU. YOLOv2 divided the input into 13x13 grid cells by removing a pooling layer that was later more helpful with larger or smaller items. The goal was also to

overcome a small object limitation in YOLOv1. Furthermore, this version of YOLO model also integrated the multi-scale training strategy that could strengthen the model robustness to various input image sizes. Every 10 batches during training process, the network was resized according to the new random image dimension which was a multiple of 32 dues to convolutional layer down sampling property. Lastly, for reducing the computation time, YOLOv2 used a more light-weighted base model architecture, so called Darknet-19, with 19 convolutional layers and 5 layers of max-pooling. In addition, YOLOv2 also had the extended version which could detect more than 9000 classes, also known as YOLO9000 [31]. YOLO9000 was a slightly modified version of YOLOv2, which considered in merging small object detection dataset with large ImageNet dataset. Since the overlapping class labels of these two datasets could not combine directly, the WordTree hierarchical model was demonstrated for this difficulty. The general labels were placed in the top closer to the root and the fine-grained labels were branched similar to leaves. The probability of each class could calculate by following the path from that node to the root.

YOLOv3 was another improved version of YOLO family which outperformed its predecessor by using the state-of-the-art algorithms and also was practical with real-time scenarios [32]. Firstly, in order to predict the objectness confidence score, YOLOv3 proposed the logistic regression for predicting the confidence score for each bounding box, while the prior YOLOs used the sum of the square errors. Next, in some dataset like ImageNet, not all the labels were mutually exclusive. In such a case, choosing the Softmax function as the activation function of the output layer might not be the best choice, since it just converted predicted scores to distributed probabilities which summed up to one. For instead, YOLOv3 used multiple independent logistic classifier. The input image could then have multiple labels and contained both mutually and non-mutually exclusive objects. The complexity of the YOLOv3 model was also decreased by getting rid of the Softmax function. Consequently, with the new activation functions, the classification loss function must be modified into the binary cross-entropy instead of the Mean Square Error (MSE). Inspired by the idea of the Feature Pyramid Network (FPN), YOLOv3 could then predict the object up to three different box scales for every location in the input image and then perform feature extraction. This technique improved the detection performance on various object scales and was a significant solution especially for detecting small things. The Feature Pyramid Network design was illustrated in Figure 8. Furthermore, YOLOv3 also adopted cross-layer connections between two prediction layers and the fine-grained feature maps. Initially, the coarse-grained feature map was up-sampled, then the concatenation was performed to combine it with the previous one. Moreover, the YOLO's feature extractor was modified from Darknet-19 to Darknet-53, which based on the 53 convolutional layers. Darknet-53 was much deeper and composed of mostly 3x3 and 1x1 kernels including skip connections just like residual network in ResNet. In comparison, it has less BFLOP (billion floating point operations) than ResNet-152, but could still maintain the same detection accuracy at two-times faster. The full YOLOv3 architecture was illustrated in Figure 9.

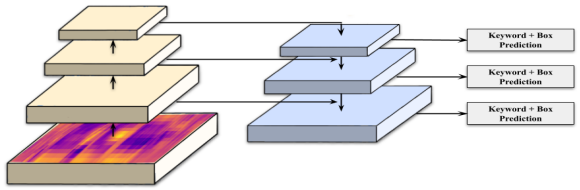


Figure 8: Feature Pyramid Network.

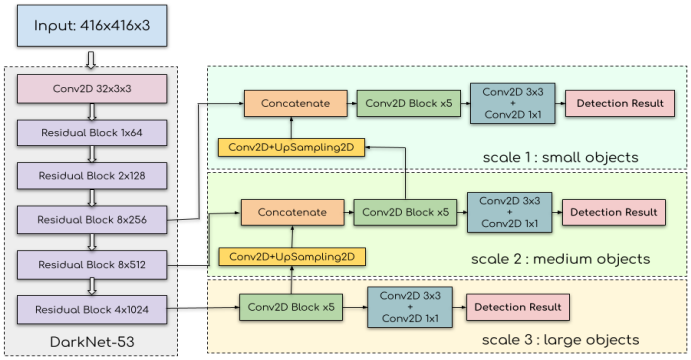


Figure 9: The YOLOv3 architecture.

YOLOv3 may therefore conduct high-detection capabilities for object detecion. In YOLOv3 paper, the researcher argued that YOLOv3 might exceed three times the speed of the Single Shot Multibox Detector (SSD). While RetinaNet was more precise, it also battled with the computational time. In small object criteria, YOLOv3 has also shown substantial progress.

YOLOv3 also had the lightweight version which was known as the tiny-YOLOv3 model. It used the same training strategy as the full version, however, the minor modifications were applied with the full YOLOv3 model architecture to gain a lot more computation speed and consume fewer computational resources for inference. This model was ideally optimized for the use of the low resource platform such as embedded systems and mobile devices. Firstly, the tiny-YOLOv3 simplified the number of convolutional layers. It used only 7 convolutional layers. This tiny version just employed a few numbers of 1x1 and 3x3 convolutional layers for extraction of features. It replaced a convolution layer with the pooling layer with a step size of 2 by focusing on summarized the complex dimension of the standard model. Although it used a slightly different architecture with the full YOLOv3, the model base structure was still the same which composed of 2D-convolutional, batch normalization and Leaky Rectified Linear Unit (ReLU).

In our research, the speech signal can now be handled as an ordinary image by transforming the recorded voice signal into the MFCC feature image. Then, Thai spoken keywords that contain with in the voice image could be compared as objects. Therefore, the object detection technique could be applied to an automatic speech recognition task. In order to understand the Thai speaker purpose, we applied this core methodology to find the corresponding Thai keywords in the voice image by feeding the whole sentence MFCC and matched them with our predefined sentence categories which each group of Thai keywords related to. Moreover, we propose that the localization ability of the object detector could be deployed especially with some input spoken sentence, which relies on the restricted keyword ordering such as

the unique number sequence in the flight number. As mentioned previously, YOLOv3 has been proven to be fast and accurate. It was the most stable and reliable version of YOLO since we started to research this work. We also compared the detection performance between the full-YOLOv3 and its tiny version. Since tiny-YOLOv3 was extremely fast, some mean average precision (mAP) was traded off. Our keyword detection system was illustrated in Figure 10.

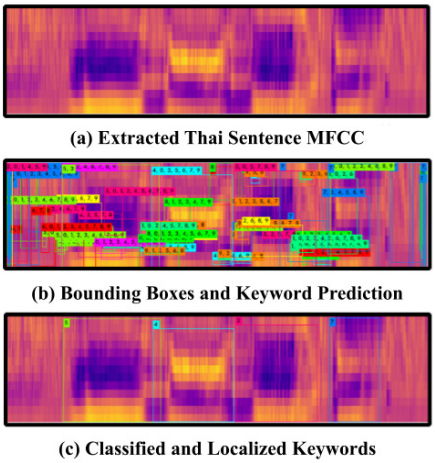
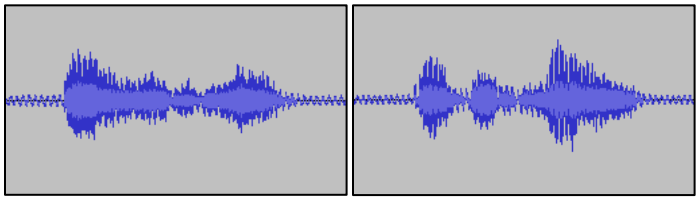


Figure 10: The keyword detection overview with Thai sentence MFCC feature image as an input.

5. Voice Dataset Collecting and Preprocessing

This article draws on speech datasets made by 60 native Thai speakers and recorded in several places proposed to simulate the variety of the background noises. Our selected participants were varied in ages and the genders were equally distributed. In terms of age, the highest number of overall participated subjects was in teenage (20s), which was 18 people (30%), and those aged between 40 and 49 were the smallest, which was 7 people (11.6%). The total datasets contained 2400 WAV files with 40 unique keywords. Table 1 summarized the baseline characteristic of our dataset. At the beginning, each recorded voice was processed using the Audacity software before converting into MFCC feature image, which was shown in Figure 11 and 12.



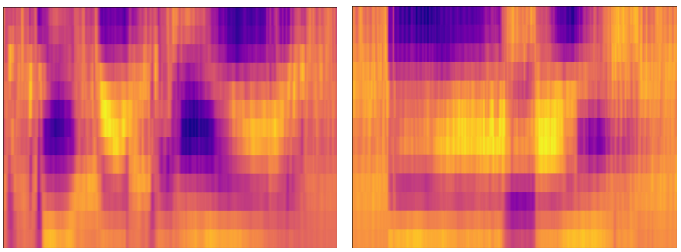
(a) (b)

Figure 11: The examples of recorded voice: (a) · (Air Asia); (b) (Thai Airways)

In order to avoid overfitting and improve the model robustness, some data augmentation methods including the tempo and pitch randomizing were applied to increase the quantity of training data, therefore finally finishing up with a total of 9600 WAV files. The datasets were obtained in accordance with the passenger information machine criteria for the research of the automatic speech recognition (ASR) system in actual settings. The automated system required to give the reliable information, which correlate to the passenger's flight number as a real-time answer to the passenger's frequently asked questions (FAQs). In the aviation industry, a flight number is composed of a two-character airline designator and numbers with a maximum length of four digits. Thus, we proposed three datasets consisted of airline names, numbers and FAQs keywords. The first dataset was selected from the ten popular airlines at the Suvarnabhumi airport in Thailand. Next, the second dataset consisted of ten numbers (0-9). The final dataset contained 20 keywords which could categorize into 7 passenger FAQ topics.

Table 1: The baseline characteristic of our dataset.

Characteristics	Training (80%)	Validation (20%)	Total
Population	48	12	60
Sex			
Male	24	6	30
Female	24	6	30
Age Group (Y)			
10-19	9	2	11
20-29	15	3	18
30-39	6	2	8
40-49	6	1	7
50-59	6	2	8
60-69	6	2	8
.WAV files			
Raw Dataset	1920	480	2400
Raw Dataset + Augmentation	7680	1920	9600



(a) (b)

Figure 12: The examples of extracted MFCC image: (a) · (Air Asia); (b) (Thai Airways).

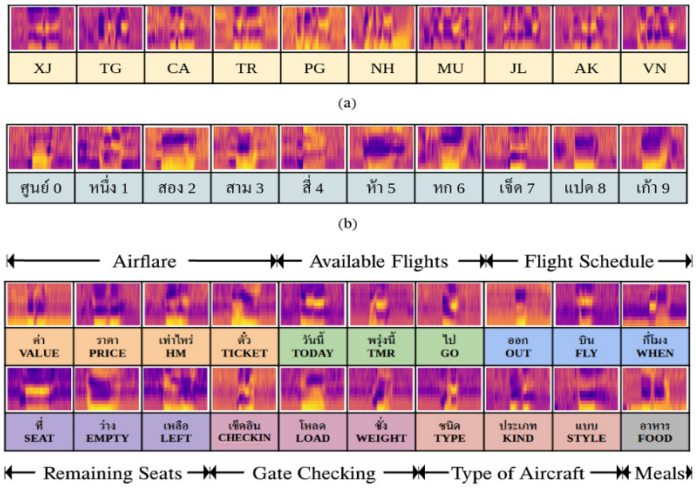
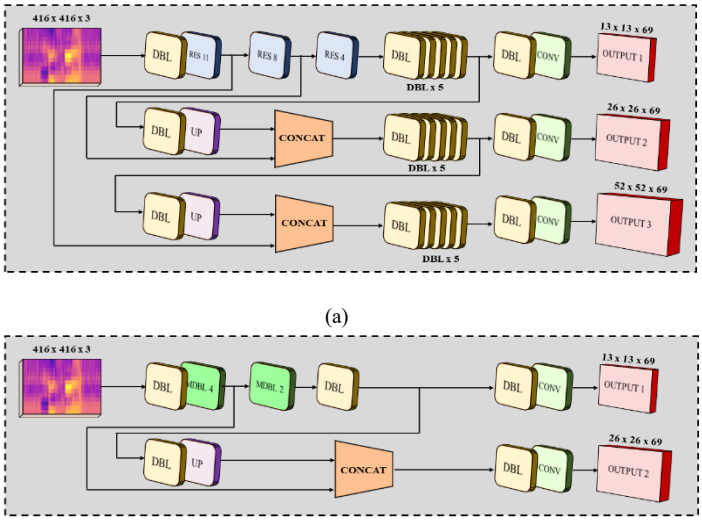


Figure 13: The overview of three datasets: (a) airline designators; (b) numbers 0-9; (c) FAQs keywords

Figure 13 showed the overview of three datasets. In every dataset, the images were randomly split with a ratio of 80:10:10, which included training, validation, and also test sets for further evaluation.

6. Highlights of the Research Work

In Thai linguistics, the absence of large speech corpora and the language morphology could inhibit development and researches in the speech recognition field as mentioned on the introduction section. Although several organizations like NEXTEC concentrated on eliminating this limitation, the complicity of the spontaneous voice input, especially in a noisy environment, seemed to be a problem for many researchers. Through the keyword spotting technique, we were interested in overcoming this problem. This technique could not only filter the sentence redundancies but also be easily maintained in data-scarce domains since we simply need only significantly keywords to train the deep learning model. In order to expand the keywords for more use cases, without retrained the large model like traditional approach, it could be done easily by create new lightweight model with small dataset and included into the latest model which would be clearly described in the experiment section.



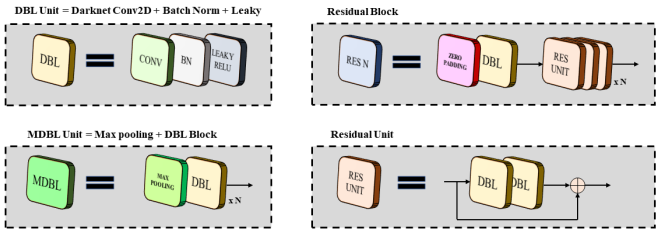


Figure 14: Model architectures: a) Full YOLOv3 architecture; b) Tiny YOLOv3 architecture; c) Legend

In more details, our manuscript has described how to develop the basic and lightweight ASR as the alternative approach of the classical ASR that might be effective in certain complex spoken sentences or languages like Thai, and could be employed in many scenarios where the real-time requirement is crucial. Our data collection is juts utilized to show the specific occurrences that can be applied to our proposed strategy. By applying computer vision technique to the linguistics field, the state-of-the-art one-stage object detector, You Only Look Once, is chosen and responsible for keyword classification and localization due to its detection performances. YOLOv3 also released the lightweight and optimized version of its full architecture as Tiny YOLOv3. Since this version of YOLOv3 is appropriate for performing the real-time computing on the resource-limited platform like edge devices, we also decided to compare this version of YOLOv3 on the same task with the regular YOLOv3. The differences between the full YOLOv3 and tiny YOLOv3 structures are described on Figure 14. The training strategy of our models are clearly explained on the following section.

7. Network Training Method

Since object detection was the supervised learning, we needed to provide the ground truth to the network. Therefore, before the training process, each target keyword on the MFCC feature images of the training set needed to be manually labeled. In this phase, the LabelImg tool has been used because the YOLO annotation format has already been provided. Consequently, the XML file according to each MFCC image would be generated to indicate the necessary labeling information for the network training step. In detail, the actual labeling steps were holding the mouse cursor, framing the target keyword region, and then double-clicking to identify the corresponding keyword class. The YOLO annotation format in XML file was composed of class index, the quotient of x-coordinate of the bounding box's center and the image width, the quotient of y-coordinate of the bounding box's center and the image height, the quotient of the bounding box width and the image width, and the quotient of the bounding box height and the image height. Due to the fact that the keyword in the MFCC feature image had the unique pattern, then it is possible to readily separate the specific keyword region from the uniform background noise region. The example of keyword labeling was described in Figure 15.

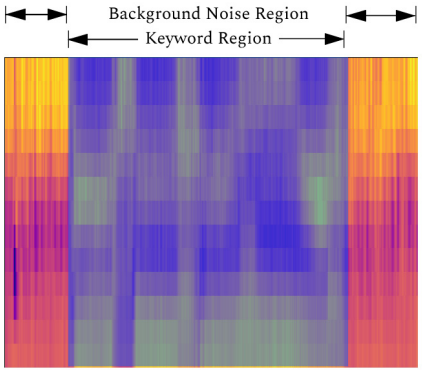


Figure 15: The example of keyword labeling.

The training platform in this article were carried out on the desktop PC (OS:Ubuntu 16.04LTS system, RAM: 16 GB, GPU: Nvidia GeForce GTX 1070Ti, CPU: Intel Xeon 8 cores). Both regular and tiny version of YOLOv3 were trained under the Darknet framework then the training results were analyzed to prove that Tiny YOLOv3 performance was acceptable for the proposed method. The stochastic gradient descent algorithm was used to perform the network training. In order to start training the model, we chose the pre-trained models of the ImageNet dataset as the initial convolutional weights. By considering the training durations, the maximum number of iterations was extended and the training process would be stopped whenever the loss was low enough. We set the Batch parameter to 64 and the Subdivision parameter to 16. As a result, the batch was split into 8 mini-batches which could not only reduce the memory usage, but also accelerate the training process and also generalized the model. Next, the learning rate was chosen to be 0.001. During training, the average loss and mean average precision (mAP) could be visualized in real-time using the graph as shown in Figure 16a. The final results were shown in Table 2.

Table 2: The models' performance of all datasets

Weights	Iterations	Loss	mean Average Precision (mAP)
Airline names			
Tiny YOLOv3	49000	0.12	0.99
Full YOLOv3	21400	0.005	0.99
Numbers (0-9)			
Tiny YOLOv3	50000	0.16	0.98
Full YOLOv3	34800	0.012	0.98
Frequently Asked Questions			
Tiny YOLOv3	287000	0.08	0.99
Full YOLOv3	41100	0.029	0.98

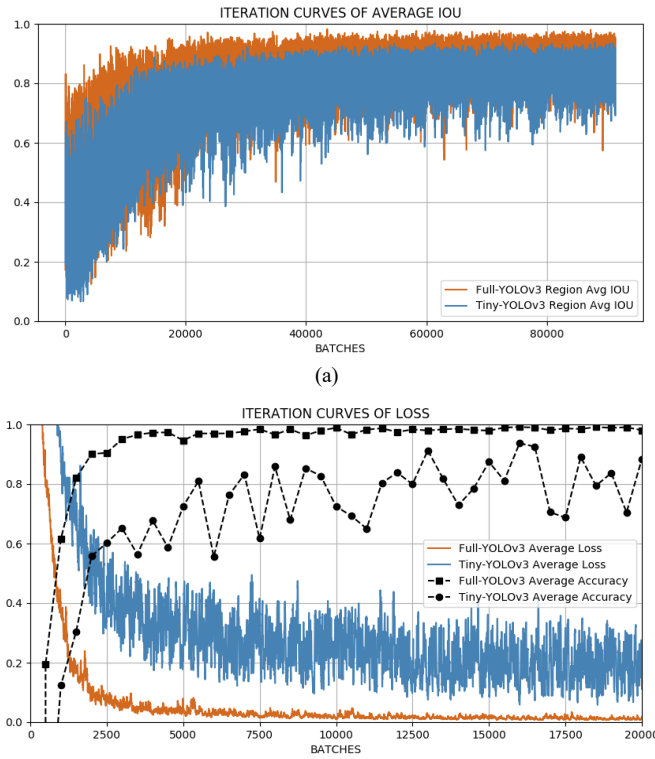


Figure 16: Iteration curves: (a) Iteration curves of loss; (b) Iteration curves of average IoU

As shown in Figure 16, although the fitting degree and the average IoU value of the full version of YOLOv3 was superior compared with the tiny version of YOLOv3, the light-weight version of YOLOv3 appeared to have well-enough training result and could be employed for further experimental evaluations.

8. Experiment and Analysis of Results

In order to investigate the true performance of the finally trained YOLOv3 weights, we considered using the unseen MFCC feature images extracted from the extra recorded voices of 20 individuals (10 males, 10 females) who didn't present in the prior training lists. This experiment was conducted to represent the airport use case of our proposed approach in real life. Whether planning to fly locally or internationally, an airport is always the first priority to asking for the flight information and useful navigation suggestion. From that point of view, we then proposed the autonomous information providing system that could respond to the frequently asked questions (FAQs) of the flight information from the passengers based on the flight number they told in real time. This service might provide passengers all the information they need on their flight only with their voice. This service is therefore straightforward and convenient to use, and additional services such as self-check-in can be integrated and the time and costs of congested and crowded airports can be reduced. For the first task, the automatic speech recognition system was used for detecting the airline names. Then, the 1-to-4-digit numbers would be recognized by the second task. The flight number was then confirmed by the combination of airline name and airline number. In Figure 17, we illustrated some examples of this task. Firstly, our proposed system expected the airline name which was "¢"ƒ‡¹‰" (kan – bin - thai) or Thai Airways (TG) on the upper

left of figure. Next, it would ask for airline number which was "Õç ÃÃ¬ ¡Ã" (haa - see - saam - jet) or 5-4-3-7 on the upper right of figure. The flight number, TG5437, was then generated by combining these two inputs which showed on the ticket.

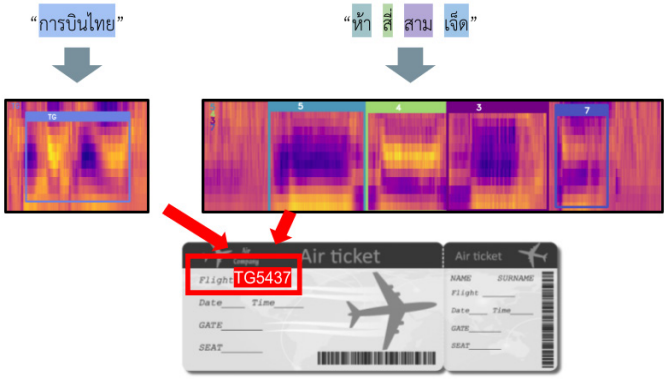


Figure 17: The example of the flight number recognition

Next, the third task was responsible for answering the frequently asked questions (FAQs) according to the provided passenger flight number. The third task example was described in Figure 18 respectively. In Figure 18, we presented some tasks from all seven the frequently asked questions (FAQs) categories of the flight information from the passengers. In the left side of figure, the input interrogative sentence "ราคาเท่าไหร่?" (la-ka-tao-lai) meant "how much is a ticket?". In this sentence, it contained two keywords that related to airflare category (described in Figure 13) which were ราคา (PRICE) and เท่าไหร่ (HM), then the system responded with all available ticket prices. Next, in the right side of figure, the input interrogative sentence "เช็คอินที่ไหน?" (check-in-tee-nai) meant "where is my check-in gate?". This sentence contained one keyword which was เช็คอิน (CHECKIN), thus the system knew that it related to gate checking question and performed the location guidance service.

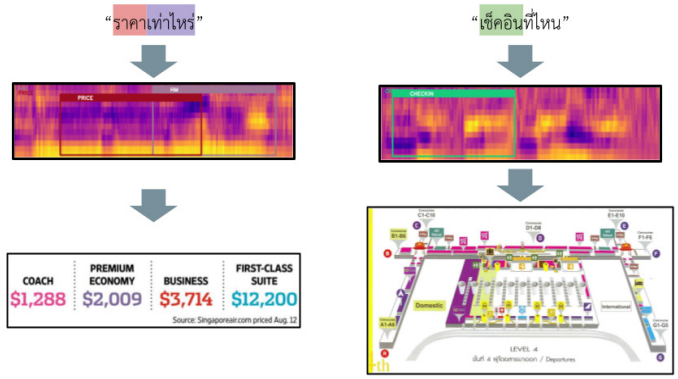


Figure 18: The example of the FAQs recognition

In this experiment, Precision, Recall and F1 score were used as the evaluation metrics. Precision is the proportion of the amount of correctly detected keywords over the entire number of detected keywords. Recall is the proportion of the amount of correctly detected keywords over the entire number of keywords in the experiment set. The F1 score is the harmonic mean of Precision and Recall at particular thresholds. The calculation formulas were shown in Equations (17)-(19).

$$Precision = \frac{TP}{TP + TF} \tag{17}$$

$$Recall = \frac{TP}{TP + FN} \tag{18}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{19}$$

Specifically, TP, or True Positive, indicates the amount of correctly detected keywords. TN, or True Negative, indicates the amount of correctly detected backgrounds. FP, or False Positive, indicates the number of incorrect detections. FN, or False Negative, indicates the number of missed detections.

The full and tiny version of YOLOv3 were trained and tested separately on three different tasks. Airline name detection was the simplest task and the FAQs recognition was the most challenging task. In detail, Airline name detection required just the one utterance, which might contain just one name, and so resulted in high precision and recall in both tiny and full YOLOv3 model. The next challenge was that of the airline number which required the sequence of 1-to-4-digit numbers, thus it could be seen that the accuracies of both models were slightly dropped. The last assignment was the most difficult, which the input sentence could be varied depending on the speaker experiences. To tackle this challenge, we used the keyword spotting idea to search for only the related terms we defined and categorized. As a result, it seemed that the detection performances of both networks have declined dramatically yet remained over 0.6. The detection threshold and IoU in this experiment were set at 0.20 and 0.50 respectively. The true positive, false positive, false negative, precision, recall, and F1-score results for all test samples were summarized in Table 3. The comparison of the F1-score of both full and tiny YOLOv3 on three different tasks could then be visualized in Figure 19. The test results showed that the first task was highly sensitive for both models, airline name detection, which was the easiest task and then experienced poor detection result on the last task, the FAQs recognition task, the hardest task as we mentioned previously. By compared with the full version of YOLOv3, the light-weight and simplified vision of YOLOv3, so-called tiny YOLOv3, could receive the comparable performance and could unusually outperform in terms of Recall by almost 0.1 on the second task. To clarify the poor performance of both models on the third task, the FAQs recognition task, two potential reasons were given as following. First, the number of classes was doubled in comparison with the first and second tasks and the non-keyword parts might be included in a spoken sentence. Second, the input MFCC feature dimension may vary by the length of the sentence and the words can be compressed when the input images have been scaled, leading to feature loss and false negative.

Table 3: Performances of both models on testing dataset

Weights	IoU = 0.50, Threshold = 0.20					
	TP	FP	FN	Precision	Recall	F1
Airline names						
Full YOLOv3	91	2	1	0.978	0.989	0.984
Tiny YOLOv3	90	6	0	0.938	1	0.968
Numbers (0-9)						
Full YOLOv3	192	39	70	0.831	0.733	0.779
Tiny YOLOv3	221	58	48	0.792	0.822	0.807
Frequently Asked Questions						
Full YOLOv3	121	39	83	0.756	0.593	0.664
Tiny YOLOv3	108	61	74	0.639	0.593	0.615

Based on the detection performances shown in Table 3, the average F1-score of the full and tiny YOLOv3 from three different tasks were 77.54% and 77.23% respectively. In terms of localization performance using the same box labels, the tiny YOLOv3 boxes seemed to have slightly shifted due to the simpler model architecture, which was shown in Figure 20, however, it was not affected the keyword detection proficiency. In conclusion, the experiment results showed that the tiny YOLOv3 could perform well enough and have competitive results with the full YOLOv3.

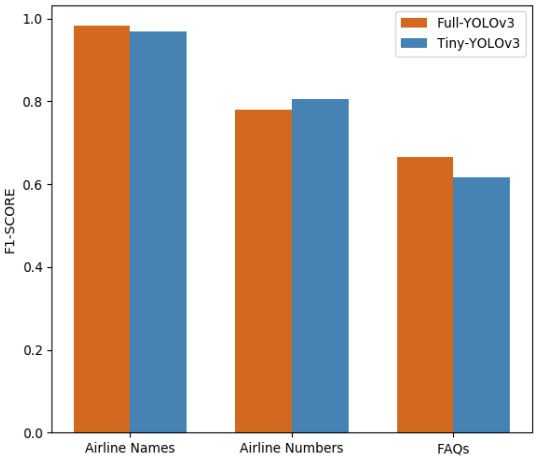
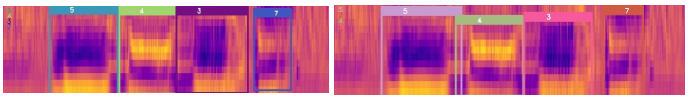


Figure 19: The comparison of three task detection results



(a) (b)

Figure 20: The comparison of the localization accuracy: a) Full YOLOv3; b) Tiny YOLOv3.

Furthermore, in order to test the model robustness, we injected the white noise with the relative amplitude α from 0.1 to 0.4 to each of the audio files using the Audacity software. In accordance with the white light which contains all wavelengths with equal intensity of the visible spectrum, white noise also contains the equally distributed energy in all audible frequencies resulted in a steady humming sound. The F1 score and accuracy measured after increasing the white noise intensity, which belonged to the regular and tiny YOLOv3 were shown in Figure 21, in which the y-axis was the measure and the x-axis was the intensity (α) of the white noise added to the audio files.

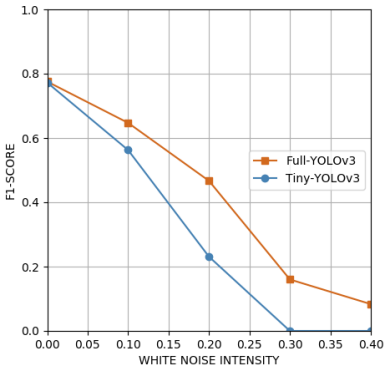
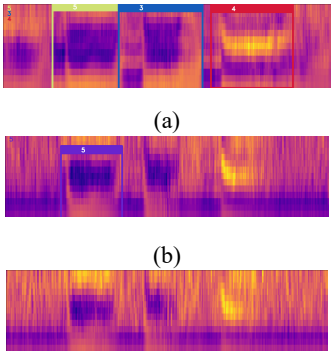


Figure 21: The performances of the full and tiny YOLOv3 when injecting the background noise.

It could be seen from Figure 22 that the full YOLOv3 model had the better adaptability than the tiny YOLOv3 to the noisy environment. Although the higher degrees of noise intensity could significantly degrade the model performances, the F1-score above 0.65 could still be obtained by both models when the variances of the white noise were below 0.05.



(a)

(b)

(c)

Figure 22: The comparison of the detection results:

a) α=0.1; b) α=0.2; c) α=0.3.

9. Conclusions and future work

This study provided the alternative approach for the automatic speech recognition (ASR) was proposed to detect Thai keywords in the spoken sentence. We next proposed the three separate real-world tasks to examine our model detection capabilities. The tasks were different in terms of difficulty, ranging from the standard test that just recognized the Thai Airway names, to the most difficult assignment that detects the specific keywords in the frequently asked questions (FAQs) using the keyword spotting technique. The core methodology is based on the Mel-frequency cepstral coefficients (MFCC), which is chosen as the feature extraction for

the speech signal, because this technique artificially implements the behavior of the human hearing mechanism based on the usage of the non-linear frequency scale of the real human, so it results in the parametrically resemblances between the extracted vectors and human sense of hearing. Then the state-of-the-art convolutional neural network object detector, You Only Look Once (YOLO), was performed as the keyword localizer and classifier. Due to the requirements of real-time speed and high accuracy of the ASR system, the tiny version of YOLOv3, which was the simplified and light-weight version of YOLOv3, was evaluated and compared with the regular version using the precision, recall, and F1-score to verify the feasibility and superiority. From the experiment, the F1-score of the full and tiny YOLOv3 from three different tasks were 77.54% and 77.23% respectively. To conclude, tiny YOLOv3, with the lower computational time and comparable detection accuracy compared to the regular YOLOv3, was proven to meet the ASR requirements and was the most suitable model for using in the low resource platforms.

Thai ASR was still the interesting and challenging topic because of the morphological richness of Thai language and the difficulties in developing Thai ASR model using the traditional technique. The YOLOv3 was applied to the keyword detection. Nevertheless, the experimental results drew back some further applications and additional improvements.

1. Compared with the results on the first and second task, the poor result on the third task told that the dataset should to be increased to produce higher performance.
2. To increase the model robustness to the background noise, injecting noise into dataset and trained together with the original dataset should be useful.
3. This work only proposed the Thai ASR application in the real-world airport scenario. Therefore, the research of the same basis should be increased to extend the application or investigate the new approaches.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] K. Sukvichai, C. Utintu, and W. Muknumporn, "Automatic Speech Recognition for Thai Sentence based on MFCC and CNNs," 2021 Second International Symposium on Instrument, Control, Artificial Intelligance, and Robotics (ICA-SYMP), 2021. doi:10.1109/ica-symp50206.2021.9358451

[2] H. Thaweesak Koanantakool, T. Karoonboonyanan, and C. Wutiwiwatchai, "Computers and the Thai Language," IEEE Annals of the History of Computing, **31**(1), 46-61, 2009. doi:10.1109/mahc.2009.5

[3] T. Pathumthan, ''Thai Speech Recognition Using Syllable Units,'' master's thesis, Chulalongkorn Univ., 1987.

[4] S. Suebvisai, P. Charoenpornsawat, A. W. Black, M. Woszczyna, and T. Schultz, ''Thai Automatic Speech Recognition,'' Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP), 2005. doi:10.1109/icassp.2005.1415249

[5] I. Thienlikit, C. Wutiwiwatchai, and S. Furui, ''Language Model Construction for Thai LVCSR,'' Reports of the Meeting of the Acoustic Soc. of Japan, ASJ, 131-132, 2004.

[6] M. Woszczyna, P. Charoenpornsawat and T. Schultz. "Spontaneous Thai Speech Recognition." The Ninth International Conference on Spoken Language Processing Multimodal Technologies, Inc. Carnegie Mellon University, Australia, 2006.

[7] S. Kasuriya, V. Sornlertlamvanich, P. Cotsomromg, S. Kanokphara, and N. Thatphithakkul, ''Thai Speech Corpus for Speech Recognition,'' Proc. Int'l Conf. Int'l Committee forthe Coordination and Standardization of Speech Databases and Assessments (O-COCOSDA), 105-111, 2003.

[8] S. Tangruamsub, P. Punyabukkana, and A. Suchato, "Thai Speech Keyword Spotting using Heterogeneous Acoustic Modeling," 2007 IEEE International Conference on Research, Innovation and Vision for the Future, 253-260, 2007. doi:10.1109/rivf.2007.369165

[9] P. Khunarsa, J. Mahawan, P. Nakjai, and N. Onkhum, "Nondestructive Determination of Maturity of the Monthong Durian by Mel-frequency Cepstral Coefficients (MFCCs) and Neural Network," Applied Mechanics and Materials, **885**, 75–81, 2016. doi:10.4028/www.scientific.net/amm.855.75

[10] Y. Segal, T. S. Fuchs, and J. Keshet, "SpeechYOLO: Detection and Localization of Speech Objects," Interspeech 2019, 2019. doi:10.21437/interspeech.2019-1749

[11] G. Chen, C. Parada, and G. Heigold, "Small-footprint Keyword Spotting using Deep Neural Networks," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4087–4091, 2014. doi:10.1109/icassp.2014.6854370

[12] J.R. Rohlicek, W. Russell, S. Roukos, and H. Gish, "Continuous Hidden Markov Modeling for Speaker-independent Word Spotting," International Conference on Acoustics, Speech and Signal Processing (ICASSP), 627–630, 1990. doi:10.1109/icassp.1989.266505

[13] D. Grangier, J. Keshet, and S. Bengio, "Discriminative Keyword Spotting," Automatic Speech and Speaker Recognition, 173–194, 2009. doi:10.1002/9780470742044.ch11

[14] S. Fernández, A. Graves, and J. Schmidhuber, "An Application of Recurrent Neural Networks to Discriminative Keyword Spotting," Artificial Neural Networks – ICANN 2007, 220–229, 2007. doi:10.1007/978-3-540-74695-9_23

[15] T. Sainath and C. Parada, "Convolutional Neural Networks for Small-Footprint Keyword Spotting," 2015.

[16] D. Namrata, "Feature Extraction Methods LPC, PLP and MFCC in Speech Recognition," International Journal for Advance Research in Engineering and Technology (ISSN 2320-6802), 1, 2013

[17] R. Ranjan, A. Thakur, "Analysis of Feature Extraction Techniques for Speech Recognition System," International Journal of Innovative Technology and Exploring Engineering (IJITEE) (ISSN 2278-3075), 8. 2019

[18] Z.Q. Zhao, P. Zheng, S.T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," IEEE Trans. Neural Networks Learn. Syst., 1-21, 2019.

[19] R.L. Galvez, A.A. Bandala, E.P. Dadios, R.R.P. Vicerra, and J.M.Z. Maningo, "Object Detection Using Convolutional Neural Networks," TENCON, 2018. doi:10.1109/tencon.2018.8650517

[20] H. Jiang and E. Learned-Miller, "Face Detection with the Faster R-CNN," 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition, 2017. doi:10.1109/fg.2017.82

[21] Z.-Q. Zhao, H. Bian, D. Hu, W. Cheng, and H. Glotin, "Pedestrian Detection Based on Fast R-CNN and Batch Normalization," Lecture Notes in Computer Science, 735-746, 2017. doi:10.1007/978-3-319-63309-1_65

[22] R. Girshick, "Fast R-CNN," 2005 IEEE International Conference on Computer Vision (ICCV), 2015. doi:10.1109/iccv.2015.169

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," IEEE Trans. PatternAnal. Mach. Intell., **37**(9), 1904–1916, 2015.

[24] J. R. R. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, "Selective Search for Object Recognition," International Journal of Computer Vision, **104**(2), 154–171, 2013. doi:10.1007/s11263-013-0620-5

[25] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, **39**(6), 1137-1149, 2017. doi:10.1109/tpami.2016.2577031

[26] T.-Y. Lin, P. Doll´ar, R. B. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. doi:10.1109/cvpr.2017.106

[27] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single Shot Multibox Detector," in Lecture Notes in Computer Science, 21-37, 2016. doi:10.1007/978-3-319-46448-0_2

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. doi:10.1109/cvpr.2016.91

[29] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov,

D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. doi:10.1109/cvpr.2015.7298594

[30] M. J. Shafiee, B. Chywl, F. Li, and A. Wong, "Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video," Journal of Computational Vision and Imaging Systems, **3**(1), 2017. doi:10.15353/vsnl.v3i1.171

[31] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition, 7263-7271, 2017. doi:10.1109/cvpr.2017.690

[32] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 279-290 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Comparative Analysis and Modern Applications of PoW, PoS, PPoS Blockchain Consensus Mechanisms and New Distributed Ledger Technologies

Caglar Arslan*, Selen Sipahioğlu, Emre Şafak, Mesut Gözütok, Tacettin Köprülü

HAVELSAN Inc., Ankara, 06510, Turkey

ARTICLE INFO

Article history:
Received: 09 July, 2021
Accepted: 01 October, 2021
Online: 14 October, 2021

Keywords:
Blockchain
Distributed Ledger Technology
Crypto Finance
Electronic Voting
Supply Chain
Digital Notary
Insurance
Records Management
Distributed File Systems
Artificial Intelligence
Federated Learning

ABSTRACT

Central authorities have registered economic transactions based on trust since the dawn of time. With the advent of paper, methods for documenting transactions and records became more difficult and detailed. With the widespread use of computers in recent years, digital recording has made record keeping easier and since then every technical advancement has made data recording more convenient and faster. Transactions have accelerated as technology has advanced, but the maintenance of records has remained under the jurisdiction of the authorities. The use of distributed ledger systems allows data to be released from the control of central authorities. A distributed ledger is a peer-to-peer network in which non-centralized data is shared by participants. The Bitcoin payment system and cryptocurrency, which arose after the 2008 financial crisis, was the first distributed ledger application. Central authorities' poor economic decisions aided in the rise of Bitcoin. The aim of Bitcoin is to make the current financial systems liberal from the influence of central authorities. With the rise in popularity of Bitcoin, the blockchain technology that underpins it has begun to garner interest. While blockchain technology was initially associated with the financial sector due to Bitcoin, research into its use in other sectors such as supply chain, records management, electoral systems, notary services, file systems, energy and artificial intelligence has begun. Several Blockchain infrastructures have been built to allow the use of blockchain technology in a variety of industries. However, it has been revealed in practical approaches that blockchain technology has limitations in terms of speed and scalability. As a result, new distributed ledger technologies with increased speed and scalability have been established. Hashgraph, Tangle, Tempo, Holochain are examples of newly developed distributed ledger technologies. Different influential features distinguish new generation distributed ledger technologies from the conventional Blockchain methods, which can yield into several and practical modern applications. In this study, Blockchain and new generation Distributed Ledger Technologies are compared and possible future applications are outlined.

1. Introduction

Blockchain is a distributed database technology in which a database is stored and maintained without the use of a central server across an entire network or among similar parties. No one on the network, inside or outside, has the right to modify or remove recorded data. Transparency is also established by disseminating the data to all chain participants by validating the transmission data via a consensus algorithm [1].

A consensus algorithm is a fault-tolerant mechanism used in computer and blockchain systems to achieve the required agreement among distributed processes or multi-agent systems on a single data value or a single network state.

A Blockchain infrastructure, as a type of distributed system composed of multiple entities, requires a variant of consensus algorithm in order to operate. The characteristics of any kind of Blockchain are determined by its internal consensus architecture.

A Blockchain, which is defined by a consensus mechanism, enables us to have an immutable database of transaction record among the participating peers. A sample transaction flow is summarized in 'Figure 1' below in order to gain a better understanding of the underlying working process of Blockchain.

*Corresponding Author: Caglar Arslan, HAVELSAN HQ Mustafa Kemal Mah. 2120. Cad. No: 39 06510 Cankaya Ankara Turkey, +903122195787, Email: caglara@havelsan.com.tr

https://dx.doi.org/10.25046/aj060531

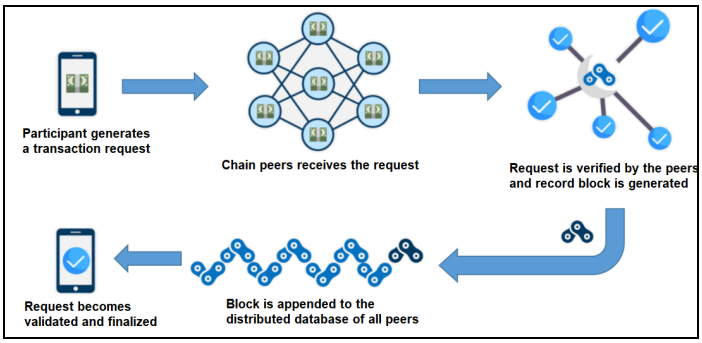


Figure 1: Sample Blockchain Transaction Flow

In summary, the information registry that exists within a group of nodes is made up of multiple blocks of data in which the data is stored and then consolidated within the network after a predefined validation process which is called the consensus.

The necessity of having an immutable distributed record database/table for a digital currency exchange system and applicability of Blockchain to this necessity gave rise to the first useful Blockchain application, which is widely known today as Bitcoin [2]. Since then the blockchain has been synonymous with financial transactions since the rise of Bitcoin, research is currently underway to see whether it can be applied to a variety of areas such as supply chain, records management, energy market, and artificial intelligence.

A variety of consensus algorithms as well as alternative distributed ledger technologies (DLTs), other than the classical Blockchain with the basic Proof-of-Work consensus, have been vastly proposed in recent years each with its own set of performance and security characteristics. It has to be noted that every application's requirements cannot be met by a single consensus algorithm [1]. Hence, it is critical to compare the consistent consensus algorithms as well as other distributed ledger technologies (DLTs) on a technical level in order to identify their strengths, weaknesses, and potential future applications.

Blockchain variants, which are defined by their accompanying consensus algorithms, can have problems regarding throughput, scalability, power consumption and storage usage which can handicap the potential usage areas besides digital currency exchanges/cryptocurrencies. As a result, consistent consensus variants and new distributed ledger technologies (DLTs) should be technically investigated in order to achieve more successful engineering designs of distributed, immutable, manageable, and scalable applications and thus allowing the widespread adoption of trustworthy distributed applications and products in everyday life.

In this study, Proof of Work (PoW), Proof of Stake (PoS), Pure Proof of Stake (PPoS) and new Distributed Ledger Technologies which are named as Hashgraph, Holochain, Tangle and Tempo are introduced and explained. After the introduction and explanations of the mentioned mechanisms, a comparison is made under the performance criteria of Throughput (transactions per second), Scalability (number of nodes), Energy Consumption (J per transaction) and Storage. After this comparison, possible modern applications that can be developed with Blockchain and DLTs are explained and then application-specific minimum performance requirements are given in order to outline the achievability of the applications. Afterwards, suitability of the explained consensus mechanisms and DLTs to each application are explained.

2. Blockchain and Consistent Consensus Mechanisms

Blockchain is a decentralized distributed database technology that acts as a trust layer for the data exchange amongst participating peers.

The following key features define any type of Blockchain: Immutability, transparency, privacy, durability, decentralism, and security.

- Immutability: No participant can change the data stored in the blockchain.
- Transparency: Transparency is provided to all participants in the Blockchain by associating the transaction data with the concerning participants. All transactions are traceable.
- Privacy: The use of private and public keys is a crucial aspect of blockchain privacy. Asymmetric cryptography is used in blockchain systems to secure transactions between users.
- Durability: A Blockchain can operate indefinitely since a copy of the data is distributed across the participating nodes.
- Decentralism: The transfer of control and decision-making from a centralized entity (individual, organization, or group thereof) to a distributed network is referred to as decentralization in blockchain.
- Security: The structure of data created by blockchain technology has built-in security properties. Trust in transactions is ensured by the use of cryptographic principles, decentralization, and consensus. Each new block in a cryptographic chain connects to all the blocks before it in such a way that tampering is nearly impossible. A consensus mechanism validates and agrees on all transactions within the blocks, ensuring that each transaction is true and correct.

It has to be noted that throughout this study, Blockchain is referred as the two most popular consensus mechanisms namely the Proof of Work (PoW) and Proof of Stake (PoS) and a new consensus algorithm that is consistent with the key Blockchain features (i.e. decentralized, secure and scalable) and named as Pure Proof of Stake (PPoS).

2.1. Proof of Work (PoW) Consensus

The PoW consensus mechanism has been popularized by Bitcoin as one of the most widely used methods in blockchain (BTC). Miners and the electricity needed to perform the calculations that verify BTC transactions are the distinguishing features of PoW systems. Miners use computer hardware to run network nodes that use computational power to solve mathematical puzzles known as proofs of work. The miner who solves the puzzle first confirms the most recent block of transactions on the blockchain.

The successful miner then broadcasts the new block to all other nodes, who confirm its accuracy and add it to their copy of the blockchain, creating a verifiable record of data for the entire network. Consensus is represented by the verification process. A new block can only be added to the network after this data has been confirmed. For being the first to validate a new block of data and

add it to the PoW blockchain, miners receive newly minted cryptocurrency, the block reward (in the case of Bitcoin, they receive BTC).

Proof of Work blockchains aim to generate blocks at regular intervals — for example, Bitcoin generates one block every ten minutes. PoW networks are limited in speed and scale because of the energy-intensive proofing process. Furthermore, PoW networks are engineered to be more or less challenging depending on the network's computing capability.

Despite their speed and scalability restrictions, PoW blockchains have typically given better security while maintaining significant decentralization. Because PoW systems are distributed, it is extremely difficult for a malicious actor to take control of the blockchain by controlling the majority of computing power on the network. Typically, the hardware, power, and computing expenses are prohibitively expensive [3].

However, the same characteristics that make PoW blockchains secure also make it difficult to participate in the network as a node. Mining on many major networks has been monopolized by large-scale mining operators who have amassed influence in network governance. Since operating a mining rig and paying for the associated hardware and electricity costs is too expensive for the average user, large-scale mining operators have collected major influence in the network governance.

2.2. Proof of Stake (PoS) Consensus

PoS is the second most often used consensus method, and it addresses many of the issues that plague PoW blockchains, such as slowness, scalability, inefficient energy usage, and expensive requirements for entering the network.

PoS blockchains use validators instead of miners to validate transactions. Validators are network node operators who validate data in the same way that PoW systems do, but there is no energy-intensive computational process required to earn the right to validate. Instead of doing proofs of work, validators "stake" part of the blockchain's native currency in order to be considered as a validator node. The prospective validator mainly uses blockchain-born crypto tokens to provide collateral. The system randomly selects a validator to confirm data when validating the data held in a transaction block on a PoS blockchain. Certain factors, such as the number of tokens staked, can increase the likelihood of a validator being chosen. The validator is rewarded with network transaction fees when a block is confirmed, and then the process starts over [3].

Proof-of-Stake blockchains keep the network secure and validators honest by demanding validators to stake their tokens. If validators are acting deliberately or incompetently, their involvement and access to the network is lost by a 'slashing' procedure. This stimulus structure makes it possible for validators to earn through legal operation rather than by breaching regulations.

Accessing the PoS blockchains is easier as the validators on the PoS blockchains do not need to invest in high electricity costs and expensive hardware. PoS blockchains are arguably better for the environment than PoW networks in terms of sustainability because they consume significantly less electricity [4].

2.3. Pure Proof of Stake (PPoS) Consensus

PPoS is a new fully decentralized and promising consensus mechanism. The amount of influence each user has over the selection of a new block is proportional to their stake (number of tokens) in the system. Users are chosen at random and in secret to propose and vote on block suggestions. Every online user has the opportunity to be chosen to propose and vote. The probability of a user being picked, as well as the weight of its proposals and votes, is proportionate to its stake.

This means that only two blocks of PPoS cannot be sent to the chain simultaneously since the required threshold of committee votes can be reached by a single block. Up to 1 block in a given round is confirmed and written on the chain. When a block is created, users may immediately depend on the transactions it includes. Thus users can rely on that the block, since it will be part of the chain forever.

PPoS does not place a small group of users responsible for generating blocks, and users do not have to transfer their voting power to a limited number of users. All users are allowed to propose and vote on blocks with a probability that is directly in proportion to their stake.

PPoS doesn't force a user to put a part of their stake aside for a consensus process, and a user can not restrict his ability to spend their stake by participating in a consensus protocol.

In a completely decentralized fashion, PPoS guarantees full participation, protection and speed [4].

3. Distributed Ledger Technologies

The authors first mentioned distributed ledger technology in their 1991 article "How to time-stamp a digital text" [5]. Distributed ledger technology (DLT) is a decentralized framework in which data is shared among network participants. Despite the fact that the approaches to distributed ledger technology are technically distinct, they all have three fundamental components in common.

Every distributed ledger technology should support peer-to-peer transactions, have participants who store the network, and use a consensus mechanism to manage it.

The ability of two or more parties to operate without the use of a central authority is known as a peer-to-peer network. Participants in a distributed ledger are independent computation nodes that record, share, and synchronize transactions.

DLTs are managed without the central authority being required [6]. With the independent conduct of transactions, the efficiency can be improved by reducing errors.

To address the speed and scalability issues in Blockchain technology, new distributed ledger technologies have emerged, such as the Tangle, Hashgraph, Tempo, and Holochain. This newly developed distributed ledger infrastructures aim to improve blockchain technology's limited features.

It has to be also noted that every Blockchain is a type of DLT but not every DLT is a Blockchain [7].

3.1. Hedera Hashgraph

Swirlds developed Hashgraph, a proprietary distributed ledger technology. The consensus algorithm and the Directed Acyclic Graph (DAG) data structure distinguish it from Blockchain technology. Hashgraph does not use conventional Blockchain consensus. Instead of conventional consensus mechanisms, it employs the random gossip protocol and the virtual voting consensus algorithm.

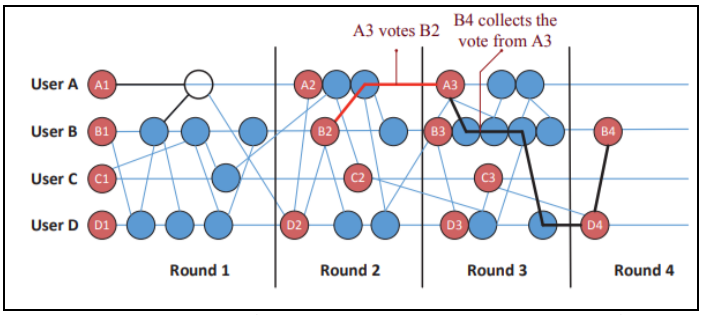


Figure 2: An example of Hashgraph's operation. Image courtesy of [8]

Any transaction that occurs in the Hashgraph network is referred to as an event. Each event includes a timestamp, a hash of the previous two events, transactions, and electronic signature information. The random gossip protocol distributes the time stamp of the verification process to neighboring nodes at random until the data is applied to the ledger [9]. Refer to the Figure 2 as an example of the Hashgraph algorithm.

In order to validate the transaction, this information must be transmitted to the majority of network participants. Transaction timestamps are documented in the case of a separate hashgraph ledger maintained by the participants. Transactions that are agreed by two-thirds of the network through virtual voting are considered legitimate and are registered by each node. Virtual voting can require several rounds in the absence of a majority of participants,

3.2. Tangle

A directed acyclic graph (DAG) is a specialized subject in mathematics, particularly graph theory and computer science. A DAG is made up of vertices and edges (also known as arcs), with each edge pointing from one vertex to another in such a way that following those directions will never result in a closed loop. This mathematical model is being considered as an alternative approach for developing alternative Blockchain structures for applications consisting of many devices [10].

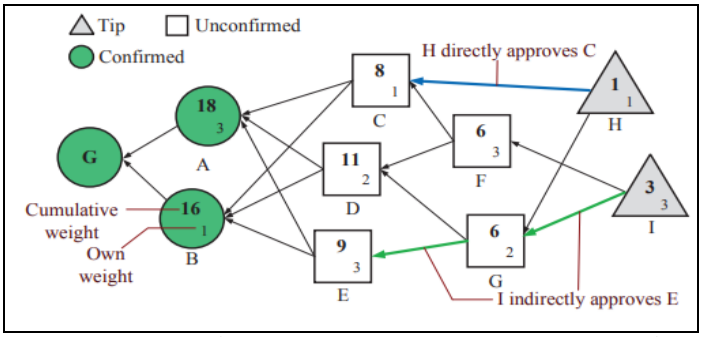


Figure 3: An example of Tangle's DAG based operation. Image courtesy of [8]

Tangle is a new type of DLT which is considered as one of the major DLT that vastly uses the DAG model. Refer to the Figure 3 for visualizing the Tangle's operation.

Since the current mainstream Blockchain infrastructures (i.e. PoW) are unable to meet the efficiency and scalability requirements of IoT systems that require a Trust Layer, Tangle DLT is becoming increasingly important in establishing trust and security for the Internet of Things (IoT) applications.

There is no transaction fee in Tangle since there are no blocks or miners to add data to the network. Transactions are not kept in a sequential order. Transactions are not recorded in a linear fashion. Each transaction in the consensus protocol verifies the previous two transactions and records the transaction data. In terms of security, it is less secure than Blockchain technology especially in the early stages of the network. When the transaction volume is low, network coordinators are used for protecting the network.

The use of network coordinators may result in the network not being fully distributed. The goal is to improve the scalability, speed, and efficiency of the Tangle protocol. The speed of the network increases in direct proportion to its size, especially as the network grows.

Tangle, which is used in IOTA, is one of the most important applications that use the DAG model [11].

It is appropriate to use this advanced DAG based DLT particularly in IoT applications or any other applications that require providing speed and scalability to a network composed of large number of devices and transactions.

3.3. Holochain

Holochain is an open-source distributed ledger technology. Unlike Blockchains, Holochain relies on individual chains rather than a single chain. While transactions in Blockchain technology are made over a single chain for all network participants, participants in Holochain store data in their local chains as part of the public network, Instead of recording all of the data on the network, participants only record their own transactions. This situation should be considered as the main reason why this DLT requires low amount of storage on participating nodes. The Distributed Hash Table (DHT) [12] ensures the correctness of local chains. DHT stores the chain data's DNA (hash) information and ensures that the chain's invariance is maintained. As indicated in Figure 4, DHT is stored in parts of the mesh topology [13]. While some users are offline, the network can continue to function. The disadvantage of this protocol is that information can be lost if a registered peer's device completely malfunctions. A backup procedure is additionally required for all peers in order to prevent data losses.

Distributed applications (dApps) built on the Holochain,which are named as Holochain Apps (hApps), can run quickly, and the network's scalability is limitless. Millions of transactions can be processed per second with Holochain. Holochain aspires to transform the current internet's center-based services.

3.4. Tempo Radix

Sharding method is used by Tempo distributed ledger technology in order to distribute data amongst many nodes in the

network. Hence, this situation eases the amount of storage required on the peers.

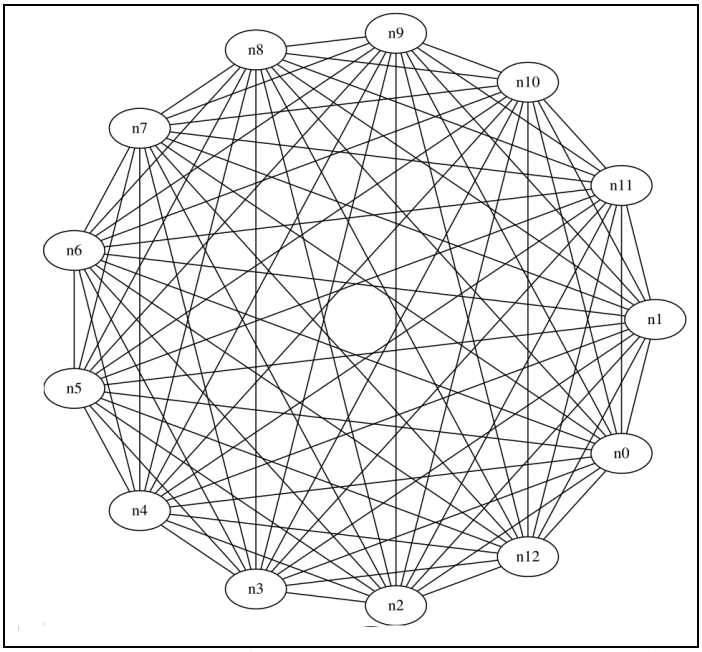


Figure 4: A representation of Holochain's DHT topology. Image courtesy of [13]

Tempo relies on Lamport's Logical Clocks, see Figure 5 for a sample representation of this concept, to come to an agreement on the distributed ledger. Any time a transaction is recorded, the logical hour number is also recorded. Each node increments its logical counter for every transaction it witnesses. Due to the unreliability of using the physical clock in distributed systems, the logical clock system is used. Tempo logical hours are smart counters that count upwards in increments of one when a participant sees a new and verified transaction in the ledger,

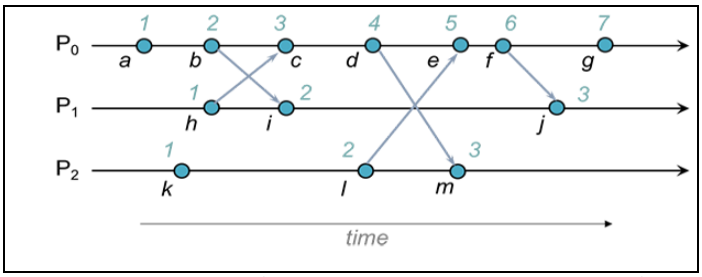


Figure 5: A representation of Lamport Logical Clocks. Image courtesy of [14]

Tempo participants use the Gossip protocol in order to synchronize the timeliness of the information in their tracks [15].

Participants in a blockchain network have access to the entire ledger, but the Tempo infrastructure uses shredding technology to limit participants' access to a subset of the ledger. Some sections of the participants' data are also being recorded by neighboring nodes as a data-loss prevention mechanism.

4. Comparison of PoW, PoS, PPoS Blockchain Consensus Mechanisms and Distributed Ledger Technologies

4.1. Blockchain versus Hedara Hashgraph

In terms of speed and scalability, Hedera Hashgraph can outperform the Blockchain technology. The transaction volume is high due to the Hedera Hashgraph's random gossip protocol, which allows transactions to be released quickly to the network. When a node conducts a transaction, it transfers transaction information to a randomly chosen neighbor node [9]. The node that discovers the transaction will also pass this information on to a random neighbor node. This procedure is repeated until the entire network is aware of the information. As a result, the network is immediately alerted when a new transaction occurs.

Hashgraph DLT uses virtual voting to validate each transaction, rather than the costly mining process in some Blockchain infrastructures. If two-thirds of the entire network accepts the transaction, the transaction is valid. Virtual voting can take place several times if the majority is not reached. In Hedera Hashgraph distributed ledger technology, a maximum of 250,000 transactions per second can be performed [16].

Hedera Hashgraph is more scalable than Blockchain. The Hedera Hashgraph network can potentially have an infinite number of participants.

4.2. Blockchain versus Tangle

Tangle distributed ledger technology has increased scalability and quantum durability compared to Blockchain technology, while reducing transaction fees in open blockchains.

To confirm a transaction in Tangle, each node must verify at least two transactions from previous transactions. The participant's transaction becomes more valid as it validates more transactions. It is the most resistant distributed ledger technology even it is also considered as quantum-safe.

Tangle arises amongst the other DLT alternatives with its lowest transaction fee. When compared with Blockchain, the security of Tangle is weaker when the network size is smaller.

The number of transactions per second in Tangle DLT his estimated between 500 and 800 [17]. Due to the nature of Tangle's operation, it is assumed to allow an unlimited number of participants.

Tangle infrastructure can be preferred in applications that require speed and scalability such as instant mobile voting applications. Also, it is allegated that Tangle can also have a good impact on the Energy Trade industry [8].

4.3. Blockchain versus Holochain

In comparison to Blockchain, resources are used efficiently and scalability is increased with Holochain. Data on the network is kept in separate chains rather than a common ledger in Holochain technology .

As part of the public network, each participant keeps its own local records. Each participant keeps a separate ledger. Distributed Hash Table (DHT) ensures the accuracy and immutability of local ledgers. Holochain DLT allows a truly distributed architecture in which everyone maintains their own ledgers.

As opposed to Blockchain technology, energy efficiency is ensured by avoiding costly mining operations via the Holochain consensus mechanism.

Since there is a central point through which all transactions must pass, Holochain does not set a TPS (transactions per second) value like other blockchain-based or blockchain-derived projects [18]. Holochain, on the other hand, is a generalized distributed computing protocol. Due to the underlying mechanisms, throughput and scalability are assumed as theoretically infinite.

4.4. Blockchain versus Tempo Radix

In terms of speed, scalability, and resource efficiency, Tempo distributed ledger technology can outperform Blockchain. As opposed to Blockchain's single chain, Tempo distributes the subportions of the entire ledger onto the network participants' storage. In other words, Tempo participants store the puzzle pieces by viewing the entire ledger as a puzzle [15].

Each node keeps a piece of its neighbor node in the fragmentation process in order to prevent data loss. Tempo distributed ledger technology's shredding process has advantages in terms of speed and resource efficiency. In Blockchain, a lot more resources are used since the entire block data / ledger is kept on all participants. Additionally, distributed ledger synchronization is slower and less scalable than Tempo distributed ledger technology.

Tempo distributed ledger technology has been tested to handle 1,400,000 transactions per second [19]. The Tempo network has a significant advantage over Blockchain technology because the maximum number of participants is theoretically unlimited.

The Gossip protocol is used in Tempo distributed ledger technology ensures that nodes have up-to-date information about the distributed ledger. The Gossip protocol allows the nodes in a network to propagate new configurations to the rest of the network. This protocol allows network nodes to communicate with one another and exchange information about their components.

4.5. Comparison of Blockchain (PoW, PoS, PPoS) and DLTs

In a typical setting, performance and evaluation metrics for developing a distributed application depends on four key parameters, Throughput (total number of transactions per second on the distributed network), Scalability (Supported number of nodes in the distributed network), Energy consumption per transaction and the level of storage requirement/usage.

'Table 1' summarizes the key performance and evaluation metrics of the consensus mechanisms and DLTs which are outlined in this study.

Throughput, scalability values and the storage usage levels are gathered from the internally conducted research and from the following references: [16], [17], [20]-[22].

For the energy consumption of Blockchain consensus protocols, in a network of 10000 nodes, which is approximately the number of active full nodes on PoS, note that this amounts to approximately 1000 J per transaction for PoS, it is many orders of magnitude less than for the current PoW blockchains which is about 1 billion J per transaction [23]. It has to be also noted that PPoS uses 30 J per transaction [24].

As for the DLT segment, Tangle uses 3.6 J per transaction [25]. It is claimed that Hashgraph uses 3600 J per transaction [26]. Holochain's energy consumption is estimated as 325 J per transaction [27]. Energy consumption of Tempo is not available in the literature, however it is considered very low since no mining is involved [28].

Table 1: Key performance comparison and evaluation metrics of Blockchain (PoW, PoS, PPoS) and DLTs

	Throughput (transactions per second)	Scalability (number of nodes)	Energy Consumption (J per transaction)	Usage of Storage
Blockchain - PoW	~7 tps	~10.000 nodes	~1 billion J/transaction	High
Blockchain - PoS	~25 tps	~10.000 nodes	1000 J/transaction	High
Blockchain - PPoS	~10000 tps	500.000 nodes	~30 J/transaction	High
DLT - Hashgraph	~250.000 tps (Theoretically) ~10.000 tps (recent practical result)	Theoretically Unlimited	3600 J/transaction	Moderate
DLT - Tangle	~500 tps	Theoretically Unlimited	3.6 J/transaction	Moderate
DLT - Holochain	Theoretically Unlimited	Theoretically Unlimited	325 J/transaction	Low
DLT - Tempo	~1.400.000 tps	Theoretically Unlimited	N/A (Considered very low)	Low

5. Modern Applications of Blockchains and Distributed Ledgers

In empirical settings, blockchain and distributed ledger technologies can be used in a variety of ways. Like all technologies, a Blockchain application must be viewed as a sociotechnical system. A sociotechnical system consists of a technological artifact and the social environment in which it is used, as well as the interactions between the two. When viewed as a sociotechnical system, the core characteristics of Blockchain and DLTs can be applied to a variety of situations, resulting in a wide range of use cases.

5.1. Cryptocurrency

A cryptocurrency is a digital or virtual currency that is mostly based on consistent Blockchain consensus mechanisms and protected by cryptography, making counterfeiting and double-spending nearly impossible. Cryptocurrencies are distinguished by the fact that they are not issued by any central authority thus allowing them to be theoretically immune to government intervention or manipulation.

Cryptocurrencies have the potential to make it easier to transfer funds between two parties without the use of a trusted third party such as a bank or credit card company. In addition, users can avoid the high fees charged by banks and financial institutions for wire transfers by preferring cryptocurrency based fund transfers with minimal processing fees [29].

Major disadvantageous situation of Cryptocurrencies is that the semi-anonymous nature of Cryptocurrency transactions makes them ideal for a variety of illegal activities, including money laundering and tax evasion. However, Bitcoin is a poor

choice for conducting illegal operations because forensic analysis of the Bitcoin has aided authorities in detecting criminal activity. It has to be noted that more privacy-oriented coins like Dash, Monero, and ZCash, on the other hand, are far more difficult to track than the Bitcoin [30].

5.2. Decentralized Finance (DeFi)

Decentralized Finance, also known as DeFi, is a movement that allows users to access financial services such as borrowing, lending, and trading without relying on centralized entities. These financial services are delivered through Decentralized Applications (Dapps), the vast majority of which are built on the PoS powered Ethereum platform [31].

DeFi is a collection of products and services that act as a replacement for institutions such as banking, insurance, bonds, and money markets, rather than a single product or company. DeFi Dapps allow users to combine their services in order to expand their options. DeFi Dapps typically require collateral to be locked into smart contracts in order to function. The Total Value Locked is a term used to describe the total collateral locked in DeFi Dapps [32].

By eliminating the need for any middlemen, DeFi Dapps have the potential to revolutionize traditional financial services. It should be noted, however, that DeFi in its current state is still very new and experimental, with many projects being improved on a daily basis. DeFi may evolve further over time and become completely unrecognizable from what it is now.

5.3. Non-Fungible Tokens (NFT)

NFTs (non-fungible tokens) are a new type of blockchain-based token that is one-of-a-kind and indivisible. They first appeared in late 2017. NFTs are digital tokens that can be used to show who owns one-of-a-kind items. Unlike PoW powered Bitcoin, where each coin is identical, NFTs are unique, with their own set of features. Tokenized items include art, valuables, and even real estate that no one can alter the ownership record. NFTs can only have one official owner at a time and are mainly secured by the PoS based Blockchain infrastructures [33].

5.4. Digital Identity

Any document that can be used to prove an individual's identity is known as an identity document. We frequently share our identity information to authenticate in our daily lives. Personal data is violated when identity information is stored, processed, and shared against the person's will. There is no universally accepted identity in the Internet.

A digital identity system can be used for information confirmation as well as authentication. Users will have complete control over their digital ID and personal information. Verification can be handled with or without sharing entire personal information in the digital identity system.

Blockchain has progressed significantly since its inception as a distributed ledger system for tracking bitcoin ownership. This technology has the potential to replace traditional identity management systems with a highly trusted mechanism. Users will be able to have more control over their own identities owing to the infrastructure provided by the Blockchains [34] and DLTs.

Organizations can only use the data with the consent of customers, and no central entity will be able to compromise a customer's identity. Self-sovereign identity, which is inherently unalterable and more secure than traditional identity systems, has been made possible by Blockchain [35]. This has the potential to completely transform how we connect to various online services using our identities. Individuals would verify their identity using their self-sovereign ID, eliminating the need for passwords.

As with any life-changing innovation, there has been a long period of evolution, with experts exchanging ideas and little agreement on what self-sovereign ID means. It is based on the idea that an individual must have control over how his identity is managed. With user consent, the ID cannot be tied to a single system and must be interoperable across multiple platforms. Experts considered combining various identifying data, such as demographic and employment-related information, as well as individual information revealed by others.

In the Blockchain domain, digital identities and user information will be stored in a digital wallet. There will be no need to carry or store many cards or documents. In a Blockchain based digital identity system, documents such as our citizenship card, bank cards, diplomas, certificates, driver's license, and passport are being planned to be stored in self-sovereign id oriented digital wallets.

5.5. Supply Chain Management

One of the most important areas in which blockchain projects are concentrated is considered as supply chain management. A trust issue arises because of the large number of parties involved in supply chain processes. Supply chain transaction information is being planned to be securely stored by Blockchain [36]. The current method of running these processes on paper makes it difficult to manage and unreliable. By using Blockchain, it is possible to conduct transactions securely on digital media and to view all stages of the processes transparently with confidence. Circulation of counterfeit products can also be prevented from the market using this approach. Preventing the sale of smuggled goods will protect citizens' rights and countries from incurring tax losses.

Consumers may not fully confirm whether the procured products were genuine or not in a traditional supply chain process. Consumers can reliably view the production and distribution history of products by using Blockchain enhanced Supply Chain Management Systems. With this understanding in mind, countries may want to look at the manufacturing, distribution, and shipping history of all imported goods in the future.

5.6. Digital Notary Services

A notary, also known as a public notary, is a state-appointed official who has the authority to certify a specific action or deed. Contracts, documents, affidavits useful in other jurisdictions, and deeds are all included. In order to perform the role of notary, an official must be objective when making various declarations. A notary's duties include everything from signing a property sale record to administering oaths and keeping accurate records of other important deeds and documents. The officer can use an

electronic or digital notary to perform notarizing functions digitally.

The authorized seal will be applied to the certified document by the electronic notary. This activity requires cryptography and a secure public key in order to manage, develop, store, and distribute the digital certificate since a digital notary must keep an up-to-date electronic record of all completed notary tasks. Remote implementation of a significant digital solution for public notary activities is also possible. This notarization process is expected to be aided by Blockchain and DLT technology. As a prospecting result, users of this service can trust the integrity of data on the Blockchain or DLT.

Nonrepudiation, tamper-resistance features and the well-known characteristics of distributed ledgers make it an excellent tool for enhancing notary capabilities [37].

5.7. Land Registry and Property Ownership Management

Land registration entails gathering information such as the ownership and property size information of a land piece. The entire process of land registry maintenance is extremely time-consuming because it demands the safekeeping of massive volumes of written registers. The current method is insecure since the majority of this procedure is opaque and continuous sales of a property must be accurately recorded. There have been several proposals to automate land register data recordings by eliminating the need for handwritten records. This is initially accomplished by storing the data in massive databases. However, in terms of data security, such a strategy is ineffective because the data contents are easily accessible, and data tampering can occur in poorly maintained databases.

A distributed ledger based solution for property registration could significantly improve efficiency and possibly eliminate fraud [38]. Contracts and ownership data can be stored in a distributed manner too. Blockchain and DLT implementations can make it easier to track data transactions because it eliminates the need for physical involvement thus enhancing the overall system security for users. Developing countries are considering Blockchain and DLTs as a tool for updating their existing land registry systems.

5.8. Health Records Management

Preventing illnesses or detecting them early would improve society's living conditions and reduce the workload of hospitals. Citizens' health status should be constantly tracked using preventive health technologies. The volume of health data will grow as a result of constant surveillance of citizens' health status. This data can be immutably stored and exchanged for clinical trials using Blockchain and DLT technology.

Health data is vital information that countries must safeguard. Theft of these details poses a threat to the country's national security. Health information is also essential to the pharmaceutical industry. As a result, health data has become an obvious priority for financial benefit.

Blockchain and DLTs can serve as a link between healthcare researchers and health data records in order to facilitate data exchange. Blockchain and DLTs can provide secure storage and exchange of health data at any time. Anonymized health data can have economic benefits to both users and organizations [39]. A health consortium with Blockchain infrastructure, comprised of organizations that generate and need data, can be created. As a result, the data can be safely exchanged with the desired entities and organizations for scientific purposes.

5.9. Insurance

The insurance industry can use Blockchain and DLT technology in order to improve core internal processes like claim-submission and processing, as well as fraud detection and prevention. Blockchain can also be used to modify and possibly create new insurance brands, as well as improve existing ones. It has also come to light that Blockchain faces a number of issues, including scalability, security, and privacy, as well as taxation and regulation. The insurance industry has yet to reap the benefits of new generation blockchain architectures and distributed ledger technologies [40].

5.10. Electronic Voting Systems

It has long been a challenge to create a secure voting structure that offers the same fairness and privacy of current voting schemes while also providing the transparency and flexibility of electronic systems. To compete with the conventional ballot system, an electronic voting system must meet the same security and privacy criteria.

A Blockchain or DLT based e-voting technology has the ability to reduce election fraud while also increasing voter access [41]. Such an e-voting system could decentralize controls, enabling electors to take on certain duties while maintaining a copy of the electoral register. Since the copies of the ballots are a voting history that cannot be reversed, there can be no null votes [42].

5.11. Energy Distribution and Trade

Given the finite nature of resources, countries must increase their production of renewable energy resources. As a result, the number of local producers who use renewable energy is expected to grow in the future. The structure of energy distribution and trade will be altered as the number of local energy producers increases. To encourage local producers, an energy trading platform would be developed, allowing producers to load their excess energy to the grid immediately.

It is necessary to ensure that energy exchange between the local producer and the consumer is secure. Blockchain and Distributed ledgers could be used to create a platform that allows local producers and consumers to transact usage certificates instantly. For such an energy distribution and trade infrastructure, all parties involved in the energy distribution process must form a distributed ledger based consortium. In the energy sector, Blockchain and DLT will bring together market regulators, distributors, consumers, home users, and producers on a single, reliable energy market and sharing platform [43].

This platform will enable real-time allocation of required and preferred types of energy, as well as the approval of energy consumption. This platform would provide incentives to renewable energy producers. Blockchain has the potential to provide an infrastructure for instant energy trade for energy consumers and local conventional or renewable energy producers [44]. Smart meters will also need to be developed in order to measure the produced or purchased energy.

Active customer engagement and energy trading would require immutable, open, and dependable smart contracts. Such digital energy market systems can ensure that energy is priced immediately and it is delivered on demand, and allows the customer to determine their own energy consumption preferences. It would also make the consumer more efficient in deciding the frequency of energy usage.

It has to be noted that especially for the energy industry, DAG based structures can have a critical role [8].

5.12. Distributed File Systems

Distributed file systems (DFS) have gained a lot of interest recently. Instability, auditing, and incentive mechanisms are all weaknesses in traditional Peer-to-Peer (P2P) distributed file systems. Although Blockchain-based DFSs successfully provide enough incentives and security assurances, a number of problems such as scalability and privacy concerns are preventing the next generation of Blockchain DFSs from being developed. As a result, it is anticipated that DLTs have much greater potential than conventional Blockchain in order to establish a globally distributed file system due to their storage-wise and scalability advantages [45].

5.13. Artificial Intelligence and Federated Learning

A new machine learning paradigm known as federated learning (FL) has arisen as a result of the increasing computing capabilities at end user devices, as well as the growing privacy concerns over sharing sensitive raw data. FL has the potential to avoid exchanging data directly by training models locally at each client and aggregating learning models at a central server, hence avoiding privacy leaks. The standard FL structure, on the other hand, is strongly reliant on a single central server, which could fail if that server acts maliciously. To address the single point of failure problem, new approaches are being developed that focus on establishing a DLT-assisted decentralized FL framework that can effectively prevent malicious clients from poisoning the learning process while also providing clients with a self-motivated and reliable learning environment.

Distributed Ledger Technologies have the potential to improve transparency and trust in the distributed machine learning process and generating consistent models by permanently logging it into a distributed ledger. Since the model is available to the public and the history is preserved, the blockchain tracks not just the current model but also all of the previous iterations. Since decentralization is inherent in both, federated learning can benefit from distributed ledgers [46].

6. Application-specific Minimum Performance Requirements for Achievability of the Applications

Minimum required levels of performance requirements for the above mentioned applications are given in Table 2. This section does not provide the desired ideal state of applications. It has to be noted that; there does not exist any universal solution for all types of applications. Hence, it is necessary to outline the needed basic and minimum performance requirements for having a better decision-making action in order to select an appropriate protocol or conduct a better protocol implementation strategy during the initial phases of a specific distributed project. All of rationals of Table-2 are given in the subsections.

A scoring system of Low, Moderate and High is used in order to outline the minimum requirements. It has to be that 'DC' denotes, as a scientific and engineering abbreviation, "Don't Care".

Table 2: Application-specific Minimum Performance Requirements for the Achievability of the Applications

	Throughput (transactions per second)	Scalability (number of nodes)	Energy Efficiency (J p/tx)	Level of Storage Requirement
Crypto Currency	Low	Low	DC	DC
DeFi	Low	Moderate	Moderate	DC
NFT	Low	Low	DC	DC
Digital ID	Moderate	DC	DC	Moderate/High
Supply Chain	Moderate	Moderate	DC	High
Digital Notary	Low	Low	DC	High
Land Registry	Low	Low	DC	Moderate
Health Records	Moderate	Low	DC	High
Insurance	Moderate	Low	DC	Low
Electronic Voting	High	High	High	Moderate
Energy Trade	Moderate	High	High	Moderate
DFS	High	Moderate	DC	High
AI/FL	Moderate	Moderate	DC	High

6.1. Cryptocurrency – Minimum Performance Requirements

Cryptocurrency applications can be achieved with low rate of transactions per second and low number of participating nodes. Any cryptocurrency system can be achieved with low levels of transaction rate with a small number of validating or participating nodes. The most basic and stable example of this situation can be given as the PoW based Bitcoin in which the performance metrics are given in Table 1.

6.2. DeFi – Minimum Performance Requirements

As in the case of Cryptocurrency, any DeFi project can be achieved with low transaction rate. Major difference is the necessity of having a more energy efficient and moderately scalable infrastructure for the purpose of having a more future-proof DeFi system. The reason is that; there is the possibility of having multiple mobile devices participating in the DeFi domain in the near future.

6.3. NFT – Minimum Performance Requirements

NFT portals or NFT systems recently does not have and need many transactions per second and participating nodes, since NFTs are used mainly for the occasional transfer of the ownership of any concept.

6.4. Digital ID – Minimum Performance Requirements

The required transactions of Digital ID records will be moderately high enough. In an interconnected e-government environment, the transactions between the official Digital ID issuers, validators and peers will be moderately high enough in a daily setting. This situation will also be true for keeping and updating the Digital ID records of the objects.

Considering the large number of files in a countrywide setting, the storage requirement will be essentially very high for the issuer and verifier sides of a Digital ID system.

As a special note; if there exits multiple IoT (Internet of Things) devices in an Digital ID system then the level of storage requirement needs to be moderately large enough in order to store the entire history of record in the IoT node's lifecycle.

6.5. Supply Chain – Minimum Performance Requirements

In a worldwide supply chain setting, the amount of transactions between suppliers, manufacturers, distributors, retailers and consumers is considered very large. As a result, the required level of storage is needed to be large enough in order to collect and store the entire decades of transaction records.

The required minimum transactions need to be moderately high enough due to the speed requirements in a global environment for a supply chain operation and considering the amount of participating nodes in this setting, the scalability level should also need to be moderate.

6.6. Digital Notary – Minimum Performance Requirements

On a daily urban life setting, the necessity and frequency of notary services is considered high. In a Blockchain/DLT based Digital Notary scenario, the notary branches in a city can be considered as validating nodes. When compared with any large scale application, the number of notary branches are not large enough and the approval time needed in a branch does not require high levels of scalability. As a result, the minimum throughput and scalability requirements are selected as low.

One major situation about the notary services is considered as the increasing intensity of the volumes of documents. Hence, the minimum required level of storage is selected as high.

6.7. Land Registry – Minimum Performance Requirements

Land registry services can be considered as a subsegment of digital notary services. The only difference between the digital notary requirements is the volume of data that needs to be stored. It is considered lesser than the digital notary. Hence, the minimum storage requirement is set as moderate.

6.8. Health Records – Minimum Performance Requirements

It is required to have fast approvals of transactions of data in a network of hospital environment due to the criticality of individual health; therefore a Blockchain based immutable health records database minimally should need to have a moderate level of transaction rate of health recordings. The healths data is defined as laboratory test results, reports, prescriptions, diagnostic info and imaging data. It has to be also noted that due to the intensity of large amount of imaging data, the minimum required level of storage level is considered as high.

6.9. Insurance – Minimum Performance Requirements

In a global or multi-clienteed realtime digital insurance scenario, daily and instant insurance contracts are being forecasted to be part of society's daily life. Anticipated participants will be mostly rental items such as automobiles. Due to the fact that; the number of registered validating nodes of an insurance company is considered as low enough. Hence, the scalability is selected as low. Whereas the speed of transactions should be moderately high enough in order to reduce the waiting durations during insurance processes. In addition to the transaction rate and scalability requirements, since the needed records data is text based, the minimum requirement for level of storage is determined as low. However, if there will be many items to be insured frequently then the requirement should be considered as moderate.

6.10. E-Voting – Minimum Performance Requirements

Electronic Voting, as one of the most suitable and politically sensitive application on the Blockchain and DLT domain, have high demanding transaction rate and scalability requirements. The setting environment in this application area is considered as fully mobile and distributed voting system. Hence, the energy efficiency requirement is selected as high. Given that; the necessity of having the voting/election as rapid as possible in a large scale voting event, the transaction rate requirement is also considered as high and due to the required mobile operation the scalability should also be considered as high.

6.11. Energy Trade – Minimum Performance Requirements

Blockchain-enabled Energy Trade sector is being regarded as a digital transformation technology in a Smart City context. In a practical sense, this concept is considered as a large scale IoT application in which the participants measure, share and certificate the energy production and consumption. With the inclusion of prosumers to the consumers and producers, the frequency of transactions of the energy certifications would be moderately high enough in a city/country wide environment. Hence, the minimum transaction rate is selected as moderate. Because of the prosumer and consumer-approval support for such an application, the minimum scalability requirement is considered as high. Because of the scenario of consumer-approvals and instant-billing and energy monitoring cases, the minimum energy efficiency should be selected as high. The data transactions are moderately high enough and the data size to be stored is essentially low. Hence, the minimum level of storage requirement should be kept as moderately high.

6.12. DFS – Minimum Performance Requirements

A globally effective Distributed File System should need to have immutable and zero-loss structure. It should be also accessible fast enough, the participating nodes should be fully decentralized and the number of nodes should be moderately enough. Energy efficiency is not considered as a requirement. Due to the variety of multiple data types such as text, audio, video and other possible large file types, the minimum required level of storage is considered as high.

6.13. AI/FL – Minimum Performance Requirements

As a promising future application, federated learning and distributed neural networks model training will require high levels of data storage and fast execution times and multiple nodes.

7. Applicability of Blockchain (PoW, PoS, PPoS) and DLTs on Modern Applications

The Table 3 above summarizes Blockchain's (Note that Blockchain is considered as PoW, PoS, PPoS in this study) and new DLTs' suitability for the explained modern applications which are outlined throughout this study by following the information of minimum performance requirements per application given in Section 6 and Table 2. For better assessment the values in Table 1 is converted into the Low, Moderate, High performance scoring.

Table 3: Performance scoring levels of Blockchain (PoW, PoS, PPoS) and DLTs

	Throughput (transactions per second)	Scalability (number of nodes)	Energy Efficiency	Storage Efficiency
Blockchain - PoW	Low	Low	Low	Low
Blockchain - PoS	Low	Low	Moderate	Low
Blockchain - PPoS	Moderate	High	High	Low
DLT - Hashgraph	Moderate	High	Moderate	Moderate
DLT - Tangle	Low	High	High	Moderate
DLT - Holochain	High	High	High	High
DLT - Tempo	High	High	High	High

Table 4 is constructed based on the scoring values of the performance criteria provided in Table 3 and minimum performance requirements provided in Table 2.

Table 4: Applicability of PoW, PoS, PPoS and DLTs on Modern Applications

	PoW	PoS	PPoS	Hashgraph	Tangle	Holochain	Tempo
Crypto Currency	✓	✓	✓	✓	✓	✓	✓
DeFi			✓	✓	✓	✓	✓
NFT	✓	✓	✓	✓	✓	✓	✓
Digital ID			✓	✓		✓	✓
Supply Chain			✓	✓		✓	✓
Digital Notary	✓	✓	✓	✓	✓	✓	✓
Land Registry	✓	✓	✓	✓	✓	✓	✓
Health Records			✓	✓		✓	✓
Insurance			✓	✓		✓	✓
Electronic Voting						✓	✓
Energy Trade			✓			✓	✓
DFS						✓	✓
AI/FL			✓	✓		✓	✓

One major factor about the mapping of the consensus/DLT protocols is the distinction of consensus/protocol specific storage efficiency and application-specific minimally required storage levels. During the selection process of the suitable protocol, designer should be aware of that the needed amount in the peer's storage will be directly dependent on the selected protocol's storage efficiency and the desired application's required minimum level of storage. As an example, if the selected application requires high levels of storage and the selected consensus/protocol has low storage efficiency, then the resultant distributed system's peers will possibly require an enormous storage space. Hence, the storage related columns are not taken into consideration for the consensus/protocol selection process. The remaining Throughput, Scalability and Energy Efficiency criteria are taken into consideration for the selection of suitable consensus/protocol for the listed applications.

8. Conclusions

By deploying DLTs and Blockchain to suitable applications, parties can perform stable transactions for the first time in human history without relying on a centralized authority. Blockchain and Distributed Ledger Technologies have the power to modify most structures which are currently under central authority's control because of the concept of trust on peer transactions.

Blockchain infrastructures are inefficient in applications with high participation and transaction volume. As a result, new distributed ledger technologies have been created that primarily address the speed and scalability issues of Blockchain. Hashgraph, Tangle, Holochain and Tempo distributed ledger technologies are the recent alternatives to Blockchain. While these new distributed ledger technologies claim to solve the problems of Blockchain technology, their effectiveness has yet to be proven.

Participants and peers should should have the rights to conduct transactions without the use of intermediaries, and all Blockchain and Distributed Ledger Technologies should ensure decentralized management and participant data security. Under the hood of this prospect, a decentralized, flat and fully democratized structure with applications on Finance, Market Exchanges, Digital ID, Voting, Energy Trade, Distributed Storages, Insurance, Records Managements, Land Registry and AI will certainly constitute the future of humanity.

The completion of necessary country-wide regulations, as well as technical studies, will be a critical step toward the global adoption of blockchain and distributed ledger technologies, as well as related critical applications, in the near future.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This study is fully supported by HAVELSAN Inc. R&D, Technology and Product Management Directorate and HAVELSAN Blockchain and Distributed Systems Laboratory.

References

[1] N. Chaudhry and M. M. Yousaf, "Consensus Algorithms in Blockchain: Comparative Analysis, Challenges and Opportunities," 2018 12th International Conference on Open Source Systems and Technologies (ICOSST), 54-63, 2018, doi: 10.1109/ICOSST.2018.8632190.

[2] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System" [Online]. Available: https://bitcoin.org/bitcoin.pdf [Accessed: 05-08-2020].

[3] L. M. Bach, B. Mihaljevic and M. Zagar, "Comparative analysis of blockchain consensus algorithms," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1545-1550, 2018, doi: 10.23919/MIPRO.2018.8400278.

[4] Cryptopedia, "Types of Blockchain: PoW, PoS, Private, and DLT" [Online], Available: https://www.gemini.com/cryptopedia/blockchain-types-pow-pos-private [Accessed: 01-Jul-2021]

[5] S.Haber, W.S.Stornetta, "How to time-stamp a digital document.", J. Cryptology 3, 99–111, 1991, doi:10.1007/BF00196791.

[6] A.Sunyaev, Distributed Ledger Technology. In: Internet Computing. Springer, Cham, 2020, doi:10.1007/978-3-030-34957-8_9.

[7] A. Panwar, V. Bhatnagar, "Distributed Ledger Technology (DLT): The Beginning of a Technological Revolution for Blockchain," 2nd International Conference on Data, Engineering and Applications (IDEA), 1-5, 2020, doi: 10.1109/IDEA49133.2020.9170699

[8] H. Alen, C. Tomislav & Z. Ivana, Demystifying Distributed Ledger Technologies: Limits, Challenges and Potentials in the Energy Sector. IEEE Access, **7**, 1-1. 2020. Doi: 10.1109/ACCESS.2020.3007935.

[9] Hedera Hashgraph, "Hedera: A Public Hashgraph Network & Governing Council"

[10] F. M. Benčić and I. Podnar Žarko, "Distributed Ledger Technology: Blockchain Compared to Directed Acyclic Graph," 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), 1569-1570,2018, doi: 10.1109/ICDCS.2018.00171.

[11] N. Zivi et al., Directed Acyclic Graph as Tangle: an IoT Alternative to Blockchains. 1-3, 2019. 10.1109/TELFOR48224.2019.8971190.

[12] Holochain, "Holochain" [Online]. Available: https://github.com/Holochain/holochain-proto/blob/whitepaper/holochain.pdf [Accessed: 06-08-2020].

[13] Ceptr, "Holochains for Distributed Data Integrity", [Online]. Available: http://ceptr.org/projects/holochain [Accessed:08-Jul-2020]

[14] Rutgers University, Logical clocks, [Online]. Available: https://www.cs.rutgers.edu/~pxk/417/notes/logical-clocks.html [Accessed:07-Jul-2021]

[15] Tempo, "Building Tempo" [Online]. Available: https://www.radixdlt.com/post/tempo-consensus-lessons-learned/ [Accessed: 05-08-2020].

[16] Cao, Bin & Li, Yixin & Zhang, Lei & Zhang, Long & Mumtaz, Shahid & Zhou, Zhenyu & Peng, Mugen. (2019). When Internet of Things Meets Blockchain: Challenges in Distributed Consensus. IEEE Network. PP. 1. 10.1109/MNET.2019.1900002.

[17] K. Enis Data Storage in the Decentralized World: Blockchain and Derivatives, 2020. Doi: 10.26650/B/ET06.2020.011.03.

[18] Holochain, "What is the TPS (Transactions Per Second) on Holochain? - Learning Library / FAQs - Holochain Forum" [Online]. Available: https://forum.holochain.org/t/what-is-the-tps-transactions-per-second-on-holochain/191 [Accessed: 08-Jul-2021]

[19] Radix DLT Tempo, "10 years of Bitcoin history, replayed in under 30 minutes | The Radix Blog | Radix DLT" [Online], Available: https://www.radixdlt.com/post/replaying-bitcoin [Accessed: 01-Jul-2021]

[20] Holochain, "The Holochain Guidebook FAQ",

[21] Hedera Hashgraph, "Hedera Hashgraph", [Online]. Available: https://hedera.com/ [Accessed: 05-Jul-2021]

[22] Lepore, Cristian & Ceria, Michela & Visconti, Andrea & Rao, Udai Pratap & Shah, Kaushal & Zanolini, Luca. (2020). A Survey on Blockchain Consensus with a Performance Comparison of PoW, PoS and Pure PoS. Mathematics. 8. 1782. 10.3390/math8101782.

[23] Sedlmeir, J., Buhl, H.U., Fridgen, G. et al. The Energy Consumption of Blockchain Technology: Beyond Myth. Bus Inf Syst Eng 62, 599–608 (2020). https://doi.org/10.1007/s12599-020-00656-x

[24] Sustainable Blockchain: Estimating the Carbon Footprint of Algorand's Pure Proof-of-Stake | Algorand, [Online]. Available: https://www.algorand.com/resources/blog/sustainable-blockchain-calculating-the-carbon-footprint . [Accessed:07-Jul-2021]

[25] IOTA, "Tangle Energy Benchmarks for IOTA", [Online]. Available: https://blog.iota.org/internal-energy-benchmarks-for-iota/ [Accessed: 06-Jul-2021]

[26] Leafscore, "Sustainable Crypto 2021", [Online]. Available: https://www.leafscore.com/blog/the-9-most-sustainable-cryptocurrencies-for-2021/ [Accessed:07-Jul-2021]

[27] Holochain, "Holoport Electricity Usage", [Online]. Available: https://help.holo.host/support/solutions/articles/36000035595-will-holo-use-lots-of-electricity-how-much-electricity-will-my-holoport-use- [Accessed: 07-Jul-2021]

[28] A. Hany & W. Gary. Intersections between IoT and distributed ledger, 2019. Doi: 10.1016/bs.adcom.2018.12.001.

[29] Hashemi Joo, M., Nishikawa, Y. and Dandapani, K., "Cryptocurrency, a successful application of blockchain technology", Managerial Finance, Vol. 46 No. 6, 715-733,2020, doi:10.1108/MF-09-2018-0451

[30] Decrypt, "What Are Privacy Coins? Monero, Zcash, and Dash Explained - Decrypt" [Online]. Available: https://decrypt.co/resources/what-are-privacy-coins-monero-zcash-and-dash-explained [Accessed: 30-Jun-2021]

[31] Ethereum, "Decentralized finance (DeFi)" [Online]. Available: https://ethereum.org/en/defi/ [Accessed: 30-Jun-2021]

[32] Investopedia, "DeFi (Decentralized Finance) Definition and Why it Matters" [Online]. Available: https://www.investopedia.com/decentralized-finance-defi-5113835 [Accessed: 29-Jun-2021]

[33] Cryptopedia, "Non-Fungible Tokens (NFTs): Crypto Collectibles | Gemini" [Online]. Available: https://www.gemini.com/cryptopedia/nft-non-fungible-token-crypto-collectibles [Accessed: 29-Jun-2021]

[34] M. Takemiya and B. Vanieiev, "Sora Identity: Secure, Digital Identity on the Blockchain," 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 582-587,2018, doi: 10.1109/COMPSAC.2018.10299.

[35] Consensys, "Zug Digital ID: Blockchain Case Study for Government Issued Identity" [Online]. Available: https://consensys.net/blockchain-use-cases/government-and-the- public-sector/zug [Accessed: 17-08-2020].

[36] N.Hackius, M.Petersen "Blockchain in Logistics and Supply Chain : Trick or Treat?" , Epubli, 23, 2017, doi:10.15480/882.1444

[37] Gao, Ying & Pan, Qiaofeng & Liu, Yangliang & Lin, Hongliang & Chen, Yijian & Wen, Quansi. (2021). The Notarial Office in E-government:A Blockchain-Based Solution. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3066184.

[38] Barbieri, Maurice, and Dominik Gassen. "Blockchain–can this new technology really revolutionize the land registry system." Responsible Land Governance: Towards an Evidence Based Approach: Proceedings of the Annual World Bank Conference on Land and Poverty. 2017

[39] Agbo, C.C.; Mahmoud, Q.H.; Eklund, J.M. Blockchain Technology in Healthcare: A Systematic Review. Healthcare 2019, 7, 56, doi:10.3390/healthcare7020056

[40] Anokye Acheampong AMPONSAH, Adebayo Felix ADEKOYA and Benjamin Asubam WEYORI, "Blockchain in Insurance: Exploratory Analysis of Prospects and Threats" International Journal of Advanced Computer Science and Applications(IJACSA), 12(1), 2021. http://dx.doi.org/10.14569/IJACSA.2021.0120153

[41] N. Kshetri, J. Voas, "Blockchain-Enabled E-Voting," in IEEE Software, 35(4), 95-99, 2018, doi: 10.1109/MS.2018.2801546.

[42] R. Hanifatunnisa, B. Rahardjo, "Blockchain based e-voting recording system design," 2017 11th International Conference on Telecommunication Systems Services and Applications (TSSA), 2017, 1-6, doi: 10.1109/TSSA.2017.8272896.

[43] Andoni, V. Robu, D.Flynn, S. Abram, D. Geach, D.Jenkins, P. McCallum, A.Peacock,"Blockchain technology in the energy sector: A systematic review of challenges and opportunities",Renewable and Sustainable Energy Reviews, 100,143-174,2019,doi:10.1016/j.rser.2018.10.014.

[44] K. Mannaro, A. Pinna, M. Marchesi, "Crypto-trading: Blockchain-oriented energy market," 2017 AEIT International Annual Conference, 1-5,2017, doi: 10.23919/AEIT.2017.8240547.

[45] H. Huawei et al., When Blockchain Meets Distributed File Systems: An Overview, Challenges, and Open Issues. IEEE Access, **7**, 2020 10.1109/ACCESS.2020.2979881.

[46] M. Chuan et al., When Federated Learning Meets Blockchain: A New Distributed Learning Paradigm, IEEE Access, **7**, 1189, 2020. Doi: 10.1109/ACCESS.2020.2971189.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 292-297 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Performance of Vertical Axis Wind Turbine Type of Slant Straight Blades

Hashem Abusannuga*, Mehmet Özkaymak

Energy Systems Engineering, Faculty of Technology, Karabuk University, Karabuk, 78050, Turkey

ARTICLE INFO

Article history:
Received: 18 April, 2021
Accepted: 25 July, 2021
Online: 03 August, 2021

Keywords:
VAWT
Multi-Stream Tube
Self-Starting Problem

ABSTRACT

There is no doubt that energy is one of the most important requirements of life, and its importance increases with the passage of time, and this is what make countries to harness the capabilities and scientists in developing energy systems of all kinds, one of the most important energy systems these days is what is known as vertical axis wind turbines. If we compare this type of system with horizontal axis wind turbines, it is characterized by a relatively lower manufacturing cost. But on the other hand, it suffers from less efficiency in addition to the problem of starting the self-movement. The idea of this research revolves around the use of an engineering design for the vertical axis wind rotor that is very rarely used in the field of wind energy. This design takes the geometric shape of two inverted trapezoids. Within the framework of this study, the term "slant straight-blade vertical axis wind turbine" (SS-VAWT) was assigned to the wind rotor. Amendments have been made to the mathematical model of Multi stream tube to make it suitable for application and work on (SS-VAWT), where, it is known that the multi-stream tube model uses primarily and only for the original Darrieus and the H-Darrieus rotors. In order to prove the efficacy of the software used, the results obtained from it were compared with the practical results of previous studies, as it proved its effectiveness in obtaining the satisfactory results that were intended for this analysis. The analyzes and investigations that were conducted on the improved SS design included changing the geometry by changing some of its dimensional parameters represented in rotor height, rotor diameter, number of rotor blades, rotor blade section length, rotor blade section type and rotor blades inclination angle on the horizontal plane. Within the scope of the case studies that were worked on in this research, the results showed that the best efficiency of the SS rotor was achieved in the range of height to radius ratio (0.66 to 1), cord line length to radius ratio about 0.12 The angle of inclination of the blades is between 45- and 65-degrees Degree. In these ranges, the value of Max power factors has reached its turn, and the energetic range of the rotor has increased as a function of the peripheral relative velocity, in addition to a relatively large solution to the problem of starting self-movement, which appears through the highest-power factor values to move away from the limits of negative values in the range Terminal forgetfulness from 1 to 3. In addition, the effect of changing Raynaud's number on the turbine aerodynamic performance has been investigated. The results showed that the higher the Reynolds value, the higher the power factor value, the higher the energy range and the lessening the problem of starting the self-movement.

1. Introduction

This paper is an extension work originally presented in "2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)" [1]**.** Due to the increase in the costs of electric energy consumption with the progression of time, we have called for the need to focus on renewable energies and invest them in order to meet the needs of humanity in terms of energy, because of the advantages of renewable energies such as their presence in nature permanently, inexhaustible, clean, multi-source and free of charge in nature. Where you need systems to harness and convert it into electrical

*Corresponding Author: Hashem Abusannuga, hashem_ebrahim@yahoo.com

https://dx.doi.org/10.25046/aj060432

energy. The most important types of wind energy and solar energy. [2], [3]. When the wind passes according to the "cut in speed Vc" of the vertical axis wind rotor, the air rotor acquires a rotational speed around the rotor shaft to make θ angle from 0^0 to 360^0, and thus the ends of the rotor blades acquire a continuous terminal speed in ωr, which there is a difference in the values and proportionality between the wind speed V in contact with the rotor, this leads to the formation of what is known relative velocity W, which make the angle of attack α with the direction of the ωr, therefore there is a strong relationship between the angle of attack α and the Tip Speed Ratio λ_0 or ($\omega R/V_1$). In the science of aerodynamics, specifically in the field of wings and blades, it is known that the angle of attack α has a direct and strong effect on the forces of lift L and Drag D, meaning that it has an effect on the ratio (L/D). Accordingly, the link between the terminal relative velocity and the ratio (L/D) becomes clear. Linking to the above, in order to achieve optimal efficiency to the rotor, must be sure, that the values of each $\omega R/V_1$, α, θ and L/D are consistent [4]-[7]. With the aim of developing the efficiency of VAWT's, many different engineering designs have emerged that compete with each other in terms of high performance and lower manufacturing cost [8]-[12]. Following some instances from VAWT's engineering designs. Original Darrieus-VAWT, its engineering shape is a parabolic, it gains its rotational movement from the lift force generated on its blades as a result of aerodynamics. On the other hand, this system suffers from a low coefficient of performance Cp, which does not exceed 35%. It also vibrates severely when rotating, which makes its manufacturing cost high due to the increase in structural supports to reduce vibrations [13]-[16]. Darrieus was developed so that his rotary shape became in the form of H, so it was called H-Darrieus, The H-Darrieus is nearly as efficient as the original Darrieus, but has lower manufacturing costs due to its straight blades [17]-[21]. The Helical-turbine is one of the most important developments of vertical axis wind turbines. It is characterized by its aerodynamic equilibrium during its rotation, as well as overcoming structural problems such as bending stress and vibrations. While its efficiency is slightly lower than other types. [22], [23]. All types of vertical axis air fans that were mentioned above and that were not mentioned also, suffer from the inability of the air rotor to start rotating. This problem is relatively addressed by adding a secondary system that relies in its rotation on obstruction to make the air rotor easier to start rotating. But this solution did not completely solve the self-starting problem [24]-[27]. All VAWT's models of vertical axis wind turbines were produced and attempted to be sold globally as HAWT's, but with little traction [28], [29]. The originality of this research appears in its geometric design, which takes the form of two inverted parabolas, and the modification of the famous mathematical method "Multi-Stream Tubes" MST so that it becomes appropriate to apply it to this design. The objective of this improved design is to increase the power factor and reduce the problem of starting movement compared to other types of VAWT's. Figure 1 illustrates the developed design of this research. The work methodology was completed so that twelve geometric shapes of the rotor SS-VAWT were selected, different among them in terms of rotor height 2H, rotor diameter 2R, number of rotor blades N, length of rotor blade cross section C or rotor blade inclination angle β. Subsequently, several analyzes were conducted with the aim of verifying the effect of SS-VAWT rotor geometry variables on its performance. Through the results of the analyses conducted by a software specially designed for this research, the optimum engineering design of the SS-VAWT rotor was reached, which was moved to other stages of this research, including analyzes by Ansys Fluent, manufacturing a prototype and testing it in a wind tunnel. This is a comprehensive research that started with a developed idea that came based on an extensive review of many previous studies and research, this paper shows theoretical aspects that have an applied research extension that is being implemented now through the manufacturing processes of a prototype which will be shown by the experiments that will be conducted on it in subsequent papers in the future.

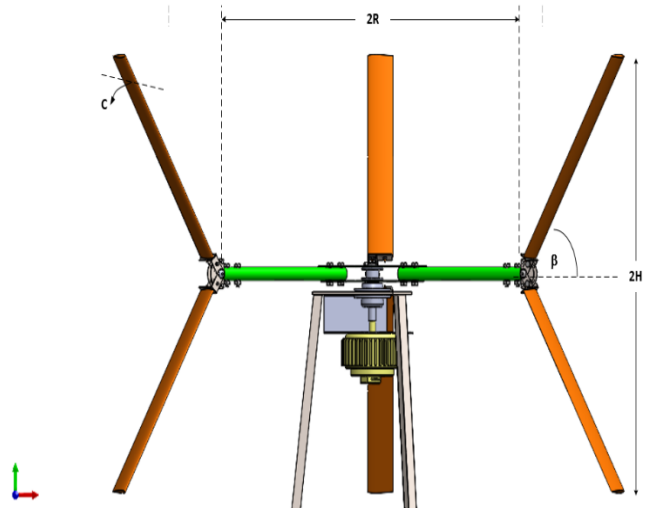


Figure 1: Geometry shape of SS-VAWT model by Solidworks

2. Method

The Multi-Stream Tube is mathematical method used in this work with certain adjustments to be optimal for apply to SS-VAWT and this model is preferred because it has a strong predictability for overall power production of the rotor and is also distinguished by efficiency and ease of use in the study of the influence of geometrical shape parameters on performance of wind rotor. With relation to details provided in [30], a brief overview of this model is offered. The rotor is substituted by the named "imaginary actuator disk." This disk was blocked by a sequence of "stream-tubes." It is an imagined tubs, the top of which consists of a stream-line, and thus the velocity vector is tangent everywhere on its surface. The velocity of the air across the disk differs from one stream-tube to another, much as the velocity of the air during the stream-tube shifts from its "free velocity" value V1 in front of the disk to its air velocity value "V" at the disk stage and even to its velocity "V2" at the wake area behind the disk. This continuous variation in disk velocity happens when part of the kinetic energy in the flow from which it passes is extracted [31]. Previous experimental data obtained in a wind tunnel was used to evaluate the software. The program was created to evaluate three different kinds of vertical axis wind turbines: original Darrieus rotor (Darrieus-VAWT), straight blades Darrieus rotor (H-VAWT), and slant straight blades rotor (SS-VAWT), the latter of which is the subject of this study. Figure 2 depicts the program approach in detail. Table 1 lists the twelve different SS-VAWT designs that have been examined and their efficiency

characteristics evaluated. The aerodynamic efficiency under the control of geometrical variables like blade inclination angle (β), wind rotor height (2H), wind rotor diameter (2R), number of rotor blades (N), and airfoil chord line was expressed using power factor curves derived from the analysis phase of the twelve configurations (C). For the presentation of the results in the shape of the Cp curves, the data for the rotor output are typically shown in non-dimensional formats, which allows such data to be utilized irrespective of the wind rotor scale while preserving geometric continuity across rotors in various dimensions. As a result, power coefficient curves as a characteristic of tip speed ratio are popular. The following diagram depicts the presentation of many important equations in the utilized model.

Table 1: Twelve configurations of SS-VAWT

"H" change configurations			
N	R	H	C
3	1	0.5	0.12
3	1	1	0.12
3	1	1.5	0.12
"R" change configurations			
N	R	H	C
3	0.5	1	0.12
3	1	1	0.12
3	1.5	1	0.12
"N" change configurations			
N	R	H	C
2	1	1	0.12
3	1	1	0.12
4	1	1	0.12
"C" change configurations			
N	R	H	C
3	1	1	0.06
3	1	1	0.12
3	1	1	0.18

2.1. Mathematical Expressions

- following SS-VAWT geometry expression was created for this research:

$$\left(\frac{r}{R}\right) = \left(\frac{\left(\frac{Z}{H}\right)}{\frac{1}{\left(\frac{H}{R}\right)} * \mathrm{Tan}\left(\beta * \frac{\pi}{180}\right)}\right) + 1 \quad (1)$$

- The power P is given by:

$$P = \frac{\rho NC}{2\pi} \int_0^H \int_0^\pi W^2 r \frac{\omega C_t}{\sin\beta} d\theta dz \quad (2)$$

- Tip speed ratio relation is:

$$\lambda_\circ = \frac{\omega R}{V_1} \quad (3)$$

- Mathematical formula of power coefficient, C_p, is:

$$C_P = \frac{P}{1/2 \rho V_1^3 A} \quad (4)$$

and

$$C_p = \frac{NC}{\pi A} \int_0^H \int_0^\pi \left(\frac{W}{V_1}\right)^2 \lambda_0 \frac{r}{R} \frac{C_t}{\sin\beta} d\theta dz \quad (5)$$

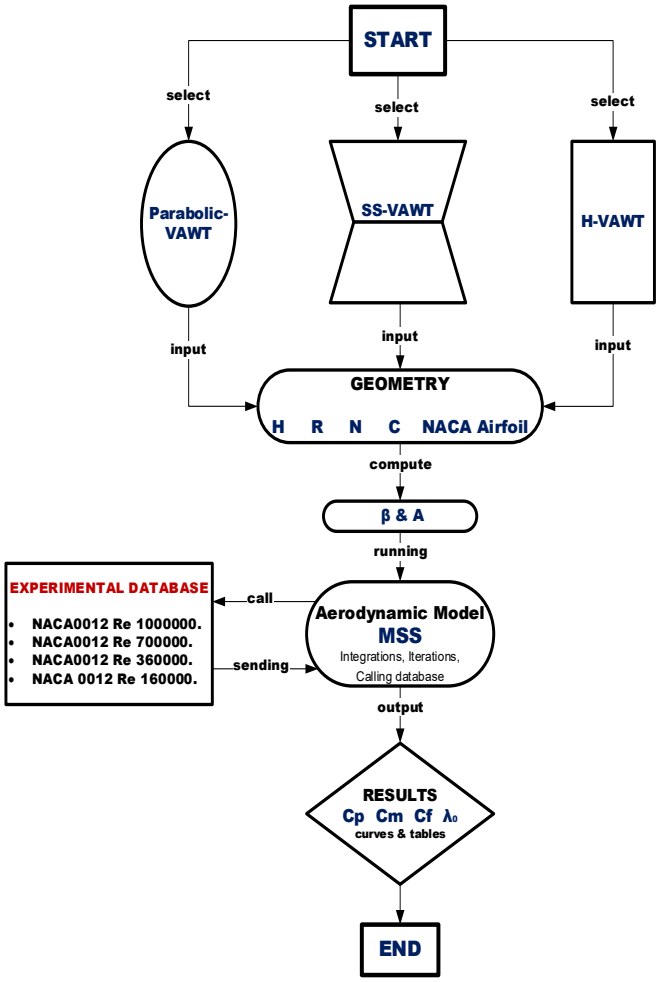


Figure 2: Flowchart of the software utilized

3. Results and Discussion

3.1. (β) Effect at Variable Values of H

The design group specialized in studying the effect of H change with β change is shown in Figure 3**.** The results are divided into four classes of curves, each group representing a different β angle, where we have four different angles as discussed above. In the same way, each category contains three curves that resulted from three different H values. The influence of the b angle on the Cp at three different values of H can be expressed as follows: at an angle of 85 degrees, the curves pattern did not vary much and there was a severe convergence between them, suggesting that there is no major effect of H value difference at this angle. The rotation began with a value of The value of CpM approached 0.3 at a value of λ_0 of around 3.8, and the values of the power factor appeared in a small λ_0 range between 3 and 5, indicating that the turbine is

inefficient in this case due to the low value of CpM, the narrow range of λ_0, and the magnitude of the self-starting issue. When b is equivalent to 65, the turbine's output increases when opposed to the situation where the angle is 85 degrees. In this scenario, the original λ_0 value is between 2 and 2.6, and the CpM at a comparable λ_0 value of 3.2 is calculated to be about 0.36. Furthermore, the value of λ_0 has improved, now varying between 2.2 and 4.8. As a consequence, these circumstances culminated in a relative dominance of 85 over the β. In comparison to the previous two cases of b equal to 85 and 65 degrees, the efficiency excellence of the third party described by β equal to 45 degrees is apparent. As the turbine rotation range increased from λ_0 (approximately 1.2) to 1 (approximately 4 and 5), the turbine rotation range also grew. Furthermore, the CpM values have been improved to approximately 0.36 and 0.42. To be more precise, the wind turbine in this case offered more control, a longer operating range, and a solution to the self-starting dilemma. When the β is equivalent to 25 degrees, the output of the SS-VAWT is investigated, the curves reveal that the turbine works poorly and is unsuitable for electrical energy generation. Where the power factor CpM highest value varied from 0.16 to 0.27.

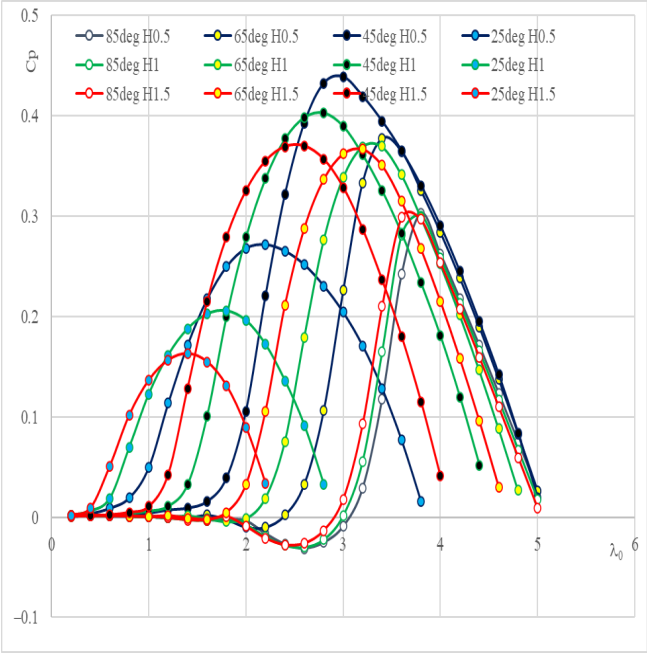


Figure 3: Power factor curves for twelve geometric shapes with change in height and angle of inclination of the blades.

3.2. (β) Effect at variable values of R

As seen in Figure 4 changing the angle and rotor radius values results in different Cp curve action. The engineering configurations that performed best were (β = 85, R = 1.5), (β = 65, R = 1.5, 1), and (β = 45, R = 1.5, 1), with CpM values ranging from 0.36 to 0.45. However, in terms of the energy continuum, (β = 65 degrees, R =1.5) and (β = 45 degrees, R =1.5) is preferred, with the rotor (β= 45 degrees, R =1.5) marked by an early start of rotational speed at the value of = 1.6. The bulk of the instances in Figure 4 had a relatively narrow energy range and a very low Cp, which led to poor performance.

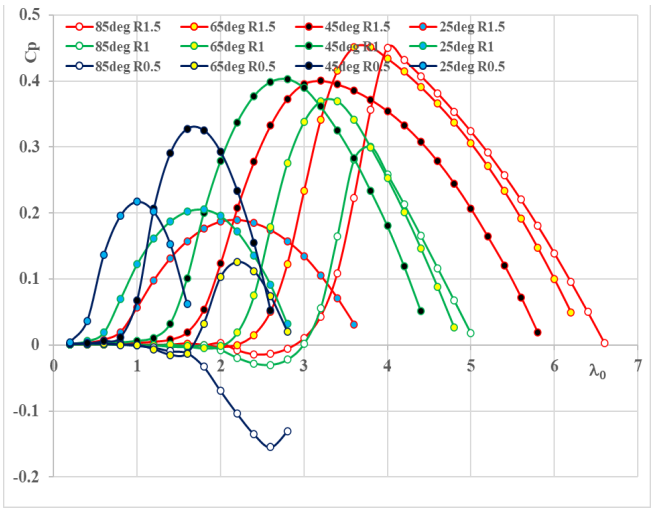


Figure 4: Power factor curves for twelve geometric shapes with change in diameter and angle of inclination of the blades.

3.3. (β) Effect at variable values of N

The three classes of curves in Figure 5 each show a specified N as 2, 3, or 4 and each include four curve angles = 85, 65, 45, and 25 degrees. It's readily evident in each category where N remains unaltered that the value of CpM increases, then falls, with the peak of CpM occurring at angles of 65 and 45 degrees. If we define the curves in another way, dividing them into four classes, each with a different fixed angle and vector N, we can see that when the beta 85 decreases dramatically, the value of the CpM decreases sharply as well, and the value of the energy spectrum decreases as well. It is obvious that modifying the conduct of curves at = 65 degrees is equivalent to changing attitudes at = 85 degrees in the preceding example. In the case of = 45 degrees, changing the value of N has no effect on the value of CpM, where It ranged between 0.37 and 0.40 in both situations, and it is balanced by values varying between 2.4 and 3, and the energy spectrum seems broad as compared to other sample cases. Finally, in the case of 25, the value of CpM rises with increasing N, but it does not surpass 0.24, considering the fact that the energy spectrum stays very small despite the shift in N.

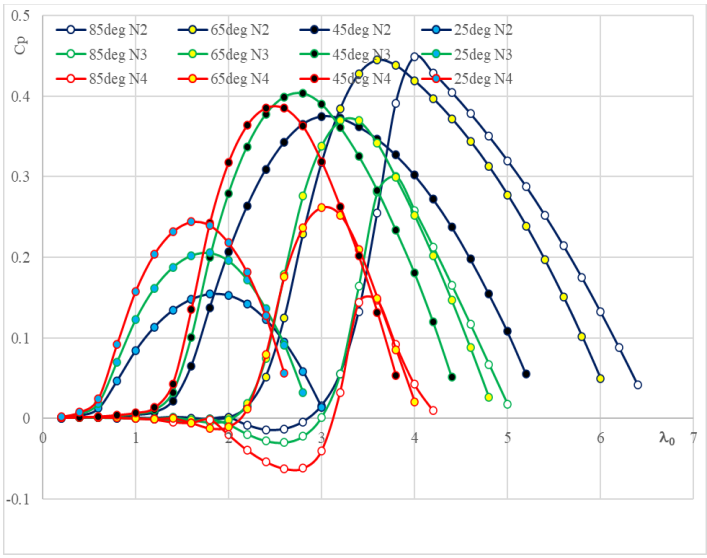


Figure 5: Power factor curves for twelve geometric shapes with change in number of blades and angle of inclination of the blades.

3.4. (β) Effect at variable values of C

Figure 6 expresses the effect of changing the blade section length on the SS-VAWT rotor in terms of power factor curves patterns. It is clear that there is a great similarity in the patterns of the Figure 6 curves with the patterns of the Figure 5 curves in paragraph 3.3. Thus, the ratio N/C can be used instead of N separately and C separately to check the efficiency of the SS-VAWT as a shortening of the analysis time.

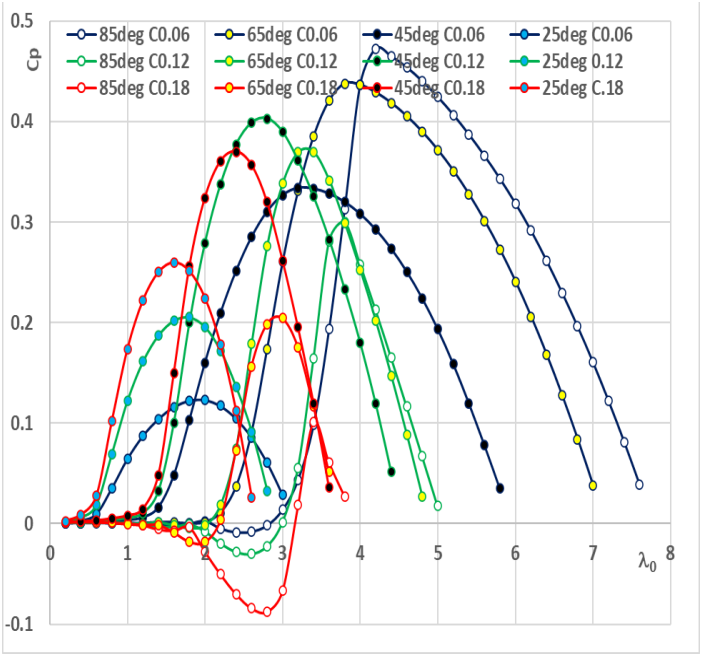


Figure 6: Power factor curves for twelve geometric shapes with change in cord line length and angle of inclination of the blades

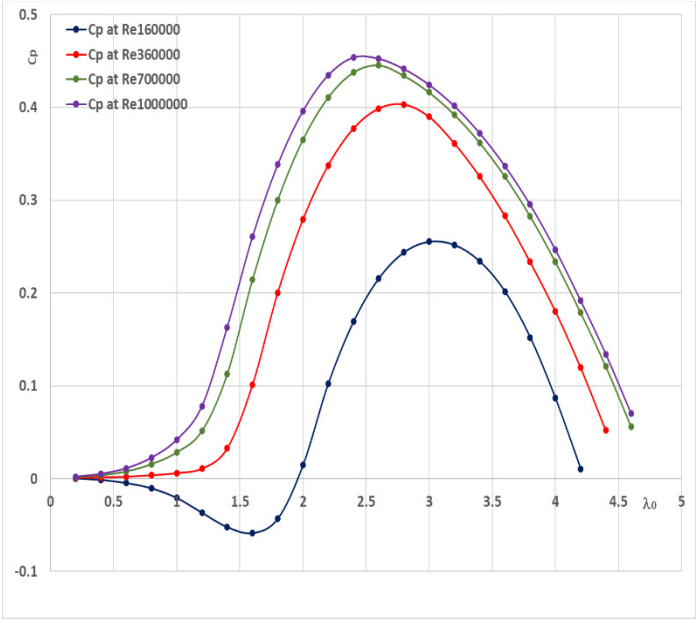


Figure 7: Influence of Reynolds Number on performance of SS-VAWT

3.5. Effect Reynolds Number on power factor of SS-VAWT

Figure 7. shows the change in the power factor values by changing the Reynolds number values. The values are 160000, 360000, 700000 and 1000000 [32]. It can be seen that in the case Reynolds number is equal to 160,000 the value of CpM is 0.25 and corresponding to λ_0 equal to 3. When Reynolds number is equal to 360,000, the CpM value is 0.4 and corresponding to λ_0 is 2.8. As for the Reynolds number equal to 700000, the CpM value is 0.44 and occurred at λ_0 equals 2.6. Finally, the Reynold Number 1000000 yielded CpM is 0.46 corresponding to a λ_0 of 2.4 value. Power factor curves also show that the power range ranges from 2 to 4.2 when the Re is 160,000, and 1.7 – 4.4 for Re equal to 360,000, and when Re = 700,000 shows the power range from 1.4 to 4.6, while at Re =1000000 the power range is 1.2 -4.7.

4. Conclusion

Within the scope of this research, the computer program achieved its objectives, as it predicted with high accuracy the power factor of the SS-VAWT rotor in many different study cases. By comparing with a previous study case that was harnessed for comparison only, the program is in great agreement with the practical results. The most important results extracted from this research are that the highest CpM values and the widest energetic range are achieved at angles of inclination of the blades between 45° and 65°. Regarding the problem of initiation of movement of the rotor, the satisfactory results were in the research cases in which the angles of inclination of the blades are between 25° and 45°. Regarding the effect of the Reynolds number on the performance of the SS-VAWT rotor, when Re value increase, the CP values increase, and consequently the CpM values increase. Moreover, as the Re values rise, the power range rises over the λ_0 values. Also, the problem of the self-starting appeared to diminish as the Re values accented.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] H. Abusannuga and M. Özkaymak, "Towards Evaluating the Performance of Vertical Axis Wind Turbine Consist of Slant Straight Blades," in 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Madrid, 2020.

[2] A. Kulshrestha, O. Mahela, M. Gupta, N. Gupta, N. Patel, T. Senjyu, M. Danish and M. Khosravy, "A Hybrid Fault Recognition Algorithm Using Stockwell Transform and Wigner Distribution Function for Power System Network with Solar Energy Penetration," Energies, **13**(14), 3519, 2020. doi: 10.3390/en13143519.

[3] A. Faramarz, T. Shiro, N. Tsutomu, N. Tomokazu, M.R. Asharif, " Analysis of Non-linear Adaptive Friction and Pitch Angle Control of Small-Scaled Wind Turbine System," (eds) Control and Automation, and Energy System Engineering. Communications in Computer and Information Science, 26-35, 2011. https://doi.org/10.1007/978-3-642-26010-0_4

[4] B. F. Blackwell, "The Vertical-Axis Wind Turbine "How It Works"," Sandia Laboratories, Albuquerque, New Mexico, 1974.

[5] R. Templin, "Aerodynamic Performance Theory for the NRC Vertical-Axis Wind Turbine," National Aeronautical Establishment, LTR-LA-160, Canada, 1974.

[6] M.J. Ralph, S.V. Maria and D. J. Ray, "Theoretical Performance Of Cross-Wind Axis Turbines With Results For A Catenary Vertical Axis Configuration," NASA Langley Research Center, Virginia, 1975.

[7] P. N. Shankar, "On The Aerodynamic Performance of a Class of Vertical Shaft Windmills," Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, **49**(16), 35-51, 1976.

[8] E. Sheldahl, B.F. Blackwell, "Free-Air Performance Tests of a 5-Metre-Diameter Darrieus Turbine," Sandia Laboratories, New Mexico, 1977.

[9] I.Paraschivoiu, "Double-Multiple Streamtube Model For Darrieus Wind Turbines," in Institut de Recherche de l'Hydro-Quebec, Quebec, 1981.

[10] P.M. Kumar, K. Sivalingam, S. Narasimalu, T.C. Lim, S. Ramakrishna and H. Wei "A Review on the Evolution of Darrieus Vertical Axis Wind Turbine: Small Wind Turbines," Journal of Power and Energy Engineering, **7**, 27-44, 2019.

[11] B. Muhammad, H. Nasir, F. Ahmed, A. Zain, J. Sh. Rehan and H. Zahid, "Vertical Axis Wind Turbine – A Review of Various Configurations And Design Techniques," Renewable and Sustainable Energy Reviews, **16**(4), 1926-1939, 2012.

[12] T. Dang, "Introduction, History, And Theory of Wind Power," in 41st North American Power Symposium, Starkville, MS, 2009.

[13] W. Tjiu, T. Marnoto, M. Sohif, H. Ruslan and K. Sopian, "Darrieus Vertical Axis Wind Turbine For Power Generation I: Assessment Of Darrieus VAWT Configurations," Renewable Energy, **75**, 50–67, 2015.

[14] A.D. Thomas, L.M. Timothy, "Developments in Blade Shape Design for a Darrieus Vertical Axis Wind Turbine," Sandia National Laboratories, Albuquerque, New Mexico, 1986.

[15] K.G. Emil, "Characteristics of Future Vertical Axis Wind Turbines," Sandia National Laboratories, Albuquerque, New Mexico, 1982.

[16] W. Tjiu, T. Marnoto, M. Sohif, H. Ruslan, K. Sopian, "Darrieus vertical axis wind turbine for power generation II: Challenges in HAWT and the opportunity of multi-megawatt Darrieus VAWT development," Renewable Energy, **75**, 560-571, 2015.

[17] W. S. Bannister, "Aerodynamic Studies Of A Straight-Bladed Vertical-Axis Wind Turbine," in 1St. B.W.E.A Wind Energy Workshop, Springfield, VA, 1979.

[18] W. S. Bannister, "A Theoretical Analysis Of Small Vertical Axis Wind Turbines," in International Synposium on "Applications of Fluid Mechanics and Heat Transfer to Energy and Environmental Problems, Greece, 1981.

[19] Y. Hara., N. Horita, S. Yoshida, H. Akimoto, T. Sumi, "Numerical Analysis of Effects of Arms with Different Cross-Sections on Straight-Bladed Vertical Axis Wind Turbine," Energies, **12**(11), 2019, https://doi.org/10.3390/en12112106

[20] J. Fadil, Soedibyo, M. Ashari, "Performance Analysis Of Vertical Axis Wind Turbine With Variable Swept Area,", International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, 2017.

[21] A. Gorlov, "Development of the Helical Reaction Hydraulic Turbine," Northeastern University, Boston, 1998.

[22] D. Han, Y. Heo, N. Choi, S. Nam, K. Choi, K. Kim, "Design, Fabrication, and Performance Test of a 100-W Helical-Blade Vertical-Axis Wind Turbine at Low Tip-Speed Ratio," Energies, **11**(6), 2018.

[23] M. Rahman, T. Salyers, , A. Mahbub, A. ElShahat, V. Soloiu, and M. Emile, "Investigation of Aerodynamic Performance of Helical Shape Vertical-Axis Wind Turbine Models With Various Number of Blades Using Wind Tunnel Testing and Computational Fluid Dynamics,", International Mechanical Engineering Congress and Exposition, Phoenix, Arizona, 2016.

[24] R. Dominy, P. Lunt, A. Bickerdyke, J. Dominy, "Self-Starting Capability Of A Darrieus Turbine," Journal of Power and Energy, **221**(1), 111–120, 2007.

[25] J. Zhu, H. Huang, & H. Shen, "Self-Starting Aerodynamics Analysis Of Vertical Axis Wind Turbine," Advances in Mechanical Engineering, **7**(12), 2015.

[26] M.S. Omar, I.A. Ahmed, E.A. Ahmed, A.A. Amr, A. M. Elbaz, "Numerical Investigation Of Darrieus Wind Turbine With Slotted Airfoil Blades," Energy Conversion and Management: X, 5, 100026, 2020.

[27] J. Krishnaraj, Sivakumar Ellappan, M. Anil Kumar, "Additive Manufacturing of a Gorlov Helical Type Vertical Axis Wind Turbine," International Journal of Engineering and Advanced Technology (IJEAT), **9**(2), 2019.

[28] IRENA, "Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects," International Renewable Energy Agency (IRENA), Abu Dhabi, 2019.

[29] J. Damota, I. Lamas, A. Couce, J. Rodríguez, "Vertical Axis Wind Turbines: Current Technologies and Future Trends," in International Conference on Renewable Energies and Power Quality (ICREPQ'15), Coruna, 2015.

[30] H.J. Strickland, "The Darrieus Turbine: A Performance Prediction Model Using Multiple Streamtubes," Sandia Laboratories, Albuquerque, New Mexico, 1975.

[31] N. Batista, R. Melício, V. Mendes, J. Figueiredo, A. Reis, "Darrieus Wind Turbine Performance Prediction: Computational Modeling,", 4th Doctoral Conference on Computing, Electrical and Industrial, Costa de Caparica, 2013.

[32] R. E. Sheldahl, Klimas, P C, "Aerodynamic Characteristics of Seven Symmetrical Airfoil Sections Through 180-Degree Angle of Attack for Use in Aerodynamic Analysis of Vertical Axis Wind Turbines," Sandia National Laboratories, United States, 1981.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 297-302 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Reading Acquisition Software for Portuguese: Preliminary Results

Ana Sucena*, Ana Filipa Silva, Cristina Garrido, Cátia Marques

Research and Intervention Reading Centre, Polytechnic Institute of Porto, Porto, 1800-412, Portugal

ARTICLE INFO

Article history:
Received: 07 May, 2021
Accepted: 15 June, 2021
Online: 19 October, 2021

Keywords:
I Read
Acquisition
Reading
Spelling

ABSTRACT

The persistent difficulties in reading and spelling acquisition are a risk factor for learning motivation. Play-like intervention tools have been developed to face these difficulties. "I read" is a software that seeks to develop and introduce systematic reading and spelling skills training in a playful and complementary way. This software is intended for children at the beginning of their school journey, as well as for those who reveal reading and/or spelling difficulties. This article intends to present the goals and the structure of this software, as well as the preliminary results of its implementation and game enjoyment with 244 children between 5 and 7 years old. Results indicate that 58% of the participants completed the activities dedicated to alphabetical decoding, and 42% were able to reach the last stage of the game, dedicated to orthographic decoding. Regarding the enjoyment with the software, 96% of the participants classify the games as fun games. In conclusion, training with this software revealed to be beneficial for reading and spelling skills promotion, as well as to increase the overall enjoyment and motivation for learning.

1. Introduction

Research highlights persistent difficulties in reading acquisition as a key risk factor for school motivation, learning and children's self-esteem [1]. These difficulties lead to long-term negative consequences, such as behavior problems, depression and other mental health issues [2]. If reading disabilities are not early addressed, difficulties tend to generalize to other domains, exposing students to consecutive experiences of failure, thereby diminishing their motivation to learn [3], jeopardizing future knowledge acquisition [4]. In turn, when these difficulties are identified early and lead to a prompt intervention, the likelihood of reversing failure trajectories is very high [5], [6].

Thus, early intervention to face learning difficulties, namely on reading and spelling [7]-[10] is of utmost importance. Early intervention can have an even more ambitious purpose: preventing the child from experiencing learning difficulties, hence the term "at-risk" (of experiencing reading and writing learning difficulties). The risk for retention, school dropout and low motivation decreases with the adoption of early intervention focused on the key competencies for reading acquisition [11]. Phonemic awareness, letter-sound knowledge and decoding skills are the key competencies for successful reading and writing acquisition [11], [12].

*Corresponding Author: Ana Sucena, asucena@ess.ipp.pt

As a result, the past decade has been witnessing the fast development of information technologies, together with the rapid development of serious games [13]. Serious games consist of entertaining tools with educational purposes, presented as digital games throughout which children are required to overcome numerous obstacles. The serious games primary goal is to contribute to increasing the players' motivation to acquire and expand knowledge and practising skills [13]. Educational aspects are included in the gameplay, which will be acquired and automatized by the players during the gaming process [13]. To ensure effective learning in serious games many aspects are important. Designers are not only faced with the challenge of creating appealing graphics and a compelling narrative, but also with the additional challenge of incorporating suitable pedagogical strategies into their narratives [14].

In sum, serious games are congruently described in the literature as digital games that have an explicit and carefully thought-out educational purpose and are not solely intended for amusement [15], [16]. The development of these types of games, in a motivating, challenging and playful environment has proven to be a more effective tool at conveying knowledge and achieving behavior change [14] than the traditional teaching and learning methods [17], [18]. Serious games also present an opportunity for proactive learning by involving players in an immersive learning experience where they can apply their knowledge, learn from

experience, and test complex or risky scenarios in a safe and play-full environment.

Player enjoyment is the term used to express the satisfaction perceived when performing tasks and activities that require high levels of concentration and skills in digital games [19]-[21]. Enjoyment is a form of fun that serious games can promote, resulting from the challenges the players face and must be able to overcome to advance in the game [22]. Player enjoyment is an important aspect for educational serious games, as they are commonly introduced to students by professionals responsible for the teaching process who are leading students to play, rather than the students starting to play willingly or through a recommendation from friends. Therefore, promoting player enjoyment can keep them interested in continuing to play along with all the levels, continuing to develop their skills in the contents covered by the game, potentially enhancing the results from the use of these educational tools in specific domains.

The game levels are used to show the progress of users. It is a common practice for serious games to adopt levels to show increasing difficulty and complexity levels. This option is based on shaping - a behaviour changing strategy. Shaping consists of breaking tasks down into small discrete steps and training in a systematic and precise way [23]. This can be linear (e.g., you level up after earning 50 points), or exponential. Creating the levels at the right difficulty and complexity (level balancing) is important to avoid drop-outs. Difficulty can also serve as a way to motivate the users as they passed a specifically difficult level the child gets points and move forward in the game [24]. The different levels give players a direction in order to reach the goal that needs to be achieved. Players have something to accomplish within their overall experience [25].

Due to these features, serious games have been successfully applied in various domains such as medicine, military, or complex processes training, among others; example success-cases include commerce, environment and ecological behaviour, cartography, machine learning, software development, innovation, health and medical issues, tourism, energy, mobility and transportation, accessibility fashion, usability, risk management, marketing [26] and with children with learning disabilities [11].

Serious games have become increasingly popular in education, supporting children with learning disabilities [27]-[32]. It can be used as a tool to reinforce the teaching process and provide enriching learning contexts for its users [26]. Most children with learning disabilities can be successful in reading acquisition with the help of a tutor or a specialized teaching program [33]. Emotional support also plays an important role for these children while dealing with learning disabilities [33]. Several studies have shown that children prefer playing a “fun” computer game than performing a “boring” exercise [34]. The idea is that the user does something useful, for his/her knowledge. If the user has fun, she/he will probably continue to play, achieving the serious goals of the game [34]. Serious games can be a great help in these situations, not substituting, but complementing the work done by therapists [35]. Moreover, improving the life quality of children with learning disabilities has a crucial impact on their learning and future path.

The results of serious games as a tool to intervene with children with learning disabilities are supported by the empirical findings of several studies [36]. There is an increasing demand for innovative and cost-effective evidence-based solutions to the educational challenges faced by teachers and students. Reading related difficulties represent 80% of the overall learning difficulties [37], which makes it all the more important to invest in developing reading acquisition serious games.

The Graphogame reading promotion software is one of the pioneer's projects regarding serious games in learning acquisition. The Graphogame provides children with training in phoneme awareness, letter-sound and early word decoding training. This play-like software was developed in a research project, originally developed in Finland, then adapted for over 20 language versions (e.g., Finnish [9], [38]-[40]; German [41], [42]; English [12], [43]; Portuguese [11], [48]; Spanish [45]). In this computer game set, the child uses headphones and hears a sound that corresponds to a letter, word or pseudoword, and he/she must select the written option corresponding to the sound.

Recently, a new Portuguese serious game was developed – I Read [46], [47]. This serious game aims to promote phonemic awareness, letter-sound knowledge and decoding. Compared to the Portuguese Graphogame the I Read has a greater focus on letter-sound knowledge as well as on spelling tasks. For each set of letter-sound, activities are focusing on monosyllabic words (CV and CVC syllabic structure) reading and spelling. The I Read serious game intends to facilitate reading and spelling acquisition in a playful and complementary environment. In this study, we describe the I Read software, along with preliminary results of its implementation on reading and spelling skills, as well as on the player enjoyment levels.

In this study we present the I Read software that aims to practise basic literacy skills, we present the main results of its implementation, as well as, assess the children enjoyment with the game.

2. Method

2.1. Participants

244 first graders participated in this study, aged between 5 and 7 years old, studying in Portuguese public schools. All children were identified as at risk of experiencing difficulties in reading. None of the participants revealed any sensory, motor or cognitive disturbances according to school registers.

2.2. Instrument

The I Read is a recently developed computer software, for reading and spelling training with Portuguese speaking children. I Read [36], [37] focus on the development of the alphabetical principle - training letter-sound and phonemic awareness - and decoding through systematic training, along with different games. This software is intended for children at the beginning of their school path and withchildren at risk of experiencing difficulties in reading acquisition. The I Read can be played on any computer or tablet, with internet access. Each child logs in with a username and password. Information about the game is recorded online, which allows the child to resume the exact game in the next session.

The general and specific goals of the I Read software are described in table 1. There are 13 game levels. Between level one and nine, the goal is to promote the alphabetic principle and alphabetic decoding, thus all stimuli correspond to simple graphemes and have simple syllabic structure. Between level 10 and 13, the focus shifts to orthographic decoding, with stimuli composed of complex graphemes and complex syllabic structure.

Table 1: I Read goals

General goals	Specific goals
Simple vocalic grapheme (level 1)	Phonemic awareness Letter-sound knowledge Alphabetical decoding
Oral diphthongs (level 2 and 3)	Phonemic awareness Alphabetical decoding Spelling decoding
Nasal diphthongs (level 4)	Phonemic awareness Alphabetical decoding Spelling decoding
Simple consonant grapheme (level 5 to 9)	Phonemic awareness Letter-sound knowledge Alphabetical decoding Spelling decoding
Complex consonant grapheme (level 10 to 13)	Phonemic awareness Letter-sound knowledge Spelling decoding

Within each level, the difficulty increases along with the games. For example, at level five (p, t, s) there are 13 sub-levels.

The 1st and 2nd sub-levels focus on phonemic awareness and letter-sound knowledge training; from the 3rd to the 13th sub-level, the goal shifts to alphabetical decoding. The same difficulty continuum applies to the syllabic length, as words/ pseudowords presented on the initial levels are monosyllabic CV, CVG– whereas two-syllable words/ pseudowords are introduced in more advanced levels - CV.CV, CVG.CV/CV.CVG. Finally, reading tasks precede spelling tasks - from the 3rd until the 8th sub-level the focus is on reading games; spelling games appear from the 9th level.

The narrative behind the software is that of a challenge to discover planets and satellites in a galaxy, along with spatial-like music, intending to contribute to a playful appearance. Each planet represents a level, whereas satellites (around each planet) represent the game sublevels. Each game is presented with a different scenario (e.g., under the ocean; in space; on a submarine) and contains multiple trials. Each trial contains multiple stimuli, always presented along with distractors.

There are two types of game: a) identification/ reading – the child hears a sound corresponding to a letter, word or pseudoword and he/she identifies the orthographic match on the screen; b) spelling – the child hears a word or pseudoword and is required to select and order the constituent letters (presented in the screen, in random order).

In all games, there is a button represented by a loudspeaker, which the child can press whenever he/she needs to hear the target stimulus again. The accuracy feedback is provided immediately after the correct or incorrect answer (through an auditory and visual stimulus). If the answer is inaccurate, the stimulus visually trembles and disappears from the screen, thus reducing the number of stimuli on the screen (until only the right option appears on the screen). The software adapts to the performance of the player, by eliminating the possibility for repeated mistakes. As the child starts improving performance, the number of distractors increases again. It was adopted the errorless learning strategy. This strategy aims to minimize opportunities for errors, increasing the frequency at which the child encounters reinforcement. Minimizing errors also reduces the likelihood that the child engages in challenging behavior [39]. When the child selects the correct answer, a green checkmark appears on the screen (along with a "correct-like" sound). At the end of each game, there is feedback on the game's performance, through the attribution of one to three stars.

I Read can be played in two modes, free or pre-set mode. Thus, it is possible to select specific levels/planets (free mode, designed for the teacher/ adult accompanying the training) or to follow the preset outline (designed for the child playing on her own) by playing the levels in increasing difficulty, with the advantage of the software adaptation to each child learning rhythm [39].

The player enjoyment with the I Read was assessed by asking the players to classify their satisfaction with the game using a three Likert scale using smiles. If the players are very satisfied they answer three smiles, more or less satisfied two smiles and if they are not satisfied one smile.

2.3. Procedures

This software was implemented with at-risk first graders. Training occurred 3 to 4 times a week in sessions between 15 to 20 minutes, in a school context, between November and May of 2020, under the supervision of an expert in reading and spelling promotion, with specific training on the software. Each child played individually, on a computer, in a noiseless environment and/or with the use of headphones. Children played in the preset mode, according to their learning pace. The information of each player gaming sessions was recorded, so the child could resume playing at the point where he/she stayed the last time he/she played, on any computer, as long as he/she has access to his/her user name and password.

3. Results

The ultimate challenge – playing all the 13 planets and respective satellites up until the last one - was reached by 48 children (15% of the total children). On average, participants took 18 sessions to reach the final game, with a frequency of play of about once per week.

On a level-by-level analysis, all participants concluded the vowels and diphthongs levels. The percentage decreases when we analyse the alphabetic decoding and the orthographic decoding (respectively, 58 % and 42%) (Table 2).

Table 2: Distribution of participants reaching each level of difficulty

Levels	Number Participants (%)	Number Participants (%) by levels of groups
1	0	0
2	0	
3	0	
4	0	
5	50 (20.5%)	142 (58.2%)
6	27 (8.6%)	
7	21 (6.7%)	
8	25 (7.9%)	
9	19 (6.03%)	
10	32 (10.2%)	102 (41.8%)
11	8 (2.5%)	
12	14 (4.4%)	
13	48 (15.2%)	

Regarding player enjoyment, 233 (95.5%) participants found fun to play the games (Figure 1).

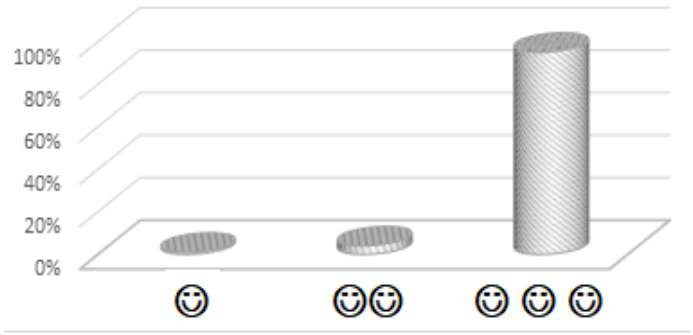


Figure 1: Player's enjoyment

4. Discussion

Ensuring literacy acquisition in early grades is a fundamental investment. In this study, we intended to (i) present the I Read software that aim to practice basic literacy skills, (ii) the main results of its implementation as well as (iii) children enjoyment with the game. The development of serious games has a large impact on education, prompting the need for upgraded policy making, curriculum development, and teacher training. Providing teachers and educational technicians with research technology empowers them to take their places as qualified experts and advisers that policymakers need for making informed choices for better education quality [4].

The results of this study show that all participants managed to finish the first stage, focused on the Alphabetic Principle - vowels, diphthongs, and oral and nasal diphthongs. Around 60% of children reached the second stage – focused on alphabetic decoding. These are promising results, as these students finish the first grade dominating the alphabetic decoding. Having started the 1st grade with weak pre-reading skills, these children conclude the 1st grade with reasonable reading skills, thus moving on to the 2nd grade without the at-risk label. The I Read software positively impacts the attentional and motivational level [11], which are, for at-risk children, often negatively affected by the growing awareness of their difficulties [8].

There is, yet, potential to increase the average results. Regarding the last stage - focusing on orthographic decoding – the results are modest. This stage was attained by 42% of the children and only 15% managed to finish all games. It is important to develop strategies to increase the percentage of children who finish all games. One possible strategy could be to count on the support of an adult during the time when the child is playing I Read. Complementary activities with the supervision of an adult in order to promote orthographic decoding are also an important strategy to help children facing difficulties. The results indicate that children at risk for reading acquisition difficulties benefit from a remedial reading intervention that is enriched with the use of the I Read software at the very beginning of the first grade.

Regarding player enjoyment, the results confirm the playful nature of the software as an effective tool to intervene with children with learning disabilities. These results also encourage us to continue the intervention, in particular with children who are facing difficulties. The focus of the software on phonemic awareness, letter-sound knowledge and decoding skills, which are the key competencies for successful reading and writing acquisition [8], [9], [11], [12], [16] along with the playful environment, proved to be highly beneficial to promote reading acquisition. Thus, a simple learning game can provide learning and playful opportunities for struggling readers.

In future studies, it is important to test the impact of the software in terms of effective learning, creating two groups of at-risk children – an intervention group that plays regularly and a comparative group that does not use the software. Also, in order to understand the role of the teacher's training in the application for the software, in future studies this variable should be controlled. Future studies could also include a thorough analysis of the intervention process to better understand the evolution of participants from session to session, along with the teachers' perspective regarding the software efficacy and a follow-up study of the participants, thus contributing for this educational tool to become even more effective for teachers and students. Since literacy is closely tied to social and economic development, quality research on this subject, along with its dissemination play a key role in the maintenance and increase of economic and social increase. Adaptations of the I Read to different Portuguese speaking countries and for other languages could be an important step for the worldwide promotion of reading acquisition with at-risk children, by focusing the training on the very specific skills, highly associated with success in reading acquisition.

5. Conclusion

Successful reading and writing acquisition is associated with an overall and long-lasting good performance in the educational, social and professional path. In the same vein, any impairment in the early reading acquisition can negatively affect the life trajectory in general. To intervene with at-risk children regarding reading acquisition is known to have high economic revenue, as a successful reading acquisition increases the odds for broad school success. Losing one year in the labour force (after school retention

caused by learning difficulties) corresponds to a loss of around $80,000, on an estimate for a college-educated male working full time and year-round, retiring at age 67 in the USA [42].

Serious games are highly effective on reading acquisition as well as, on broader terms, on the learning process. Serious games are a strong complement to more conventional instruction methods, with a higher capacity for motivating children in result of the game-like and technology-like appearance. Research has shown that serious games have the potential to prevent and remediate reading difficulties. The use of the playful-like I Read software, focusing on reading and spelling skills, has proven to be an important complementary tool for early intervention with at-risk children. It is the author's belief (and expectation) that by captivating the children's attention for reading-related training, the I Read will play an important role in the foundations of the overall enjoyment and motivation for learning.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by European Horizon 2020, under OPERAÇÃO NORTE-08-5266-FSE349 000095.

References

[1] R. Lovio, A.Halttunen, H.Lyytinen, R Näätänen, T. Kujala, "Reading skill and neural processing accuracy improvement after a 3-hour intervention in preschoolers with difficulties in reading-related skills," Brain Research, **11**, 42-55, 2012, doi:10.1016/j.brainres.2012.01.071

[2] A. M.Undheim, L.Wichstrøm, A. M.Sund, "Emotional and behavioral problems among school adolescents with and without reading difficulties as measured by the youth self-report: a one-year followup study," Scand. J. Educ. Res. **55**, 291–305, 2011, doi: 10.1080/00313831.2011. 576879

[3] H. Lyytinen, J.Erskine, "Early identification and prevention of reading problems," Encyclopedia on Early Childhood Development, 1–5, 2016. Retrieved from http://www.childencyclopedia.com/documents/Lyytinen-ErskineANGxp.pdf

[4] S.Raspin, R.Smallwood, S.Hatfield, L.Boesley, "Exploring the use of the ARROW literacy intervention for looked after children in a UK local authority," Educational Psychology in Practice, **35**(4), 1-13, 2019, doi:10.1080/02667363.2019.1632172

[5] M. Hall, M. Burns, "Meta-analysis of targeted small-group reading interventions," Journal of School Psychology, **66**(1), 54-66, 2018. doi:10.1016/j.jsp.2017.11.002

[6] H. Lyytinen, "New Technologies and interventions for learning difficulties: Dyslexia in Finnish as a case study". In Foresight Mental Capital and Wellbeing Project: The Government Office for Science. Government Office for Science, London: UK, 2008.

[7] P.Hatcher, C.Hulme, M.Snowling, "Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure" Journal of Child Psychology and Psychiatry, **45**(2), 338–358, 2004.

[8] U. Richardson, H. Lyytinen, "The GraphoGame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read", Human Technology, **10**, 39-60, 2014. doi:10.17011/ht/urn.201405281859.

[9] N.L. Saine, M.K.Lerkkanen, T.Ahonen, A.Tolvanen, H.Lyytinen, "Computer-assisted remedial reading intervention for school beginners at risk for reading disability", Child Development, **82**, 1013-28, 2011doi:10.1111/j.1467-8624.2011.01580.x

[10] H. Wimmer, H.Mayringer, "Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies", Journal of Educational Psychology, **94**(2), 272-277, 2002.

[11] A.Sucena, , J.Cruz, F.L.Viana, A.F.Silva, "Graphogame português alicerce: software de apoio a crianças disléxicas. In M. J. Gomes, A. J. Osório & L. Valente (Eds.), Atas da IX Conferência Internacional de TIC na Educação (pp. 396-405). Braga: Universidade do Minho/Centro de Competência em TIC na Educação, 2015.

[12] F.E. Kyle, J. Kujala, U. Richardson, H. Lyytinen, U.Goswami, "Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG rime and GG phoneme", Reading Research Quarterly, **48**, 61-76, 2013, doi:10.1002/rrq.038.

[13] Y.Zhonggen, "A Meta-Analysis of Use of Serious Games in Education over a Decade. Hindawi" International Journal of Computer Games Technology 2019, 4797032, doi:10.1155/2019/4797032

[14] O.Troyer, F.Van Broeckhoven, , J.Vlieghe, "Linking serious game narratives with pedagogical theories and pedagogical design strategies," J Comput High Educ **29**, 549–573, 2017, doi:10.1007/s12528-017-9142-4.

[15] C.C. Abt Serious Games University Press of America, 1987.

[16] M.A. Khenissi, F. Essalmi, M. Jemni, "Comparison between serious games and learning version of existing games," Procedia - Social and Behavioral Sciences, **191**, 487-494, 2015, doi:10.1016/j.sbspro.2015.04.380.

[17] N. Bäulke, M.Eckerlein, M. Dresel, "Interrelations between motivational regulation, procrastination and college dropout intentions" Unterrichtswissenschaft, **46**(4) 461-479, 2018, doi: 10.1007/s42010-018-0029-5.

[18] Y. Du, S. Chaaban, A.M. Sabah, L. Al-Thani, Wang, Active learning engagement in teacher preparation programmes—A comparative study from Qatar, Lebanon and China Asia Pacific Journal of Education, **40**(3), 283-298, 2020, doi:10.1080/02188791.2020.1717436

[19] M. Csikszentmihalyi, "Finding flow: the psychology of engagement with everyday life," Psychol. Today, 1-7, 1997.

[20] P. Sweetser, P. Wyeth Gameflow, "A model for evaluating player enjoyment in games," Comput. Entertain., **3** (2005), 1-3, doi:10.1145/1077246.1077253.

[21] J. Almeida, L. Machado, "Design requirements for educational serious games with focus on player enjoyment," Entertainment Computing, **38**(100413), 2021.

[22] M. Prensky, "Digital game-based learning," Comput. Entertain., **1**, 1-21, 2003.

[23] M. Spreckley, R. Boyd, "Efficacy of Applied Behavioral Intervention in Preschool Children with Autism for Improving Cognitive, Language, and Adaptive Behavior: A Systematic Review and Meta-analysis", The Journal of Pediatrics, 2019, doi:10.1016/j.jpeds.2008.09.012

[24] V.Insley, D. Nunan, "Gamification and the online retail experience," International Journal of Retail & Distribution Management, **42**(5), 340–351, 2014, doi:10.1108/IJRDM-01-2013-0030

[25] L.M. Lubbe, C.Gerritsen, M. Klein, K. Hindriks, "Empowering vulnerable target groups with serious games and gamification," Entertainment Computing, **38**(100402), 1-27, 2021, doi 10.1016/j.entcom.2020.100402

[26] Baptista, G., Oliveira, T., "Gamification and serious games: A literature meta-analysis and integrative model", Comput. Hum. Behav., **92**, 306–315, 2019. doi:10.1016/j.chb.2018.11.030

[27] M. Ronimus, K. Eklund, J. Westerholm, R. Ketonen, H. Lyytinen, "A mobile game as a support tool for children with severe difficulties in reading and spelling", Journal of computer assisted learning, **36**, 1011–1025, 2020, doi: 10.1111/jcal.12456.

[28] M. Niemelä, T. Kärkkäinen, S. Äyrämö, M. Ronimus, U. Richardson, H. Lyytinen, "Game learning analytics for understanding reading skills in transparent writing systemritish", Journal of Educational Technology, **51**(6), 2376–2390, 2020. doi:10.1111/bjet.12916

[29] K. Khowaja, S. Salim, "A framework to design vocabulary-based serious games for children with autism spectrum disorder (ASD)", Universal Access in the Information Society **19**, 739–781, 2020 doi:10.1007/s10209-019-00689-4.

[30] K.D. Godinez, L.S. Gaytán-Lugo, P.A. Alcaraz-Valencia, Rocío Maciel Arellano, "Evaluation of a Low Fidelity Prototype of a Serious Game to Encourage Reading in Elementary School Children", CLIHC '17, November 8–10, 2017, doi:10.1145/3151470.3156640.

[31] S. Serret, S. Hun, S. Thümmler, P. Pierron, A. Santos, J. Bourgeois, F Askenazy, "Teaching Literacy Skills to French Minimally Verbal School-Aged Children with Autism Spectrum Disorders with the Serious Game SEMA-TIC: An Exploratory Study", Frontiers Psychology, **8**(1523), 2017, doi: 10.3389/fpsyg.2017.01523.

[32] J.M. Thomson, N. Foldnesb, P.H. Uppstadc , M. Njåc , O. J. Solheimc , K. Lundetræ, "Can children's instructional gameplay activity be used as a predictive indicator of reading skills?", Learning and Instruction, **668**, 101348, 2020 doi:10.1016/j.learninstruc.2020.101348.

[33] A. J. Alcázar, E. Venegas, S. Criollo, S.L. Mora An Approach to Accessible

Serious Games for People with Dyslexia Sustainability 2021, **13**, 2507. doi:10.3390/su13052507.

[34] O. Gaggi, C. E. Palazzi, M. Ciman, G. Galiazzo, S. Franceschini, M. Rufino, S. Gorgi, A. Facoetti, Serious Games for Early Identification of Developmental Dyslexia, Comput. Entertain **15**, 2, Article 4 (February 2017), 1-24. doi:10.1145/2629558.

[35] R.M. Tomé, J.M. Pereira, M. Oliveira, "Using serious games for cognitive disabilities". In Proceedings of the International Conference on Serious Games Development and Applications, Berlin, Germany, 9–10, October 2014.

[36] M. J. Dondlinger, "Educational video game design: A review of the literature", Journal of Applied Educational Technology, **4**(1), 21–31, 2007.

[37] S. Shaywitz, "Entendendo a dislexia: um novo e completo programa para todos os níveis de problemas de leitura", Porto Alegre: Artmed. 2006.

[38] N.L. Saine, M.K. Lerkkanen, T.Ahonen, A. Tolvanen, H. Lyytinen, "Predicting word-level reading fluency outcomes in three contrastive groups: Remedial and computer-assisted remedial reading intervention, and mainstream instruction", Learning and Individual Differences **20**, 402-414, 2010 doi:10.1016/j.lindif.2010.06.004

[39] V. Markham, A. Giles, G. Davies, V. Adshead, G. Tamiaki, R. May, Applications of within-stimulus errorless learning methods for teaching discrimination skills to individuals with intellectual and developmental disabilities: A systematic review", Research in Developmental Disabilities, **97**, 103521, 2020, doi: 10.1016/j.ridd.2019.103521

[40] M.Ronimus, H.Lyytinen, "Is school a better environment than home for digital game-based learning? The case of GraphoGame". Human Technology, **11**(2), 123–147, 2015 doi:10.17011/ht/urn.201511113637

[41] S.Brem, S. Bach, K.Kucian, J. V. Kujala, T. K.Guttorm, E. Martin, et al."Brain sensitivity to print emerges when children learn letter-speechbsound correspondences", Proc. Natl. Acad. Sci. U.S.A. **107**, 7939–7944, 2010, doi:10.1073/pnas.0904402107

[42] S. Bach, U. Richardson, D. Brandeis, E. Martin, S. Brem, "Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade". NeuroImage, **82**, 605–615, 2013. doi: 10.1016/j.neuroimage.2013.05.062

[43] J.Worth, J. Nelson, J. Harland, D. Bernardinelli, B. Styles, "GraphoGame Rime: Evaluation report and executive summary", Slough, UK: National Foundation for Educational Research, 2018.

[44] A. Sucena, A. F.Silva, F.L.Viana, "Intervenção precoce nas dificuldades de aprendizagem da leitura com recurso ao software Graphogame", Revista Digital do Programa de Pós-Graduação em Letras da PUCRS Porto Alegre, **9**(2), 200-212, 2016.

[45] R. Rosas, J.P.Escobar, M.P. Ramírez, A.Meneses, A. Guajardo, "Impact of a computer-based intervention in Chilean children at risk of manifesting reading difficulties", Infancia y Aprendizaje, **40**(1), 158–188, 2017. doi:10.1080/02103702.2016.1263451

[46] A. Sucena, Eu Leio, Escola Virtual, 2015. Retrived from https://www.escolavirtual.pt/Pagina-Especial/euleio.htm

[47] A. Sucena, C. Marques, A.F. Silva, A. Ramalho, R. Machado, R. Santos, C. Santos, C. Garrido, H. Freitas, I. Santos, I. Rangel, M. J. Mata, "I Read: Reading and writing skills promotion software". In P. Arantes, & J. V. Sá (Eds.), ARTECH 2019 Proceedings of the: 9th International Conference on Digital and Interactive Arts. Braga: Portugal, 2019. ISBN 978-1-4503-7250-3

[48] A. Sucena, J.F. Carneiro, M.T. Restivo, "Serious games for reading acquisition: A tentative prototype", Teaching and Learning in a Digital World, **2**, 686-692, 2018. doi:10.1007/978-3-319-73204-6_75

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 298-305 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Evaluation of Information Competencies in the School Setting in Santiago de Chile

Jorge Joo-Nagata[*,1], Fernando Martínez-Abad[2]

[1]*Universidad Metropolitana de Ciencias de la Educación, Departamento de Historia y Geografía, Ñuñoa, 7760197, Chile*

[2]*Universidad de Salamanca, Instituto Universitario de Ciencias de la Educación, Salamanca, 37008, Spain*

ARTICLE INFO

Article history:
Received: 07 June, 2021
Accepted: 19 July, 2021
Online: 16 August, 2021

Keywords:
Competencies for life
Blended learning
Primary education

ABSTRACT

This study evaluated the competencies related to digital information use through technological tools aiming to acquire applicable knowledge by searching and retrieving information. Methodologically, a quasi-experimental design without a control group was applied to a sample of primary education students from Chile (n=266). First, a diagnosis of the digital-informational skills is performed, and, later, the results of a course in a blended learning context -b-learning- (treatment) are shown. The results show significant differences between the participant groups, confirming the learning in information competencies and distinguishing an initial level from a posterior intermediate level.

1. Introduction

In the present context of Information and Communications Technology (ICT), the development of basic skills, and contexts of data overstocking, they have gradually become the three greatest fields for process of teaching-learning in different schooling levels. Additionally, expectations and objectives for information use in a widely intertwined world are set. This way, key skills in the ICT context are search, use, evaluation, and information processing [1,2].

Thus, it is understood that the integration of new data and communication systems, mainly from the set of technologies such as the Internet, corresponds to a social reality that is challenging the global educational system [3,4] and, particularly, the educational system in Chile. The concern of the educational administration is well reflected in the successive efforts to inquire about and propose improvements in material coverage (facilities, hardware and software material coverage, connectivity at a national level), in training and innovation of teachers or the organization of teachers at a management or curriculum development level. However, the scientific knowledge on acquisition and development of digital and information competencies requires the application of systemic methodologies for collection, analysis, and validation of the information that allows drawing applicable conclusions, with a capacity of generalization to advance in this new curricular content.

The quality of education in Chile, in its different levels, is the reason why it is so stressed on considering including key competencies in the foundations of the mandatory educational curriculum and, a variable that considerably impacts in this improvement is the training level performed at a teacher level and the school education stage [5,6]. The evaluation and training are included in the implementation of ICT innovative actions at educational centers.

The objective of this article is to evaluate digital and information competencies along with the implementation and evaluation of an ICT innovative project leading to the training in searching and processing of information in students at the school educational system. Specifically, a diagnosis of the digital-informational competencies level in school students was performed to later propose an experimental design able to verify the effectiveness of a training program of blended learning for the development of digital-informational education, competencies in the permanent training of school students. Finally, a variation level in the evaluated dimensions is set - search, evaluation, processing, and information communication - at the group of analyzed students.

The rest of the article is organized as follows: firstly, this document provides information on digital-informational skills in a formal educational setting and the current situation of ICT. Then, the applied methodology, and the sample characteristics are detailed. Hereafter, the pre-test and post-test results are summarized. Lastly, discussion and conclusions are shown..

[*]Corresponding Author: Jorge Joo-Nagata, jorge.joo@umce.cl

2. Theoretical framework

2.1. Digital-informational skills and the school setting

The fact of using ICT as learning objects, vehicles of information literacy and tools of permanent training of students from the school system is going to provide an added value to the research field, showing themselves not only as methodological elements but also as findings and achievements. The ICT use in the development of school education is often presented as a positive impact on the development of digital skills that belong to the 21st century and skills in general [7,8]. Thus, digital literacy can be understood as the development of a wide range of skills derived from the use of applied technology, allowing students to research, create contents and communicate digitally, and therefore, to constantly participate in the development of a society with a strong digital component [7–9].

Despite the classic denomination of "digital natives" [10], where it is established that these digital skills are in current students by default, several studies have determined that their knowledge is in medium and low levels [11–15]. Thus, considering this information, it is necessary to develop these skills aiming to prevent differences in the use of digital technologies. In the particular case of Chile, several studies have shown that students are capable of solving tasks related to the use of the information as consumers, as well as organizing and managing digital information. On the other hand, very few students can have success in skills related to the use of information as producers and creators of contents [9,16,17].

Skills related to information management through digital tools –particularly free access resources found on the Internet– with the objective of acquiring valid and verified knowledge through the search and collection of information, along with the interpretation of text information from the reading of new data, reflection, and its evaluation, will be understood as information literacy, also known as ALFIN, by its Spanish initials [18–20].

Beyond the existing controversy on the inportance of terminology used in the very definition of concepts such as "Processing of Information" and "Digital Skill" [21–23], the fact is that "Informational Skill" can be placed within the scope of action of this key competency. At the center of the generic skills the information literacy is found, which has become a new paradigm at the ICT scene, which is understood as the cognitive–affective fabric that allows to people not only recognizing their informational needs but also acting, understanding, finding, evaluating, and using information of the most diverse nature and sources.

2.2. Digital learning in a blending learning setting

For the development of digital and information literacy, the adoption of online modality is increasingly massive, in particular when the courses are integrated into academic study plans, or in contexts where this type of education is required [24]. Besides, the experience of adopting online or b-learning modalities has been strengthened in extreme situations like confinement derived from COVID-19 disease, where teachers understand the context of online learning. However, during the implementation, a variety of problems have arised, like facilities availability, Internet access, the need for new forms of planning, implementation, and evaluation of learning, and collaboration with the parents [25–27]. Notwithstanding the above, digital technologies represent an educational resource for a better adaptation to different types of students and their different personal situations. Elements such as focusing on diversity and encouraging the development of key skills are taken into consideration, while they offer multiple possibilities for collaboration between teachers and partners in new communicative scenes of online character.

Thus, the challenge of providing an interactive learning experience for students in big classmate groups and the concern on the quality of teaching in this type of milieu have been key catalysts for the reconsideration of educational approaches in the different educational levels [28].

Blended learning (b-learning), also known as hybrid learning, constitutes an evolving field of research within the wider dimension of electronic learning (e-learning), and it related to instruction practices combining traditional presential approaches with online learning or mediated by technology with the Internet as the main tool used [24,29]. From the beginning of the 21st century, a significant number of studies in this field, with their diverse implications in educational fields have been developed. However, there is still limited evidence that b-learning improves the results of learning in students [24,30].

3. General context

The interest of organizations such as the European Union, the OECD, and UNESCO for this type of research is relevant [31–33], since the circle of evaluation-formation and educational innovation constitutes one of the strategic axes in the education field for the development, not only of key areas in all the productive fabric but also in critical thought in a globalized world. This way, the formation of citizens regarding skills focused on demands of work environment in the 21st century should constitute a fundamental concern for current governments.

Currently, a great impact in the study of information competencies in research has been reached [34–38]. However, efforts on the progress of evaluation tools of the proficiency level in digital and information skills are not wider enough. Most of the research carried out have established and implemented measure scales of self-perception of their own digital and information skill, of their elaboration, and in a setting of the general evaluation of digital skills [39–47]. It is important to point out that, while most of the research previously analyzed and cited keep a precise and concrete application, only a few of them present an integration of the evaluation and the use of digital and information competencies within the educational curriculum [47–49]. Although there is a specific increase of these digital skills in teachers from diverse areas, particularly teachers who teach one course, the development focused on the digital and information competencies in Chile has not been produced, and rather, they have been settled from the general implementation of digital skills [9,50–54].

Additionally, a significant part of the research carried out in the formal education field, from national as well as international contexts, and especially in primary school and university educational levels, establish training programs that develop themselves from specific curricular or disciplinary aspects

[39,42,55–60]. Nevertheless, due to the multidimensional structure that digital and information skills raise, the experiences within these training programs tend to aim at the specific dimension of technology, collection, and processing of data, rather than a complex, global and holistic view of them.

4. Objectives and hypothesis

Analyze the level of ICT competencies of the school education body to, secondly, present an experimental design able to check the efficacy of a teaching program for the increase of information competencies in the permanent formation of students from school education. Finally, the variation level of the measured dimensions –search, evaluation, processing, and communication of the information– is established in the group of students that were analyzed.

5. Materials and methods

5.1. Research design

This is an information analysis study based on investigation data in students of the last stage of primary school [61], where a quasi-experimental test type design without a control group was used [62,63]. In this study, an educational intervention with ICT and information competencies was performed to support innovation and integration models in the teaching and learning process..

5.2. Participants

The study sample is students of both sexes who belong to the last year of primary education in Chile, in the Province of Maipo area. The sample were stratified according to the school levels they belonged to, along with their source schools.

The sample was composed of 266 students and, according to the socio-demographic features, distribution can be recapitulated as:

- By variable sex of the sample was 50.4% of female students and 49,6% of male students.
- By age of students was established between 12 to 17 years, with an average of 13.89 years.
- The distribution based on education level was made according to sex: 63.53% of students belonged to the 8º grade of primary school, where 50.88% were women and 49.11% were men. 36.46% of students belonged to the 7º grade of primary school, where 50.51% were men and 49.48% were women.

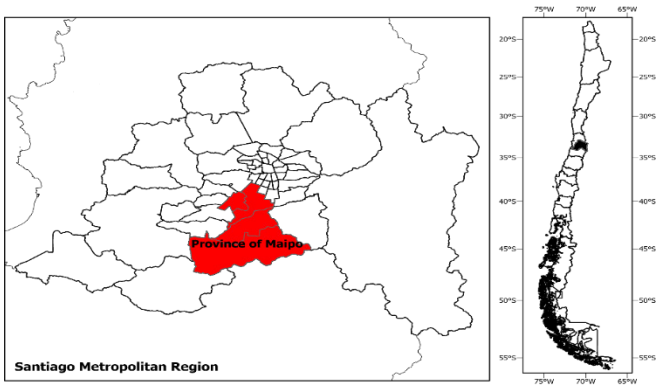


Figure 1: Study area: Province of Maipo

The sample derivates from the population of primary school students, whose ages range between 12 and 17 years old, belonging to the Province of Maipo in the Santiago Metropolitan Region (Figure 1), equivalent to N=106,745 [64], with a level of significance of $\alpha = 0.05$ and a maximum of homogeneity $p = q = 0.5$ for the sample of 266 participants, where the sampling error obtained is equal to 8.2%

5.3. Data collection instrument (test)

The dependent variable has been defined as the level of acquired competency of the students on digital and information skills, measured before (pre-test) and after (post-test) the implementation of the educational intervention (treatment).

The instrument was applied in the pre-test stages as a diagnosis of the competencies, and the post-test was applied for the evaluation of the information competencies reached by the participants students, measuring the level of performance obtained after the treatment. The test is composed of 29 elements of dichotomous nature, from 37 questions of single and multiple selection. 7 questions in the information search dimension, 10 questions in an evaluation dimension, 5 questions in the information processing dimension and other 9 questions on communication and dissemination of the information are presented (table 1).

Table 1: Dimensions and variables used in the instrument (test)

Dimension	Source	Metric	Description
Predictor variables			
Socio-demographic	- Test. - Data provided by the educational center or the Ministry of Education.	- Sex. - Age. - Grade - Municipality	Socio-demographic data of the participants of the study (sample).
Criterion variables			
Digital and information skills focused on the informational part.	Test	-Search of the information.	Grade obtained in the survey (instrument) about the item. Questions 1 to 17
		- Evaluation of the information	Grade obtained in the survey (instrument) about the item. 9 questions going from items 8 to 9.
		Analysis and selection of information (processing)	Data obtained through the survey (instrument) about item. 9 questions going from items 10 to 14.

		Communication of the information.	Grade obtained in the survey (instrument) about item. 9 questions going from items 15 to 18.

5.4. Treatment

The independent or treatment variable involved the teaching intervention applied to the participant during 30 training hours in a blended context [65–67], and adapted from the proposal described in [68]. This way, the teaching intervention applied in the training of sample corresponds to an adaptation of the instrument for the development of information competencies of secondary education [65,68] which has been tested and validated in that area.

5.5. Data analysis

Concerning the data analysis, after the exploratory initial analysis of the distributions of the variables and the evenness of the variances and covariances structures, parametric techniques, ANOVA with repeated measures are applied. Intra-subjects factors (pre-test/post-test) and inter-subjects factors (type of school) are incorporated. After the study with repeated measures, other techniques complementing the results are applied, such as the t-test. Additionally, all scores have been calculated so that each of the items is valued with a maximum score of 1 and a minimum score of 0. Besides, the scores of dimensions such as the average score of the set of items of the n dimension multiplied for 10 have been calculated. Thus, all dimensions are ranged from 0 to 10 points, which facilitates interpretation and contrast. Finally, the total score is calculated as the sum of the scores in the 4 dimensions. To ensure the internal consistency of the test results, the Cronbach's alpha test was performed, whose value was equal to 0.731, considered as adequate or acceptable [69,70]. All these statistical processes have been carried out in the SPSS 25 and R 4.0 software.

5.6. Procedure

The data in tests, in all cases, was collected mainly through online questionnaires. Responsible teachers were contacted, and the students participated by performing both pre-test and post-test online.

6. Results

6.1. General and descriptive results

Firstly, it is possible to determine values that approximate an average of 5 points (table 2) for intra-subjects characteristics (sex, grade and school). In parallel, there may be an important dispersion of values, which increases in post-test results.

It is observed that at the sample's level, the average reaches higher scores in Search, Processing, and Communication dimensions, while lower scores are registered in the Evaluation dimension (table 3). At a general level, the average score is greater in the student samples in the post-test. The dispersion of values increases in the post-test, so the teaching e-learning process has created higher levels of inequity between the measured students.

Table 2: total descriptive scores (n = 266)

		Pre-test		Post-test	
		Mean	**S.D**	**Mean**	**S.D.**
Sex	**Male**	4.809	1.263	5.207	1.413
	Female	4.694	1.213	5.331	1.449
Grade	**7º**	4.778	1.089	4.998	1.446
	8º	4.736	1.317	5.425	1.401
School	**School 1**	4.682	1.236	5.171	1.399
	School 2	5.051	1.207	5.697	1.494

Table 3: Descriptive statistics per dimension (n = 266)

		Mean	**S_x**	**P_{25}**	**P_{50}**	**P_{75}**
PRE-TEST	**SEARCH**	4.540	1.532	3.654	4.615	5.385
	EVALUATION	5.079	1.939	3.333	5.556	6.667
	PROCESSING	4.511	1.692	3.636	4.545	5.455
	COMUNICATION	5.000	2.044	4.000	5.000	6.000
	TOTAL	**4.752**	**1.237**	**3.953**	**4.651**	**5.581**
POST-TEST	**SEARCH**	5.552	1.824	3.846	5.385	6.923
	EVALUATION	5.046	1.985	3.333	5.000	6.667
	PROCESSING	4.863	1.876	3.636	4.545	6.364
	COMUNICATION	5.550	2.128	4.000	6.000	7.000
	TOTAL	**5.270**	**1.430**	**4.419**	**5.233**	**6.337**

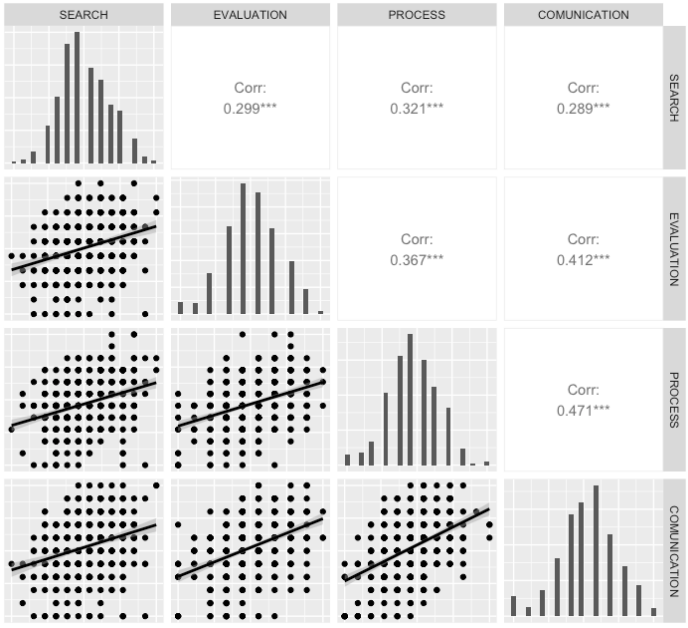


Figure 2: relation and distribution between dimensions

Moreover, it is observed how the distribution of pretest and post-test is similar in both groups, and their progress is similarly reached, while in the Evaluation dimension there is a decline or no

progress at all for both cases. It does not seem to be interaction according to the school; therefore, it is not considered a covariable. By obtaining correlations between dimensions, it is possible to establish that all values are significant and positive (figure 2).

By comparing between groups and the results obtained in the post-test, it is possible to perceive differences between the course and the school groups. However, based on the sex variable, values in the scores in different dimensions do not show great differences (figure 3).

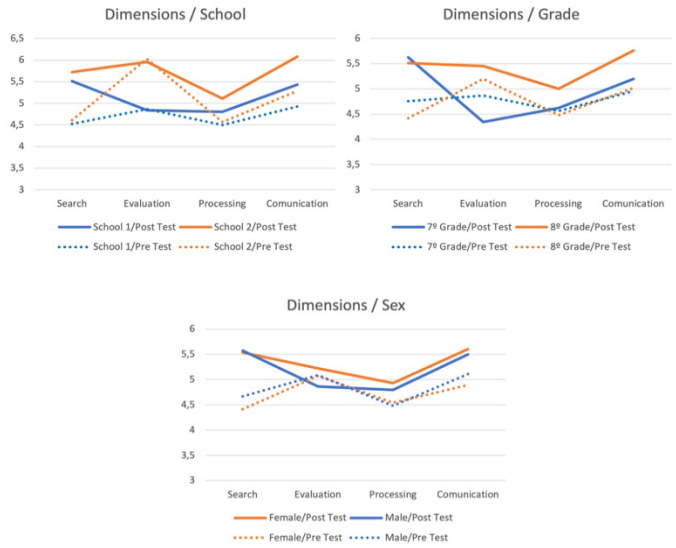


Figure 3: Post-test results based on dimensions and groups (n=266)

It is also possible to establish that in total scores, from all the variables analyzed in the contrasts, the sex variable is not significant in an inter-subject level nor an intra-subject interaction (table 4).

Table 4: Comparison per variables (n = 266)

		Pre-test		Post-test	
		Mean Dif.	t (p.)	Mean Dif.	t (p.)
Sex		-0.115	-0.758 (0.449)	0.123	-0.704 (0.482)
Grade		0.041	- 0.264 (0.792)	-0.427	-2.365 (0.019)
School		-0.368	-1.909 (0.057)	-0.526	-2.366 (0.019)
			Post-test – Pre-test		
			Mean Dif.		t (p.)
Sex	**Male**		-0.398		-3.997 (<.001)
	Female		-0.636		-5.614 (<.001)
Grade	**7º**		-0.220		-1.667 (0.099)
	8º		-0.689		-7.688 (<0.001)
School	**School 1**		-0.488		-5.631 (<0.001)
	School 2		-0.646		-4.359 (<0.001)
Complete sample			-0.518		-6.841 (<0.001)

6.2. ANOVA results of repeated measures

A hypothesis in the use of the ANOVA test with repeated measures is the matrix homogeneity of the covariances of the dependent variables [71,72]. This hypothesis is determined through the Box test, with the following results (table 5):

Table 5: Box test of the equality of covariances matrix

Box M.	**11.194**
F	**1.842**
df1	**6**
df2	**268376.644**
Sig.	**0.087**

Thus, the significance level showed that in this test a value of 0.087 was obtained, which exceeds 0.05 and, therefore, with a probability grade of 95% that the hypothesis on covariances matrix observed of the dependent variables are equal between groups.

Descriptive statistics and the results of the ANOVA test of repeated measures applied to data (table 6) showed that there is a significant interaction depending on the Grade, so it was decided to keep this variable, regardless that it is not significant at an individual level.

Table 6: Intra-subjects effects

INTRA-SUBJECT EFFECTS	F	p.
PRE-POST	**30.92**	**<.001**
PRE-POST * Grade	**8.518**	**0.004**
PRE-POST * School	**0.088**	**0.767**

The training action was significantly more effective in the eighth-grade group course than in the seventh-grade course. While the levels are similar in the pre-test scores - slightly higher in the eighth-grade group -, the training action had better results in the eighth-grade students. Nevertheless, the School variable does not have interaction in pre-post test scores (Figure 4).

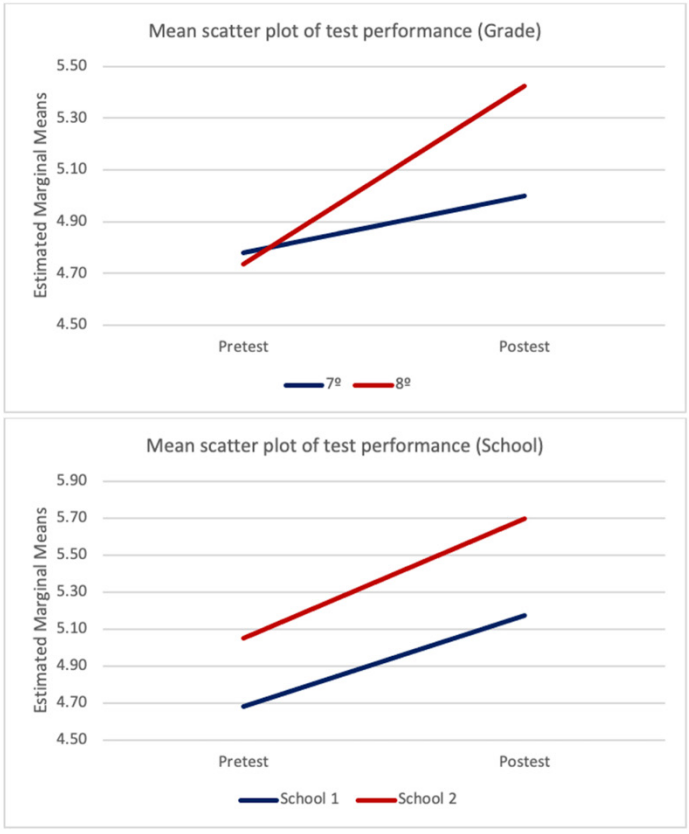


Figure 4: interactions between pre-test and post-test scores in grade and school variables

In table 7 it is possible to see the effects caused by the characteristics of the subjects as a transversal measure. Thus, it is possible to establish that the School variable is significant, but that is not the case for the Grade variable. This implies that at a general level, and calculating the average score of each subject in the pre-test and post-test, and also considering the criteria of the average score variable, there are significant differences for the School variable, but not for the Grade variable.

Table 7: Inter-subjects effect

INTER-SUBJECT EFFECTS	F	p.
Grade	0.186	.666
School	4.408	.037
Grade * School	-	-

Although all students in different schools and courses reached substantial progress in their digital competencies, there are also constant differences between the schools participating in the research.

7. Discussion

The students participating in the study, despite belonging to two different schools and two different teaching levels, in which there are different methodologies and contents within the elementary school, behave as a heterogeneous group in the information competencies area, with statistically significant differences between groups compared in the totals as much as in the analyzed dimensions (tables 1-7 and figures 1-2), which is complemented with the average age of the test responding students along with the sex variable, factors that do not have great incidence in the results due to the developed b-learning educational intervention.

Regarding the diverse analyzed variables, students behave in a heterogeneous manner, despite all the participants increasing their scores in the post-test measures. Within the established dimensions, the differences existing in the evaluation and the processing of the digital information areas are highlighted, a situation that has been stated in other investigations [16]. This way, particular uses in the application of technologies skills are presented. Regarding values on indicators of digital competencies and forms in which they show, participant students keep medium to low digital competencies indexes, which coincide with what has been carried out in others similar research [4,49,73].

8. Conclusions

Within the complexity of the Information Society, which is constantly mediated by the impact of ICT, it demands to the educational processes the incorporation of key skills, where the digital and information skills are prominent, related to the treatment of information in a virtual setting and the competency of digital processing. From this aspect, the training of the students from the first cycles of teaching gains importance. Thus, this study evaluated the efficacy of an educational intervention in a b-learning setting for the training in information competencies at an elementary educational level (seventh and eighth grades), obtaining progress considered significant but not enough for the ICT context where we live.

Empirical values obtained in the different dimensions confirm an appropriation in the digital competencies learning, where it is possible to differentiate an initial level from a posterior level after the applying of the educational intervention, and also between other variables as grade from school variables, establishing different realities in each educational context.

In the conclusions, the importance of the evaluation and training in digital and information skills is considered, addressing the fundamental dimensions that compose them, and establishing the efficacy of the implemented educational intervention. On the other hand, the research contributes to the study area according to the scale of reference - Santiago de Chile - and the training context - elementary students. Even though in Chile there are efforts to know the use of technologies in secondary teaching students [9], in the case of elementary students there are no studies that establish the characteristics and the levels concerning the information competencies in a digital setting.

Finally, after an analysis of the research contributions, weaknesses are established, which are focused on the design and development of the evaluation instrument and the experimental level of the applied design.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was supported by the Research Direction of the Universidad Metropolitana de Ciencias de la Educación, Chile [Proyecto FIPEA 02-18].

References

[1] M. Bielva Calvo, F. Martinez Abad, M.E. Herrera García, M.J. Rodríguez Conde, "Diseño de un instrumento de evaluación de competencias informacionales en Educación Secundaria Obligatoria a través de la selección de indicadores clave," Education in the Knowledge Society (EKS), **16**(3), 124–143, 2015, doi:10.14201/eks2015163124143.

[2] UNESCO, A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2, 2018.

[3] A. Aslan, C. Zhu, "Pre-service teachers' perceptions of ICT integration in teacher education in Turkey," Turkish Online Journal of Educational Technology, **14**, 97–110, 2015.

[4] V. Dagiene, Development of ICT competency in pre-service teacher education, IGI Global, Lazio: 65–75, 2013.

[5] E. Palma Gajardo, "Percepción y Valoración de la Calidad Educativa de Alumnos y Padres en 14 Centros Escolares de la Región Metropolitana de Santiago de Chile," REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, **6**(1), 85–13, 2008.

[6] M. Perticará, M. Román, Los desafíos de mejorar la calidad y la equidad de la educación básica y media en Chile, Konrad-Adenauer-Stiftung, Santiago, Chile: 95–122, 2014.

[7] J. Fraillon, J. Ainley, W. Schulz, T. Friedman, E. Gebhardt, Preparing for life in a digital age: The IEA international computer and information literacy study. International report, Springer International Publishing, London: 15–25, 2014, doi:10.1007/978-3-319-14222-7_1.

[8] R. Schmid, D. Petko, "Does the use of educational technology in personalized learning environments correlate with self-reported digital skills and beliefs of secondary-school students?," Computers & Education, **136**, 75–86, 2019, doi:10.1016/j.compedu.2019.03.006.

[9] M. Claro, D.D. Preiss, E. San Martín, I. Jara, J.E. Hinostroza, S. Valenzuela, F. Cortes, M. Nussbaum, "Assessment of 21st century ICT skills in Chile: Test design and results from high school level students," Computers & Education, **59**(3), 1042–1053, 2012, doi:10.1016/j.compedu.2012.04.004.

[10] M. Prensky, "Digital natives, digital inmigrants," On the Horizon, **9**(5), 1–6, 2001.

[11] C.J. Asarta, J.R. Schmidt, "Comparing student performance in blended and traditional courses: Does prior academic achievement matter?," The Internet and Higher Education, **32**, 29–38, 2017, doi:10.1016/j.iheduc.2016.08.002.

[12] F.D. Guillen-Gamez, M.J. Mayorga-Fernández, M.T.D. Moral, "Comparative research in the digital competence of the pre-service education teacher: face-to-face vs blended education and gender," Journal of E-Learning and Knowledge Society, **16**(3), 1–9, 2020, doi:10.20368/1971-8829/1135214.

[13] M.A. Harjoto, "Blended versus face-to-face: Evidence from a graduate corporate finance class," Journal of Education for Business, **92**(3), 129–137, 2017, doi:10.1080/08832323.2017.1299082.

[14] L. Mohebi, Leaders' Perception of ICT Integration in Private Schools: An Exploratory Study from Dubai (UAE), SSRN Scholarly Paper ID 3401811, Social Science Research Network, Rochester, NY, 2019, doi:10.2139/ssrn.3401811.

[15] P. Slechtova, "Attitudes of Undergraduate Students to the Use of ICT in Education," Procedia - Social and Behavioral Sciences, **171**, 1128–1134, 2015, doi:10.1016/j.sbspro.2015.01.218.

[16] T. Ayale-Pérez, J. Joo-Nagata, "The digital culture of students of pedagogy specialising in the humanities in Santiago de Chile," Computers & Education, **133**, 1–12, 2019, doi:10.1016/j.compedu.2019.01.002.

[17] J. Joo Nagata, P. Humanante Ramos, M.Á. Conde González, J.R. García-Bermejo, F. García Peñalvo, "Comparison of the Use of Personal Learning Environments (PLE) Between Students from Chile and Ecuador: An Approach," in Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, ACM, New York, NY, USA: 75–80, 2014, doi:10.1145/2669711.2669882.

[18] D. Bawden, "Revisión de los conceptos de alfabetización informacional y alfabetización digital," Anales de Documentación, **5**, 361–408, 2002.

[19] I. Szőköl, K. Horváth, "Introducing of New Teaching Methods in Teaching Informatics," in: Auer, M. E. and Tsiatsos, T., eds., in The Challenges of the Digital Transformation in Education, Springer International Publishing, Cham: 542–551, 2020, doi:10.1007/978-3-030-11932-4_52.

[20] S. Virkus, "Information literacy in Europe: a literature review," Information Research, **8**(4), 2003.

[21] J. Castaño, J.M. Duart, T. Sancho, "A second digital divide among university students," Cultura y Educación, **24**(3), 363–377, 2012, doi:10.1174/113564012802845695.

[22] J. Castaño-Muñoz, "Digital inequality among university students in developed countries and its relation to academic performance.," International Journal of Educational Technology in Higher Education, **7**(1), 43–52, 2010.

[23] L. Starkey, "A review of research exploring teacher preparation for the digital age," Cambridge Journal of Education, **50**(1), 37–56, 2020, doi:10.1080/0305764X.2019.1625867.

[24] C. McGuinness, C. Fulton, "Digital Literacy in Higher Education: A Case Study of Student Engagement with E-Tutorials Using Blended Learning," Journal of Information Technology Education: Innovations in Practice, **18**, 001–028, 2019.

[25] S. Dhawan, "Online Learning: A Panacea in the Time of COVID-19 Crisis," Journal of Educational Technology Systems, **49**(1), 5–22, 2020, doi:10.1177/0047239520934018.

[26] I. Fauzi, I.H.S. Khusuma, "Teachers' Elementary School in Online Learning of COVID-19 Pandemic Conditions," Jurnal Iqra' : Kajian Ilmu Pendidikan, **5**(1), 58–70, 2020, doi:10.25217/ji.v5i1.914.

[27] D. Moszkowicz, H. Duboc, C. Dubertret, D. Roux, F. Bretagnol, "Daily medical education for confined students during coronavirus disease 2019 pandemic: A simple videoconference solution," Clinical Anatomy, **33**(6), 927–928, 2020, doi:10.1002/ca.23601.

[28] D.J. Hornsby, R. Osman, "Massification in higher education: large classes and student learning," Higher Education, **67**(6), 711–719, 2014, doi:10.1007/s10734-014-9733-1.

[29] D.R. Garrison, H. Kanuka, "Blended learning: Uncovering its transformative potential in higher education," Internet and Higher Education, **7**(2), 95–105, 2004, doi:10.1016/j.iheduc.2004.02.001.

[30] G. Siemens, D. Gašević, S. Dawson, Preparing for the digital university: A review of the history and current state of distance, blended, and online learning., Athabasca University, Arlington: Link Research Lab.: 234, 2015.

[31] Eurydice, La profesión docente en Europa. Prácticas, percepciones y políticas, Oficina de Publicaciones de la Unión Europea, Luxemburgo, 2015.

[32] OCDE, La definición y selección de las competencias clave. Resumen ejecutivo., DeSeCo, OCDE: 20, 2005.

[33] T. Valencia-Molina, A. Serna-Collazos, S. Ochoa-Angrino, A.M. Caicedo-Tamayo, J.A. Montes-González, J.D. Chávez-Vescance, Competencias y estándares TIC desde la dimensión pedagógica: una perspectiva desde los niveles de apropiación de las TIC en la práctica educativa docente., Cali, 2016.

[34] M. Bouckaert, "Designing a materials development course for EFL student teachers: principles and pitfalls," Innovation in Language Learning and Teaching, **10**(2), 90–105, 2016, doi:10.1080/17501229.2015.1090994.

[35] C. Flores-Lueg, R. Roig-Vila, "Percepción de estudiantes de Pedagogía sobre el desarrollo de su competencia digital a lo largo de su proceso formativo," Estudios pedagógicos, **42**(3), 7, 2016.

[36] J. Garrido M, D. Contreras G, C. Miranda J, "Analysis of the pedagogical available of preservices teacher to use ICT," Estudios Pedagógicos (Valdivia), **39**(ESPECIAL), 59–74, 2013, doi:10.4067/S0718-07052013000300005.

[37] A. Martínez, J.G. Cegarra Navarro, J.A. Rubio Sánchez, "Aprendizaje basado en competencias: Una propuesta para la autoevaluación del docente," Profesorado. Revista de Currículum y Formación de Profesorado, **16**(2), 325–338, 2012.

[38] P. Martínez Clares, B. Echeverría Samanes, "Formación basada en competencias," Revista de Investigación Educativa, **27**(1), 125–147, 2009.

[39] A. Blasco Olivares, G. Durban Roca, "La competencia informacional en la enseñanza obligatoria a partir de la articulación de un modelo específico," Revista española de Documentación Científica, **35**(Monográfico), 100–135, 2012, doi:10.3989/redc.2012.mono.979.

[40] L. González Niño, G.P. Marciales Vivas, H.A. Castañeda Peña, J.W. Barbosa Chacón, J.C. Barbosa Herrera, "Competencia informacional: desarrollo de un instrumento para su observación," Lenguaje, **41**(1), 105–131, 2013.

[41] S.U. Kim, D. Shumaker, "Student, Librarian, and Instructor Perceptions of Information Literacy Instruction and Skills in a First Year Experience Program: A Case Study," The Journal of Academic Librarianship, **4**(41), 449–456, 2015, doi:10.1016/j.acalib.2015.04.005.

[42] E. Kuiper, M. Volman, J. Terwel, "Developing Web literacy in collaborative inquiry activities," Computers & Education, **52**, 668–680, 2009, doi:10.1016/j.compedu.2008.11.010.

[43] F. Martinez-Abad, P. Torrijos-Fincias, M.J. Rodríguez-Conde, "The eAssessment of Key Competences and their Relationship with Academic Performance:," Journal of Information Technology Research, **9**(4), 16–27, 2016, doi:10.4018/JITR.2016100102.

[44] E. Resnis, K. Gibson, A. Hartsell-Gundy, M. Misco, "Information literacy assessment: a case study at Miami University," New Library World, **111**(7/8), 287–301, 2010, doi:10.1108/03074801011059920.

[45] H. Saito, K. Miwa, "Construction of a learning environment supporting learners' reflection: A case of information seeking on the Web," Computers & Education, **49**(2), 214–229, 2007, doi:10.1016/j.compedu.2005.07.001.

[46] S. Santharooban, "Analyzing the level of information literacy skills of medical undergraduate of Eastern University, Sri Lanka," Journal of the University Librarians Association of Sri Lanka, **19**(2), 2016.

[47] S. Santharooban, P.G. Premadasa, "Development of an information literacy model for problem based learning," Annals of Library and Information Studies (ALIS), **62**(3), 138–144, 2015.

[48] S.C. Kong, "A curriculum framework for implementing information technology in school education to foster information literacy," Computers & Education, **51**, 129–141, 2008, doi:10.1016/j.compedu.2007.04.005.

[49] A. Pérez Escoda, M.J. Rodríguez Conde, "Evaluación de las competencias digitales autopercibidas del profesorado de Educación Primaria en Castilla y León (España)," Revista de Investigación Educativa, **34**(2), 399–415, 2016, doi:10.6018/rie.34.2.215121.

[50] M. Claro, A. Salinas, T. Cabello-Hutt, E. San Martín, D.D. Preiss, S. Valenzuela, I. Jara, "Teaching in a Digital Environment (TIDE): Defining and measuring teachers' capacity to develop students' digital information and communication skills," Computers & Education, **121**, 162–174, 2018, doi:10.1016/j.compedu.2018.03.001.

[51] Enlaces, Competencias TIC en la Profesión Docente, Ministerio de Educación, Chile, Santiago de Chile, 2007.

[52] Enlaces, Estándares en Tecnología de la Información y la Comunicación para la Formación Inicial Docente, Ministerio de Educación, Chile, Santiago de Chile, 2007.

[53] Enlaces, Matriz de habilidades TIC para el Aprendizaje, 2013.

[54] I. Jara, M. Claro, J.E. Hinostroza, E. San Martín, P. Rodríguez, T. Cabello, A. Ibieta, C. Labbé, "Understanding factors related to Chilean students' digital skills: A mixed methods analysis," Computers & Education, **88**, 387–398, 2015, doi:10.1016/j.compedu.2015.07.016.

[55] I. Aguaded, I. Marín-Gutiérrez, E. Díaz-Pareja, "La alfabetización mediática entre estudiantes de primaria y secundaria en Andalucía (España)," RIED. Revista Iberoamericana de Educación a Distancia, **18**(2), 275–298, 2015, doi:10.5944/ried.18.2.13407.

[56] M. Fuentes Agustí, C. Monereo Font, "Cómo buscan información en Internet los adolescentes," Investigación en la escuela, (64), 45–58, 2008.
[57] M.J. Grant, A.J. Brettle, "Developing and evaluating an interactive information skills tutorial," Health Information & Libraries Journal, **23**(2), 79–88, 2006, doi:10.1111/j.1471-1842.2006.00655.x.
[58] N. Landry, J. Basque, "L'éducation aux médias : contributions, pratiques et perspectives de recherche en sciences de la communication," Communiquer. Revue de communication sociale et publique, (15), 47–63, 2015, doi:10.4000/communiquer.1664.
[59] G.P. Marciales Vivas, H.A. Castañeda-Peña, J.W. Barbosa-Chacón, I. Barreto, L. Melo, "Fenomenografía de las competencias informacionales: perfiles y transiciones," Revista Latinoamericana de Psicología, 58–68, 2016, doi:10.1016/j.rlp.2015.09.007.
[60] M.L. Tiscareño Arroyo, J. de J. Cortés-Vera, "Competencias informacionales de estudiantes universitarios:: una responsabilidad compartida. Una revisión de la literatura en países latinoamericanos de habla hispana.," Revista Interamericana de Bibliotecología, **37**(2), 117–126, 2014.
[61] J.H. McMillan, S. Schumacher, Research in Education: Evidence-Based Inquiry, 7th Edition, Pearson, New Jersey, 2010.
[62] C. Anwar, A. Saregar, Y. Yuberti, N. Zellia, W. Widayanti, R. Diani, I.S. Wekke, "Effect size test of learning model arias and PBL: Concept mastery of temperature and heat on senior high school students," Eurasia Journal of Mathematics, Science and Technology Education, **15**(3), 2019, doi:10.29333/ejmste/103032.
[63] R. Sagala, R. Umam, A. Thahir, A. Saregar, I. Wardani, "The effectiveness of stem-based on gender differences: The impact of physics concept understanding," European Journal of Educational Research, **8**(3), 753–761, 2019, doi:10.12973/eu-jer.8.3.753.
[64] Asociación Chilena de Municipalidades, Universidad San Sebastián, COMUNAS Y EDUCACIÓN: Una aproximación a la oferta educativa comunal, Dirección de Estudios AMUCH Asociación de Municipalidades de Chile, Santiago (Chile): 34, 2017.
[65] M. Bielva Calvo, F.M. Martínez Abad, M.J. Rodríguez Conde, "Validación psicométrica de un instrumento de evaluación de competencias informacionales en la educación secundaria," Bordón. Revista de Pedagogía, **69**(1), 27–43, 2016, doi:10.13042/Bordon.2016.48593.
[66] J. Cabero Almenara, "Formación del profesorado universitario en TIC. Aplicación del método Delphi para la selección de los contenidos formativos," Educación XX1, **17**(1), 111–132, 2013, doi:10.5944/educxx1.17.1.10707.
[67] J. Cabero Almenara, M. Llorente Cejudo, V. Marín Díaz, "Hacia el diseño de un instrumento de diagnóstico de 'competencias tecnológicas del profesorado' universitario," 2010.
[68] F. Martínez Abad, S. Olmos Migueláñez, M.J. Rodríguez Conde, María J., "Evaluación de un programa de formación en competencias informacionales para el futuro profesorado de E.S.O.," Revista de Educacion, **370**, 45–70, 2015.
[69] D. George, P. Mallery, SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update, Allyn and Bacon, 2003.
[70] J.C. Nunnally, Psychometric theory, 2nd ed., McGraw-Hill, New York, 1978.
[71] M.E. Legget, M. Toh, A. Meintjes, S. Fitzsimons, G. Gamble, R.N. Doughty, "Digital devices for teaching cardiac auscultation - a randomized pilot study," Medical Education Online, **23**(1), 1524688, 2018, doi:10.1080/10872981.2018.1524688.
[72] A. Tahriri, M.D. Tous, S. MovahedFar, "The Impact of Digital Storytelling on EFL Learners' Oracy Skills and Motivation," International Journal of Applied Linguistics and English Literature, **4**(3), 144–153, 2015, doi:10.7575/aiac.ijalel.v.4n.3p.144.
[73] J. García-Martín, J.-N. García-Sánchez, "Pre-service teachers' perceptions of the competence dimensions of digital literacy and of psychological and educational measures," Computers & Education, **107**(Supplement C), 54–67, 2017, doi:10.1016/j.compedu.2016.12.010.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 306-319 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Segmentation of Stocks: Dynamic Dimensioning and Space Allocation, using an Algorithm based on Consumption Policy, Case Study

Anas Laassiri*, Abdelfettah Sedqui

Innovative Technologies Laboratory, Abdelmalek Essaadi University, ENSAT, Tangier, 90000, Morocco

ARTICLE INFO

Article history:
Received: 07 June, 2021
Accepted: 08 August, 2021
Online: 16 August, 2021

Keywords:
Stock management
Warehousing
Dynamic dimensioning
Storage location assignment
Bin packing

ABSTRACT

This paper addresses the stock management aspect. Through this work, we provide a dynamic model of dimensioning and allocation of stocks to storage location for the automotive industry field. This model takes into consideration all constraints of the supply chain (24 constraints) from the suppliers passing by production, storage up to customer delivery and transport. At the end of this paper we will be able to specify the stock replenishment policy, particularly the definition of stock alerts (minimum, nominal, maximum) in quantity and days of stock, in space occupied, and in financial value. These stock thresholds will be integrated in material resource planning, storage allocation procedure and financial budget follow-up. The tool developed is decision-making support for logisticians. The algorithm proposed has solved a real instance and ensure a balanced stock fill rate (99%) in 1200 seconds.

1. Introduction

In every supply chain, the total costs that affect the logistic cost center are very high compared to all production costs. This has a direct impact on companies' margin, and therefore, on the competitiveness of the products in the market knowing that a super high selling price cannot be accepted by all customers. In this context, companies are trying to set a demarche to control its logistics costs. The general framework of this work concerns a key element for the success of this demarche; we are talking about the immobilized cash (stocks) as well as the charges related to the surface occupied by these stocks. Inventories are the basic and main logistical data for making decisions, which can affect the operational, tactical, and sometimes strategic level. Moreover, defining a replenishment policy is fixing a minimum, average, and maximum thresholds, this implies the necessary surface area and the allocation of products to the available storage surface. The main goal is to guarantee cost control, align with budget and at the same time simplify and streamline the internal physical flow. Therefore, inventory management becomes a compulsory.

In the literature, stock management arouses the interest of researchers and manufacturers who have defined procurement policies for all contexts and types of products. In this extended paper of research, originally presented in the 5th international conference on logistics operations management (GOL'20) [1], we will define, by considering the constraints of logistics management, a dynamic stock model which will be used to define stock thresholds and serves as a decision support tool and facilitate the allocation of products in the warehouse. We will define the minimum and maximum stock alarms as well as the corresponding storage area in order to better manage both space and the cash of the company and to follow the perspectives announced in the conference paper presented in the 5th international conference on logistics operations management (GOL'20), regarding the storage location assignment problem which can be modeled as a bin packing problem based on segmentation phase results.

The rest of this article will be organized as follows: we will mention different stock policies in the state of the art. Then, we will detail the constraints of logistic management which formulate our problem of stock policy definition in addition to the assignment of stocks demarche, develop this model on Excel VBA. Finally, a real case study from an industrial background will be studied (Moroccan automotive company).

2. State of art

The stock management in a literature point of view depicts a multitude of replenishment policies. The most addressed are the reorder point methods (Q, R) (figure 1, part 1 taken from [2]), this demarche assumes that within a lead-time of a product, the quantity ordered should be equal to the total quantity consumed. The second method was introduced by [3] and controls stocks through up to level ordering intervals (T, S) (figure 1, part 2 taken from [2]), the Third method defines the economic order quantity

*Corresponding Author: Anas LAASSIRI, laassiri.anas@gmail.com

https://dx.doi.org/10.25046/aj060434

(EOQ) (figure 1, part 3) as part of Wilson model for inventory management [4] considering the ordering and holding cost. Supposing that the ordering cost and delivery time is constant, moreover, the minimum stock alert is considered null, and the maximum stock that can be reached is equal to the EOQ.

There are other policies, based on the ECR: efficient consumer response. This model proposes a collaborative policy in an efficient way between the manufacturers and distributors to boost the gains and achieve high level of productivity from the supplier up to the consumer among of them we can address:

- Continuous replenishment planning (CRP), it is a base that supports the efficient consumer response, which refers to a program that triggers the production and the flow of a product through the supply chain as soon as an identical product have been purchased by the end consumer. As shown by [5] brings varying benefits in terms of inventory cost savings.
- SSM (GPA in French) which means the shared supply management and includes two types: vendor managed inventory (VMI) with or without consignment, the co-managed inventory (CMI).
- The consignment stock (CS) which is an innovative approach to manage inventories in which the vendor removes his inventory and maintains a stock of materials at the buyer's plant. The reference [6] treats the consignment stock policy considering economical and logistics constraints and define the minimum and maximum stock levels of a vendor managed inventory.
- We find also the deterministic and stochastic demand assumption, and stock out assumptions if we lose the customer or there is a stock out cost to pay or assume.

To the best of our knowledge, none of the models addressed in the literature, have treated the constraints of inventory management that we will approach all together.

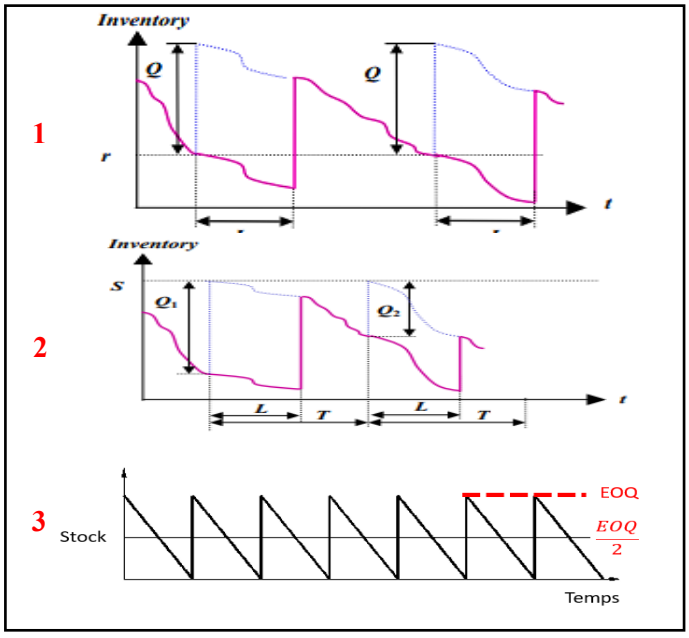


Figure 1: Part 1 (Q, R) policy, part 2 (T, S) policy and part 3 (EOQ) policy

3. Problematic

In an automotive industry environment, between the most uncertain goals to attain we can cite inventory objective level; a supply manager is facing difficulties to determine the policy of procurement in such an uncertain environment. Defining the minimum and maximum replenishment level to ensure production and customer needs, and meet the financial targets set by the top management. This target is calculated in function of the turnover and working days and without integrating the real logistical constraints. To avoid the problem of shortage, the stock managers are obliged to put more and more days in stock, for most cases in a non-efficient way. The model we propose will stands for a decision making support tool that assists stock managers, to determine fixed and variable stock parameters: in order to calculate the minimum and maximum alerts of stock in function of the EOQ. These parameters will be integrated in the material resource planning (MRP) which involves the calculation of requirements, but in an efficient and dynamic way, considering also the main variables impacting the EOQ : type of stocks, the daily average consumption and other vital parameters such as: transit time, minimum ordered quantity. Then assigning the defined stock targets to the storage location areas.

3.1. Objectives

By means of this research, we intent to:

- Define the minimum, nominal, maximum stock alerts, in value and in quantity and in days of stock taking into account the various parameters of stock.
- Define the correct storage area for products in the warehouse in function of the stock alerts defined: dynamic storage.
- Assign products to storage location by maximizing the net surface used while minimizing the distance between the same families of product (multi-purpose).
- Check the correspondence of the defined stock with the available surface.
- Guarantee the Fluidity of goods through the different process of the supply chain.

4. Segmentation of stock: dynamic model

4.1. Constraints of management

In a supply chain, the flow from the supplier up to the customer involves many parameters of management specific to each stage or part of this chain; these parameters directly affect the policy of stock management and make it complicated to define.

If we take the procurement phase as an example, the ordered quantity by the purchaser must abide by the supplier terms defined in the supplier schedule agreement, particularly the minimum batch supply. The ordered quantity should be greater to the minimum batch defined by the supplier to maintain the price defined in the contract; otherwise, a new price will be applied because of the new batch order which will be greater than the defined one. In the other hand, if the gap between the two quantities is really huge compared to the weekly average consumption, the purchaser have to assume the gap in stock even if it will not be consumed immediately, moreover, he has to reserve a space for this gap while waiting for its consumption.

This kind of constraints can have less impact in some cases, for example if the daily average consumption is equal or is a multiple of the supplier batch size, in this case we can order only the required quantity and respect the targeted days of coverage set by the logistic manager. To deal with these constraints and to define the stock management policy, we must start by listing and classifying those constraints all over the supply chain, then taking them into consideration while dimensioning the stocks.

The identified constraints in a supply chain while managing the procurement or the sale of products are defined below and was collected from the industrial sector (several meetings were conducted with logistic managers, engineering pilots, purchasing teams were done in order to regroup all the constraints):

For the raw materials we define:

- Packaging: define the parts in one package, it can be expressed in Kilograms, Meters or Pieces; this data could be found in the logistic contract of each product.
- Batch supply: minimum quantity ordered from a supplier with a normal price.
- Transport Frequency: procurement delivery frequency and represents the number of inbounds in a period (week, two weeks, one month) divided by the period expressed in days.
- Scrap Rate: rejection due to the process (production) and all rejected parts by quality controllers
- DPM: defective parts per million due to the supplier quality of delivered pieces.
- Supplier service rate in full: quantity of parts received compared to the ordered quantity in the firm horizon.
- Supplier service rate on time: quantity of supplies on time divided by total quantity supplied in a specific period.
- Production Reliability rate: the volume declared in production divided by total volume scheduled by logistics.
- Daily average consumption: average consumption of a product due to production and related to the explosion of bill of materials.
- Supplier Firm lead time: number of working days separating the day following the order and the day of the delivery.
- Supplies Frequency: the procurement program update frequency
- Supplier Transit Time: duration in working days between the supplier's plant and the customer point of delivery.

Price: unit price of the product in monetary value.

- Stock for Specific issues: safety stock resulting of a managerial Decision or imposed by the supply policy.
- Packaging details: dimensions authorized stacking level.

For the finished goods we define:

- Packaging: define the parts in one package.
- Minimum production batch size: minimum quantity to launch in production lines.
- Delivery Frequency: customer delivery frequency, it represents the number of trucks loaded to deliver in a period divided by the period expressed in days.
- Rejection Rate and Scrap rate: scrap rate at the end of production and rejected parts following a quality control.
- Production Reliability Rate: the ratio between the volumes produced and the planned volumes.
- Service rate on time in full: the ratio between deliveries on time and the total expected deliveries on a specific period.
- Customer DPM: Defective parts per million delivered to the customer.
- Customer variability: maximum tolerated variability of delivery instructions (firm orders compared to forecast)
- Daily average deliveries. : the mean daily deliveries.
- Customer firm lead time: number of working days separating the day following the release of orders, and the expected day of the delivery.

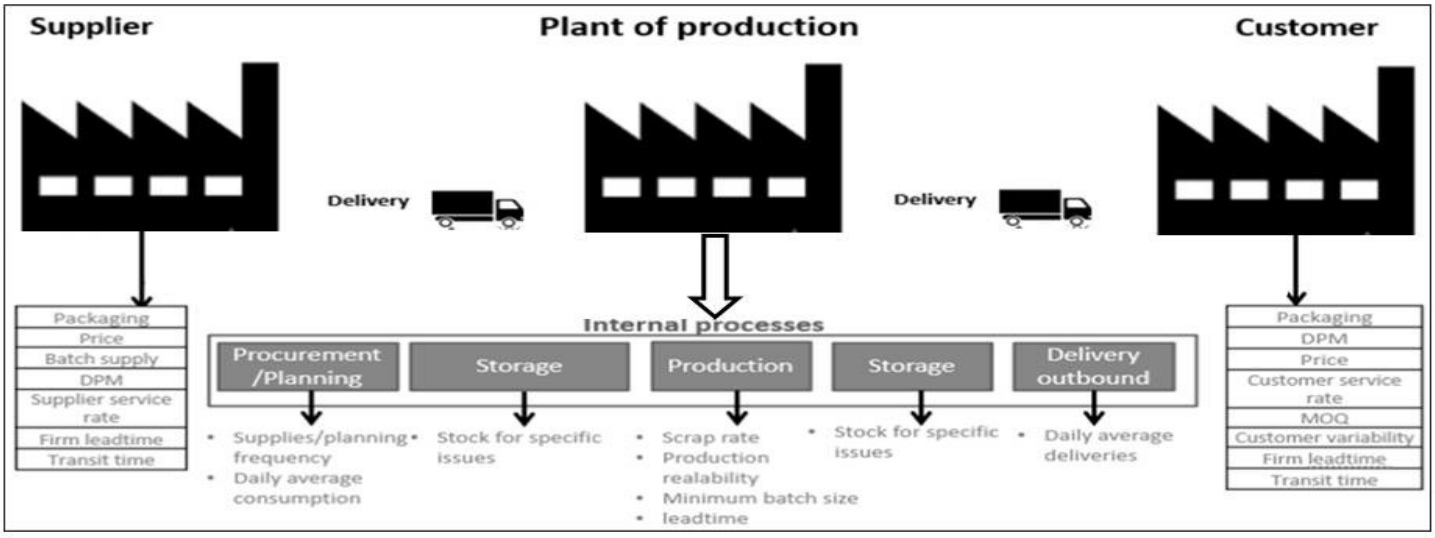


Figure 2: Constraints of management of products in a supply chain

- Planning frequency: the master Production Schedule update frequency.
- Production lead time: time between the Raw Materials out of stock and Finished Goods produced. Must be expressed in working days.
- Transit time: the time between the warehouse and the agreed delivery place.
- Stock for specific issues: safety stock due to management or customer decision.
- Price: unit price in monetary value.

4.2. Proposed model of stock

The operating stocks in the model proposed are defined as follows:

- *Static stock*

For raw materials / finished goods:

- The stock covering the non-respect of the production schedule.
- The stock covering the non-quality, non-compliance, rejects.
- The defined stock for specific issues: decision of top management committee or customer.

- *Dynamic stock*

For raw materials:

- The stock due to deliveries, in function of supply delivery frequency.
- The stock covering the gap between the order batch and the delivery frequency in case they are not synchronized.
- The stock due to variability covering the non-respect of the production schedule.

For finished goods:

- The stock due to deliveries, which depends on the customer delivery frequency.
- The stock covering the gap between the production schedule and the delivery frequency in case they are not synchronized.
- The stock covering the variability of customer orders.

- *Stock alerts*

- Minimum stock: stands for the stock related to total service rate and quality level, in addition to the stock for specific issues. This stock includes internal constraints related to both production and quality processes. It will integrated in the master production schedule MPS and supply program SP as the minimum alert to launch either the replenishment of raw materials, or production orders to maintain the level of stock for finished goods.
- Nominal stock: stock covering all the constraints mentioned before and ensures the availability of material for production, the total stock hovers around this nominal stock.
- Maximum stock: represents the nominal stock added to stock due to the supplier batch size (for raw materials) and production batch (for finished goods) in addition to the stock due to variability. This maximum stock will be integrated also in the MPS/SP program and stands for the upper limit of total stock. This maximum stock could be reached and tolerated by the top management in these two cases :
- While receiving a new inbound batch (new reception).
- Before an outbound delivery (stock assigned to the temporary preparation area/ Bin waiting to be loaded in customer truck).

- *Dynamic surfaces:*

Dynamic surfaces have a direct link with the minimum, nominal and maximum stock, and are generated to follow the variability of these stocks:

- Minimum surface: stands for the minimum space that would accommodate the minimum stock defined expressed in handling units (number of pallets for example), integrating the unit area of one handling unit and the possible stacking level as well as the type of storage (flat or shelving).
- Nominal surface: stands for the nominal space that would accommodate the nominal level of stock defined expressed in handling units, integrating the unit area of one handling unit and the possible stacking level as well as the type of storage (flat or shelving).
- Maximum surface: stands for the maximum area that would accommodate the maximum stock defined expressed in handling units, integrating the unit surface of the handling unit and the number of stacking levels as well as the type of storage space (flat, which means massive or shelving using metallic structures).

- *Storage location allocation:*

The problem of allocating materials to the storage location area can be modeled as a task assignment problem as schematized below:

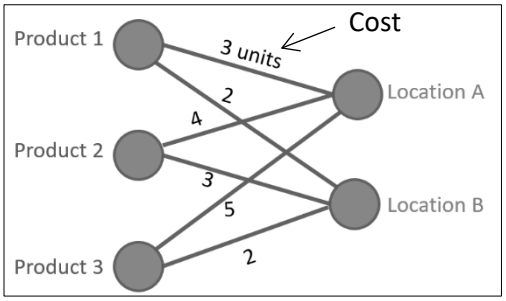


Figure 3: modeling of task assignment problem

As shown in the figure above, we will assign products (represented with handling units) to the storage location, which is divided to sub storage location named bins. Our problem can also be considered as a bin packing problem. It is about finding the most economical storage possible for a set of items into bins.

In our case, it is about finding the closest optimal assignment for a set of products that are assigned to storage location. Respecting lean flow directives (raw materials and finished goods should be stored separately to avoid flow congestion). The storage locations will be represented by boxes (Bins) and the products will be represented by items. To solve this NP hard problem we will use tree dimensional Bin packing algorithm, and to improve results we will pair the bin packing algorithm to an algorithm of list (first fit decreasing problem [7].

The table below summarizes the analogies between our assignment problem and the core model of bin packing problem modified from [8] considering the 3 dimensions and the constraints related to our context :

Table 1: Analogies between bin packing and our problem

Criteria	**Bin packing problem**	**Stock allocation problem**
Data	Article Bin Volume of the article	products Storage location Dimensions of the product
Objective function	Assigning items to bins	Assign stocks to volume
Constraints	The volume capacity of the bin	The volume of the storage location
Objective	Minimize the number of bins used	Minimize the volume used
Other		Compatibility constraint between raw materials and final products, they must be separated to better flow organization, and between bins and products in function of storage type

4.3. Parameters and Equations

- Notation

Table 2: Stock parameters

Raw Materials	
Pck_i	Parts per packaging of the raw material (RM) i
B_i	Batch Supply of the RM i
$Sfreq_i$	Supply Frequency of the RM i
S_i	Scrap Rate of the RM i
DPM_i	DPM of the RM i
Sf_i	Supplier service rate in full of the RM i
St_i	Supplier service rate on time of the RM i
PR_i	Production Reliability rate of the RM i
DAC_i	Daily average consumption of the RM i
F_i	Supplier Firm lead time of the RM i
$Freq_i$	Supplies Frequency of the RM i
TT_i	Supplier Transit Time of the RM i
SSI_i	Stock for Specific issues of the RM i
p_i	Price of the RM i
L_i	Length of the RM i
l_i	Width of the RM i
G_i	Stacking level of the RM i
Finished Goods	
Pck_j	Packaging of the finished good (FG) j
B_j	Minimum production batch size of the FG j
$Dfreq_j$	Delivery Frequency of the FG j
S_j	Rejection Rate and Scrap rate of the FG j
PR_j	Production Reliability Rate of the FG j
Sr_j	Service rate on time in full of the FG j
Var_j	Customer variability of the FG j
DAD_j	Daily average deliveries of the FG j
C_j	Customer firm lead time of the FG j
$Plan_j$	Planning frequency of the FG j
PLT_j	Production lead time of the FG j
DPM_j	DPM of the FG j
SSI_j	Stock for specific issues of the FG j
TT_j	Customer Transit time of the FG j
p_j	Price of the FG j
L_j	Length of the FG j
l_j	Width of the FG j
G_j	Stacking level of the FG j
DAC	Daily average consumption
	Stock alerts in quantity
$Minstock_k$	Minimum stock of the part number k
$Nomstock_k$	Nominal stock of the part number k
$Maxstock_k$	Maximum stock the part number k
	Dynamic surface
l_k	Length of the part number k
w_k	Width of the part number k
$Hustock_k$	Number of packages in part in a handling unit for the part number k
$Stacking_k$	Authorized stacking level for the part number k
$Minar_k$	Minimum surface for the part number k
$Momar_k$	Nominal surface for the part number k
$Maxar_k$	Maximum surface for the part number k
	Stock alerts value
$MinstockV_k$	Minimum stock of the part number k
$NomstockV_k$	Nominal stock of the part number k
$MaxstockV_k$	Maximum stock the part number k
$Cost_k$	Unit cost of the part number k

The graphic bellow synthetizes the static and dynamic stocks taken into consideration in our model:

- ✓ Dynamic stock: represents the level of stock made available to the planner for smoothing.
- ✓ Static stock: represents the incompressible level of stock.

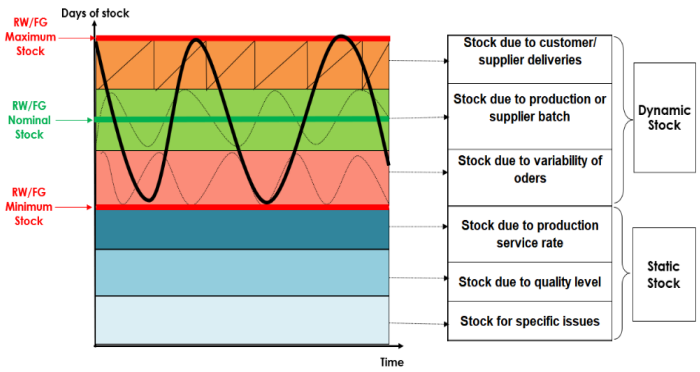


Figure 4: Static and Dynamic Stocks

- *Static stock*

RM

$$(S_i + DPM_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i \quad (1)$$

$$1 - Sf_i \times St_i \times \sqrt{\frac{F_i}{Freq_i}} \quad (2)$$

$$SSI_i \quad (3)$$

FG

$$(S_j + DPM_j) \times \sqrt{\frac{C_j}{Plan_j}} \times Plan_j \quad (4)$$

$$1 - PR_j \times Sr_j \times \sqrt{\frac{C_j}{Plan_j}} \quad (5)$$

$$SSI_j \quad (6)$$

- *Dynamic stock:*

RM

$$\frac{Freq_i}{2} + TT_i \quad (7)$$

$$(1 - PR_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i \quad (8)$$

$$\frac{B_i}{DAC_i \times 2} + TT_i - \left(\frac{Freq_i}{2} + TT_i\right) - (S_i + \frac{DPM_i}{1\,000\,000}) \times \sqrt{\frac{F_i}{Freq_i}} \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i - \left(1 - Sf_i \times St_i \times \sqrt{\frac{F_i}{Freq_i}}\right) - (1 - PR_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i \quad (9)$$

FG

$$\frac{DFreq_j}{2} + TT_j \quad (10)$$

$$Var_j \times \sqrt{\frac{C_j}{Plan_j}} \times Plan_j \quad (11)$$

$$\frac{B_j}{DAC_j \times 2} + TT_j - \left(\frac{DFreq_j}{2} + TT_j\right) - (S_j + DPM_j) \times \sqrt{\frac{C_j}{Plan_j}} \times Plan_j - (1 - PR_j \times Sr_j \times \sqrt{\frac{C_j}{Plan_j}}) - Var_j \times \sqrt{\frac{C_j}{Plan_j}} \times Plan_j \quad (12)$$

- ✓ (1): Stock due to quality level.
- ✓ (2): Stock due to production reliability.
- ✓ (3): Stock for specific issues.
- ✓ (4): Stock due to quality level.
- ✓ (5): Stock due to production reliability.
- ✓ (6): Stock for specific issues.
- ✓ (7): Stock due to supply frequency.
- ✓ (8): Stock due to production variability.
- ✓ (9): Stock due to supplier batch.
- ✓ (10): Stock due to delivery frequency.
- ✓ (11): Stock due to customer variability.
- ✓ (12): Stock due to production batch.

- *Stock alerts in quantity:*

For RM:

$$Minstock_k = (S_i + DPM_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i + 1 - Sf_i \times St_i \times \sqrt{\frac{F_i}{Freq_i}} + SSI_i \quad (13)$$

$$Nomstock_k = Minstock_k + \frac{Freq_i}{2} + TT_i + (1 - PR_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i + \frac{B_i}{DAC_i \times 2} + TT_i - \left(\frac{Freq_i}{2} + TT_i\right) - (S_i + \frac{DPM_i}{1\,000\,000}) \times \sqrt{\frac{F_i}{Freq_i}} \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i - (1 - Sf_i \times St_i \times \sqrt{\frac{F_i}{Freq_i}}) - (1 - PR_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i \quad (14)$$

$$Maxstock_k = SSI_i + (S_i + DPM_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i + 1 - Sf_i \times St_i \times \sqrt{\frac{F_i}{Freq_i}} + (1 - PR_i) \times \sqrt{\frac{F_i}{Freq_i}} \times Freq_i + TT_i + Max\left(Freq_i, \frac{B_i}{DAC_i}\right) \quad (15)$$

- *Stock alerts in value:*

$$MinstockV_k = Minstock_k \times Cost_k \quad (16)$$

$$NomstockV_k = Nomstock_k \times Cost_k \quad (17)$$

$$MaxstockV_k = Maxstock_k \times Cost_k \quad (18)$$

- *Dynamic surface:*

$$Minar_k = Minstock_k \times (Hustock_k \div Stacking_k) \times l_k \times w_k \quad (19)$$

$$\mathrm{Nomar_k} = \mathrm{Nomstock_k} \times (Hustock_k \div \mathrm{Stacking_k}) \times \mathrm{l_k} \times \mathrm{w_k}) \quad (20)$$

$$\mathrm{Maxar_k} = \mathrm{Maxstock_k} \times (Hustock_k \div \mathrm{Stacking_k}) \times \mathrm{l_k} \times \mathrm{w_k}) \quad (21)$$

- *Allocation problem: bin packing model*

The bin packing problem is NP hard in strong sense, this has been demonstrated using the reduction with 2-Partition, to solve medium to large instances we need to adopt heuristics developed particularly to solve this problem. The most popular heuristics in literature are first fit, Next fit and best fit, which have proven its effectiveness. The most adapted heuristic to our context is First fit, since the articles are assigned in a given order (in our case the order is in function of the daily average consumption). Then each article is placed in the first bin that can contain it.
A new box is considered only if this article does not fit in any box (considering the remaining empty volume in each box).The box remains open until the end of the execution of the algorithm, in case we engage the constraints of compatibility between articles, if an article is not compatible with a bin or an article or with another article already assigned in the current box, it will be assigned to the next box that ensure compatibility first then ensure the volume constraint contrary to the next fit heuristic where box is closed permanently if the volume of the current article cannot be contained in the box. The box is permanently closed, and it is impossible to assign a new article even if the remaining empty volume is enough to contain a new article. A new box is considered and becomes the current box. The reference [9] have proven the result of First fit:

$$First\ Fit\ (O) \leq Optimal(O) + 2\ \forall O \quad (22)$$

Research in the article [10] have demonstrated that the sequential coupling of list algorithms and bin packing heuristics guarantee better results, the first fit decreasing is the one that provides the best results, at a factor 2 of the optimal as proved by the reference [2]. Our assignment problem involves the allocation of stocks (total handling units) to the storage location (divided into sub-storage locations) with a flow constraint, that the raw materials and final products should be stored separately to remain an organized physical flow. It serves as a simulation tool to check if the policy of replenishment proposed sticks with the space that we have in the warehouse considering that the layout and the storage location are already defined.

Next, we will adopt the first fit decreasing heuristic to resolve the storage location assigning problem based on the daily average consumption of the handling unit, which refers to the daily consumption of the product marked in the HU label, sorted in the descending order (the HU of high runner products will be assigned first). We will use the following formalization in addition to some modifications related to our context:

- i: The index that represents the handling unit i.
- j: The index that represents the storage location bin j.
- S: The set of storage location bins in the warehouse.
- HU: The set of handling units.
- V_i: The volume occupied by the handling unit i (3D).
- V_j : The volume of the bin j.
- Dac_i: The daily average consumption of the handling unit i.

The decision variables are described below:

$$x_{i,j} = \begin{cases} 1, if\ the\ Hu\ i \in HU\ is\ assigned\ to\ the\ storage\ bin\ j \in S \\ 0, else \end{cases}$$

$$C_{i,j} = \begin{cases} 1, if\ the\ HU\ i \in HU\ is\ compatible\ with\ the\ Bin\ j \in S \\ 0, else \end{cases}$$

The algorithm bellow describes the first fit decreasing algorithm adopted which is the sequential coupling of the list algorithm based on the consumption of productions and the first fit algorithm:

Algorithm 1: The principle of the first fit decreasing algorithm

Result: $A_j \forall j \in S$

Input data:
- HU, $i \in \{1, \dots, \mathrm{card(HU)}\}$, V_i , Dac_i $\forall \mathrm{i} \in \mathrm{HU}$
- S , $j \in \{1, \dots, \mathrm{card(S)}\}$, V_j , $\forall j \in S$
- $C_{i,\mathrm{j}}$, $\forall$ j $\in$ S

Initialization:
- A_j =0, the initial volume of the bin j $\forall$j$\in$ S, all bins is empty at the beginning.
- $Chart\ A_{(3,card(HU))}$, the chart with card (HU) columns and 3 lines that contains respectively: Hu_i, V_i, Dac_i
- $Chart\ B_{(3,Card(HU))}$, the chart with 3 columns that contains: Hu_i, V_i, Dac_i after the step of sorting.

Sort $Dac_i, \forall i \in HU$ in a descending order, then place them with the corresponding HU and volume in $Chart\ B$

for k = 1 ... $card(HU)$ **do**
  j : = 1, Assigned: = false
  while *(assigned= false)* & $j \leq$ card(S)) **do**
    if the HU $Chart\ B_{1,k}$ is compatible with the bin j ($C_{chart\ B_{(1,k)},j} = 1$) **then**
      if $Chart\ B_{(2,k)}$ holds in j, ($A_j + Chart\ B_{(2,k)} < V_j$) **then**
        Assign $Chart\ B_{(1,k)}$ to the bin j:
        $x_{Chart\ B_{(1,k)},j} = 1$
        $A_j = A_j + Tab\ B_{(2,k)}$
        Assigned: = True
        j : = j + 1
      end
    end
  end
end

4.4. Resolution method: Segmentation and assignment of stocks

The demarche of segmentation starts from the integration of input data, until the calculation of the stock alerts: in quantity, in number of packages (HU), in financial value and in square meters.

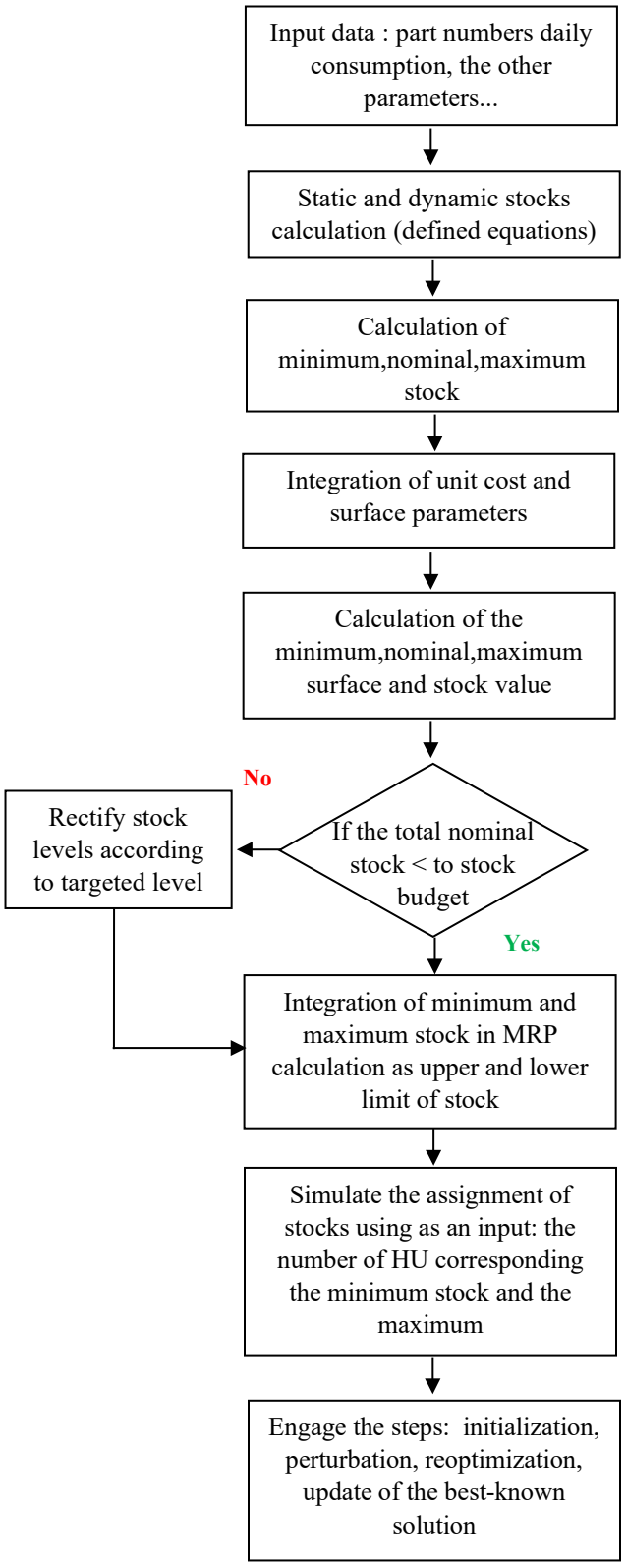


Figure 5: Segmentation and assignement demarche

The output of the segmentation demarche will be the input data for the storage location assignment demarche, the demarche starts with the initialization (input data), then engaging the first fit decreasing algorithm for the construction step (best known solution) based on the daily average consumption. Next, we will use a large neighborhood search as described below:

- Step 1 (perturbation): we randomly remove handling units from bins, and randomly empty a few bins completely. Then we sort the bins in a decreasing order in function of the current volume occupied in each one A_j.
- Step 2 (optimization): we select the removed HU, and reassign them into bin using a constructive heuristic, in our case we have chosen the wall building heuristic because of the flow constraints, we should place similar HUs next to each other as if we build a brick wall.
- Step 3 (update of the best-known solution): if the new solution obtained is better than the best known solution, we replace the current best known solution, if not, we go back to the first step and engage a new cycle of perturbation, reoptimization.

  The demarche of stock segmentation and storage location assignment is described below:

The action plan mentioned in the segmentation demarche tends to optimize six variables:

- Stock delivering (adapting the frequencies of delivery to minimize the number of days of finished good stock).
- Batch of production or supply (to synchronize the daily average consumption or daily average delivery).
- Variability of the demand (that must be contractual).
- Service Rate (of both customer and supplier).
- Quality level (of both supplier and customer delivered parts) an finally the safety stock.

For the case study, we will use the script code developed by Güneş Erdoğan (CLP Spreadsheet Solver) with some modification related to our context.

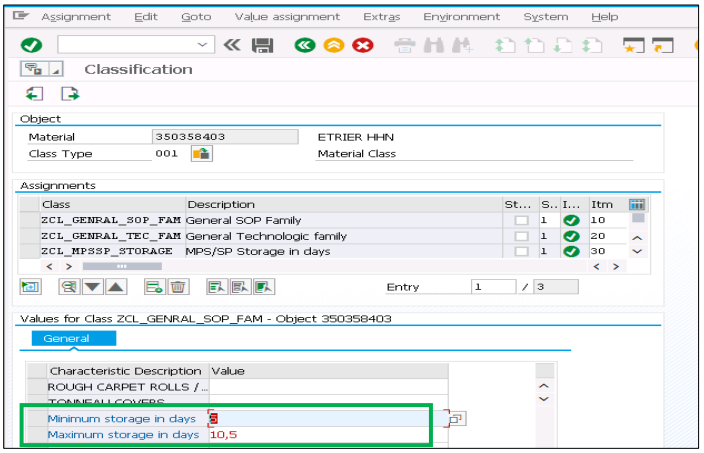


Figure 6: Master data transaction (classification tab)

							Av/Re	W13						W14	W15
Part number	Description	Minimum Batch size	Alarm Min.	Alarm Max.	Type of info.	Stock (SKU)	<=28/03	29/03	30/03	31/03	01/04	02/04	03/04	05/04 - 11/04	12/04 - 18/04
350658746	PART A	400	0	11,5	1.Stock Quantity	801,3	801	801	801	801	467	467	467	533	599
350658746	PART A	400			1.Stock Cover		14,4	13,4	12,4	11,4	10,4	9,4	8,4	9,94	11,31
350658746	PART A	400			1.Stock Value	76 235	76 235	76 235	76 235	76 235	44 462	44 462	44 462	50 745	57 028
350658746	PART A	400			2.Requirement						334			334	334
350658746	PART A	400			7.Purchase Init. Qty									400	400
350658746	PART A	400			7.Purchase /PU									1	1
350658746	PART A	400			7.Purchase New Qty									400	400

							Av/Re	W13					W14		
Part number	Part description	Minimum lot size	Alarm Min.	Alarm Max.	Type of info.	Stock (SKU)	<=30/03	31/03	01/04	02/04	03/04	05/04	06/04	07/04
351264852	Part B	120	2	11	1.Stock Quantity	1320	1 560	1 800	1 320	1 440	1 560	1 800	2 040	2 280
351264852	Part B	120			1.Stock Cover		10,6	10,8	11,6	11,2	10,8	11	10,86	10,71
351264852	Part B	120			1.Stock Value	51 249	60 567	69 885	51 249	55 908	60 567	69 885	79 202	88 520
351264852	Part B	120			2.Requirement				840					
351264852	Part B	120			3.Prod. init. Qty		240	240	360	120	120	240	240	240
351264852	Part B	120			3.Prod. /PU		2	2	3	1	1	2	2	2
351264852	Part B	120			3.Prod. New Qty		240	240	360	120	120	240	240	240

Figure 7: SP (1) and MPS (2)

4.5. Integration of alerts in MRP calculation

After the calculation of minimum and maximum alerts of stock, we integrate theme in SAP following the steps below:

- ✓ Access to the transaction of master data.
- ✓ Choose the classification tab.
- ✓ Add the alerts of stock in the corresponding field.

This data in the classification tab will be used to run the Material Resource Planning in automatic mode (MRP job) for the frozen period or Manuel smoothing of orders for the forecast horizon.

Our alerts were integrated in the supply VBA program (SP) and Master Production Schedule VBA program (MPS) already developed by the company using EXCEL VBA. The datasheet of SP includes the following data (Figure 7):

- The stock quantity and cover and value extracted from SAP.
- The requirement of the part in question into the frozen and forecast horizon (named "Part A" in the figure 7).
- The cover of the part in question in the end of the period (day, week).
- The purchase quantity in function of the batch size or its multiples.
- In each period (day, week) the quantity proposed by the system ensures that the coverage after acquisition of this quantity is between the minimum and maximum alert.
- If the cover exceeds the maximum alert of stock, the system doesn't propose any purchase quantity and jump to the next period.

The datasheet of MPS includes the following data (figure 7):

- All the data mentioned in SP.
- It works the same as SP, but instead of purchase quantity, we find the production quantity and instead of supplier batch size we find the production minimum size.

4.6. Limits of the proposed method

- ✓ If there is an error in the input data, consequently the results will not be consistent and right.
- ✓ If there is any discrepancy between the physical and the stock injected in the information system, the quantities proposed by MRP (either planned orders or purchase order) won't ensure the coverage of stock shown in the information system.
- ✓ The alerts of stock should be updated in a regular basis (every three months for example) knowing that the volumes could change in horizon forecast and the daily average consumption must follow the new volumes.
- ✓ The model of assignment considers zero distance between two HUs, and the volume and number of HU are already known before engaging the solver (off-line).

5. Case study: Moroccan industrial compan

4.7. Segmentation of stocks

The segmentation model developed was tested in a real case study conducted using data of a Moroccan multinational company specialized in automotive parts.

We have chosen 5 references of each family to show the details of each parameters and how we can interpret the results obtained using equations and logistic parameters defined before.

After integrating the ten references in the spreadsheet developed, we obtain the result below:

After realizing the simulation with the current data we conclude the following point by family of productions

For raw materials:

- To see the impact of stock due to transit time, stock to cover the frequency of delivery, and stock to covering quality of service let's take an example of the part number A : the nominal stock is equivalent to 8.2 days this stock include also the transit time equivalent to four days.

TOTAL MINIMUM STOCK	TOTAL NOMINAL STOCK	TOTAL MAXIMAL STOCK
in days	in days	in days
1,6	5,9	7,8

TOTAL MINIMUM STOCK	TOTAL NOMINAL STOCK	TOTAL MAXIMAL STOCK
in packaging	in packaging	in packaging
82	343	465

REFERENCE	DESIGNATION	Customer	Packaging	Production Batch	Transport Frequency	Quality Level		Demand Variability			Daily consumption	Firm lead time	Planning frequency	Production lead time+Transit time	Stock for specific issues	Stock due to delivering	Stock due to production batch	Stock due to quality level	Stock due to Service Rate	Stock due to customer variability	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK
			in parts	in parts	in days	scrap rate	Customer PPM	Production Reliability %	Service rate on time in full	Customer variability	in parts	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in qty	in qty	in qty	in packaging	in packaging	in packaging
350103233	Ref G	xxxxx	120	240	2,50	3,0%	80	60,00%	97,00%	10%	2200	6	5	1,5		2,8	0,0	0,2	2,3	0,5	2,5	5,8	7,0	5 399	12 654	15 404	45,0	105,5	128,4
350107523	Ref H	xxxxx	100	100	5,00	3,0%	120	98,00%	97,00%	20%	2000	6	5	1,5	1,0	4,0	0,0	0,2	0,3	1,1	1,4	6,5	9,0	2 871	13 062	18 062	28,7	130,6	180,6
300438218	Ref I	xxxxx	180	180	5,00	3,0%	100	98,00%	97,00%	20%	1150	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	501	6 361	9 236	2,8	35,3	51,3
300453929	Ref J	xxxxx	100	100	5,00	3,0%	100	98,00%	97,00%	20%	820	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	357	4 535	6 585	3,6	45,4	65,9
313008732	Ref K	xxxxx	100	200	5,00	3,0%	100	98,00%	97,00%	20%	480	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	209	2 655	3 855	2,1	26,5	38,5

MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK
in value	in value	in value
191 339 MAD	967 442 MAD	1 256 308 MAD

REFERENCE	Unit cost	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	PACKAGING Length (L)	PACKAGING Width (w)	PACKAGING High (H)	Stacking	Storage Type (Racks or Flat Storage)	Stock height	Surface per Packaging	MINIMUM SURFACE	NOMINAL SURFACE	MAXIMAL SURFACE	MINIMUM Rack Locations	NOMINAL Rack Locations	MAXIMAL Rack Locations
	in value	in value	in value	in value	in (mm)	in (mm)	in (mm)	in packaging		in m	in m²	in m²	in m²	in m²			
332233112	82	19 603	79 991	116 763	1200	1000	1100	5,0	Racks	6	1				1	2	3
332233122	128	38 296	127 653	200 990	1200	800	1000	5,0	FLAT	5	1	1	1	2			
332233132	58	93 239	190 144	219 273	1200	1000	1200	5,0	FLAT	6	1	5	8	10			
332233142	59	26 525	409 688	542 232	1200	1000	1400	5,0	FLAT	7	1	1	7	10			
332233152	59	13 675	159 966	177 050	1200	1000	1400	5,0	FLAT	7	1	1	4	4			

Figure 8: The datasheet of raw materials

TOTAL MINIMUM STOCK	TOTAL NOMINAL STOCK	TOTAL MAXIMAL STOCK
in days	in days	in days
1,6	5,9	7,8

TOTAL MINIMUM STOCK	TOTAL NOMINAL STOCK	TOTAL MAXIMAL STOCK
in packaging	in packaging	in packaging
82	343	465

REFERENCE	DESIGNATION	Customer	Packaging	Production Batch	Transport Frequency	Quality Level		Demand Variability			Daily consumption	Firm lead time	Planning frequency	Production lead time+Transit time	Stock for specific issues	Stock due to delivering	Stock due to production batch	Stock due to quality level	Stock due to Service Rate	Stock due to customer variability	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK
			in parts	in parts	in days	scrap rate	Customer PPM	Production Reliability %	Service rate on time in full	Customer variability	in parts	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in days	in qty	in qty	in qty	in packaging	in packaging	in packaging
350103233	Ref A	xxxxx	120	240	2,50	3,0%	80	60,00%	97,00%	10%	2200	6	5	1,5		2,8	0,0	0,2	2,3	0,5	2,5	5,8	7,0	5 399	12 654	15 404	45	105	128,4
350107523	Ref B	xxxxx	100	100	5,00	3,0%	120	98,00%	97,00%	20%	2000	6	5	1,5	1,0	4,0	0,0	0,2	0,3	1,1	1,4	6,5	9,0	2 871	13 062	18 062	29	131	180,6
300438218	Ref C	xxxxx	180	180	5,00	3,0%	100	98,00%	97,00%	20%	1150	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	501	6 361	9 236	3	35	51,3
300453929	Ref D	xxxxx	100	100	5,00	3,0%	100	98,00%	97,00%	20%	820	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	357	4 535	6 585	4	45	65,9
313008732	Ref E	xxxxx	100	200	5,00	3,0%	100	98,00%	97,00%	20%	480	6	5	1,5		4,0	0,0	0,2	0,3	1,1	0,4	5,5	8,0	209	2 655	3 855	2	27	38,5

TOTAL MINIMUM STOCK	TOTAL NOMINAL STOCK	TOTAL MAXIMAL STOCK
in value	in value	in value
509 441 MAD	1 795 461 MAD	2 373 334 MAD

REFERENCE	Unit cost	MINIMUM STOCK	NOMINAL STOCK	MAXIMAL STOCK	PACKAGING Length (L)	PACKAGING Width (w)	PACKAGING High (H)	Stacking	Storage Type (Racks or Flat Storage)	Stock height	Surface per Packaging	MINIMUM SURFACE	NOMINAL SURFACE	MAXIMAL SURFACE	MINIMUM Rack Locations	NOMINAL Rack Locations	MAXIMAL Rack Locations
	in value	in value	in value	in value	in (mm)	in (mm)	in (mm)	in packaging		in m	in m²	in m²	in m²	in m²			
350103233	65,6	354 097,6	829 893,6	1 010 243,9	1200	800	1200	6,0	FLAT	7,2	1,0	7,7	17,3	21,1			
350107523	43,0	123 594,1	562 288,2	777 526,5	1200	800	1200	6,0	Racks	7,2	1,0				29	131	181
300438218	37,1	18 598,6	236 237,0	343 017,8	1200	800	1200	6,0	Racks	7,2	1,0				3	36	52
300453929	36,8	13 151,0	167 041,8	242 546,0	1200	800	1200	6,0	Racks	7,2	1,0				4	46	66
313008732	54,5	11 393,9	144 723,9	210 140,1	1200	800	1200	6,0	Racks	7,2	1,0				3	27	39

Figure 9: The datasheet of final products

- This duration could be reduced if we consider local suppliers of customers, that the distance between the place of order and the place of delivery is closer. This decision could have an impact on the order cost. Then, we have 5.3 days in stock because of the frequency of delivery to customer, if there is a possibility to raise the frequency from 2 trucks to 3 trucks per week, this quantity of stock days will be largely optimized. Consequently, the overall value of stock will be optimized reducing the impact on immobilized cash. Moreover, if we have a contract with suppliers that stipulate a global service rate greater than 90%, the stock related to service rate will be improved (currently 1,7 days). In the other hand, a minimum order quantity synchronized with daily consumption will solve many problems of returns of stocks non-consumed to the warehouse after production. This constraint was taken into consideration while developing the program under Excel VBA, so if we have a non-synchronization between these two values; this means that the resulting rate is greater than one week (generally 5 working days); the program underlines in red the minimum supplier batch. In the current simulation data, the resultant rate doesn't exceed one week or production. Regarding the space of storage required if the storage type required is: RACK (shelving structures), the spreadsheet gives the number of modules equivalent to nominal stock for the part number A is about 2 modules. For the flat type of stock. If the storage type required is: FLAT, the spreadsheet specifies the net surface needed in square meters (for the part number B and C, the surface needed is respectively 4 and 8 square meters). For the

total space needed (includes the stock net space and services aisles space), we apply a mark-up of 2.4.

- Another example of optimization we found, for the reference B. The stock due to the batch supply is 1.3 days, knowing that we consume 60 parts per day, and the batch supply is 800 parts. So, if we reduce the supplier batch, we could gain 1.3 days in stock of this reference.
- For the reference D, if we increase the supply frequency from 1 truck per month to 1 truck per week, we could gain 7.5 days in stock and the stock due to supply frequency will be 20.5 days. Furthermore, if we find a forwarder who can ensure the delivery in less than 18 days, for example 14 days, the stock due to supply frequency will be 16.5 stock days.

For finished goods:

- The important stock for the references H,I,J,K is the stock due to the frequency of customer deliveries, that implies a stock to cover the time between two deliveries, we have 4 days for each reference, if we increase the delivery frequency for example from one truck per week to 2 trucks per week, we could gain 2 days in stock
- The stock due to production service is huge for the reference G, if we oblige the production team to respect the production plan communicated by logistics, we could improve the reliability rate and gain 2.3 days in stock caused by the current production reliability rate.

Considering all this remarks, the action plan to reduce stocks should start by eliminating the important stock mentioned above. The spreadsheet also gives the corresponding space that should be revered for each reference (minimum, nominal and maximum); all these informations will serve as an input data for the assignment phase.

4.8. Assignment of stocks in storage locations

For the assignment phase, we define the following data structure according to the bin packing notation:

$$
Article \begin{cases} - HU(handling\ unit) \begin{cases} - Volume \begin{cases} Length \\ Width \\ Height \end{cases} \\ - Quantity \\ - Type \begin{cases} RW: raw\ materials \\ FG: finished\ goods \end{cases} \end{cases} \\ - DAC(daily\ average\ consumption) \\ - Storage\ type \begin{cases} Flat \\ Rack\ or\ shelving(RCK) \end{cases} \end{cases}
$$

$$
Bin \begin{cases} - Bin\ type \begin{cases} - Flat \begin{cases} - 1 \\ - 2 \\ - .. \end{cases} \\ - RCK \begin{cases} - 1 \\ - 2 \\ - .. \end{cases} \end{cases} \\ - Volume\ capacity \begin{cases} - Lenth \\ - Width \\ - Height \end{cases} \\ - Number\ of\ bins \end{cases}
$$

Figure 10: Simulation data structure

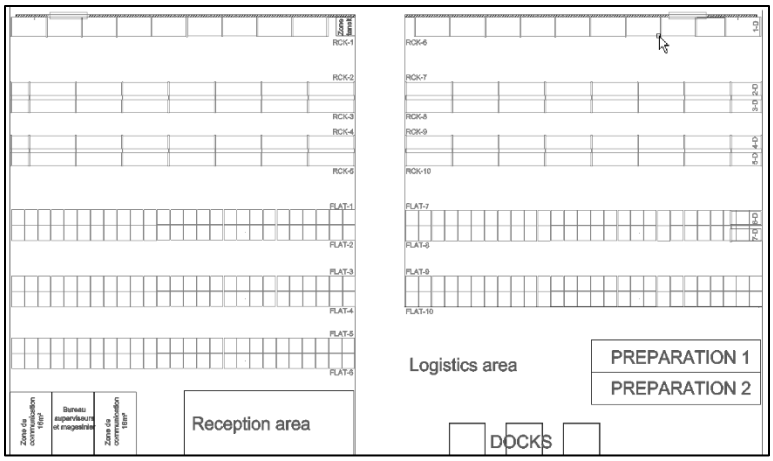


Figure 11: The layout of the warehouse

We consider the following instances:

- Considering the warehouse of the company (figure 11).
- The set of storage bin (Table 3).
- 10 reference of RW and FG (table 4 and 5) as an example. The two family of products are not compatible (can't be stored in the same bin). Compatibility between bins and references (storage type needed).

Table 3 : The set of Bins

Bin Type	Number of BINs	Width in meter (x)	Height in meter (y)	Length in meter(z)
FLAT	10	28	6	1,2
RCK	10	28	6	1,2

Table 4: The set of RW references

RW	Storage type	DAC	Number of HU	Width (meter) (x)	Height (meter) (y)	Length (meter) (z)
RW-1-A	FLAT	120	2	1	1,1	1,2
RW-2-B	RCK	60	5	0,8	1	1,2
RW-3-C	RCK	400	33	1	1,2	1,2
RW-4-D	FLAT	225	28	1	1,4	1,2
RW-5-E	FLAT	116	11	1	1,4	1,2
RW-6-F	RCK	1200	80	1	1,1	1,2
RW-7-G	FLAT	700	56	1	1,1	1,2

RW-8-H	FLAT	865	40	1	1,1	1,2
RW-9-I	FLAT	923	21	1	1,1	1,2
RW-10-J	RCK	1032	84	1	1,1	1,2

Table 5 : The set of FG references

RW	**Storage type**	**DAC**	**Number of HU**	**Width (meter) (x)**	**Height (meter) (y)**	**Length (meter) (z)**
FG-1-A	FLAT	2200	105	0,8	1,2	1,2
FG-2-B	RCK	2000	131	0,8	1,2	1,2
FG-3-C	RCK	1150	35	0,8	1,2	1,2
FG-4-D	RCK	820	45	0,8	1,2	1,2
FG-5-E	RCK	480	27	0,8	1,2	1,2
FG-6-F	FLAT	640	50	0,8	1,2	1,2
FG-7-G	FLAT	500	42	0,8	1,2	1,2
FG-8-H	FLAT	140	64	0,8	1,2	1,2
FG-9-I	FLAT	980	99	0,8	1,2	1,2
FG-10-J	FLAT	655	83	0,8	1,2	1,2

Table 6: Simulation results

Bin	**Volume occupied**	**References Assigned**	**Number of HU assigned**
Flat 1	100%	FG-9-I FG-10-J	175
Flat 2	100%	FG-6-F FG-1-A FG-10-J FG-7-G	175
Flat 3	99.94%	RW-1-A RW-4-D RW-7-G RW-8-H RW-9-I	145
Flat 4	54%	FG-6-F FG-8-H	94
Flat 5	10.48%	RW-5-E RW-9-I	13
Flat 6	0%	-	-
Flat 7	0%	-	-
Flat 8	0%	-	-
Flat 9	0%	-	-
Flat 10	0%	-	-
Rack 1	99%	FG-5-E FG-2-B FG-3-C	173
Rack 2	91%	RW-2-B RW-6-F RW-10-J	141
Rack 3	42%	RW-10-J RW-3-C	61
Rack 4	38%	FG-5-E FG-4-D	67
Rack 5 to Rack 10	0%	-	-

We take as an example the progress of assignment of BIN 1 Flat to show how the algorithm used works:

- Phase 1, 2: we place HU of the reference FG-9-I next to each other, knowing that the storage type required is “Flat”.
- Phase 3: the algorithm moves to the reference FG-10-J which is compatible with the FG-9-I and also requires a flat storage type.
- Phase 4, 5, and 6: the algorithm place similar items next to each other in order to form stock piles with the same references.

After engaging the assignment VBA solver developed initially by Güneş Erdoğan and modified by us for 1200 seconds and 67978 iterations, we obtain the following results:

After analyzing the simulation results, we have concluded the following points:

- The number of bins used: 9 (20 in total).
- Number of HU assigned: 1044 (all the HU were assigned correctly).

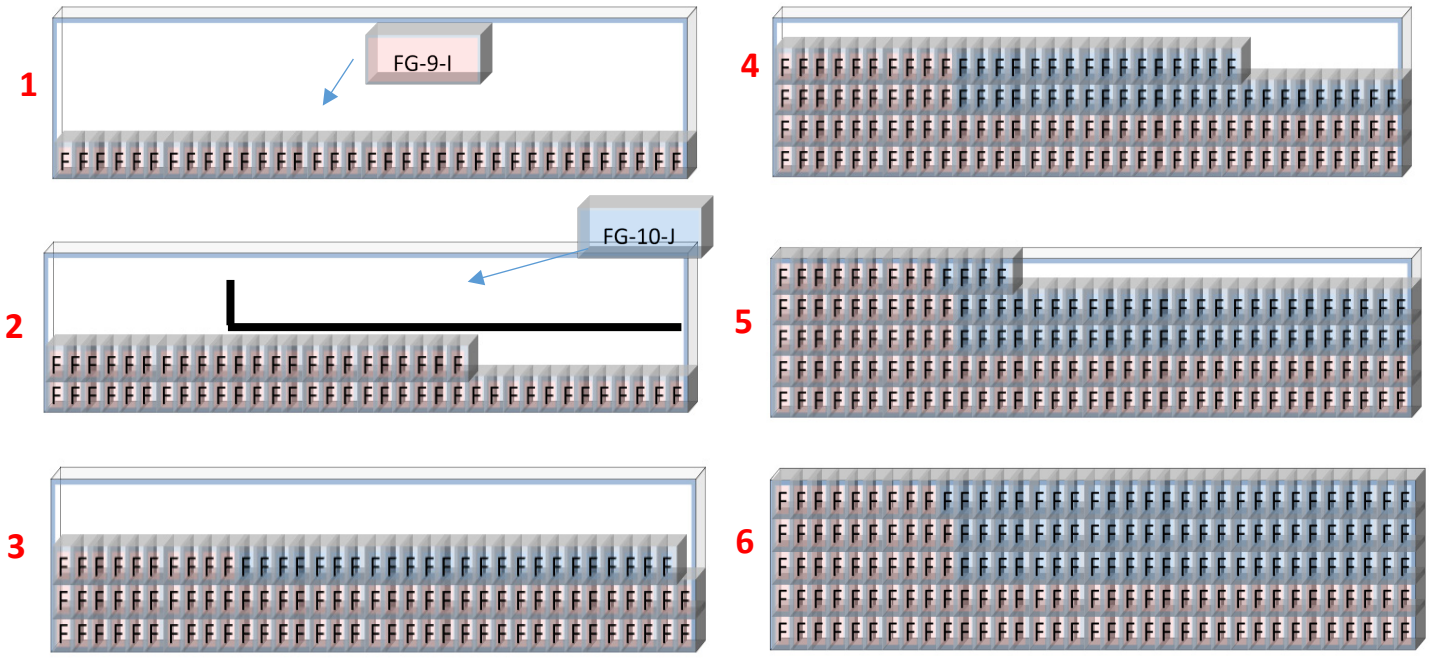


Figure 12: The assignment progress of the Bin 1 Flat

- The constraints of compatibility between raw materials and finished goods have been respected; in each bin we found only one family of products; Raw materials or Finished goods but not both. This means that there will be no congestion while preparing raw materials to be transferred to the linefeed of finished goods to be transferred to preparation area.
- The constraints of compatibility between the references to assign and the storage bin type (Flat or Rack) have been respected.
- The filling rate of bins is maximized, where there is no constraint of compatibility.
- The stocks are assigned following the physical flow, starting from the first bin, then the second, then the third… which keeps the warehouse organized.
- The wall building heuristic has given good quality results, if we look at the first part of assignment progress (figure 13), we observe that the assignment is done by placing HU's of similar articles next to each other to form the first level of stock (phase 1) then the second level (phase 2). In the third level, the heuristic moves to the article FG-10-j which is compatible with FG-9-I and place HU in the third level. Moving to the next levels, the piles formed respond to the notion: brick wall, we found a harmony of articles in each pile, which is a part of the visual management; it is easy to identify the references stored in the bin FLAT 1 as an example when we are present physically in the warehouse.

4.9. Synthesis

The spreadsheet developed defines the alerts of stocks in addition to the nominal stock in quantity, in days of stock and in monetary value for each part number. Furthermore, the required space for each reference equivalent to each alert of stock is defined considering the requested type of storage either massive or rack stock, and the tolerated stacking level. In the MRP calculation, we have integrated the upper and lower limit of stock in order to propose quantities of orders inside this interval of stock when running automatically (daily Job of MRP)

This tool facilitates the management of procurement and production planning by providing an interface including all logistics parameters throughout the supply chain impacting the EQO. The tool we have developed will facilitate the work of logistic decision-makers as it integrates the static and dynamic aspect of the stock while considering logistic management constraints. The assignment part allows to manager to check the feasibility of the stock replenishment policy defined in the segmentation phase compared to the storage locations available and the total budget authorized. Furthermore, it facilitates the management of physical flow; the spreadsheet is a logistic decision support tool. Among these advantages:

- Defining adequate procurement policy and checking the alignment of the stock value calculated with the value monthly budgeted by the top management.
- Defining a dynamic storage area for each item in the warehouse, if there is any change in the values of nominal stock, we anticipate the actions to be taken regarding the storage area.
- Detecting the anomalies and the discrepancies regarding the various logistic parameters of supply or sale (example: supplier minimum order and daily average consumption of products) and see the impact on the number of days in stock in real time mode.

- Ensuring a better interpretation of the stock constraints (FG and RM), by segmenting them into six variables, and then prioritizing the actions plan to get the optimal stock:
  - Stock delivering: Increase the frequency of the delivery downstream (PF) and upstream (RM).
  - Batch: Decrease the launches quantities (FG) and the minimum order (RM).
  - Variability of the demand: To master the variability of customer's needs (FG) / to respect and smooth the production schedule (RM).
  - Service Rate: To respect the production schedule and follow the reliability of the customer's loading (FG)/ Master the suppliers with the supplier service rate (RM).
  - Quality level: Quality action (decrease the defective parts rate…).
  - Safety stock: Safety stock defined following to a management decision or a customer contract decision.
- Check the correspondence of the financial target defined by the top management and the stock management policy defined.

6. Conclusion and perspectives

The research work in our hand is very complex from a modeling point of view; the context of stocks in the industrial domain; and integrating all the constraints of operational management of raw materials and also finished goods. In terms of the initial objectives set, we successfully determine a static and dynamic model of stock calculation. Particularly, define the optimal order quantities in a detailed mathematical and algorithmic demarche. Then propose an algorithm to solve the storage location assignment problem known as NP-Hard problem and include the allocation of space to references in function of the alerts of stock defined in the segmentation phase within a Moroccan multinational company respecting the constraint of flow. Moreover, we propose a multi criteria decision tool to support decision makers in order to take decisions based on current policy of stock replenishment, and show the variables that impact the level of stocks and should be revised or negotiated. Finally, the model proposed could be replicated in any automotive industry context since it encompasses all the constraints adopted in such an exacting context. For other industries, where there is less variables we can use the same model using a relaxation approach of some constraints for each context

For further research, after solving the space allocation problem in both types rack and flat, we have detected a new constraint in the case of flat stocks. The algorithm proposed doesn't include the order of consumption of products when assigning them in the available space; this gap provokes one of the main Muda. It includes the operations of movement of products to withdraw the requested product since all the products are stacked, and then returning the non-needed products to the stockpiles. For all these reasons, we need to solve the assignment problem of flat stock type references to the space in the warehouse considering the picking order (retrieval) of handling units while constituting stock piles and optimize rehandling operations. This problem is also NP-Hard in a strong sense and can be modeled also as a bin packing problem tree dimensions with constraints of retrieval.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

I would like to express my gratitude to my primary supervisor, of thesis Abdelfettah Sedqui, who guided and assisted me throughout this work of research. I would also like to thank the logistics manager of the multinational company for the data, support, and the trust. A special thanks for my wife and family who supported me during this work till the redaction of these lines. I would like to thank also the two anonymous reviewers whose constructive and helpful comments have helped considerably in improving the contents as well as the presentation of the paper.

References

[1] A. Laassiri, A. Sedqui, "Dynamic dimensioning of logistic resources: Case of stocks Moroccan multinational company," in 2020 5th International Conference on Logistics Operations Management (GOL), 1–7, 2020, doi:10.1109/GOL49479.2020.9314755.

[2] M.Z. Babai, Y. Dallery, "Inventory Management: Forecast Based Approach vs. Standard Approach," International Conference on Industrial Engineering and Systems Management (IESM05), 2005.

[3] G. Hadley, T.M. Whitin, Analysis Of Inventory Systems, 1963.

[4] F.W. Harris, "How Many Parts to Make at Once," Operations Research, **38**(6), 947–950, 1990, doi:10.1287/opre.38.6.947.

[5] Y. Yao, M. Dresner, "The inventory value of information sharing, continuous replenishment, and vendor-managed inventory," Transportation Research Part E: Logistics and Transportation Review, 44(3), 361–378, 2008, doi: 10.1016/j.tre.2006.12.001.

[6] D. Battini, A. Gunasekaran, M. Faccio, A. Persona, F. Sgarbossa, "Consignment stock inventory model in an integrated supply chain," International Journal of Production Research, **48**(2), 477–500, 2010, doi: 10.1080/00207540903174981.

[7] B. Korte, J. Vygen, J. Fonlupt, A. Skoda, Le probléme du bin-packing, Springer, Paris : 475–492, 2010, doi : 10.1007/978-2-287-99037-3_18.

[8] A. Laassiri, A. Sedqui, "Dynamic dimensioning of logistics resources: Case of production workshop: Analogy with the problem of bin-packing Moroccan multinational company," in 2019 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA), 1–8, 2019, doi:10.1109/LOGISTIQUA.2019.8907270.

[9] K. Cuthbertson, D. Gasparro, "The Determinants of Manufacturing Inventories in the UK," The Economic Journal, **103**(421), 1479–1492, 1993, doi: 10.2307/2234478.

[10] K.A. Cook, G.R. Huston, M. Kinney, Managing Earnings by Manipulating Inventory: The Effects of Cost Structure and Valuation Method, SSRN Scholarly Paper ID 997437, Social Science Research Network, Rochester, NY, 2012, doi:10.2139/ssrn.997437.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 311-318 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Technique to Simulate an Oscillator Ensemble Represented by the Kuramoto Model

Mark Gourary*, Sergey Rusakov

CAD Department, Institute for Design Problems in Microelectronics of Russian Academy of Sciences, Moscow, 124365, Russia

ARTICLE INFO

Article history:
Received: 02 February, 2021
Accepted: 19 April, 2021
Online: 10 June, 2021

Keywords:
Oscillator ensemble
Kuramoto model
Coupling function
Circuit simulator

ABSTRACT

The paper presents the technique for the user-friendly numerical simulation of coupled oscillators described by the Kuramoto model. Oscillators couplings are defined as arbitrary 2π-periodic functions given by the Fourier series. Matlab procedure was developed to generate netlist for the equivalent electrical circuit diagram of the Kuramoto model. The input data of the procedure include the natural frequencies of oscillators and the amplitudes of the couplings harmonics. Kirchhoff equations of the equivalent circuit coincide with the equations of the Kuramoto model. The generated netlists provide obtaining the simulation results using standard circuit simulator. These results numerically coincide with the transients computed using the original Kuramoto model. The presented examples confirm the convenience and effectiveness of the proposed approach.

1. Introduction

This paper is an extension of work originally presented at the 24th European Conference on Circuit Theory and Design, ECCTD 2020 [1] and this is refinement of work where we proposed simulation tool for analyzing ensembles of oscillators.

Studying the collective behavior of coupled oscillators is fundamental problem that has wide application in nature, science, and technology. Simulation of large networks of coupled oscillators is actual multidisciplinary problem due to high demand in various research areas of including biology, chemistry, physics, electronics etc. [2].

The relevance of this issue has been constantly increasing recently due to the growing interest in the study of oscillatory neural networks [3, 4], to the problem of mutual injection locking of oscillators under parasitic couplings in integrated circuits [5], and to various other applications.

The ensemble of physical oscillators is described by system of ordinary differential equations (ODE). Respectively, the analysis of the ensemble can be performed by numerical methods for ODE solving. This approach is effective for an ensemble with small number of simple oscillators [6]. Direct ODE solving of large oscillator ensembles requires too high computational effort. Therefore, simplified models are required for effective simulation of large oscillator systems [7].

The simplification techniques assume that the coupling strength between the oscillators is rather weak [8, 9]. In this case the coupled oscillator waveform can be considered to coincide with the waveform of this oscillator in free running mode. Then the set of system variables includes only the phases of the oscillators.

The most widespread simplified model is the Kuramoto model (KM) [4, 9, 10] represented by ODE system with respect to phase variables. Kuramoto model describes each oscillator with a single phase variable. The oscillatory interconnections are described in Kuramoto model by 2π-periodic coupling functions.

The standard mathematical packages are usually used for the KM numerical solving. Most often Matlab software [11] is applied (see, for example, [12, 13]). Also, Python codes can be used for solving KM ODE system [14, 15].

However, using standard software packages leads in practice to some difficulties and does not provide effective simulation of an ensemble of oscillators in many applied cases. The user needs to think through the names of variables, repeatedly include the same operators in the code and describe the necessary operations for each specific task. This increases the amount of data preparation, takes a lot of time and leads to additional user errors.

For this reason, the development of universal approach and software tool to improve numerical efficiency of Kuramoto model applications and provide convenient analysis of the simulation results is the actual problem of simulation of ensemble of physical oscillators.

*Corresponding Author Mark Gourary, Email: gourary@yandex.ru

The above limitations are the motivation for the development of new numerical approaches that significantly increase the usability of applied tools.

The paper [16] is devoted to this topic. The paper is directed to reduce the arisen difficulties in process of the numerical solution while KM application. The approach [16] is based on representing the KM in the form of the state equations for an equivalent electrical circuit and applying an electrical simulation tool to this circuit. The obtained circuit simulation results are numerically equal to the results of KM simulation.

However, the implementation of the approach presented [16] has the following disadvantages:

- an equivalent circuit is constructed under the wave digital concept [17] which is run on digital signal processors and is not suitable for standard circuit simulators;
- coupling functions in KM can only be represented by sinusoids with the same amplitude, arbitrary functions are not allowed;
- considered sinusoidal coupling functions cannot include phase delay that leads to the impossibility to analyze some important types of coupled oscillators.
- The contribution of the presented paper is connected with new approach that allows to eliminate the limitations of [16]. We propose the following principles for constructing KM equivalent circuit:
- the equivalent circuit should be suitable to analyze by means of standard circuit simulators. Such a circuit consists of standard electrical components which can be described by input netlist;
- coupling functions of arbitrary 2π-periodic form can be considered in KM and specified by truncated Fourier series with the given harmonics amplitudes and phases;
- special-purpose technique to generate the equivalent circuit netlist should be applied.

In comparison with paper [1] the contribution involves additional simulation example to illustrate capabilities of the developed approach for analyzing oscillator ensembles.

The rest of the paper is organized as follows. Section 2 describes the known mathematical forms of the KM equations. Section 3 explains the forming the equivalent circuit for a given KM. Section 4 outlines the principles for the netlist generation procedure. Numerical experiments are presented in Section 5.

2. Kuramoto Model

2.1. Basic Model Equations

KM defines the behavior of a system of N weakly coupled oscillators. Each oscillator is characterized by its natural frequency (fundamental) ω_m and the time-varying instantaneous phase $\theta_m(t)$. In its simplest form, corresponding to the initial Kuramoto proposal (1) KM is represented as an ODE system [18, 19]:

$$\frac{d\theta_m}{dt} = \omega_m + \frac{K}{N}\sum_{n=1}^{N} sin(\theta_n - \theta_m), \qquad m = 1 \ldots N. \quad (1)$$

here K is the coupling strength assumed to be the same for all couplings. This form made it possible to obtain estimates of the behavior of the synchronized ensemble of oscillators. However, the assumption about the same couplings is not met in most cases, so more general form of KM was proposed [20, 21]

$$\frac{d\theta_m}{dt} = \omega_m + \sum_{n=1}^{N} A_{mn}\, sin(\theta_n - \theta_m), \qquad m = 1 \ldots N. \quad (2)$$

here A_{mn} is the individual coupling strength between m-th and n-th oscillators. Form (2) of KM is more flexible tool to represent real sets of coupled oscillators.

The natural extension of KM also considers the external periodic force [22, 23] as the excitation with a given phase θ_e applied to internal oscillators. In such case the Right-Hand Side of m-th equation (2) is supplemented by external coupling functions of the form $A_{m,e}(\theta_e - \theta_m)$.

In the simplest case when one oscillator is excited by a single stimulus, the KM equation has the form:

$$\frac{d\theta}{dt} = \omega + A_e\, sin(\theta - \theta_e). \quad (3)$$

The phase of the periodic excitation is $\theta_e = \omega_e t$ where ω_e is the excitation frequency that in the synchronized mode coincides with the oscillator frequency. The oscillator phase is $\theta = \omega_e t + \vartheta$ where φ is the initial phase of the synchronized oscillator. After substituting expressions for phases θ_e, θ into (3) we obtain the algebraic equation with respect to ϑ and the condition of its solution existence defining the locking range of the oscillator

$$\omega_e - \omega = A_e\, sin(\vartheta)\,, |\omega_e - \omega| \le A_e. \quad (4)$$

In general case the locking range ω_{mn} of m-th oscillator under the excitation by n-th oscillator is equal to the coupling factor A_{mn}.

$$\omega_{mn} = A_{mn}. \quad (5)$$

System (2) with antisymmetric sinusoidal coupling functions always results in a synchronized behavior for a set of identical oscillators. However, starting from [24], it was found that some dynamical systems of identical oscillators can demonstrate unsynchronized states with quasiperiodic oscillations. Such states, defined as chimeras [25, 26], can be represented by KMs with additional phase shifts in sinusoidal arguments. Using notation (5), we can write:

$$\frac{d\theta_m}{dt} = \omega_m + \sum_{n=1}^{N} \omega_{mn}\, sin(\theta_n - \theta_m + \varphi_{mn}), \quad (6)$$

that can also be presented as the sum of sines and cosines due to $\sin(\theta_n - \theta_m + \varphi_{mn}) = a\sin(\theta_n - \theta_m) + b\cos(\theta_n - \theta_m)$.

KM in the form (6) does not capture some of the effects inherent in coupled oscillators. E.g., superharmonic synchronization [27] of m-th oscillator under the excitation by k-th harmonic of n-th oscillator cannot be taken into account. Thus, more complicated coupling functions are needed, and the most general form of KM was defined as [10, 22]

$$\frac{d\theta_m}{dt} = \omega_m + \sum_{n=1}^{N} h_{mn}(\theta_n - \theta_m). \quad (7)$$

Here $h_{mn}(\theta_n - \theta_m)$ are 2π-periodic coupling functions that are often approximated to desired accuracy using truncated Fourier series with sufficiently large number of terms:

$$h_{mn}(\theta_n - \theta_m) = \sum_{k=1}^{K_m} A_{mn}^{(k)} \sin(k(\theta_n - \theta_m) + \varphi_{mn}^{(k)}). \quad (8)$$

Constant terms in the Fourier series (5) are zeroes ($A_{mn}^{(0)} = 0$) due to the compensation of nonzero terms by the change of fundamentals ω_m in (1). Similarly, diagonal entries are also assumed to be $A_{mm}^{(k)} = 0$ for all *k*.

Under notations (5) one can represent (8) as

$$\frac{d\theta_m}{dt} = \omega_m + \sum_{n=1}^{N} \sum_{k=1}^{K_m} \omega_{mn}^{(k)} \sin\left(k(\theta_n - \theta_m) + \varphi_{mn}^{(k)}\right). \quad (9)$$

Here $\omega_{mn}^{(k)}$ is equal to the locking range of superharmonic synchronization [13] of *m*-th oscillator under the excitation by *k*-th harmonic of *n*-th oscillator.

2.2. Modified Model Equations

To enhance the user-friendliness when simulating coupled oscillators, we have made some modifications to the KM (9).

Couplings activation moments were included in the model assuming all couplings were initially disabled.

The activation moment is defined by adding to the coupling function time-dependent multiplier $u(t - \tau_{mn})$ were $u(t)$ is the unit step function, τ_{mn} is the coupling activation moment. Then coupling function from KM (6) has the form:

$$g_{mn}(\theta_n - \theta_m, t) = u(t - \tau_{mn})\omega_{mn} \sin(\theta_n - \theta_m + \varphi_{mn}). \quad (10)$$

For the coupling functions defined by Fourier series (8), (9) we indicate the activation moment $\tau_{mn}^{(k)}$ for each *k*-th Fourier term.

Main results of KM simulations are presented by oscillator phases θ_m. However, the phase waveforms are often inconvenient for analyzing the simulation results due to include a linear component. More user-friendly data is represented by instantaneous frequency waveforms, that indicate the synchronization mode by their constant values. The instantaneous frequency in angular ($\omega_m^{\mathbf{inst}}$) or regular ($f_m^{\mathbf{inst}}$) forms is determined as following:

$$f_m^{\boldsymbol{inst}} = \frac{\omega_m^{\boldsymbol{inst}}}{2\pi} = \frac{1}{2\pi}\frac{d\theta_m}{dt}. \quad (11)$$

Note that the angular form of the instantaneous frequency $\omega_m^{\mathbf{inst}}$ (7) is used in any KM form (1), (2), (6) and (9). This also corresponds to the Right-Hand Side of KM. But usually the user needs to analyze the frequency in the regular form $f_m^{\mathbf{inst}}$, so to avoid superfluous transforms one can convert KM to a form with respect to regular frequencies by dividing it by 2π. Then KM (9) is represented as

$$\frac{1}{2\pi}\frac{d\theta_m}{dt} = \\ = f_m + \sum_{n=1}^{N} \sum_{k=1}^{K_n} u(-\tau_{mn}^{(k)}) f_{mn}^{(k)} \sin\left(k(\theta_n - \theta_m) + \varphi_{mn}^{(k)}\right). \quad (12)$$

where $f_m = \omega_m/2\pi$, $f_{mn}^{(k)} = \omega_{mn}^{(k)}/2\pi$.

3. Representation of Kuramoto Model by Equivalent Electrical Circuit

The behavior of the oscillator ensemble can be obtained by the simulation of the electrical circuit if state equations of the circuit coincide with the KM ODE system (12) of the ensemble. Thus, we proposed the following structure of the equivalent circuit.

Each *m*-th oscillator from KM corresponds to the node in the equivalent circuit with the same index *m*. The voltage of the node is equal to the phase of KM (12): $\theta_m(t) \equiv V_m(t)$.

The devices connected in parallel to *m*-th node include:

1. Capacitor C_m=1/(2π) representing the phase differentiation in the left-hand side of (12). The capacitor current is numerically equal to the instantaneous frequency of *m*-th oscillator:

$$I_m^{inst}(t) = C_m \cdot dV_m/dt = 1/2\pi\, d\theta_m/dt = f_m^{inst}(t). \quad (13)$$

2. Independent current source numerically equal to the fundamental of *m*-th oscillator

$$I_m^{fund} = f_m. \quad (14)$$

3. Nonlinear current sources correspond to Fourier terms in the Right-Hand Side of (9). Each source represents *k*-th Fourier harmonic of a coupling function from *m*-th to *n*-th oscillator:

$$I_{m,n.k}^{harm}(t, v_n, v_m) = \\ = \left(t - \tau_{mn}^{(k)}\right) f_{mn}^{(k)} \sin\left(k(V_n - V_m) + \varphi_{mn}^{(k)}\right). \quad (15)$$

here parameters $f_m, \tau_{mn}^{(k)}, f_{mn}^{(k)}, \varphi_{mn}^{(k)}$ are defined in KM (12).

One can see that Kirchhoff Current Law (KCL) equations for the equivalent circuit with *M* nodes corresponding to all oscillators of KM represent the same ODE system as the initial KM:

- the KCL equation for the node of *m*-th capacitor coincides with *m*-th equation of KM (12),
- the voltage (V) of *m*-th node is numerically equal to the instantaneous phase (rad) of *m*-th oscillator,
- the current (A) through *m*-th capacitor is numerically equal to the instantaneous frequency (Hz) of *m*-th oscillator due to (11).

Note that m-th node in the equivalent circuit (13-15) is floating due to the capacitor connection to the current sources. Hence the initial capacitor voltage equal to the initial oscillator phase should be defined to provide transient simulation.

The example in Figure 1 represents the equivalent circuit for KM of the pair of coupled oscillators.

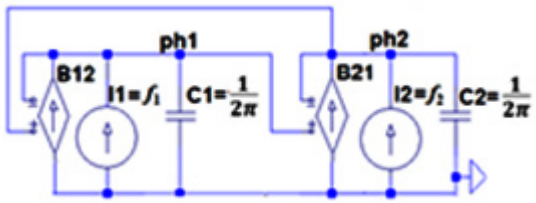


Figure 1: The equivalent circuit for KM of the pair of coupled oscillators.

Input netlist of the circuit in Figure 1 for oscillators' fundamentals f_1=100Hz, f_2=102Hz and couplings parameters, f_{12}=f_{21}=1Hz, τ_{12}=1sec., τ_{21}=5sec. has the form:

Input netlist for KM of the pair of coupled oscillators.

```
.params inv2pi=0.5/pi    *parameter to define constant 1/2π
* oscillators (13)
C1   0   ph1  {1/2π}
C2   0   ph2  {1/2π}
* fundamentals (14)
I1   ph1  0  100
I2   ph2  0  102
* couplings  (15)
B12 ph1 0 I = 1*sin(v(ph2,ph1))*u(time - 1)
B21 ph1 0 I = 1*sin(v(ph1,ph2))*u(time - 5)
* initial phases of oscillators
.IC  v(ph1) = 0
.IC  v(ph1) = 0
.end
```

After processing the netlist by LTspice circuit simulator [28] one can obtain instantaneous frequencies of oscillators as current waveforms of capacitors C1 and C2 (13). Examples of the waveforms I(C1), I(C2) are presented in Figure 2.

The waveforms in Figure 2a are obtained for parameters given in the presented netlist. While both couplings at t < 1sec are disabled the oscillators produce their natural frequencies f_1=100Hz and f_2=102Hz. At t=1sec coupling 1-2 is activated but the coupling factor f_{12}=1Hz is insufficient ($|f_1 - f_2| = 2 > f_{12}$) to provide synchronization of the oscillator #2 by the oscillator #1. So, beats appear in the instantaneous frequency waveform at 1 < t < 5sec. After the activation of both couplings at t > 5sec the oscillators are mutually synchronized with common frequency 101Hz.

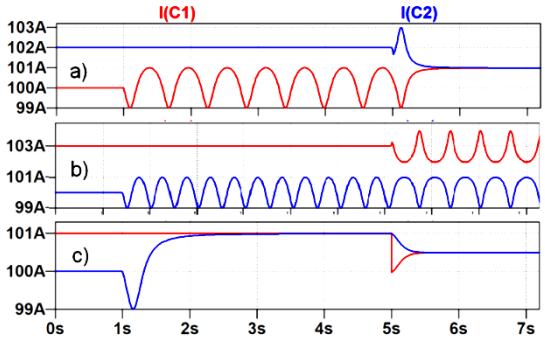


Figure 2: Simulated waveforms of instantaneous frequencies of capacitances currents obtained for f1=100Hz and a) f2=102Hz, b) f2=103Hz, c) f2=101Hz

Figure 2b demonstrates simulation results for an increased value of the second fundamental to f_2 = 103Hz. results for an increased value of the second fundamental to f_2 = 103Hz. The discrepancy of fundamentals is too large ($|f_1-f_2|$ = 3Hz) to provide synchronization even under the activation of both couplings.

The decrease of the second fundamental to f_2=101Hz (($|f_1 - f_2|$ =1Hz) leads to waveforms in Figure 2c with synchronization under both unidirectional and bidirectional couplings.

4. Automatic Forming of Input netlist by Matlab Description of Kuramoto Model

For the convenience of forming a Input netlist of an equivalent electrical circuit, we introduce the naming rules for nodes and components of the circuit that is shown in Table 1.

Table 1: Netlist names corresponding to m-th oscillator.

m-th circuit components (KM variable)	Nota tion	Netlist name
m-th node (m-th oscillator)	m	phm
Node voltage (oscillator phase)	v_m	v(phm)
m-th capacitor	C_m	Cm
Capacitor current (instantaneous frequency)	I_m^C	I(Cm)
m-th DC current source (natural frequency)	I_m^{fund}	Im
Behavioral current source (coupling harmonic) k - harmonic index, m,n – oscillators indexes	$I_{m,n.k}^{harm}$	Bm_n_k

```
Cm      0   phm   {1/2π};  differentiation operator
Im      phm  0    fm;      natural frequency
……
Bm_n_k  phm  0             k-th harmonic of mn coupling
+ I = u(t − τ_mn^(k)) f_mn^(k) sin(k(v(phn) − v(phm)) + φ_mn^(k))
……..
……..
.IC  v(phm) = PHm;         initial phase of m-th oscillator
```

One can see from this fragment that the netlist of the equivalent circuit contains significant amount of duplicate or redundant text data because the same oscillator indexes can be repeated in many component names. For example, the oscillator index *m* is included in the name of the node representing the oscillator phase and in the name of each component connected to the node.

Redundancy increases for the netlists with many Fourier harmonics (Bm_n_k). In this case the same indexes of oscillators must be presented in each behavioral source (Bm_n_k) both as the part of its name and names of reference nodal voltages v(phn).

To eliminate the redundancy of the input data, we propose to generate the Input netlist using the procedure based on the description of interacting oscillators in compact form without data duplication. The form is convenient for filling out by the user.

When developing the procedure, we firstly changed the set of primary parameters in favor of relative values. Considering that KM is applicable to the ensemble of oscillators with close natural frequencies one can introduce the "base frequency" parameter f_b, which is close to average fundamentals of all oscillators.

Individual natural frequencies are set by defining its relative deviations (r_m) from the base frequency:

$$r_m = (f_m - f_b)/f_b \text{ or } f_m = (r_m + 1) \cdot f_b. \quad (16)$$

The coupling factors $f_{mn}^{(k)}$ in (10) are determined by relative values $r_{mn}^{(k)}$ with respect either to the base frequency:

$$r_{mn}^{(k)} = f_{mn}^{(k)}/f_b, \quad (17)$$

or to the natural frequency of the excited oscillator,

$$r_{mn}^{(k)} = f_{mn}^{(k)}/f_m = f_{mn}^{(k)}/\left((r_m + 1) \cdot f_b\right). \quad (18)$$

The choice between options (17), (18) depends on the control logical variable mr=0/1 that is set for all KM couplings.

Phases in (10) are set relatively to 2π: $\phi_m^k = \varphi_m^k/2\pi$.

$$\phi_m^k = \varphi_m^k/2\pi. \quad (19)$$

The procedure to generate an Input netlist is developed using Matlab software [7], so parameters (13) – (15) are grouped in the Matlab data types that include:

1. Simple numeric variables:

- **fb** – basic frequency.

- **mr** – type of relative parameter.

- **Tsim** – simulation time.

2. Cell array **osc** that contains oscillators' parameters. Each array element contains parameters of one (m-th) oscillator:

- the relative deviations of natural frequency r_m (16);

- the initial phase $\theta_m(0)$.

The oscillator parameters are set by assigning them to the array element with the oscillator index, e.g.

osc{n1} =[r_{n1} $\theta_{n1}(0)$];

osc{n2}=[r_{n2}];

……

osc{nL}=[r_{nL} θ_{nL} (0)];

Oscillator indices (n1, n2, …nL) are arbitrary non-repeating integers. The value of zero initial phase can be omitted (see n2).

3. Cell matrix **coup** with parameters of couplings. Each matrix entry (m, n) defines cell array of parameters of harmonics $I_{m,n.k}^{harm}(t, v_n, v_m)$ (15) for given indexes of excited (m) and exciting (n) oscillators. The sequential number of the parameter vector presents the harmonic index k and components of the vector define coupling factor, phase shift and switching moment of the harmonic. Thus, the coupling function (m, n) with k harmonics is represented as matrix entry with parameters of all harmonics:

coup{m,n}={[$r_{mn}^{(1)}$ $\phi_{mn}^{(1)}$ $\tau_{mn}^{(1)}$][$r_{mn}^{(2)}$ $\phi_{mn}^{(2)}$ $\tau_{mn}^{(2)}$] … [$r_{mn}^{(k)}$ $\phi_{mn}^{(k)}$ $\tau_{mn}^{(k)}$] };

For example, the pair of coupled harmonic oscillators in Figure 1 is defined by the following description:

fb=100;

osc{1}=[0];

osc{2}=[0.06];

coup{1,2}={[0.01 0 1]};

coup{2,1}={[0.01 0 5]};

One can see that this description is more compact than presented above Input netlist of the equivalent circuit in Figure 1. More essential reducing of the size of input data can be achieved for KM containing many coupling functions with large number of Fourier harmonics.

Thus, to simulate the oscillator ensemble defined by the given KM one should perform the following operations:

- form Matlab description of the KM by filling-in cell array **osc** with parameters of oscillators and filling-in cell matrix **coup** with parameters of couplings;

- generate Input netlist by applying the developed procedure to the formed KM description;

- apply standard circuit simulator to the generated netlist;

- analyze time-domain dependences of phases and frequencies by applying the simulator waveform viewer to the corresponding nodal voltages and capacitances currents.

5. Numerical Examples

Here we present two numerical examples that illustrate the application of the developed approach to the analysis of ensembles of oscillators. To simulate the generated equivalent circuit, the LTspice circuit simulator [28] was used.

The first example presents 8 ensembles of 9 oscillators with natural frequencies evenly distributed in the interval [98 – 102] Hz:

$$f_n = 95 + n, n = 1 \ldots 9. \quad (20)$$

The oscillators are sequentially coupled in forward direction and have one or two reverse couplings forming different versions of closed sequence that are shown in Figure 3.a1-a8. The figure 3aN represents the version with the index N of the intermediate oscillator between reverse couplings (Figure 3a1 is the version with one coupling).

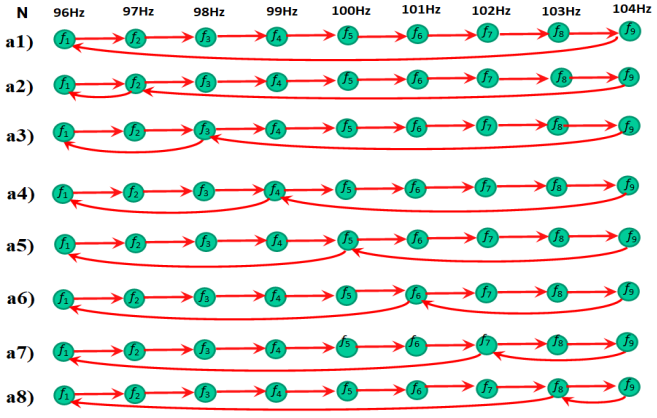


Figure 3: Series of oscillators 1-8 with reverse couplings through N-th oscillator: aN), N=1…8.

All couplings in each ensemble a1-a8 are defined by sine functions (2) with the same coupling factor $r_{mn}^{(1)} = \rho$ (for all m, n).

The aim of the experiments is to determine for each version the value of the common coupling factor providing the ensemble synchronization - ρ_{synch} and to compare obtained values for different versions. To obtain the results the ensembles in Figure a1-a8 must be simulated for a range of overall coupling factors.

To perform the simulation the developed procedure was applied. As an example we present below Matlab descriptions of KM for a1, a5 and corresponding simulation results.

```
% Matlab description for KM a1
fb = 100; mr = 0;  Tsim=15;
osc{1}=[-0.08]; osc{2}=[-0.06]; osc{3}=[-0.04];
osc{4}=[-0.02]; osc{5}=[0.0]; osc{6}=[0.02];
osc{7}=[0.04]; osc{8}=[0.06]; osc{ 9}=[0.08];
coup{2,1}={[ρ 0 1]}; coup{3,2}={[ρ 0 1]};
coup{4,3}={[ρ 0 1]}; coup{5,4}={[ρ 0 1]};
coup{6,5}={[ρ 0 1]}; coup{7,6}={[ρ 0 1]};
coup{8,7}={[ρ 0 1]}; coup{9,8}={[ρ 0 1]};
coup{1,9}={[ρ 0 1]};
```

Simulation results are represented in Figure 4a-d, where the waveforms of the instantaneous frequencies $f_1(t)$ and $f_9(t)$ are shown for the different values of coupling factors:

1. Small value ρ = 1.3Hz ≈ 0.125ρ_{synch} (Figure 4a). The waveforms are close to sine curves.

2. Moderate value ρ = 5.4Hz ≈ 0.5ρ_{synch} (Figure 4b). The distortions of sinusoids increase.

3. The value approaching the synchronization value ρ = 10.8Hz ≈ 0.99·ρ_{synch} (Figure 4c). The waveforms are represented by pulsed curves both for $f_1(t)$ and $f_9(t)$.

4. The minimal value for the synchronization ρ = ρ_{synch} = 10.8Hz (Figure 4d). Initial pulses disappear and after that both instantaneous frequencies are equal to the constant synchronization frequency $f_1(t) = f_9(t) = f_{synch}$

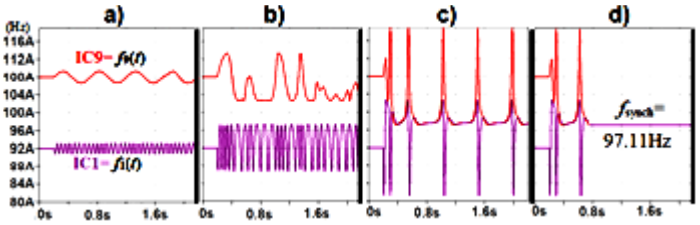


Figure 4: Simulation results for KM a1, The instantaneous frequencies $f_1(t)$ and $f_9(t)$ for the values of the common coupling factor ρ: a) ρ = 1.3Hz; b) ρ = 5Hz; c) ρ = 10.8Hz; d) ρ = ρ_{synch} = 10.8Hz.

Matlab description for KM a5 differ from KM a1 by replacing coupling coup{1,9}={[ρ 0 1]}; by 2 couplings:

```
coup{1,5}={[ρ 0 1]};  coup{5,9}={[ρ 0 1]};
```

Experimental waveforms of the instantaneous frequencies $f_1(t)$, $f_5(t)$ and $f_9(t)$ are shown in Figure 5, for the similar coupling factors: ρ = 1.05Hz, ρ = 4.2Hz, ρ = 8.35Hz, ρ_{synch} = 8.41Hz.

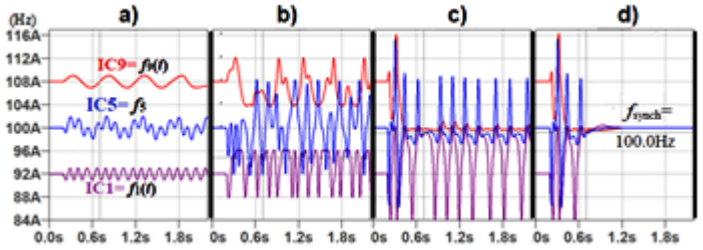


Figure 5: Simulation results for KM a5, The instantaneous frequencies $f_1(t)$ and $f_9(t)$ for the values of the coupling factor ρ: a) ρ = 1.05Hz; b) ρ = 4.2Hz; c) ρ = 8.4Hz; d) ρ = 8.41Hz.

Simulations like Figure 4, 5 were performed for all oscillator ensembles a1-a8 (Fig 3). The obtained numerical characteristics are presented in Table 2. Each column of the table contains:

- the name of the ensemble (N),
- the common coupling factor providing synchronizing (f_{ij}),
- the deviation of the synchronization frequency from the basic frequency (Δf_{syn}).

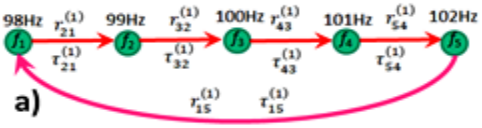
Table 2 Numerical characteristics of oscillator ensembles a1-a8

N	a1	a2	a3	a4	a5	a6	a7	a8
f_{ij}	10.86	16.49	8.01	14.36	8.411	13.97	9.15	12.41
f_s	97.11	99.79	100.0	99.23	100.0	100.0	100.3	97.08

The second example is a ring sequence (Figure 6a) of 5 oscillators with couplings, which are sequentially switched on with a delay of 5 seconds. The natural frequencies of the oscillators are evenly distributed in the interval [98 - 102] Hz (base frequency - 100 Hz). Each coupling function in an open loop 1–2–3–4–5 contains only the first harmonic with zero phase shift $\phi_{mn}^{(1)} = 0$ and coupling factors are $f_{mn}^{(1)}$=3Hz ($r_{mn}^{(1)}$=0.03).

The aim of the numerical experiments is to analyze the dependence of the synchronization process on the magnitude of the coupling function between the first and fifth oscillators ($g_{51}(\theta_{15})$).

The initial description of the ring ensemble is presented in Figure 6b for 5-1 coupling factor $r_{15}^{(1)} = 0.005$. Waveforms of instantaneous frequencies of oscillators obtained as simulation results are presented in Figure 7.




```
b)
fb=100;
osc{1}=[-0.02];  coup{2,1}=[0.030  0  5];
osc{2}=[-0.01];  coup{3,2}=[0.030  0  10];
osc{3}=[ 0.00];  coup{4,3}=[0.030  0  15];
osc{4}=[ 0.01];  coup{5,4}=[0.030  0  20];
osc{5}=[ 0.02];  coup{1,5}=[0.005  0  25];
```

Figure 6: a) The ring of 5 coupled harmonic oscillators, b) matlab description of the ring.

At t < 5sec the instantaneous frequencies of all oscillators coincide with their natural frequencies because all couplings are disabled. After coupling 1-2 is switched on at t = 5sec the second oscillator is synchronized by the natural frequency of the first one. Similarly the synchronization of oscillators 3 , 4 occurs at t = 10sec

and at t = 15sec correspondingly. But enabling coupling 4-5 at t = 20sec does not lead to the synchronization of oscillator 5 because the discrepancy of frequencies ($f_5 - f_1$ = 4Hz) exceeds the coupling factor $r_{54}^{(1)} = r_{54}^{(1)} \cdot f_b$ = 3Hz. As a result, the oscillator 5 instantaneous frequency at 20 < t < 25sec is a periodic waveform. Enabling coupling 5-1 closes the ring and interrupts all synchronizations at t > 25 sec where instantaneous frequencies of all oscillators are time varying waveforms.

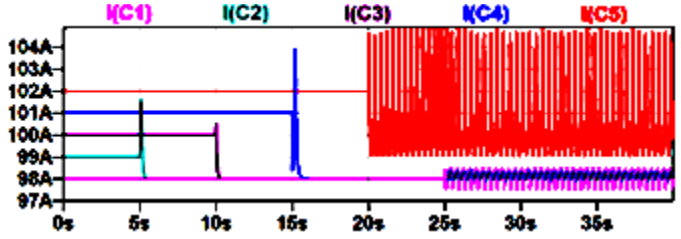


Figure 7: Simulation results obtained using matlab description in Figure 6b.

In more details the waveforms at t = 24-28 sec. can be seen in Figure 8a which is the first one of the plots in Figure 8a-e representing simulation results for some values of 5-1 coupling factor $r_{15}^{(1)}$. The unsynchronized behavior of the oscillators is seen in two intervals: $r_{15}^{(1)} \leq 0.041$ (Figure 6a,b) and $r_{15}^{(1)} \geq 0.078$ (Figure 6e). In the interval $0.0042 \leq r_{15}^{(1)} \leq 0.077$ (Figure 6c-d) all oscillators are mutually synchronized.

Thus, two bifurcation points can be defined in the axis of values of the coupling factor. At $r_{15}^{(1)} = 0.00415$ the transition from unsynchronized (Figure 8b) to synchronized (Figure 8c) mode occurs. The inverse transition (from Figure 8d to Figure 8e) occurs at $r_{15}^{(1)} = 0.0775$.

A similar analysis of the impact of the coupling factor was performed for the 5-1coupling function with a non-zero phase shift ($\phi_{15}^{(1)} = \pi/2$). The results are shown in Figure 9.

The input description of the 5-1 coupling is presented as

coup{1,5}={[0.025 0.25 25]};

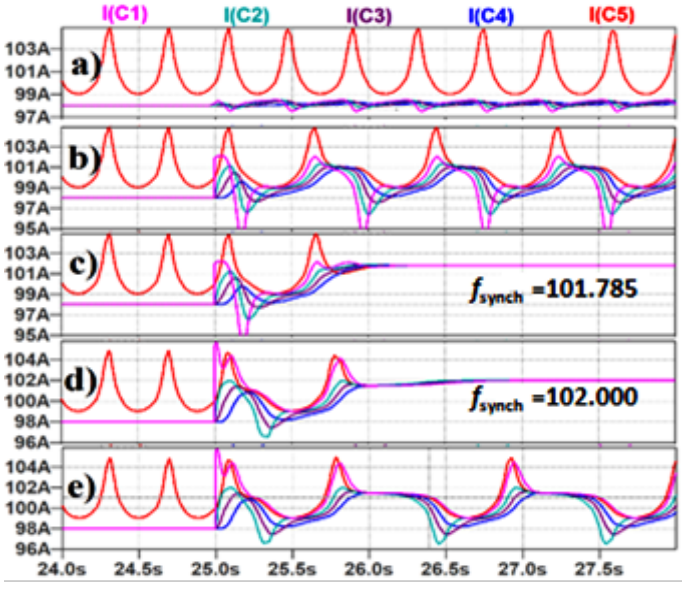


Figure 8: Simulation results for various 5-1 coupling factors a) $r_{15}^{(1)}$=0.005, b) $r_{15}^{(1)}$=0.041, c) $r_{15}^{(1)}$=0.042, d) $r_{15}^{(1)}$=0.077 e) $r_{15}^{(1)}$=0.078.

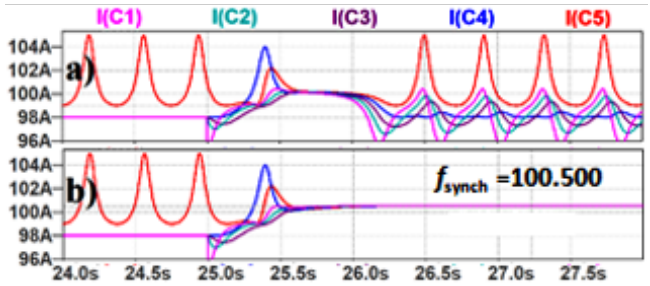


Figure 9: Simulation results for various 5-1 coupling factors under phase shift $\phi_{15}^{(1)} = \pi/2$: a) $r_{15}^{(1)}$=0.024, b) $r_{15}^{(1)}$=0.025.

Some obtained waveforms in Figure 9 demonstrate the transition from unsynchronized (Figure 9a) to synchronized (Figure 9b) mode in the single bifurcation point $r_{15}^{(1)} \approx 0.0245$.

The simulated waveforms in Figure 10 are obtained under varying the second harmonic magnitude ($r_{15}^{(2)}$) of 5-1 coupling function. The phase shift and all other harmonics of the coupling function are zeroes.

coup{1,5}={[ ] [0.014 0 25]};

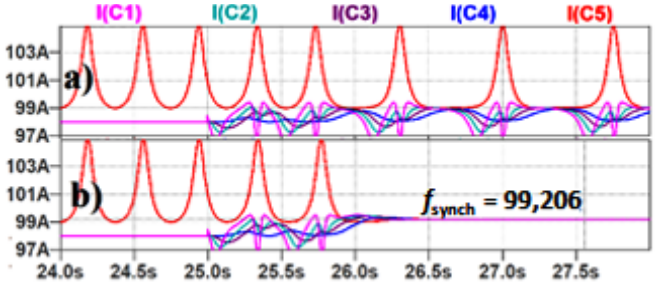


Figure 10: Simulation results for 5-1 for coupling function with the second harmonic only ($r_{15}^{(1)}$=0.0): a) $r_{15}^{(2)}$=0.014, b) $r_{15}^{(2)}$=0.015.

Here similarly to Figure 9 the domains of unsynchronized and synchronized modes are separated by the sole bifurcation point of the second harmonic $r_{15}^{(2)} \approx 0.0145$.

Thus, the presented experiments showed the useful application of the proposed approach to simulating coupled oscillators.

6. Conclusions

The paper presents the numerical approach to simulate ensembles of arbitrary oscillators. The principles of operation of the required software based on the circuit simulator and the Matlab system are developed.

The proposed approach for analyzing an ensemble of oscillators is based on the construction of model of electrical circuit. The Kirchhoff equations of this model coincide with the Kuramoto model equations of the considered ensemble. The unambiguous correspondence between the currents and voltages of the equivalent circuit and the phases and frequencies of the Kuramoto model is provided. As a result, simulating the equivalent electrical circuit performs the computation of the transients in the ensemble including computation of oscillatory phases and frequencies. The analysis of obtained transients is supported by plotting tools of modern electrical simulators.

The set of simple rules provides constructing of the equivalent circuit for coupling functions defined by their Fourier series. The numerical procedure has been developed to automatically create the input netlist. The routine generates netlist by processing the compact form of user-friendly ensemble description. The description includes lists of oscillator parameters and communication parameters represented as a Matlab structure.

The presented numerical examples have demonstrated the efficiency and convenience of the proposed approach.

Acknowledgment

The reported study was funded by RFBR, project number 19-29-03012

References

[1] M.M. Gourary, S.G. Rusakov, "A Simulation Tool for Analysis of Oscillator Ensembles Defined by Kuramoto Model," 2020 European Conference on Circuit Theory and Design (ECCTD), Sofia, Bulgaria, 1-4, 2020, doi: 10.1109/ECCTD49232.2020.9218382.

[2] S.H. Strogatz, Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry, and Engineering, Westview Press, 2015. 2, 528. ISBN 978-0-813-34910-7.

[3] M. Bonnin, F. Corinto, M. Gilli, "Periodic Oscillations in Weakly Connected Cellular Nonlinear Networks," IEEE Transactions on Circuits and Systems I: Regular Papers, **55**(6), 1671-1684, 2008, doi: 10.1109/TCSI.2008.916460.

[4] P. Ashwin, S. Coombes, R.J. Nicks, "Mathematical Frameworks for Oscillatory Network Dynamics in Neuroscience," Journal of Mathematical Neuroscience 6 (2), 1-92, 2016, doi: 10.1186/s13408-015-0033-6.

[5] P. Bhansali, J. Roychowdhury, "Injection Locking Analysis and Simulation of Weakly Coupled Oscillator Networks," in Simulation and Verification of Electronic and Biological Systems, 71-93, Springer Science+Business Media B.V., 2011, doi: 10.1007/978-94-007-0149-6_4.

[6] K.S.T. Alain, F.H. Bertrand, "A Secure Communication Scheme using Generalized Modified Projective Synchronization of Coupled Colpitts Oscillators," IJMSC, **4** (1), 56-70, 2018, doi: 10.5815/ijmsc.2018.01.04

[7] P. Maffezzoni, B. Bahr, Z. Zhang, L. Daniel, "Reducing Phase Noise in Multi-Phase Oscillators," IEEE Transactions on Circuits and Systems I: Regular Papers, **63**(3), 379-388, 2016, doi: 10.1109/TCSI.2016.2525078.

[8] C. Gong, A. Pikovsky, "Low-Dimensional Dynamics for Higher Order Harmonic Globally Coupled Phase Oscillator Ensemble", Physical Review E. **100** (6-1), 062210, 2019, doi: 10.1103/PhysRevE.100.062210.

[9] F. Dörfler, F. Bullo, "Synchronization in complex networks of phase oscillators: A survey," Automatica, **50**(6), 1539-1564, 2014. doi: 10.1016/j.automatica.2014.04.012.

[10] J.A. Acebron, L.L. Bonilla, C.J.P. Pérez Vicente, F. Ritort and R. Spigler, "The Kuramoto model: A simple paradigm for synchronization phenomena," Reviews of Modern Physics **77**(1), 137 – 185, 2005, doi: https://doi.org/10.1103/RevModPhys.77.137

[11] A. Gilat, Matlab: An Introduction with Applications, 7th Edition Pod for Student Choice, John Wiley & Sons, Incorporated, 2017.

[12] J. Kim, "Kuramoto model Numerical code (MATLAB)," in https://appmath.wordpress.com/2017/07/23/kuramoto-model-numerical-code-matlab

[13] "Cleve's Corner: Cleve Moler on Mathematics and Computing" in https://blogs.mathworks.com/cleve/2019/08/26/kuramoto-model-of-synchronized-oscillators/

[14] "Python implementation of the Kuramoto model on graphs" in https://github.com/fabridamicelli/kuramoto_model

[15] O'Keeffe, K.P., Hong, H. Strogatz, S.H. Oscillators that sync and swarm. Nat Commun 8, 1504 (2017). doi: 10.1038/s41467-017-01190-3

[16] K. Ochs, D. Michaelis and J. Roggendorf, "Circuit Synthesis and Electrical Interpretation of Synchronization in the Kuramoto Model," 2019 30th Irish Signals and Systems Conference (ISSC), Maynooth, Ireland, 1-5, 2019, doi: 10.1109/ISSC.2019.8904942.

[17] A. Fettweis, "Wave Digital Filters: Theory and Practice," Proceedings of the IEEE, **74**(2), 270–327, 1986, doi: 10.1109/PROC.1986.13458.

[18] Y. Kuramoto "Self-entrainment of a population of coupled non-linear oscillators," In: Araki H. (eds) International Symposium on Mathematical Problems in Theoretical Physics. Lecture Notes in Physics, **39**. Springer, Berlin, Heidelberg, 1975, doi: 10.1007/BFb0013365.

[19] S.N. Dorogovtsev, J.F.F. Mendes "Evolution of networks," Advances in Physics, **51**(4), 1079-1187, 2002, doi: 10.1080/00018730110112519

[20] F. A. Rodrigues, T.K.DM. Peron, P.Ji, "The Kuramoto model in complex networks," Physics Reports, **610**, 1-98 2016. doi: 10.1016/j.physrep.2015.10.008ted, 2017.

[21] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, "Synchronization in complex networks," Physics Reports, **469**(3), 93-153, 2008, doi: 10.1016/j.physrep.2008.09.002.

[22] ben-Avraham, D.A. Barrát, M. Barthélemy, A. Vespignani, "Dynamical Processes on Complex Networks," Journal Statically Physics, **135,** 773–774, 2009, doi: 10.1007/s10955-009-9761-x

[23] A. Bohn, J. García-Ojalvo, "Synchronization of coupled biological oscillators under spatially heterogeneous environmental forcing," Journal of Theoretical Biology," **250**(1), 37-47, 2008, doi: 10.1016/j.jtbi.2007.09.036.

[24] Y. Kuramoto, D. Battogtokh, "Coexistence of Coherence and Incoherence in Nonlocally Coupled Phase Oscillators," Nonlinear Phenomena in Complex Systems, **5**(4), 380–385, 2002.

[25] P. Kumar, D. Verma, D.: Parmananda, P. "Partially synchronized states in an ensemble of chemo-mechanical oscillators," Physics Letters A. **381**(29), 2337-2343, 2017. doi: 10.1016/j.physleta.2017.05.032.

[26] M.J. Panaggi and D.M. Abrams, "Chimera states: coexistence of coherence and incoherence in networks of coupled oscillators," Nonlinearity, **28**(3), 2015, doi: 10.1088/0951-7715/28/3/R67.

[27] F. Yuan, Injection-locking in mixed-mode signal processing, Springer, New York, 2019.

[28] LTspice, in https://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 320-326 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Designs of Frequency Reconfigurable Planar Bow-tie Antenna Integrated with PIN, varactor diodes and Parasitic Elements

Mabrouki Mariem*, Gharsallah Ali

Physical Department, Faculty of Mathematical, Physical and Natural Sciences of Tunis, Tunis El Manar University, Campus Universities Tunis - El Manar, Tunis, 2092, Tunisia

ARTICLE INFO

Article history:
Received: 11 February, 2021
Accepted: 11 July, 2021
Online: 16 August, 2021

Keywords:
Bow-tie antenna
Frequency- reconfigurable antenna
Multi-band frequency antenna

ABSTRACT

This paper presents the designs and the simulations of proposed structures of electronically frequency reconfigurable planar bow-tie antenna. In the first part, a modified wide band self-complementary bow-tie antenna is designed and implemented. In the second part, varactor and PIN diodes are integrated in top side to adjust electronically the modified structure of bow-tie antenna over multi-band frequency. By adjusting PIN diode between the two states and by tuning the varactor diode inside these two states; the proposed antenna demonstrates two different operational frequencies. In ON state, the antenna covers a narrow frequency bands and in OFF state the antenna demonstrates a wide-band operational frequency.
Furthermore, a new structure of reconfigurable antenna implemented with PIN diode and two hexagonal parasitic elements is developed to realize a multi-band operational frequency band and to cover GPS and GMS bands. Simulated results show a return loss less than -10dB with a gain varied between 0.5 and 3.5dB.

1. Introduction

In the modern communication system, frequency reconfigurable antenna has been achieved an important extension as for as adjusted operated band and low cost.

Indeed, various structures designs are implemented to improve configurability. In [1], mechanical control method is employed to creat a different operated frequency band. Nevertheless, this methode proposals several problems in comparison with the electronical control method which can achieve an important exactitude and a very fast speed. Electronic control method is more commonly used [2]-[4] based to RF switches such as PIN diodes, MEMS or varactor diodes. Those RF switches can adjust the effective length of the antenna which can tune the frequency band of the antenna when the control bias is varied.

A frequency reconfigurable microstrip slot antenna is designed and implemented in [5]. The proposed antenna structure can switch between six different operation frequency bands using five RF PIN diodes. Those switches are integrated in the slot of the antenna to adjust the effective length of the slot. Measured results demonstrate a return loss less than -10dB at the different frequency bands, a gain equal to 4dB and a bidirectional radiation patterns.

In [6], A compact PIFA antenna is realized to cover UHF DVB-H frequency band. The antenna is implemented using three PIN and a varactor diodes. Three metallic strips are integrated between the radiation element and the ground plane via PIN diodes when the largest metallic strips is directly soldered to the ground plane. Moreover, a lateral short circuit is related to the ground plane via varactor. Simulation results demonstrate four configurations with the appeared of different frequency bands in each configuration.

Varactor and PIN RF switches were employed to adjust the frequency of reconfigurable bow-tie antenna operate over (3-6GHz). Measured results show a reflexion coefficient less than -10dB over the different wide operational frequency bands and a stable radiation patterns with a gain varied between 3.21dB and 5.42 dB [7].

In [8], authors proposed a new configuration which consist to integrate PIN diodes over the bow-tie arms to switch the antenna

*Corresponding Author: Mariem Mabrouki et al, Tunis-Tunisia, Mariem.Mabrouki.fst@gmail.com

between Bluetooth, Wimax and Wlan bands. Measurements demonstrated an important results.

In this paper, we propose different structures of frequency reconfigurable antenna which can switch between wideband and narrow band operational frequency. The proposed structures of reconfigurable antenna have been demonstrated a multi-bands operation in two states with acceptable results.

In the second section, we present the implementation, the simulation and the measurements of the proposed structure of SCBT antenna.

In section 3, a PIN and a varactor diodes are added in the top side of the SCBT antenna to realize a frequency reconfigurable antenna which is switched between a simple frequency band and two dual frequency bands in ON state and a wide-band operational frequency in OFF state.

Then, a proposed structure based on PIN diode and two hexagonal parasitic elements is developed. The two parasitic elements are integrated in the front and in the back side of the antenna structure to produce a new frequency bands over (1.49-1.7GHz) in ON state and over (1.77-1.85GHz) in OFF state.

2. Implementation of SCBT antenna

A wide-band performance is required in the modern wireless communication systems. To obtain this goal, several wideband antennas are implemented in the literature such as Bow-tie antenna which demonstrates an important performances over a wideband frequency [9]-[11].

In this section, a proposed structure of the SCBT antenna is designed and realized using FR4 substrate.

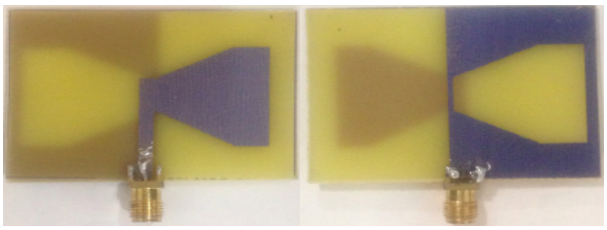


Figure 1: Prototype of proposed bow-tie antenna

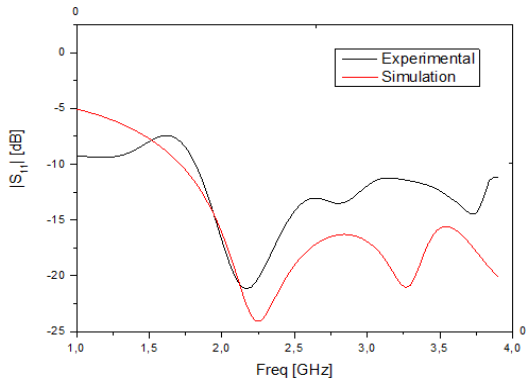
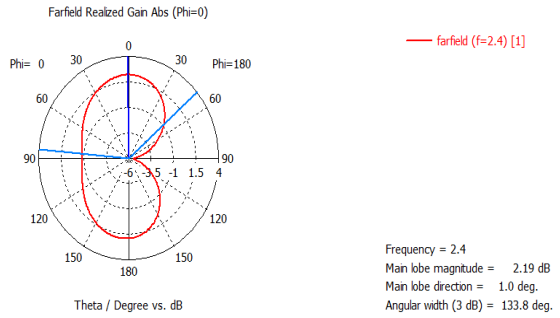


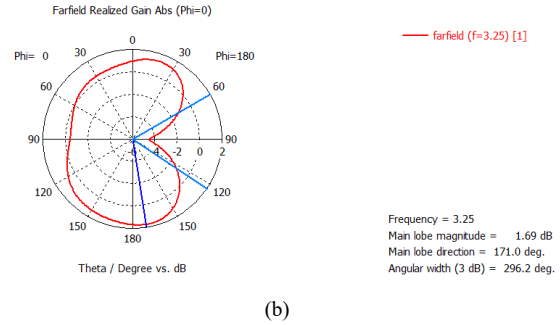
Figure 2: Bow-tie antenna Measurement and simulation

This proposed structure presents good performances such as wideband operation and simple implementation and feeding [12]. The antenna dimensions are calculated using the formulas in [12] and they are optimized via CST: L=54.4mm; W=41.07mm; Wt =2.4mm; Lt =23mm; Wp =25.6mm; Lp =21.5mm; Ls =28.6mm; Wr =0.6mm. Therefore, the size of the modified antenna structure is (50×40 mm2). The top and bottom views of the structure design are shown in figure 1.

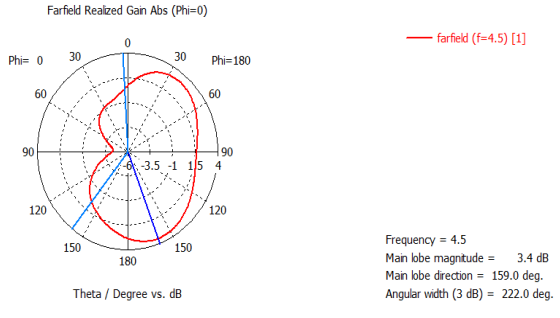
The simulation results are checked through experimental measures figure 2. Simulation return loss is validated with the experimental result and it shows a reflection coefficient less than -10 dB over (2-4GHz). The simulation results of gain patterns are presented in figure.3.



(a)



(b)



(c)

Figure 3: Gain patterns at f= 2.4, 3.25 and 4.5GHz

Three frequencies are chosen to determine the generated gain inside the operational band. Therefore, simulation results of modified SCBT antenna demonstrated a gain equal to 2.19dB, 1.69dB, 3.4dB at 2.4GHz, 3.25GHz and 4.5GHz respectively, with a bidirectional radiation patterns in E plane.

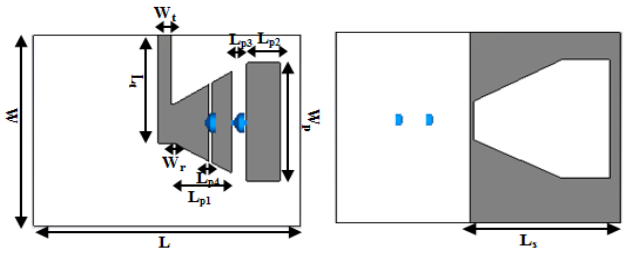
3. Frequency reconfigurable SCBT antenna using PIN and varactor diodes

3.1. Design and Simulation

Varactor and PIN diodes are implemented in the top side of the antenna structure to switch the antenna among different operational frequency bands figure 4. Therefore, the BAR 64-05w

PIN diode is used to control the antenna between the two states. When the ON state is selected, the antenna can achieve a narrow operational frequency band. Otherwise, when the OFF state is selected we can show a wideband operational frequency.

The varactor diode used is MA46585 with a capacitance varied between (0.14 pF to 2.2 pF). By adjusting the reverse bias of the varactor diode between (0V-30V), the varactor capacitances values are improved and the antenna can select the looked-for band frequency. The varactor is modeled in the simulation as a capacitor in series with a forward resistance and inductor. Antenna size is equal to (50x40) mm^2.



(a) (b)

Figure 4: Proposed frequency reconfigurable bow-tie antenna implemented with PIN and varactor diodes: (a) Front view (b) Back view

When the applied Bias control is (V=0.82V), PIN diode is modeled by the resistance R=2.7Ω. Thus, when the diode is tuned ON and by adjusting the varactor capacitances value, different narrow frequency bands are appeared. when the value of the varactor diode capacitance is equal to 0.14pf or 0.19pf, we can show a dual band operational frequency. However, when the capacitance value is fixed to 0.75 pf the antenna is pointed to a simple narrow frequency band.

Otherwise, when the no bias voltage is applied (V=0); the diode is modeled by the capacitor C= 0.22pf. Indeed, three wide bands are obtained when the varactor capacitances value is equal to 0.14pf, 0.19pf and 0.75pf.

The determined simulation results in the two states are presented in figure 5.

When the diode is switched ON and the varactor capacitance is adjusted to 0.75pf, simulation results show a reflection coefficient less than -10dB over a narrow frequency band (1.83-2GHz). Then, a dual band frequency is shown while the varactor capacitance is tuned to 0.19pf. The first band is varied among (2.2-2.5GHz) and the second band occupies the frequencies (3.5-4.2GHz), reflection coefficient in resonance frequencies is equal to -35dB and -18dB at 2.2GHz and 3.6GHz respectively. Two others dual band are achieved when the varactor capacitance is tuned to 0.14pf, the first band is varied between (2.2-2.5GHz) and the second band occupies the frequencies (4.1-4.4GHz). Return loss in resonance frequencies is equal to -60dB and -24dB at 2.1GHz and 4.3GHz respectively. However, when the PIN diode is switched OFF and the varactor capacitance is adjusted to 0.75pf, the reflection coefficient is less than -10dB over (3.5-5GHz). When the varactor capacitance is tuned to 0.19pf the antenna shows a wideband operation with a reflection coefficient less than -10dB. The frequency band is varied among (2.1-5GHz). When the varactor capacitance is adjusted to 0.14pf, the antenna presents wideband frequency operation over (2.1-4.4GHz). Figure 6 proves the gain patterns in ON state with varactor capacitances 0.75 pf, 0.19 pf and 0.14 pf and the gain patterns in OFF state with varactor capacitances 0.75 pf, 0.19 pf and 0.14 pf is illustrated in figure 7.

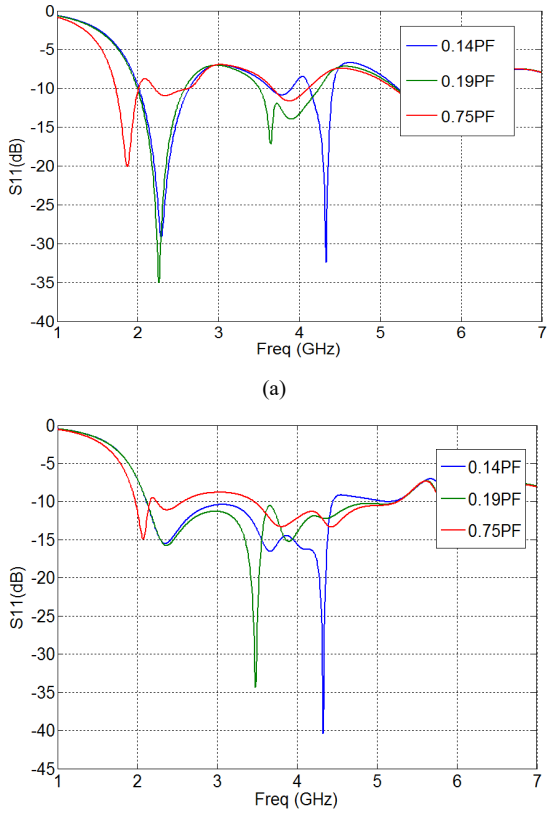
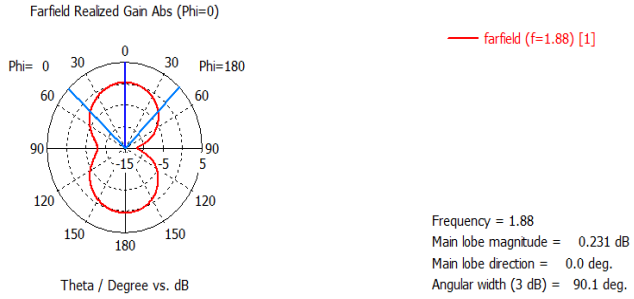
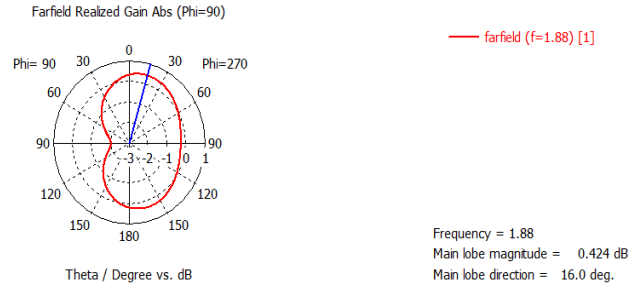


Figure 5: Return loss of reconfigurable antenna with different capacities values: (a) in ON state, (b) in OFF states

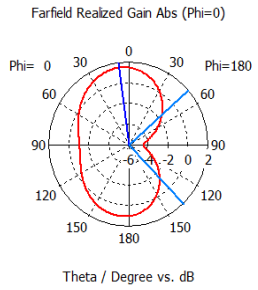


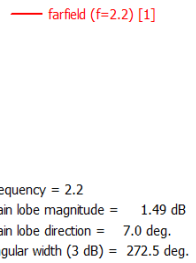
E Plane



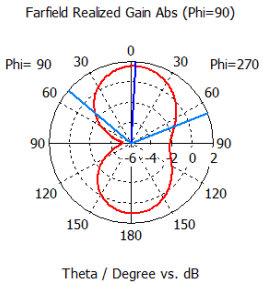
H Plane

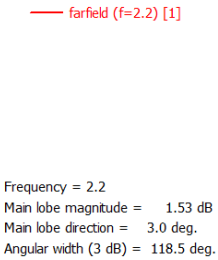
(a)



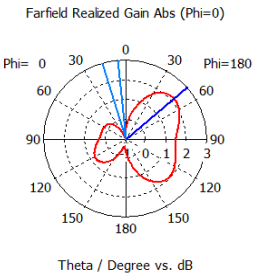


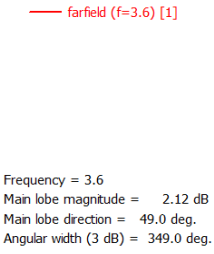
E Plane





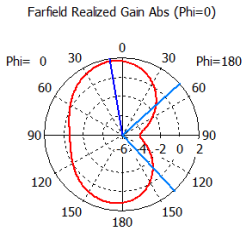
H Plane

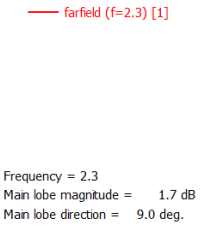




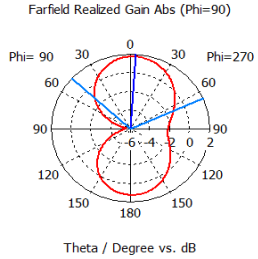
E Plane

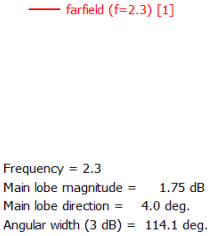
(b)



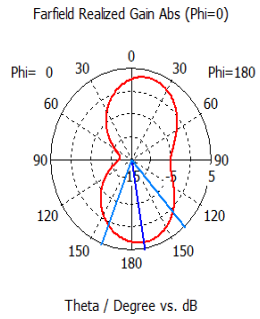


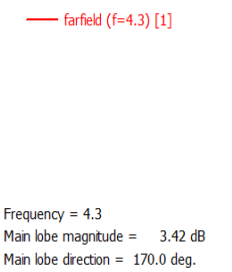
E Plane



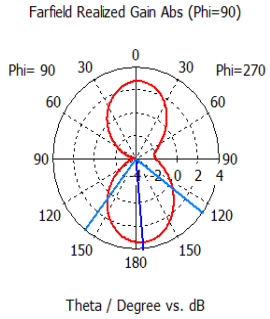


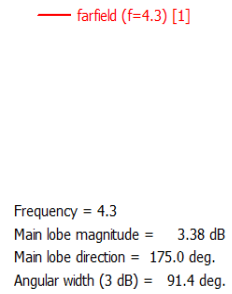
H Plane





E Plane

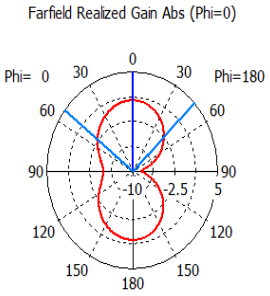


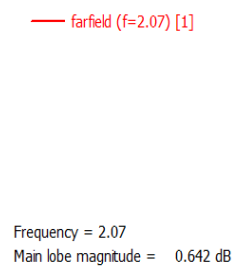


H Plane

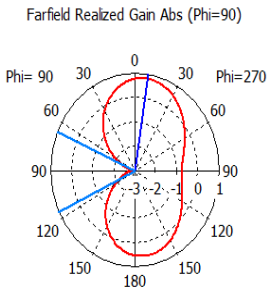
(c)

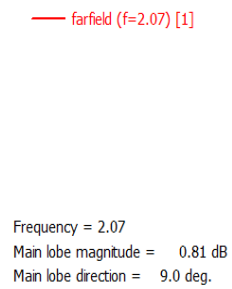
Figure 6: Gain patterns in ON state (a) C=0.75pf, (b) C=0.19pf, (d) C=0.14pf.





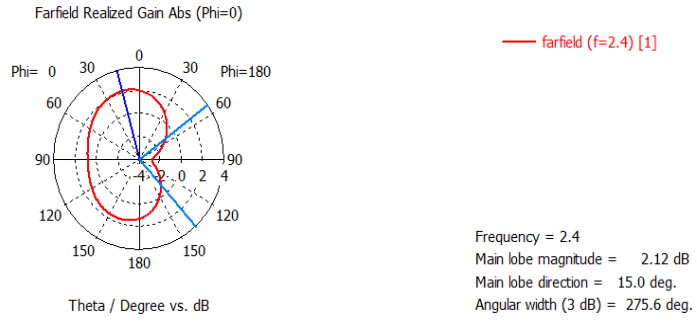
E Plane



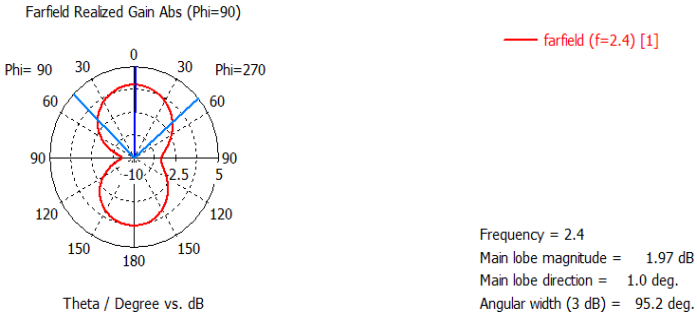


H Plane

(a)

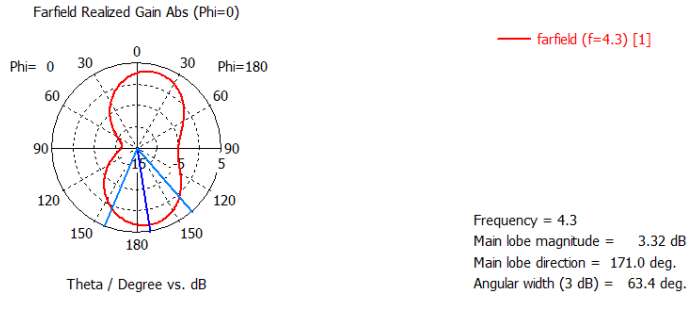


E Plane



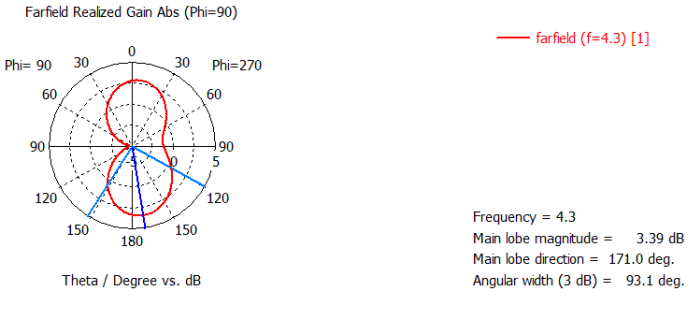
H Plane

(b)



(c)

E Plane



H Plane

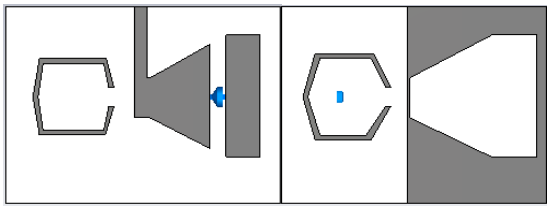
Figure 7: Gain patterns in OFF state (a) C=0.75pf, (b) C=0.19pf and (c) C=0.14pf

Gain patterns results in ON state present a gain equal to 0.231dB at 1.88GHz when the varactor capacitance is on 0.75pf, a gain varied between 1dB and 2.8dB when varactor capacitance is on 0.19pf and a gain around 3.42dB when varactor capacitance is on 0.14pf. In OFF state, the gain is varied between 0.64dB and 3.32 dB over the three wideband frequency. Consequently, bidirectional pattern has been shown and circular polarization has been produced.

4. Frequency reconfigurable SCBT antenna using PIN diode and two parasitic elements

4.1. Design and Simulation

In this section, the idea is to integrate two hexagonal parasitic elements in the front and in the back antenna structure which provide a new operational frequency bands. Using this proposed configuration, GPS and GMS band have been covered. Antenna structure is shown in figure.8.



(a) (b)

Figure 8: Proposed reconfigurable bow-tie antenna integrated with PIN diode and two parasitic elements: (a) Top view (b) Bottom view

We can show Multi-bands frequency operation when the two states are commuted. In figure 9, the reflection coefficients of the proposed antenna are illustrated.

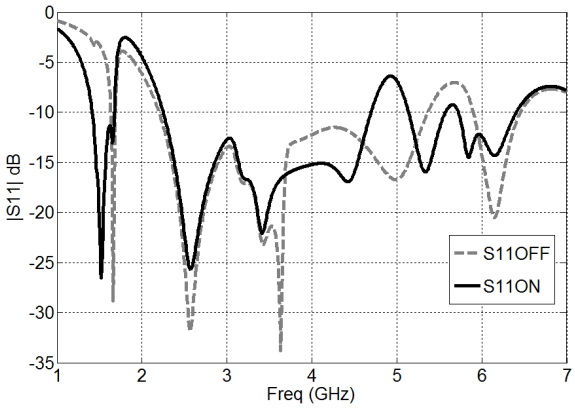
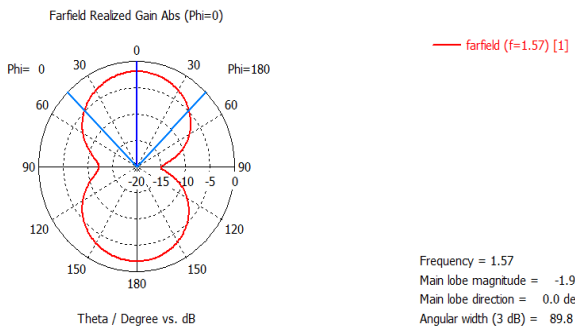


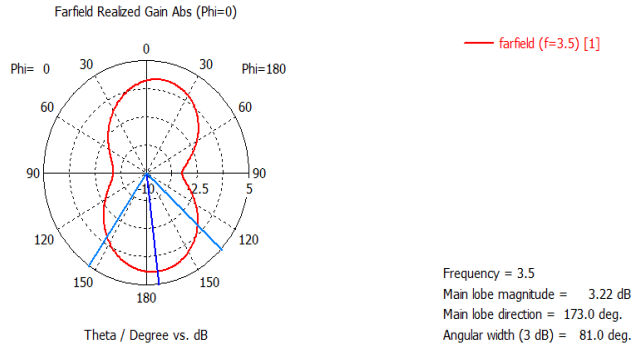
Figure 9: Reflection coefficient results

Simulation results demonstrate two different states with multi-bands operation frequency. The simulated reflection coefficients are less than -10dB at all frequency bands of the two states. In ON state, the improved frequency bands are (1.5-1.75GHz), (2.3-4.61GHz) and (5.15-5.95GHz). At the resonance frequencies 1.58GHz, 3.51GHz and 5.71GHz, the reflection coefficients are -25dB -22,5dB, -34dB respectively. When the OFF state is selected, the obtained frequency bands are (1.77-1.85GHz), (2.24-5.1GHz) and (5.8-6.5GHz) and the reflection coefficients at the resonances frequencies ,1.81 GHz, 2.41 GHz and 5.81 GHz, are -28dB, -32dB and -15dB.

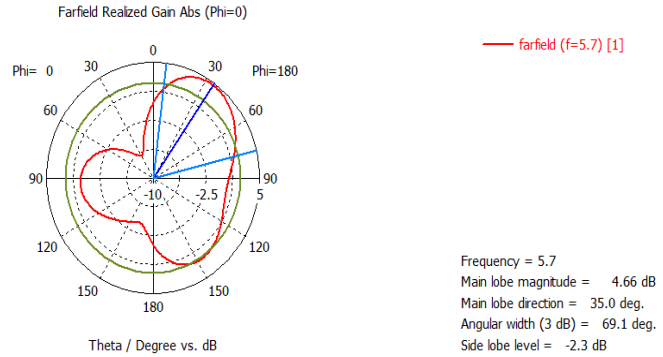
Simulated gain patterns in the two states are presented in figures 10-11.



(a)

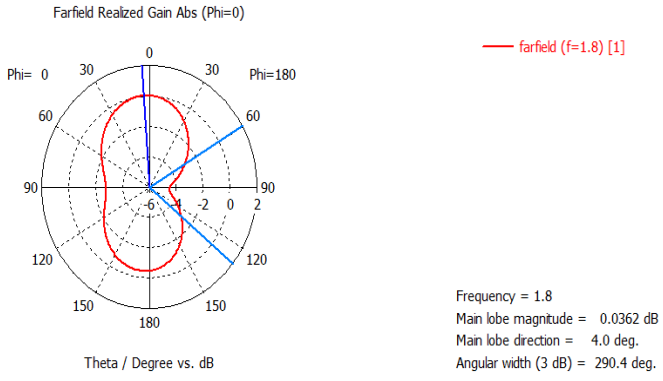


(b)

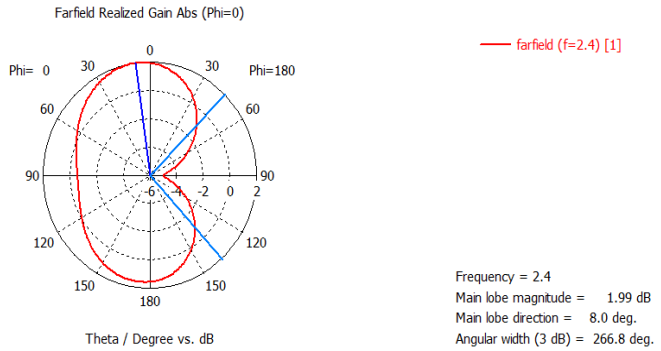


(c)

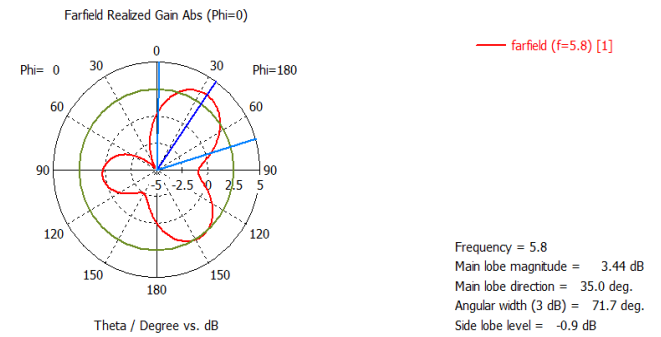
Figure 10: Gain patterns in ON state at 1.58GHz (a), (b) 3.51GHz, (c) 5.71GHz



(a)



(b)



(c)

Figure 11: Gain patterns in OFF state at 1.81GHz (a), (b) 2.41GHz, (c) 5.81GHz

Simulated gain patterns demonstrate the disparity of gain between 2.2dB and 4.7dB and a gain equal to -1.9dB at f=1.58GHz in ON state. when the OFF state is selected, generated gain is restricted to 3.55dB at all operational frequencies and a bidirectional pattern has been shown. Consequently, the proposed antenna provides a circular polarization.

5. Comparison study

Using the parasitic elements, we can show two states with a multi-band operational frequency; the structure is characterized by the simple design and a low control voltage (0v-3v). However, antenna structure using PIN and varactor diodes demonstrates six states which show a simple narrow band, dual-band and wide band operational frequency. The structure requires a two bias voltage with control voltage of (0v- 3v and 15v).

Table 1: Results Table

STATES	Varactor			Parasitic elements
	C1=0.14pf	C2=0.19pf	C3=0.75pf	
ON:	Dual-band [2.2-2.5 GHz] [4.1-4.4 GHz]	Dual-band [2.2-2.5 GHz] [3.5-4.2 GHz]	[1.83-2 GHz]	[1.49-1.7GHz] [2.29-4.6GHz] [5.1-5.9GHz]
OFF:	[2.1-4.4GHz]	[2.1-5GHz]	[3.5-5GHz]	[1.77-1.85GHz] [2.24-5.1GHz] [5.8-6.5GHz]

Reference	Type of switches	Number of switches	Antenna size (mm2)	Number of achieved Bands	Bandwidth of Each Subband
[5]	RF PIN	5	50x46	6 narrow bands	≈ 500MHz
[13]	RF PIN	5	50x100	5 narrow bands	≈ 300MHz
[14]	*PIN &varactor*	2	*30x70*	*6 narrow bands*	*100MHz<BW< 700MHz*
Proposed	*PIN &varactor*	2	*50x40*	*3narrow bands & 3wide bands*	*200MHz<BW< 3GHz*
	PIN& Parasiticelements	1	*50x40*	*2 states with multi-bands in each state*	*200MHz<BW< 3GHz*

6. Conclusion

In this paper, two structures of electronically reconfigurable antenna are designed and simulated. First proposed structure is developed based on PIN and varactor diodes. Thus, three frequency bands are appeared in each state with a good return loss and acceptable gain (1 to 3.5dB). Using this configuration, antenna can commute between a narrow band and a wide band operational frequency. Proposed antenna size is (50×40mm). The second structure of reconfigurable antenna is based on simple PIN diode and two hexagonal parasitical elements implemented in the top and the bottom side of the substrate. The proposed structure can cover GMS and GPS band and it can realize a multi-band operational frequency in each state. Simulation results show an important performance corresponding to the gain and reflection coefficient. Structure size is (50×40mm).

References

[1] Y. Tawk and C.G Christodoulou, "A new reconfigurable antenna design for cognitive radio," IEEE Antenna wireless prog.lett., **8**, 1378-1381, 2009, DOI: 10.1109/LAWP.2009.2039461.

[2] A. C. K. Mak, C. R. Rowell, R. D. Murch and C. L. Mak, "Reconfigurable Multiband Antenna Designs for Wireless Communication Devices, " IEEE Transactions on Antennas and Propagation, **55**(7), 1919-1928, 2007, doi: 10.1109/TAP.2007.895634.

[3] D.Peroulis, K. Sarabandi, and L. P. B. Katehi, "Design of reconfigurable slot antennas," IEEE Trans. Antennas Propag., **53**(2), 645–654, 2005, doi: 10.1109/TAP.2004.841339.

[4] W.H. Weedon, W.J. Payne and G. M. Rebeiz, "MEMS Switched reconfigurable antenna," In Procceeding of the IEEE International Symposuim on Antenna And Propagation, **3**, 654-657, 2001, doi: 10.1109/APS.2001.960181.

[5] H.A. Majid, M. Kamal A. Rahim, M. Rijal Hamid and M. F. Ismail, "A Compact Frequency-Reconfigurable Narrowband Microstrip Slot Antenna," IEEE Antennas And Wireless Propagation Letters, **11**, 616 - 619, 2012, doi: 10.1109/LAWP.2012.2202869 .

[6] F. Canneva. F. Ferrero, J.M. Ribero, R. Staraj, "Reconfigurable miniature antenna for DVB-H standard," IEEE Antennas and Propagation Society International Symposium," 2010, doi: 10.1109/APS.2010.5561955.

[7] T. Li, Huiquing Zhai, Long Li and changhong Liang, "Frequency – Reconfigurable Bow-tie antenna with a wide tuning Range," IEEE Antennas and Wireless Propagation Letters, **13**, 1549 - 1552, 2014, doi: 10.1109/LAWP.2014.2344676.

[8] T. Li, Huiquing Zhai, Long Li and changhong Liang, "Frequency-Reconfigurable Bow-Tie antenna for Bluetooth, Wimax and Wlan application," IEEE Antennas and Wireless Propagation Letters, **14**, 171 - 174 , 2015, DOI: 10.1109/LAWP.2014.2359199.

[9] A.A. Eldek, Atef Z. Elsherbeni, and Charles E. Smith, "Wideband Microstrip-fed Printed Bow-tie Antenna For Phased Array Systems," Microwave And Optical Technology Letters, **43**, 123–126, 2004, doi: 10.1002/mop.20396.

[10] T. Karacolak and Erdem Topsakal , "A Double-Sided Rounded Bow-Tie Antenna (DSRBA) for UWB Communication," IEEE Antennas And Wireless Propagation Letters, **5**, 446 - 449, 2006, doi: 10.1109/LAWP.2006.885013.

[11] K. P. Ray, "Design Aspects of Printed Monopole Antennas for Ultra-Wide Band Applications," International Journal of Antenna and Propagation, 1-8, 2008, doi: 10.1155/2008/713858.

[12] K.H. Sayidmarie, Yasser A. Fadhel, "A Planar Self-Complementary Bow-Tie Antenna for UWB Applications ," Progress In Electromagnetics ResearchC, **35**, 253-267, 2013, doi:10.2528/PIERC12103109

[13] Y. Choi , Ji-Hun Hong , Jong-Myung Woo, "Electrically and Frequency-Tunable Printed Inverted-F Antenna with a Perturbed Parasitic Element," Journal Of Electromagnetic Engineering And Science, **20**(3), 164-168, 2020, doi:10.26866/jees.2020.20.3.164.

[14] J. Lim, Gyu-Tae Back, Young-Il Ko, Chang-Wook Song, and Tae-Yeoul Yun, "A Reconfigurable PIFA Using a Switchable PIN-Diodeand a Fine-Tuning Varactor for USPCS/WCDMAM-WiMAX/WLAN," IEEE Transactions On Antennas And Propagation, **58**(7), 2404 - 2411, 2010, doi: 10.1109/TAP.2010.2048849.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 327-333 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Real Time RSSI Compensation for Precise Distance Calculation using Sensor Fusion for Smart Wearables

Kumar Rahul Tiwari*, Indar Singhal, Alok Mittal

STMicroelectronics, SRA-SAIL, Greater Noida, 201308, India

ARTICLE INFO

Article history:
Received: 14 July, 2021
Accepted: 02 August, 2021
Online: 16 August, 2021

Keywords:
Bluetooth® Low Energy
RSSI
Antenna Orientation
STEVAL-BCN002V1
Inertial Measurement Unit
Yaw Angle
Sensor Fusion

ABSTRACT

To effectively implement the social distancing or digital contact tracing in epidemic using an RSSI-based localization approach through Bluetooth beacon is one of the most widely used technologies, but simply using RSSI measurement is not more relevant because the RF signal is affected by several factors and the environment of usage. Traditional distance or positioning algorithms have large-ranging errors when applied for moving objects because they do not account for the device orientation and use fixed path loss models. Hence, the distance between the nodes cannot be obtained accurately by RSSI measurement in a dynamic environment. In this paper, we propose a solution to compensate for the RSSI loss in real-time by filtering out the noise and then accounting for the antenna orientation using a Beacon Packet. Antenna Orientation is determined using 9DoF (9 Degrees of freedom) IMU (Inertial Measurement Unit). The nodes simultaneously advertise their presence and scan for the presence of other similar beacons in their range. These nodes also deploy Low Power techniques during periods of inactivity to conserve battery power. Advertising is performed on three Bluetooth channels and no connection or response packet is required between the devices during advertising and scanning activities (ADV_NONCONN_IND). The addition of the Motion Sensor could also be used to optimize the battery life of the device.

1. Introduction

This paper is an extension of work originally presented in ICCCS Patna 2020 [1]. Smart Contact tracing with social distancing is one of the efficient ways to avoid the spread of contagious pandemics, such as COVID-19 and future pandemics. In such cases, it is an advantage to alert people to maintain a safe distance. Additionally, it is advantageous to record their physical contact wherever they are at work, in public places, or at a relative's home, etc. RSSI (Received Signal Strength Indicator) based distance calculation between two BLE (Bluetooth® Low Energy) based movable devices (nodes) are widely used technologies. However, RSSI measurements give lower accuracy due to variable attenuation (path loss) and fading effects with high variance [2]. Another factor that affects the RSSI is the antenna orientation and thus can affect the calculated distance between two transceivers. In an ideal scenario, the RSSI of the received signal does not very much. However, in a practical scenario the RSSI is affected by different factors: e.g., physical distance, reflections of objects, environmental parameters, movement of objects or change in the environment, antenna position and polarization etc.

In this paper, we describe the algorithm and later experimentally verified how effectively distance calculation can be done by real-time compensation of the RSSI loss using nested mathematical filters and by accounting for the relative antenna orientation between transmitter and receiver for digital contact tracing or social distancing related applications.

The solution developed is based on ST BlueNRG-2 SoC (Bluetooth Low Energy (BLE) system-on-chip), LSM6DSO and LIS2MDL (9DoF IMU Motion Sensor) for estimating device orientation. The assumption is being taken that orientation of the device is the orientation of antenna embedded in device from a fixed plane. Bluetooth® Low Energy technology is used due to several advantages like wide deployments in wearables and Low-Power consumption. etc. It can communicate with a smartphone which can configure device parameters from the app. This solution will also help to avoid huge dependency on smartphone(s) even for pausing/ resuming the logging of the beacon.

Antenna radiation pattern changes in a different orientation and this impact the RSSI of the signal, impacting parameters like distance based on RSSI. In most of the known solutions or papers,

*Corresponding Author: Kumar Rahul Tiwari, rahul.tiwari@st.com

orientation of antenna in RF devices are static, and relative antenna direction is not accounted.

2. Design Architecture

The concept has been proven using a system solution using the STEVAL-BCN002V1B development kit based on BlueNRG-2 SoC [3] [4]. This development kit hosts multiple sensor which includes gyroscope, magnetometer, accelerometer, Time-of-Flight, humidity, pressure, and microphone sensors. It can be powered by a common CR2032 coin cell. The development kit communicates with a Bluetooth® Low Energy enabled smartphone running the ST BLE Sensor app [5], available on iTunes and Google Play and stores. Following are major components in STEVAL-BCN002V1 as shown in Figure 1:

- BlueNRG-2: Programmable Bluetooth® LE 5.2 Wireless SoC
- LSM6DSO: 3D gyroscope and 3D accelerometer [6]
- LIS2MDL: high-performance, ultra-low power, 3-axis digital magnetic sensor [7]

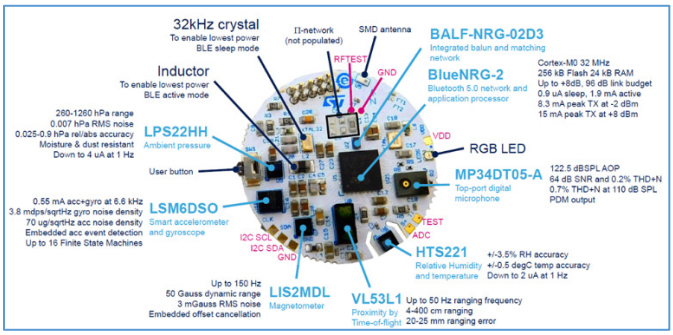


Figure 1: STEVAL-BCN002V1 known as BlueNRG-Tile

The MEMS sensor section of the STEVAL-BCN002V1 sensor node includes inertial and environmental MEMS sensors connected with the BlueNRG-2 via an I2C bus operating at 400 kHz. All sensors can generate interrupts, but only the interrupts from the LIS2MDL magnetometer and the LSM6DSO accelerometer and gyroscope are connected with the BlueNRG-2 through dedicated and independent lines.

BlueNRG-2 integrates a Bluetooth Low Energy radio (BLE), an ARM® Cortex®-M0 core, 24 kB of static RAM memory, 256 kB of Flash memory, SPI (max 1 MHz in slave and 8 MHz in master mode), two I2Cs (standard 100 kHz or fast 400 kHz), UART interfaces; two multi-function timers (MFT), a DMA controller, RTC and watchdog, and an ADC with PDM stream processor.

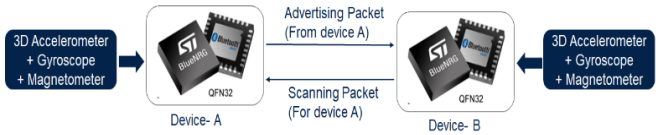


Figure 2: Sequential Advertising and Scanning by Transceiver

Once the BLE stack is initialized in STEVAL-BCN002V1 node, it will advertise non-connectable undirected beacon (ADV_NONCONN_IND) packets and sequentially scans packets from neighbor wearables as shown in Figure 2. RSSI approximation for the relative distance between two or more BLE devices/Nodes can be improved using sensor fusion of Accelerometer, Gyroscope and Magnetometer. The solution is useful in several applications especially in social distancing or digital contact tracing applications.

The impact of the RSSI variation by rotating antenna on same plane can be compensated by sending the Euler angle (Roll, Pitch, Yaw) of antenna obtained from sensor fusion through advertising packet as shown in Figure 3. In the proposed solution, only Yaw axis data (direction of the antenna) is broadcasted.

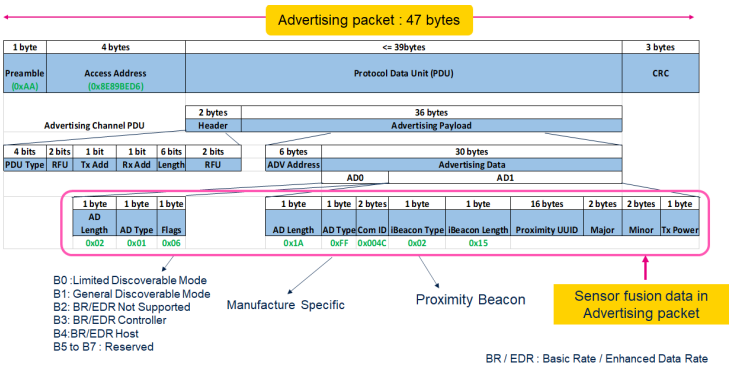


Figure 3: Advertising packet format with sensor fusion data

It is important to know that the radiation pattern is not only a function of the antenna itself, but it depends also on the overall system including PCB layout, ground plane size, space and mechanical surroundings. The Radiation pattern of the antenna as shown in Figure 4 is taken as a reference to observe the change in RSSI with different Euler angle, which is taken on X and Y plane keeping Z axis constant with reference to IMU.

The solution will overcome the drawback in which RSSI variations are huge even if the node is at same position but in a different direction or orientation. The radio frequency section of the STEVAL-BCN002V1 includes the following elements:

- BALF-NRG-02D3 ultra-miniature balun which integrates matching network and harmonics filter.
- A Pi-network which allows additional filtering and provides access points for testing. Note: This network is not populated, as the integrated balun provides the necessary matching.
- An SMD 2.4 GHz (ANT016008LCS2442MA1) antenna, which requires a certain clearance area on the PCB and specific passives for precise tuning.

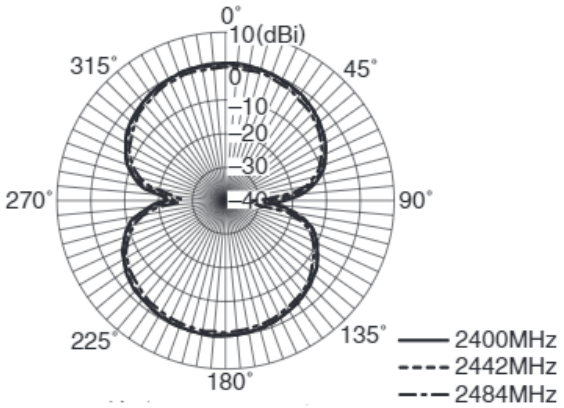


Figure 4: Antenna Radiation Pattern

3. Proximity Sensing and RSSI Compensation Algorithm using Sensor Fusion

Following paragraph explains in detail about the algorithm that is developed for calculating the proximity of two Bluetooth/RF Nodes using RSSI Measurement and compensating RSSI loss in real-time by embedding the dynamic node orientation in advertising packet. Node orientation will be calculated using 9DoF IMU Motion sensor

- Initially sensor fusion with 9 axis (3 axis each for accelerometer, gyroscope and magnetometer) is initialized, and sensor orientation is set, default orientation is ENU (East: X, North : Y, UP : Z)
- The output data rate for gyroscope and accelerometer should be equal to or greater than 100Hz, the magnetometer can be 40Hz
- Magnetometer and Gyroscope calibration is done as shown in Figure-5, Accelerometer calibration is not necessary for sensor fusion except for applications demanding very high orientation precision. Calibration image is captured from ST BLE Sensor App



Figure 5: Node rotation pattern for Calibration

- Roll, Pitch and Yaw as shown in Figure 6 is obtained as an output of sensor fusion. Next, only the Compass angle on the Yaw axis (antenna direction from a fixed point) is calculated using sensor fusion and is embedded in the last 2 bytes of advertising packet to advertise it regularly. Measurement of compass angle depends on the Output data rate of sensors too

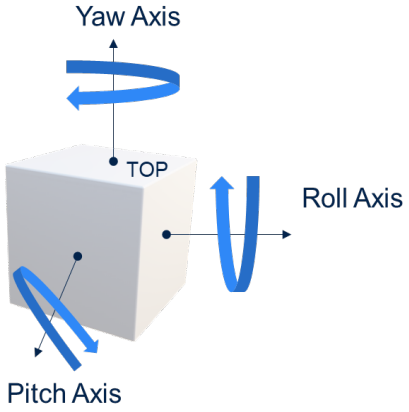


Figure 6: Roll, Pitch and Yaw axis

- Each BLE transceiver advertises and scans sequentially. Tuning the scan window, scan interval and advertising interval parameters can significantly impact battery life. The interval must be an integer multiple of 0.625ms (Time = N * 0.625ms) from 20ms to 10.24s. For ADV_NONCONN_IND (non-connectable undirected event), the minimum advertising interval can be 20ms. There is also a random delay generated by the Link Layer between 0 to 10ms in each advertising packet

The Controller maintains the list of these advertising data and provides the relevant information in one Advertising Report event from multiple devices

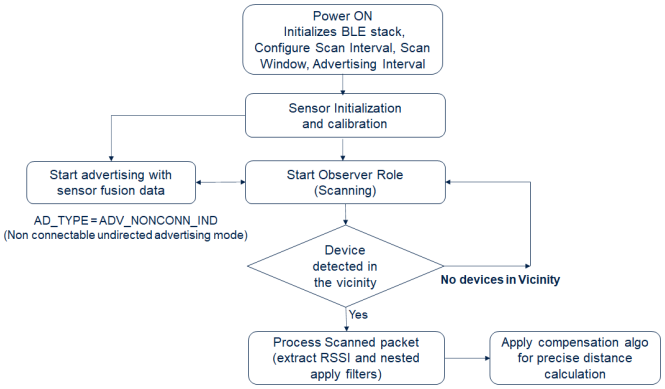


Figure 7: Flow chart of event generation during passive scanning

- RSSI strength is taken for the beacons received during scanning and nested filtering algorithm is implemented to filter out the variations. One of the filters applied in this application is Weighted Mean filter as in equation (1), where 0< αlpha <1. To determine the best, optimize constant (αlpha), several tests were performed with different constant values between 0 to 1. The results and known conditions were analyzed, and best optimized value was achieved between 0.6 to 0.7

$$\text{FILTER_RSSI} = (int8_t)(((1 - alpha) \times aCurrent_RSSI) + (alpha \times aPrevious_RSSI)); \quad (1)$$

- RSSI samples are filtered using nested filtering mechanism then transmit power and filter RSSI values are used to calculate distance [8] by using equation (2):

$$\text{Distance} = 10^{(Measured\ Power - RSSI)/(10 \times N)} \quad (2)$$

Measured Power: Factory-calibrated constant, it indicates average RSSI value at 1 meter

N: Environment dependent constant (or path loss exponent) which is in the range of 2 to 6

The path loss exponent indicates the rate at which the path loss increases with distance. Normally, N is defined between 2 to 6 as shown in Table 1 [9]

Table 1: Path loss exponent for different environments

Environment	Path loss exponent (n)
Free space	2
Urban area cellular radio	2.7 to 3.5

Shadowed urban cellular radio	3 to 5
Inside a building - line-of-sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factory	2 to 3

- During the scanning, the transceiver extracts the yaw or compass angle of the neighbor transceiver and calculates the relative yaw angle
- Slope of RSSI vs Yaw angle is calculated as shown in Figure 12, assumed to be linear. Equation (3) for real-time compensation of RSSI is derived based on the slope and Relative Yaw angle

$$CompensatedRSSI = (int8_t)(Filtered_{RSSI} + (Relative\ yaw \times Slope)) \quad (3)$$

Based on Compensated RSSI value, Compensated distance which incorporates the error due to mismatch antenna orientation and is calculated as in equation (4)

$$\text{Compensated distance} = 10^{(Measured\ Power - RSSI)/(10 \times N)} \quad (4)$$

- The Look-up table is maintained for 2m, 4m, 6m for yaw-angles ranging between 0 to 360 degrees as shown in Figure 8. The look-up table will be taken as a reference to calculate the Variation of RSSI for relative yaw angle from 0 to 90.

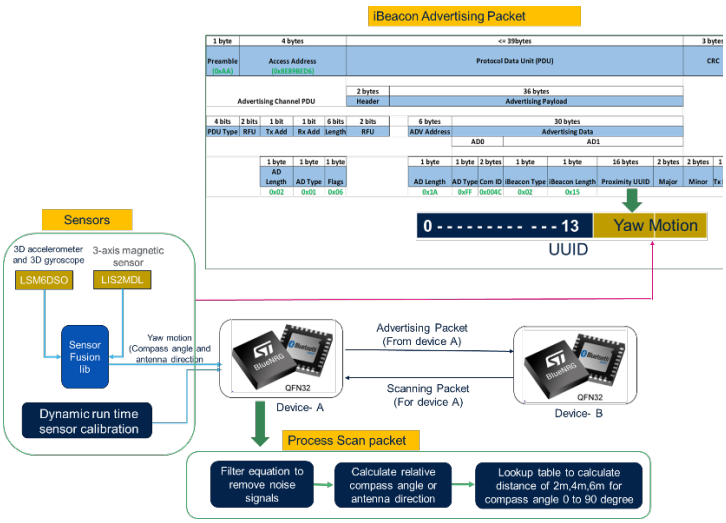


Figure 8: Flow chart of Proposed Solution

4. Adaptive Filtering of RSSI Values

Fluctuation in RSSI is influenced by environmental noise and values are varied with time due to multipath reflections [10]. Variations in RSSI on same distance can be optimized by applying nested filtering algorithms.

4.1. Weighted Mean Filter

Optimized constant value αlpha derived from experiments is multiplied with previous measured RSSI value and 1 – αlpha [11] to the current RSSI value received from same receiver or node. αlpha is for deciding the weightage between previous RSSI and current measurement, αlpha can be in between zero to one. Several tests were performed with different constant values and best results were obtained for αlpha between 0.6 to 0.7

$$Filtered\ RSSI = (1 - \alpha\text{lpha}) \times A + \alpha\text{lpha} \times B \quad (5)$$

A: Current RSSI value

B: Previous RSSI value from same receiver

4.2. Moving Average and Feedback Filter

After applying Weighted Mean filter on RSSI, output values are updated with current values. Feedback is provided to the input and parallelly moving average methodology are applied on the number of samples saved in RAM (Random Access Memory of the SoC), which further filtered the noise. Nested filtering algorithm reduces the variations to a large extent.

5. Power Optimization using Sensor Fusion

Another advantage of using Motion Sensor in Digital Contact Tracing application is to optimize the power for longer battery life. The Wearable will go automatically into low power mode if there is no movement or activity sensed by the motion sensor for certain time. Whenever there is some activity sensed, for example, wearable is worn by a person again; Device will immediately switch from sleep mode to normal mode using the Wake-up event generated by the LSM6DSO sensor. As the wakeup interrupt line is connected with motion sensor interrupt line and the sensor generates an interrupt as soon as the device moves from it's position. Action will be to stop the radio (scanning + advertisement) automatically when the devices are not being worn (for example at home) as shown in Figure 9. Similarly, the radio can be switched On if the devices are being worn.

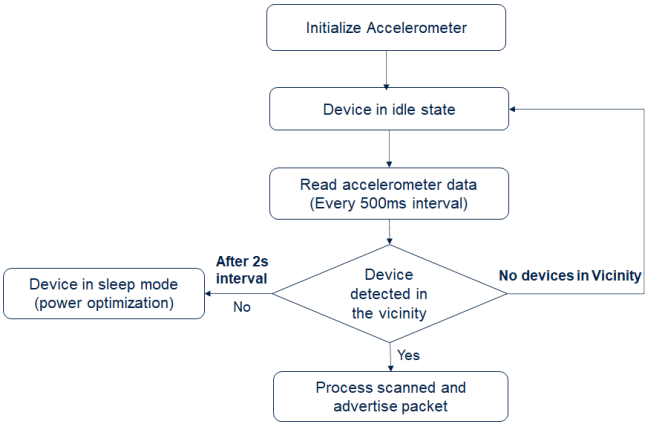


Figure 9: Device Power state in wake up or sleep mode

6. Test Analysis and Results

Various tests analysis performed on multi nodes in the network to observe RSSI variations due to antenna orientation [12]. In this paper, Sensor fusion impact is accounted to compensate RSSI and Distance between two nodes. Node-A is fixed at a center of the circle while Node-B is at 2m distance and rotated in approximately 2m radius as shown in Figure 10. STEVAL-BCN002V1 is taken as nodes shown in Figure 11. At each quarter movement, hundreds of filtered RSSI samples were measured to calculate the mean distance and error variance. With the change in

direction of Node B, it is assumed that antenna direction also changes as the or an antenna is embedded in the Node. The Direction of Antenna is measured in Yaw angle also can be counted as Compass angle. Compass angle measurement is logged as shown in Figure 12.

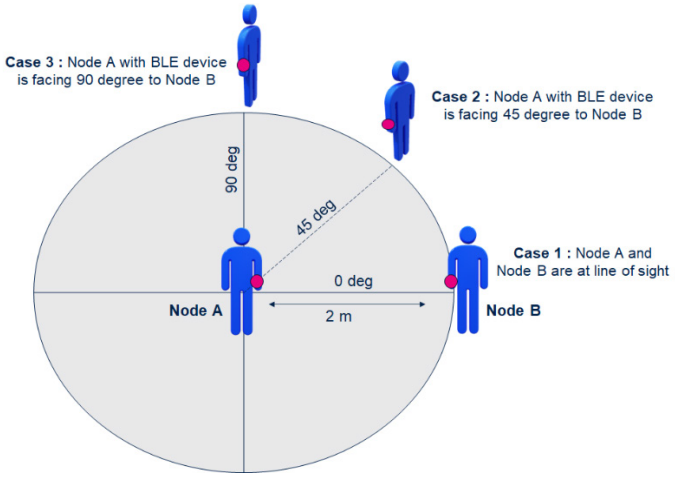


Figure 10: Node B direction with respect to fixed Node A

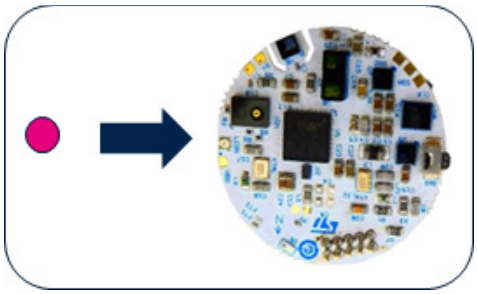


Figure 11: STEVAL-BCN002V1 as Node

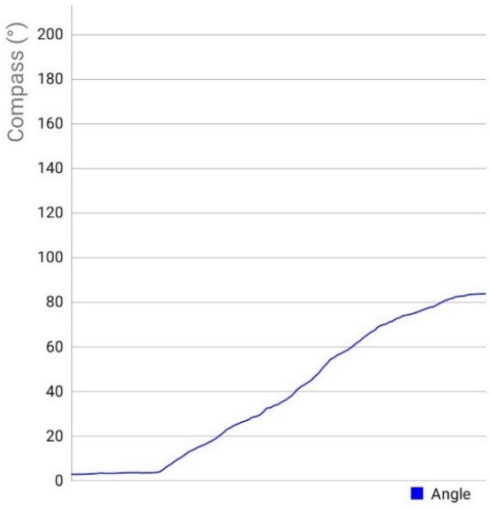


Figure 12: Compass Angle on Yaw axis is measured from 0 to 90 degree

Relative compass angle or Relative Yaw angle is the difference between the direction of two antennas in terms of angle. Even the nodes are at a fixed distance, RSSI strength is weakening as the orientation between two antennas are not aligned as shown in Figure 13.

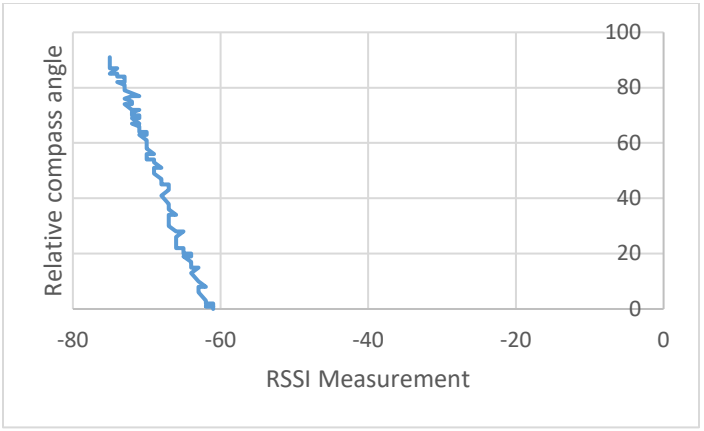


Figure 13: RSSI Measurement vs Relative compass angle on fixed position

Two experiment cases were defined to observe the impact on RSSI measurement and distance calculation, one before accounting Sensor Fusion in advertising packet and one after accounting it in packet. Total 100 samples were taken in each case.

6.1. RSSI Measurement on Node B before applying Compensation Algorithm

As shown in Figure 14 that before applying the compensation algorithm there is a huge impact in RSSI Measurement observed on Node B by just changing the relative antenna orientation or antenna direction between Node B and Node A, which further leads to calculating the wrong distance as per the path loss model, even if the actual distance between Node B and Node A is 1m. Similarly, Table 2 shows that as the Node B antenna direction is moving away from the line of sight of Node A, the calculated distance is increasing accordingly

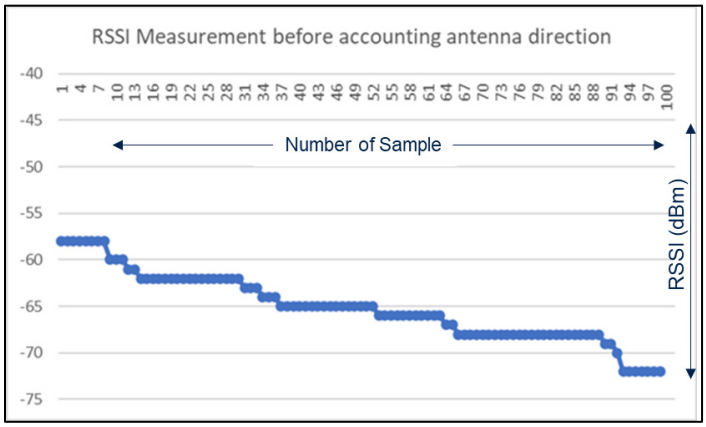


Figure 14: RSSI samples measurement without accounting relative antenna direction

6.2. RSSI Measurement on Node B after applying Compensation Algorithm

In this experiment, the Antenna direction in terms of compass angle is included in the advertising packet of Bluetooth® Low Energy node. Node-B extracts the antenna direction of Node-A to calculate the relative direction as shown in Figure 14, that after applying the compensation algorithm the impact in RSSI measurement observed on Node B by changing the relative

antenna orientation or antenna direction between Node B and Node A is compensated, which further leads to calculating precise distance calculation as per the path loss model. Similarly, Table 3 shows that calculated distance based on RSSI after accounting the effect of Antenna direction is close to 1m only

Table 2: Distance based on RSSI without compensation

Relative Yaw Angle (B - A) in degree	RSSI (without compensation) Observed	Distance based on RSSI (in meter)
0	-58	1.10201845
20	-60	1.33253754
45	-62	1.41827941
90	-64	1.58188643

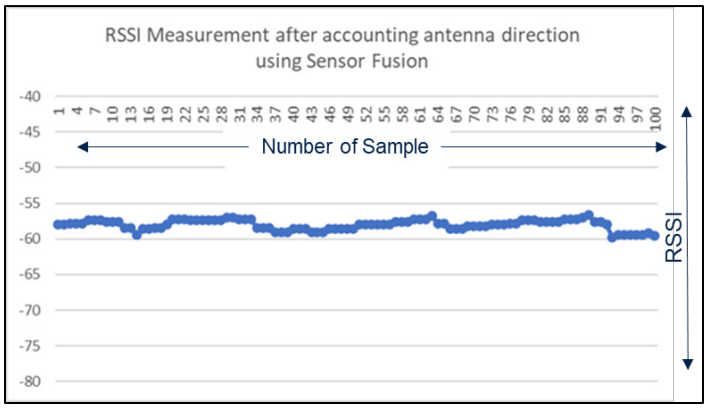


Figure 15: RSSI samples after accounting relative antenna direction

Table 3: Distance based on RSSI after compensation

Relative Yaw Angle (B - A) in degree	Compensated RSSI (dBm)	Compensated Distance (mm)
0	-58	1.10018454
20	-58	1.12301845
45	-57.2	1.07151930
90	-56	1.21618600

7. Current Consumption in Proposed Solution

The current consumption of the BlueNRG-2 can be accurately predicted under different conditions using the BlueNRG consumption tool [13]. As mentioned in paper that BlueNRG-2 SoC (Bluetooth® Low Energy application processor), LSM6DSO and LIS2MDL (9DoF IMU Motion Sensor) are the main component used in STEVAL-BCN002V1 for the proving the concept. Hence active phase and inactive phase current consumption of BlueNRG-2, LSM6DSO (accelerometer and gyroscope), LIS2MDL (magnetometer) is shown in Table 4.

Table 4: Current Consumption

Device	*Active Phase*	*Inactive phase (power not gated by MCU)*
BlueNRG-2	1.9 mA (Active mode)	0.9 µA (Sleep mode)
LSM6DSO	280 µA (50Hz)	3 µA (power-down)
LIS2MDL	475 µA (50Hz)	1.5 µA (power-down)

8. Conclusion

The use of Sensor Fusion for accounting Antenna Orientation in Real-time Location system is quite helpful to obtain the optimal design for digital contact tracing or social distancing related applications. Additionally, sensor fusion helps to optimize battery life by sensing the motion. If the node is stationary, it will direct the controller to disable all the radio activities and in the same way it will generate an event if any motion is sensed to wake up the device from sleep mode. The Antenna direction or orientation of one device is advertised using a Bluetooth beacon at regular intervals to other devices in the vicinity, so that the device while acting as a receiver can estimate the distance using RSSI by also accounting whether the device antenna is perfectly aligned or not. The Solution will be able to estimate the approximate RSSI for a specific position in a dynamic environment. The Solution will also help to avoid snooping of the beacon as the data which is being advertised is changing at regular intervals due to a change in antenna orientation.

Acknowledgment

We would like to thank STMicroelectronics for providing all the resources to accomplish this work.

References

[1] K.R. Tiwari, I. Singhal, A. Mittal, "Smart Social Distancing Solution Using Bluetooth® Low Energy," in 2020 5th International Conference on Computing, Communication and Security (ICCCS), 1–5, 2020, doi:10.1109/ICCCS49678.2020.9277175.
[2] M. Barralet, X. Huang, D. Sharma, "Effects of antenna polarization on RSSI based location identification," in 2009 11th International Conference on Advanced Communication Technology, 260–265, 2009.
[3] STMicroelectronics, "BlueNRG-2, Bluetooth® Low Energy wireless system-on-chip," (December), 2020.
[4] STMicroelectronics, "How to use the BlueNRG-Tile Bluetooth LE enabled sensor node development kit," (June), 1–44, 2020.
[5] STMicroelectronics, "STBLESensor Data brief BLE sensor application for Android and iOS," (May), 2019.
[6] STMicroelectronics, "LSM6DSO Datasheet," (January), 2019.

[7] STMicroelectronics, "LIS2MDIL Datasheet," (November), 1–41, 2018.
[8] F. Awad, A. Omar, M. Naserllah, A. Abu-Hantash, A. Al-Taj, "Access point localization using autonomous mobile robot," in 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), 1–5, 2017, doi:10.1109/AEECT.2017.8257754.
[9] Saverio Grutta, "Application Note : Radio communication range estimation in ISM band," (November), 1–47, 2020.
[10] M. Ayadi, A. Ben Zineb, "Body Shadowing and Furniture Effects for Accuracy Improvement of Indoor Wave Propagation Models," IEEE Transactions on Wireless Communications, **13**(11), 5999–6006, 2014, doi:10.1109/TWC.2014.2339275.
[11] N. Chithirala, B. Natasha, N. Rubini, A. Radhakrishnan, "Weighted Mean Filter for removal of high density Salt and Pepper noise," in ICACCS 2016 - 3rd International Conference on Advanced Computing and Communication Systems: Bringing to the Table, Futuristic Technologies from Arround the Globe, 1–4, 2016, doi:10.1109/ICACCS.2016.7586326.
[12] A. Kishk, "Fundamentals of Antennas," 2009.
[13] STMicroelectronics, "STSW-BNRG001 Data brief BlueNRG current consumption estimation tool STSW-BNRG001," (December), 1–5, 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 339-346 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Analyzing the Application of Two Main Microcontrollers in Engineering Education – A Case Study of three IEEE Conferences Focusing on Education

Arthur James Swart*

Central University of Technology, Bloemfontein, 9301, South Africa

ARTICLE INFO

Article history:
Received: 01 March, 2021
Accepted: 10 May, 2021
Online: 15 June, 2021

Keywords:
Arduino
Raspberry Pi
EDUCON
FIE
TALE

ABSTRACT

Microcontrollers have revolutionized the field of Engineering Education. Their popularity and cost-effectiveness have opened a large door of activity for innovative projects at both the undergraduate and graduate levels. The purpose of this article is to review this activity in terms of where two of these microcontrollers have been used in Engineering Education so as to recommend further possible applications. Focus is limited to papers presented at three IEEE-based conferences over the past 10 years that mentioned the Arduino or Raspberry Pi. Documentary analysis is thus used where the abstracts of the conference papers were reviewed. Results indicate that EDUCON dominated the field of microcontroller education from 2013 to 2016, while the last three years have seen more papers dedicated to this topic being presented at the FIE series of conferences. A total of nine papers relating to microcontroller education has also been presented at TALE between 2012 and 2019. The main application of these microcontrollers has been in the field of Robotics, with general electronics and design-based learning following suit. At least 11 conference papers focusing on the use of these microcontrollers at school level were found. Overall, the Arduino outranks the Raspberry Pi by almost 4:1, with the most cited papers relating to Robotics education, to helping students at home to complete science and technology experiments and to programming. Further applications can extend to energy monitoring and academic development workshops.

1. Introduction

"When I was playing the game we never had the benefit of TV or video to analyze our techniques or look at faults, we depended on other cricketers to watch us and then tell us what they thought we were doing wrong." These words by a British athlete and cricket commentator, Sir Geoffrey Boycott, well highlight the importance of analyzing a technique in order to identify any possible faults or opportunities for further improvement. Of course, this would necessitate an adjustment of a current action in order to improve the given technique that can lead to greater success. Similarly, when it comes to Engineering Education, it is good to analyze what has been published over a period of time with regard to specific topics in order to identify any trends or opportunities for further research.

The surfacing of trends from data collections, such as user-generated content streams and news articles, is a popular and important data analysis activity, used in applications such as business intelligence, quantitative stock trading and social media exploration [1]. This has been extended to many different applications and fields of study, including Education. Identified trends in a specific field can be used to create awareness of what type of research has dominated a given topic over time and in formulating further recommendations with regard to that topic. For example, consider research on the trends relating to microcontroller education. A study noted the widespread dissemination of the Do-It-Yourself (DIY) culture and how modern-day microcontrollers have helped to fuel this culture [2]. Based on the trends identified in that study, a recommendation was made to use a multi-platform method using 8-bit microcontrollers and embedded C programming to maintain the effectiveness of microcontroller education. However, that study did not indicate how the number of publications relating to microcontroller education had increased over time, neither were the specific applications listed. This would help to identify further opportunities for the future application of microcontroller education.

*Corresponding Author: Arthur James Swart, Email: drjamesswart@gmail.com

Previous research has contrasted conference papers, journal articles and student dissertations to obtain bibliometric data [3-5]. However, the novelty of this study lies in its purpose, which is to contrast papers from three international IEEE conferences with regard to the application of microcontrollers in Engineering Education, so as to identify possible further applications. The use of citation analysis in this study is not for purposes of bibliometric analysis, but rather to determine which applications have attracted the most interest from fellow researchers around the globe. The article starts with a discussion on the rise of microcontrollers in Education. The context is then given along with the methodology. Results and conclusions follow at the end.

2. The rise of microcontrollers in Education

Microcontrollers were first considered at Intel in 1969 when a Japanese company approached Intel to build some integrated circuits for calculators [6]. Their history thus spans half a century, yet their impact has been extraordinary over the past two decades. They have indeed revolutionized electronic data acquisition systems [7], which is one of their primary applications.

Microcontrollers are used in many applications that require data acquisition, including industrial automation, control, measuring and consumer electronics [8]. They can be used in everything from relatively simple systems such as washing machines and vacuum cleaners, to much more advanced systems such as cars and robots [9]. In fact, they have become ubiquitous in our daily lives [10]. Its widespread adoption has necessitated changes to educational curricula, as institutions of higher learning move to make their programmes more relevant to the ever-changing technological landscape. These institutions seek to gain a competitive edge in the market of student recruitment, as they claim that their graduates will be better prepared to meet the demands of industry in the Fourth Industrial Revolution. Indeed, this revolution has been called "disruptive", as many changes have followed in its wake [11]. These changes led to the enhancement of new competence requirements (often termed "graduate attributes") and foster adjustments and the updating of qualifications. It also leads to the emergence and extinction of qualifications.

These changes are also evident in many school programmes, as governments seek to strengthen Science, Technology, Engineering and Mathematics (STEM) education among its youth. For example, in South Africa, President Cyril Ramaphosa has indicated in 2019 that government will introduce Coding and Robotics in primary schools (first years of schooling termed Grades R to 3) with a plan to implement it fully by 2022. This is one of the steps taken by government to improve education in the country. Indeed, in many western countries, Robotics is becoming increasingly common in primary and secondary education, both as a specific discipline and a tool to make STEM subjects more appealing to children [12].

3. Context of this study

The context of this study is limited to conference papers presented at three IEEE conferences that focused on Engineering Education. These conferences are [13]:

- EDUCON: This Global Engineering Education Conference provides a forum for academic, research and industrial collaboration on global Engineering Education. It is the flagship conference of the IEEE Education Society for Europe, the Middle East and North Africa (IEEE Region 8).
- FIE: Frontiers in Education has become the premiere conference for innovative curricula, Engineering Education and research excellence. This conference is usually held in various locations around the United States (IEEE Regions 1-6) and, occasionally, in international venues.
- TALE: This International Conference on Engineering, Technology and Education is the IEEE Education Society's flagship in the Asia-Pacific region (IEEE Region 10). It caters for researchers and practitioners with an interest in engineering, technology, and integrated STEM education, as well as those interested in the innovative use of digital technologies for learning, teaching, and assessment in any discipline.

The reason for focusing only on these conferences is the fact that they are rated by Microsoft Academic [14] under the top ten Engineering Education conferences in the world, and they focus on all fields of Engineering. Many of the top ten conferences on the Microsoft Academic list are field specific, such as Instrumentation, Cybernetics, Data Mining, Computer Science and Collaborative Learning.

The topic of interest is limited to the use of two microcontrollers, namely the Arduino and Raspberry Pi. The main reason for this is due to the dominance of conference papers on these two microcontrollers at the three conferences. Figure 1 highlights the total number of conference papers per year that included the word "microcontroller" for these three conferences, as found on the IEEE Xplore website.

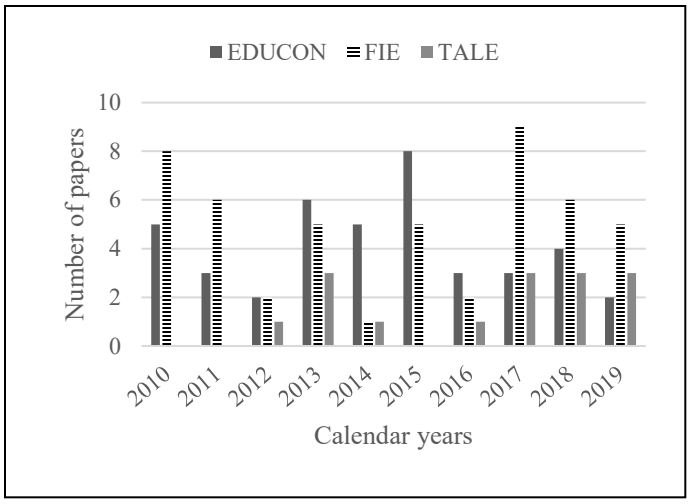


Figure 1: Conference papers using the word "microcontroller" from 2010 to 2019

More than 50% of these conference papers were related to the Arduino and Raspberry Pi. A secondary reason relates to the availability of communities of developers in which the participants share codes and experiences in projects relating to these platforms [14], which would make them easier to implement in Engineering Education.

It is noteworthy that EDUCON dominated the period of 2013 to 2016, while FIE dominated the last three years relating to the number of papers presented that mentioned the words "Arduino" and "Raspberry Pi". This fluctuation can negate forecasting in terms of which conference will dominate microcontroller education over the coming years. However, this "perceived" competition does bode well for the growth of microcontroller Education, especially within other fields of study or in other applications.

4. Methodology used

Documentary analysis were used where the abstracts of papers from the three conferences were reviewed. Documentary analysis is an important source of information, because it is non-relative, meaning that the information found in documents will remain the same over time [15]. These abstracts are retrieved from the IEEE Xplore database by limiting the search to three IEEE-based conferences, and by then searching within the results. For example, EDUCON is inserted into the search bar available on the home page. Conferences are then applied on the next page. Search within results is then used to first search for the word "Arduino" and then for the words "Raspberry Pi". The same procedure is then repeated for FIE and TALE. The search is limited to the time period of 2010 to 2019, as many conferences were impacted negatively in 2020 due to the global COVID-19 pandemic.

Abstracts are downloaded into a MS EXCEL sheet, where they are reviewed in terms of the field of application, number of authors, level of education and number of citations obtained since the publication. The citations can help to evaluate the impact of the publications [16], and also point to fields of application that are important to other researchers in Engineering Education. A large number of co-authors may indicate that more cross-disciplinary research is taking place. Within science, many studies have shown that papers with more co-authors tend to attract more citations [17].

5. Results

5.1. EDUCON results

Figure 2 contrasts the level of education where the microcontrollers have been applied. Arduino has dominated the discourse on the use of microcontrollers in Engineering Education as presented at EDUCON. A total of 15 papers were dedicated to it, with eight applied in Higher Education. It is encouraging to see that it has also been applied to STEM education at school level, where three papers were found in this regard. The Raspberry Pi has only been reported on in five papers relating to Higher Education. The generic column indicates that three papers reported on different levels of education where the Arduino has been implemented. Figure 3 indicates the various applications for the two microcontrollers that were presented since 2010 at EDUCON.

The main application relates to Robotics when considering the Arduino in Figure 3. This is not surprising, as Robotics research has increased at a rapid pace around the globe. More industries are searching for graduates with Robotics experience, and institutions of Higher Education are introducing new curriculums that include Robotics education. In fact, this is even being introduced at school level in South Africa, as mentioned earlier.

Electronics (A/D) feature second on the list of applications. Both basic analogue and digital electronics are covered in this number (three), indicating that the Arduino microcontroller has also been used to convey fundamental principles to Engineering students. Some principles may seem to be abstract to certain students. Seeing the "theory" in "practice", or in action, does indeed help students to grasp and retain key principles. This may also be applied to the application of design (two papers dedicated to this), where students need to understand various concepts that need to be synthesized to obtain a working model.

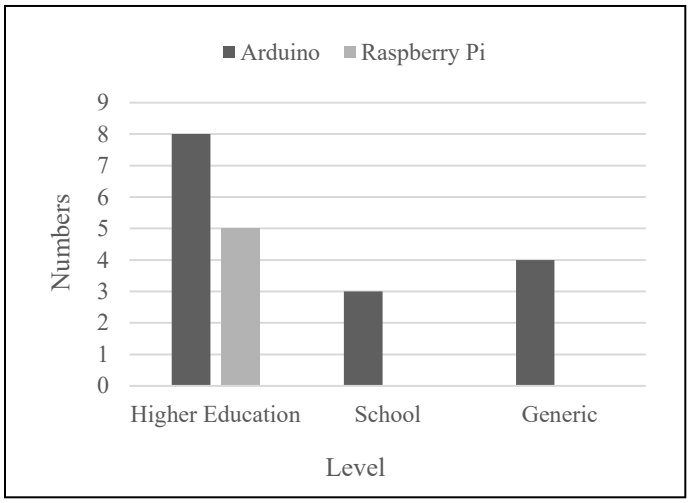


Figure 2: Level of education noted for papers presented at EDUCON

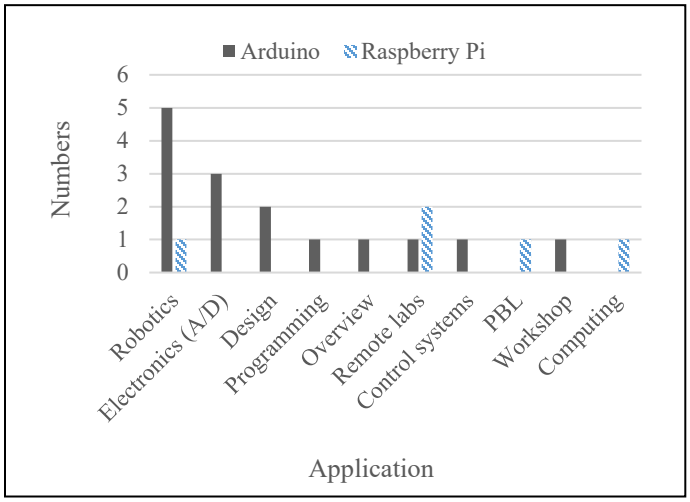


Figure 3: Applications discerned in EDUCON papers from 2010 onwards

The application "Design" covers the creation of an electronic circuit to fulfill specific requirements. "Programming" involves learning a new computer language, such as C++, while "Remote labs" involve online work were students can access an online remote laboratory from their place of residence to complete practical experiments. "Control systems" or "Control" require students to control electrical or mechanical equipment by using a microcontroller, while "PBL" (problem-based learning) requires students to make use of microcontrollers to solve an engineering-related problem. The "Workshop" application indicates a short course in which academics or students are trained in the basic use of a microcontroller, while "Computing" calls on students to

implement a microcontroller to perform mathematical or procedural tasks. The rest of the applications featured in only one paper each, with the dominant Raspberry Pi application being remote laboratories. This is expected, as the Raspberry Pi is a low-cost single-board computer that can function as a web server.

Figure 4 shows the number of authors for the 15 conference papers relating to Arduino as presented at EDUCON over the past ten years, along with their citation count. Figure 5 provides this same analysis for the Raspberry Pi.

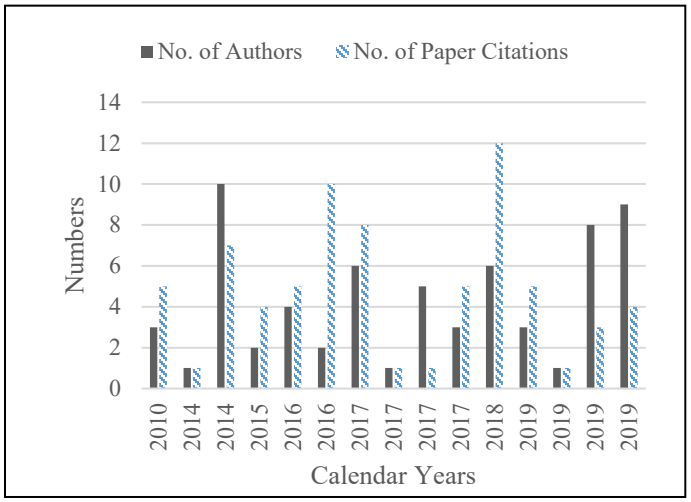


Figure 4: Number of authors versus their citations for EDUCON papers relating to Arduino as determined on 1 November 2020

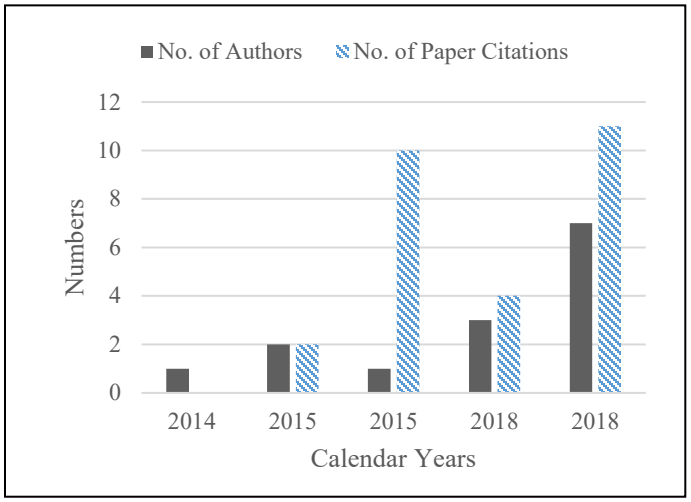


Figure 5: Number of authors versus their citations for EDUCON papers relating to the Raspberry Pi as determined on 1 November 2020

Two papers were presented in 2014 (see Figure 4), and four papers in 2019 (2 x 2014 and 4 x 2019). The maximum number of authors on a paper occurred in 2014, which was ten. That specific paper applied the Arduino microcontroller to the field of Robotics at school level. Two papers that have attracted the highest citation account were presented in 2016 (it focused on the design of electronic projects in Higher Education) and in 2018 (it focused on the use of Arduino in Robotics education at school level). Two papers were presented in 2015 (see Figure 5) and another two in 2018. The paper with the highest citation count was published in 2018 and related to the field of Robotics education in Higher Education.

5.2. FIE results

Figure 6 contrasts the level of education, where the two microcontrollers have been applied in Engineering Education papers that were presented at FIE. Again, the Arduino dominated the discourse over the past ten years at this international conference, with the majority of papers reporting on its implementation in Higher Education. A higher number of papers have been published at FIE, compared to EDUCON. However, the trend seems to be similar between the two conferences, with the Arduino again dominating in Higher Education. Figure 7 highlights the main applications for the two microcontrollers in Engineering Education.

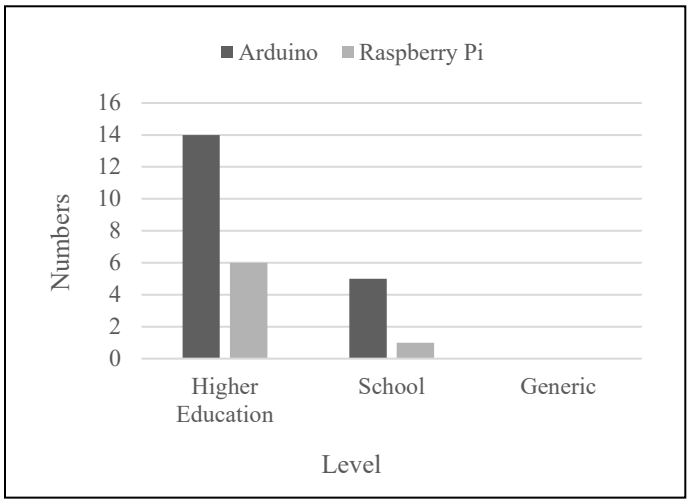


Figure 6: Level of education noted for papers presented at FIE

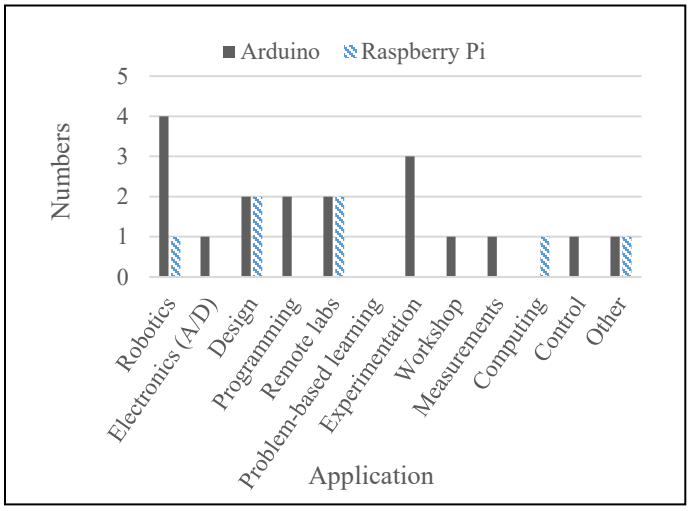


Figure 7: Applications discerned in FIE papers from 2010 onwards

Again, the field of Robotics education tops the list. However, the Arduino microcontroller has also been well applied in the field of experimentation, where two papers are focused on improving Science and Technology education at school level through laboratory experiments. The Raspberry Pi finds equal representation in the fields of design and remote laboratories.

"Experimentation" relates to the use of a microcontroller to complete a basic science experiment, while "Measurements" infer the measuring and recording of different parameters, such as temperature.

Figure 8 presents the number of authors for the papers along with their citation counts. An outstanding result is the citation of 2010. That particular conference paper reported on the use of an Arduino microcontroller to help students at home to complete science and technology experiments. This research would especially be applicable to 2020, where many students were forced to stay home due to the COVID-19 pandemic. Figure 9 provides this same analysis for the Raspberry Pi.

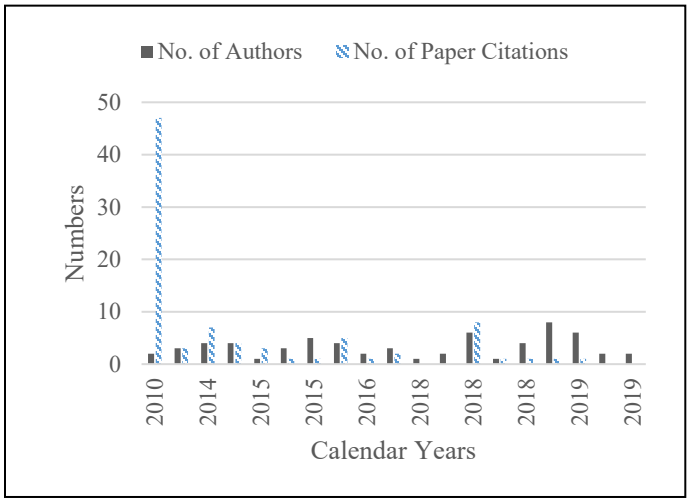


Figure 8: Number of authors versus their citations for FIE papers relating to Arduino as determined on 1 November 2020

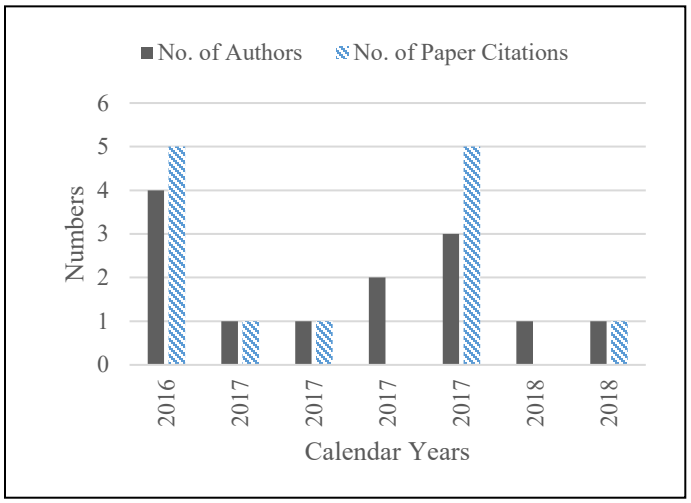


Figure 9: Number of authors versus their citations for FIE papers relating to the Raspberry Pi as determined on 1 November 2020

Four papers were presented in 2017 and another two in 2018 (see Figure 9). Two papers with the highest citation count were published in 2016 (related to the field of Robotics education in Higher Education) and in 2017 (related to remote laboratories).

Five papers were also found that reported on the use of both microcontrollers in the same course or module. All five papers reported on work done in Higher Education, where the fields of application included project-based learning, design, modelling and programming. A paper presented in 2015 has a current citation count of 29, and reported on the use of a range of microcontrollers that were being used in a number of courses relating to project-based learning. Another paper presented in 2016 has a current citation count of 26, and reported on the use of multiple microcontrollers in design-based learning.

5.3. TALE results

Figure 10 contrasts the level of education, where the two microcontrollers have been applied in Engineering Education papers that were presented at TALE. Figure 11 shows the main applications for the two microcontrollers in Engineering Education, which include Robotics education, Design and Control.

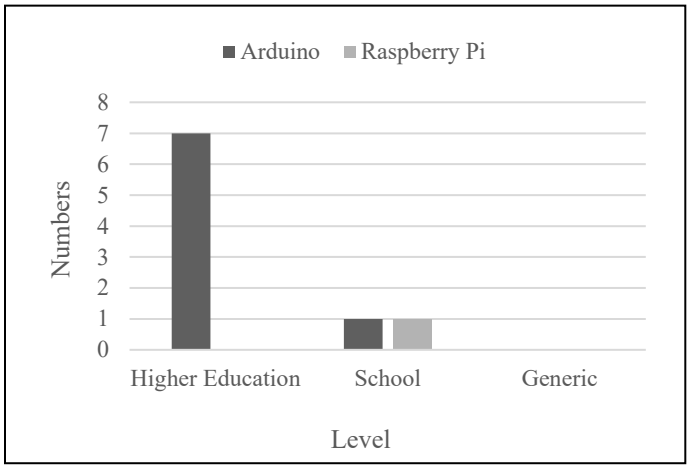


Figure 10: Level of education noted for papers presented at TALE

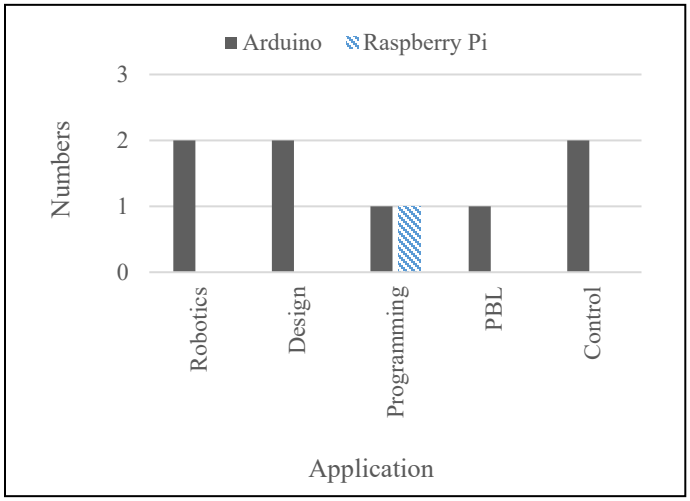


Figure 11: Applications discerned in TALE papers from 2010 onwards

Again, the use of Arduino has been well reported on in Higher Education (see Figure 10), with only two papers dedicated to this at school level for both the Arduino and Raspberry Pi. Nine papers have reported on the use of the Arduino or the Raspberry Pi over the past eight years of the TALE conference that was first held in 2012.

Figure 12 presents the number of authors for the papers along with their citation counts. Two papers were presented in 2014 and

three in 2018. One paper presented in 2014 has a current citation count of 9, and relates to project-based learning in Higher Education. Another paper presented in 2016 currently has an equal citation count, and relates to the use of Arduino in the field of Programming for freshmen Engineering students.

Only one paper relating to the use of the Raspberry Pi in the field of Programming at school level was found. Six authors contributed to this paper that was presented in 2019, and which has no citation count as yet. Table 1 contrasts the key results between the three international conferences.

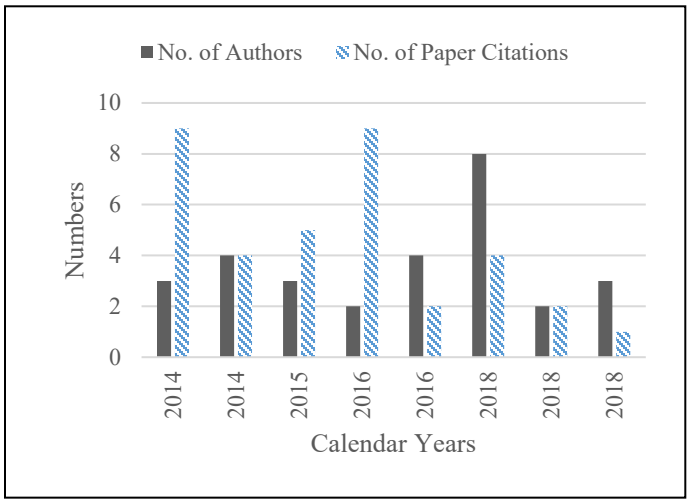


Figure 12: Number of authors versus their citations for TALE papers relating to the Arduino as determined on 1 November 2020

Table 1: Key results

Conference	EDUCON		FIE		TALE	
Microcontroller	Arduino	Rasp-berry Pi	Arduino	Rasp-berry Pi	Arduino	Rasp-berry Pi
Dominant application	Robotics	Remote labs	Robotics	Remote labs	Robotics	Program-ming
No. Papers	15	5	19	7	8	1
Max. No. Authors on a paper	10	7	8	4	8	6
Ave. Authors	4,3	2,8	3,3	1,9	3,6	6,0
Ave. Citations	4,8	5,4	4,5	1,9	4,5	0,0
Dominant microcontroller	3,0		2,7		8,0	

FIE has a higher number of papers (19 + 7 = 26) than EDUCON or TALE. However, the ratio between papers related to the Arduino and to the Raspberry Pi is very similar, being 3:1 for EDUCON and 2.7:1 for FIE. TALE has a much higher ratio, being 8:1. All three conferences had papers with multiple authors, while the average citation count is the highest for EDUCON (4.8 and 5.4 respectively).

6. Possible difficulties in implementing microcontroller education at school/college/university level

Based on the results, it can be seen that more papers have reported on the use of microcontrollers in Higher Education as opposed to school level. This may be due to a number of difficulties or challenges that are experienced at school level.

Firstly, it has been reported that schools have limited funding and resources [18] that would impact on the purchase of different technologies. Partnering with an institution of Higher Education could overcome this challenge, as many universities are mandated to engage in community development where they should make use of their resources to empower and uplift their communities.

Secondly, limited online resources to help learners or students to further their engagement with specific types of technology were noted [19]. Opting for open source software that is freely available to the public, and especially when it comes to programming that is required with microcontrollers, can help institutions or schools to decide on which option they wish to take. Both the Arduino and Raspberry Pi feature open source software. However, the Arduino is aimed at quick programming and circuit prototyping, while the Raspberry Pi acts as a learning tool for computer programming. The Arduino consumes less power than a Raspberry Pi [20], and is a cheaper board that also supports PWM and analog signals [21].

Thirdly, teaching programming is a challenge since the failure and dropout rate can be high in introductory programming courses [22]. Both teachers and learners could struggle to learn a new programming language. However, this can also be mitigated when numerous online resources and programming examples are available from which to learn. Partnering with an institution of Higher Education can also help in this regard, as academics proficient in programming can then volunteer their time to engage with learners at their schools.

7. Opportunities for further application in Engineering Education

Based on the EDUCON results, it can be forecasted that this international conference may see more papers in the future dedicated to using the Arduino microcontroller in Robotics education at school level. A rise in conference papers relating to its application in design-based modules may also be evident in the future. As regards the Raspberry Pi, more remote laboratories can be set up to help students complete practical experiments from their own home. However, further applications can be recommend with regard to the Arduino that can be presented at future conferences.

One of the past papers at FIE reported on the use of Arduino in the field of Measurements [23], in terms of educating students about remote environmental monitoring (e.g. temperature and humidity). However, the field of Energy Monitoring is also growing rapidly, as more countries seek to invest in renewable energy. Institutions should seek to incorporate energy monitoring techniques into their curriculum so as to better prepare graduates for this specific industry. An example of such a technique can relate to the use of an innovative jig to stimulate awareness among Engineering students about the principle of operation of solar panels. An Arduino can be used as part of that jig to help students visualize the results of their experiments on a personal computer.

A second application relates to the remote control of electronic or mechanical equipment. The advent of the Fourth Industrial revolution has enabled the rollout of many different sensors that

can now communicate with each other as part of the Internet of Things. One such application can focus on the remote control of garden plantation pumps using Arduino and mobile communications. Real-time monitoring of specific environmental conditions can help both academics and students to implement similar systems for other applications, such as an aquaponics system.

A third application relates to personal security. Globally, a drive exists to create awareness of gender-based violence and to empower communities to take a stand against domestic violence. Devices can be developed using microcontroller technology that can generate messages to a mobile number of a relative stipulating a victim's location and relevant surrounding information. This type of application may help students to not only think of engineering-related problems, but to also become aware of social problems that could be mitigated by the use of microcontroller technology.

A fourth application relates to workshops for academic staff. A workshop was reported on in a conference paper presented at FIE in 2019. This workshop was presented to school learners in an effort to stimulate awareness of engineering design principles [24]. Similar workshops could be designed to help promote academic development among lecturers at colleges or teachers in schools. This can form part of lifelong learning and enable individuals to remain registered with international bodies as part of their continuous professional development. An example of such a workshop can relate to helping lecturers from a technical and vocational college to improve their technical and programming skills by using the Arduino platform.

8. Conclusions

The purpose of this article was to analyze the application of two main microcontrollers in Engineering Education by reviewing what has been published over the past ten years at three international IEEE conferences that focus primarily on Engineering Education. Results indicate that EDUCON dominated the field of microcontroller education from 2013 to 2016, while more papers were dedicated to this topic being presented at the FIE series of conferences from 2017 to 2019.

Over the past ten years, 26 conference papers were presented at FIE relating to the Arduino and Raspberry Pi. Twenty conference papers were presented at EDUCON, while nine papers were presented at TALE. The main application of these microcontrollers has been in the field of Robotics, with general electronics and design-based learning following suit. At least eleven papers focusing on the use of these microcontrollers at school level were found. Overall, the Arduino outranks the Raspberry Pi by almost 4:1, with the most cited papers relating to Robotics education, to helping students at home to complete science and technology experiments, and to programming.

Further applications can extend to energy monitoring and academic development workshops. The importance of energy monitoring has increased with a global drive to install more renewable energy systems in the world. Continuous professional development is also of key importance, as professionals seek to maintain their registration with official Engineering bodies around the world.

The future of using microcontrollers in Engineering Education is indeed bright, as more academics are seeking to promote student engagement by fusing theory and practice. Only time will tell to what heights academics will rise in their innovative use of these unique microcontrollers, as future analysis will reveal further trends and opportunities in this regard.

Conflict of interest

The authors declare no conflict of interest.

References

[1] N. Kamat, E. Wu, A. Nandi, "TrendQuery: a system for interactive exploration of trends," in Proceedings of the Workshop on Human-In-the-Loop Data Analytics, 1-4, 2016. doi.org/10.1145/2939502.2939514.

[2] D. E. Bolanakis, A. K. Rachioti, E. Glavas, "Nowadays trends in microcontroller education: Do we educate engineers or electronic hobbyists? Recommendation on a multi-platform method and system for lab training activities," in 2017 IEEE Global Engineering Education Conference (EDUCON), 73-77, 2017. doi.org/10.1109/EDUCON.2017.7942826.

[3] A. J. Swart, "Citation Analysis of Master Dissertations at the Central University of Technology, South Africa," AJLAIS, African Journal of Library, Archives and Information Sciences, **29**(2), 115-129, 2019. DOI not available.

[4] S. Van Herck, A. M. Fiscarelli, "Mind the Gap Gender and Computer Science Conferences," in IFIP International Conference on Human Choice and Computers, 232-249, 2018.

[5] P. Sebo, "General internal medicine and family medicine journals: Comparative study of published articles using bibliometric data," Medicine, **99**(24), 1-6, 2020. doi.org/10.1097/MD.0000000000020586.

[6] F. H. Siddig, "Designing and Implementing a Model for a Microcontroller Controlled Pick and Place," Masters in Engineering, Sudan University of Science and Technology, 2014.

[7] A. Busari Sherif, F. Dunmoye Abibat, F. Akingbade Kayode, "Development of Arduino-based Data Acquisition System for Environmental Monitoring using Zigbee Communication Protocol," APTIKOM Journal on Computer Science and Information Technologies, **1**(3), 109-118, 2016. doi.org/10.34306/csit.v1i3.56.

[8] D. R. da S. Medeiros, M. F. Torquato, M. A. Fernandes, "Embedded genetic algorithm for low-power, low-cost, and low-size-memory devices," Engineering Reports, **2**(9), 12231, 2020. doi.org/10.1002/eng2.12231.

[9] O. H. Graven, J. Bjørk, "The use of an Arduino pocket lab to increase motivation in Electrical engineering students for programming," in 2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), 239-243, 2016. doi.org/10.1109/TALE.2016.7851800.

[10] W. Yu, O. Farook, J. P. Agrawal, A. Ahmed, "Board# 63: Teaching Microcontrollers with Emphasis on Control Applications in the Undergraduate Engineering Technology Program," in 2017 ASEE Annual Conference & Exposition2017. doi.org/10.18260/1-2--27895.

[11] G. Spoettl, V. Tūtlys, "Education and Training for the Fourth Industrial Revolution," Jurnal Pendidikan Teknologi dan Kejuruan, **26**(1), 1-6, 2020. doi.org/10.21831/jptk.v26i1.29848.

[12] J. Vega, J. M. Cañas, "PyBoKids: an innovative python-based educational framework using real and simulated arduino robots," Electronics, **8**(8), 899, 2019. doi.org/10.3390/electronics8080899.

[13] V. R. Paragarino, I. F. Silveira, M. Llamas-Nistal, "Open educational resources: A brief vision from IEEE topics," in 2018 IEEE Global Engineering Education Conference (EDUCON), 2076-2081, 2018. doi.org/10.1109/EDUCON.2018.8363495.

[14] J. Á. Ariza, "A proposal for teaching programming languages through open hardware tools," in 2016 IEEE 8th International conference on engineering education (ICEED), 202-207, 2016. doi.org/10.1109/ICEED.2016.7856072.

[15] F. E. A. Batista, G. S. Junior, I. F. Oliveira, "Precariousness, flexibility, and labor relations in the Fashion Industry," Research, Society and Development, **10**(1), e11510110841-e11510110841, 2021.

[16] Y. Xie, L. Ji, B. Zhang, G. Huang, "Evolution of the scientific literature on input–output analysis: A bibliometric analysis of 1990–2017," Sustainability, **10**(9), 3135, 2018. doi.org/10.3390/su10093135.

[17] J. M. Levitt, M. Thelwall, "Long term productivity and collaboration in information science," Scientometrics, **108**(3), 1103-1117, 2016. doi.org/10.1007/s11192-016-2061-8.

[18] S. A. Nadira, L. Shixiang, X. Chen, "SWOT Analysis of the Management and Operation of Secondary Education System in Dhaka City of Bangladesh," International Journal of Science, Technology and Society, **8**(4), 80, 2020. doi.org/10.11648/j.ijsts.20200804.11

[19] N. R. Hussein, Z. S. M Saleem, D. H. Musa, N. Ibrahim, I. A. Naqid, "Impact of COVID-19 on the medical education: experience from Kurdistan region of Iraq," Journal of Medical Education, **19**(1), 1, 2020. doi.org/10.5812/jme.106889.

[20] C. A. Okigbo, A. Seeam, S. P. Guness, X. Bellekens, G. Bekaroo, V. Ramsurrun, "Low cost air quality monitoring: comparing the energy consumption of an arduino against a raspberry Pi based system," in Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications, 1-8, 2020.

[21] E. Lemos, A. Ghoshal, A. Aspat, "Hardware Architecture and Implementation of an AI Pet Robot," International Journal of Applied Sciences and Smart Technologies, **2**(2), 21-46, 2020. doi.org/10.24071/ijasst.v2i2.2792.

[22] Z. Mabni, N. Shamsudin, S. Aliman, R. A. Latif, "Factors Influencing Students' Performance in the First Computer Programming Course Taught Using Blended Learning Approach," Environment-Behaviour Proceedings Journal, 5(SI3), 181-186, 2020. doi.org/10.21834/ebpj.v5iSI3.2559.

[23] D. Călinoiu, R. Ionel, M. Lascu, A. Cioablă, "Arduino and LabVIEW in educational remote monitoring applications," in 2014 IEEE Frontiers in Education Conference (FIE) Proceedings, 1-5, 2014. doi.org/10.1109/FIE.2014.7044027.

[24] A. J. Penn, A. Baynes, "Exploring How Affinity toward Engineering Increases in Underserved Youth After Summer Physical Computing Workshop," in 2019 IEEE Frontiers in Education Conference (FIE), 1-5, 2019. doi.org/10.1109/FIE43999.2019.9028706.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 347-354 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Recognition of Emotion from Emoticon with Text in Microblog Using LSTM

Juyana Islam[1], M. A. H. Akhand[1,*], Md. Ahsan Habib[1], Md Abdus Samad Kamal[2], Nazmul Siddique[3]

[1]Dept. of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh

[2]Graduate School of Science and Technology, Gunma University, Kiryu 376-8515, Japan

[3]School of Computing, Engineering and Intelligent Systems, Ulster University, Northern Ireland, UK

ARTICLE INFO

Article history:
Received: 01 March, 2021
Accepted: 18 May, 2021
Online: 15 June, 2021

Keywords:
Microblog
Emoticon
Emotion Recognition
Long Short-Term Memory

ABSTRACT

With the advent of internet technology and social media, patterns of social communication in daily lives have changed whereby people use different social networking platforms. Microblog is a new platform for sharing opinions by means of emblematic expressions, which has become a resource for research on emotion analysis. Recognition of emotion from microblogs (REM) is an emerging research area in machine learning as the graphical emotional icons, known as emoticons, are becoming widespread with texts in microblogs. Studies hitherto have ignored emoticons for REM, which led to the current study where emoticons are translated into relevant emotional words and a REM method is proposed preserving the semantic relationship between texts and emoticons. The recognition is implemented using a Long-Short-Term Memory (LSTM) for the classification of emotions. The proposed REM method is verified on Twitter data and the recognition performances are compared with existing methods. The higher recognition accuracy unveils the potential of the emoticon-based REM for Microblogs applications.

1. Introduction

Expressions of emotion are fundamental features of intelligent beings, especially humans, that play important roles in social communication [1, 2]. In simple words, emotion represents a person's state of mind exposing into an expression such as happiness, sadness, anger, disgust, and fear. Humans express their emotions in verbal and nonverbal modes, such as speech [3], facial expression [4], body language [5], and expression using text [6]. Emotion has a strong correlation with mental health measured by positive and negative affect and plays a vital role in human social and personal life. With the advent of internet technology, people use various social networking platforms (e.g., Facebook, Twitter, Whatsapp) for social communication. People share their thoughts, feelings, and emotions on different socio-economic, politico-cultural issues using social media. Social media contents appear to emerge as a potential resource for research in human emotional and social behaviors.

Microblog is a common platform to share opinions; hence, it is an important source of emotion analysis of individuals. Microblog makes communication more convenient in our daily life. The most popular microblogs include Twitter [7], Facebook, Instagram, LinkedIn, and Tumblr. A microblog-post can be reached a vast number of audiences within a short time through these platforms. Moreover, a post may reflect one's emotion or sentiment. Thus, sentiment analysis and emotion recognition are two critical tasks from microblog data [8–13]. Depression is the world's fourth major disease, which is deeply related to emotions [14] and often leads to suicidal tendencies. In the United States, suicide is the 10th major cause of death [15]. Social communications and microblog messages may reflect one's mental state. Thus, a potential application of microblog analysis is to take quick necessary actions against deeply depressed people (who might commit suicide) based on his microblog comments.

Analysis of social media contents, especially microblogs, has become very important in different prospects in the present internet era. Tracking and analyzing social media contents are advantageous for understanding public sentiment on any current socio, cultural, and political issue. Researchers have explored different techniques of extracting information from social media data, which have a direct impact on customer services, market research, public issues, and politics. Customer review analysis through such techniques plays an important role in improving the quality of the products and services for retaining customers and attract more [16]. The developed techniques are expected to play

*Corresponding Author: M. A. H. Akhand, Email: akhand@cse.kuet.ac.bd, and Tel.: +8801714087205

an essential role in the study of patients' psychology. Furthermore, emotion analysis is being considered as an emerging research domain for the assessment of mass opinion [17].

Automatic recognition of emotion from microblogs (REM) is a challenging task in the machine learning and computational intelligence domain. There are two main approaches to recognize emotions from microblogs: the knowledge-based approach and the machine learning (ML) approach [18]. In the knowledge-based approach, the task is to develop a set of rules analyzing the given data and then detect emotion using the rules [19]. In the ML approach, a dataset, which consists of patterns based on the features generated from the microblog data, is used to train an ML model, and then the model is used to predict the emotion for unseen data [20]. Typically, ML-based approaches are expected to perform better than knowledge-based approaches [29], [31]. Recently, deep learning (DL)-based methods, which work on the preprocessed data and do not require explicit features, are investigated for REM and found to be promising results [21], [22].

The existing REM methods ignored emoticons and other signs or symbols in the microblogs. These researches only considered texts for the recognition of emotion [23–25]. Nowadays, emoticons, the pictorial representations of facial expressions using characters and related symbols, are commonly used on social media sites. It is found that emoticons are becoming the most important features of online textual languages [26]. Among the few studies that dealt with emoticons is [22] using Convolutional Neural Network (CNN). In the study, words and emoticons from microblogs are processed separately in two different vectors and projected into the emotional space to classify using CNN. Emoticon consideration independent of the text seems not appropriate as emoticons embedded within the text fabricates a semantic or contextual meaning, which is important in emotion analysis. Placement of the emoticon within the text is also important as the different arrangement of emoticon within text may change the meaning. However, emoticon-based REM development is the motivation behind the present study.

This study aims to develop an improved REM method to keep the semantic links between emoticons and the relevant texts. Acknowledging emoticons as particular expressions of emotions, they are represented by suitable emotional words. The original sequence of emoticons in the microblog is unchanged since their sequence may have a vital role in expressing the appropriate emotion. With the necessary prepossessing of microblog data, a machine learning model suitable for examining the sequential or time-series information, known as the Long Short-Term Memory (LSTM), is employed to classify emotions. The recognition performances are compared with the existing method that uses only text expressions (i.e., ignores emoticons) in the recognition process. An initial version of the LSTM-based REM considering emoticons has been presented in a conference [1]; and, the present study is an extended version. The current REM presents the detailed theoretical analysis and experimental results. The higher recognition accuracy of the proposed REM justifies its use in emerging microblog applications.

The rest of the paper is organized as follows. Section 2 presents a brief survey of existing REM methods. Section 3 explains the proposed REM method. Section 4 provides detailed experimental results and analysis. Finally, the conclusion is presented in Section 5.

2. Related Works

Microblog analysis for REM is explored with the rapid growth of social media communication. Several studies were conducted in the last decade for REM from microblogs employing different ML methods, including Naive Bayes (NB) and Support Vector Machine (SVM). The DL-based techniques have also emerged remarkably in the recent REM studies.

Pre-processing of blog tests and distinguishable feature extraction with appropriate techniques are the two important tasks to apply any ML method for REM. Chaffar and Inkpen [18] extracted features from diary-like blog posts (called Aman's Dataset) using bags of words and N-grams. They used decision trees, NB, and SVM to recognize the six fundamental emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) using the features. The SVM is found best among the other classifiers. Silva and Haddela [27] also used Aman's data set and applied the SVM for REM purposes. But they investigated a concept called term weighting to enhance the conventional Term Frequency Inverse Document Frequency (TF-IDF) for feature extraction. Chirawichitchai [28] studied a feature selection technique by information gain and REM by SVM on Thai language blog texts from various social networking sites (e.g., Facebook).

In [29], the authors examined semi-supervised learning with SVM, called distant supervision, for REM from the Chinese tweets in Weibo using a large corpus with 1,027,853 Weibo statuses with emotion labels. Their proposed system predicted happiness emotion most accurately (90% accuracy rate) and worked well for anger. However, the system was less effective for detecting other emotions, e.g., fear, sadness, disgust, and surprise.

In [30], the authors used emoticons in their proposed REM method called the emoticon spaced model (ESM). The ESM learns a sentiment representation of words with the help of emoticons using a heuristic. Words with similar sentiments have similar coordinates in the emoticon space. The coordinates of words are fed into Multinomial naive Bayes (MNB) and SVM for classification. They applied their method on the Chinese microblog benchmark corpus NLP&CC2013 with 14,000 posts with the four most common emotion types (happiness, like, sadness, and disgust).

In [31], the authors performed REM from Twitter's data using NB; in preprocessing stage, they removed URL, special characters, stop-words, and few other things. In [32], the authors extracted features using different methods (e.g., Unigram, Bigram) on the collected 1200 Twitter emotional data and classified emotions using MNB. The large number of features combining Unigram and Bigram is shown to outperformed others with an accuracy of 95.3%.

In [14], the authors adapted emotional-related Chinese microblog (Sina Weibo) data for depression recognition adding "depression" as a new class and excluding the "surprise" class. They developed an emotion feature dictionary with seven types of emotions, namely depression, good, happiness, fear, sadness, disgust, and anger, for depression recognition using 1381 emotional words or phrases. In their study, Multi-kernel SVM is found better than KNN, NB, and standard SVM for depression

recognition from the combination of features from the user profile and user behavior and the features from blog texts.

Among different DL methods for REM, CNN and LSTM are the most well-known ones found in prominent studies recently. In [22], the authors proposed a CNN-based REM, called enhanced CNN (ECNN), that examines both texts and emoticons. Specifically, by placing the emoticons and words in two different vectors and projecting them into one emotional space, CNN is employed to classify emotion. They viewed emoticons as independent of the text, i.e., ignored the emoticon's order in the description. Such consideration might be misleading because emoticon placement or sequence in the text may have a specific meaning. ECNN applied on the Chinese Sina Weibo, NLPCC2013, and Twitter datasets (SEMEVAL). The experimental results on Chinese Sina Weibo, NLPCC2013, and Twitter microblog datasets showed that ECNN outperformed other methods, including SVM, bidirectional LSTM (BiLSTM).

On the other hand, in [21], the authors proposed a hybrid DL model, called Semantic-Emotion Neural Network (SENN), with BiLSTM and CNN for REM. BiLSTM is used to capture contextual information and focuses on the semantic relationship, and CNN is used to extract emotional features and focuses on the emotional connection between words. SENN was applied on Twitter and other social media data, but the use of emoticon is not clear in the decision process.

3. Recognition of Emotion from Microblog (REM) Managing Emoticon with Text

Recently, social media has become a dominant and popular platform for expressing and sharing emotion [3] using microblogs, photos, and videos. Remarkably, the microblog is the hot favorite choice, where one directly writes personal thoughts (e.g., own status, reactions to others, and opinions). Facebook and Twitter are examples of the most popular social media for expressing and communicating such personal thoughts in microblogs. Microblogs contain words, emoticons, hashtags, and various signs with distinct meanings. Since emoticons have become more popular elements besides the text than ever, they should be given proper attention in any microblog-based scheme of emotion recognition.

In this study, emphasis is given to emoticons and their association with texts in microblogs, considering that both are equally valuable to identify proper emotion. Some studies excluded emoticons in the preprocessing step considering those as noisy inputs [33]. But, in this study, emoticons are altered with emotional words and fused with texts for emotion recognition.

Algorithm 1: Recognition of emotion from microblogs (REM)

Input: Microblog M with W Words
Output: Emotion Category $EC \in$ (Happy, Sad, Angry, Love)

// Process 1: Emoticon alteration with textual meaning
For all $w \in W$ **do**
  If $M[w]$ is emoticon ***then***
    $M[w] = Emoticon.meaning\ (M[w])$
  End If
End For

//Process 2: Tokenization with integer encoding
For all $w \in W$ **do**
  $IW[w] = Tokenizer\ (M[w])$
End For

// Process 3: Padding with 0 to render fixed S size
For all $w \in [S - W]$ **do**
  $P[w] = 0$ // Consider 0 padding
End For
For all $w \in [S - W + 1 : S]$ **do**
  $P[w] = IW[w]$ // Copy the rest values
End For

// Process 4: Emotion recognition using LSTM
//Embedding integer to 2D vector
For all $w \in S$ **do**
  $V\ [x, y] = Embedding\ (P[w])$
End For
$EC = LSTM\ (V)$

These interpreted emotional words and other texts presented in the proposed REM help the model perform improved emotion classification.

Figure 1 illustrates the framework of REM proposed in this study for a sample microblog containing an emoticon in the text. The REM consists of four sequential processes. In Process 1, emoticons are converted into relevant emotional words according to a predefined lookup table. The words are transformed into a sequence of integer numbers in Process 2. In Process 3, padding is conducted to form a vector containing the sequence of words with equal length. Finally, in Process 4, the LSTM is employed for classification of emotions into Happy, Sad, Angry, or Love.

Algorithm 1 shows the proposed REM where individual processes are marked. It takes microblog M with W words as input and provide emotion category EC. The whole method is broadly divided into two major parts: processing microblog data using processes 1, 2, and 3, and recognition with the LSTM network. The processes are briefly described in the following subsections.

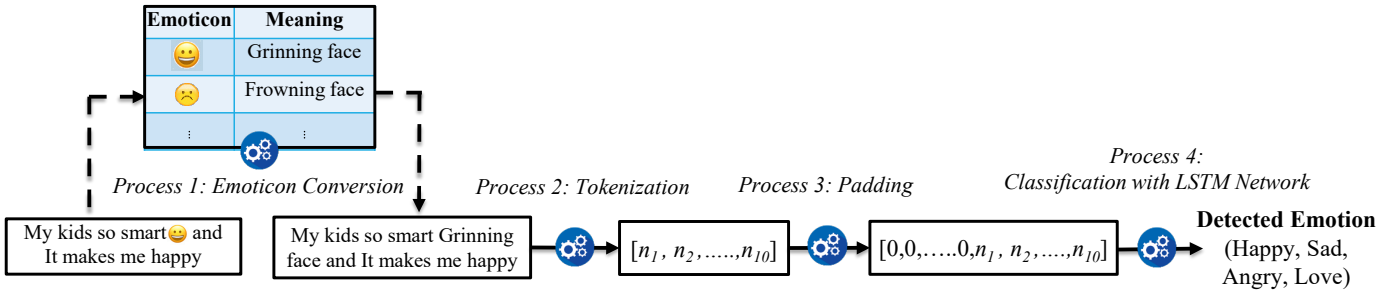


Figure 1: The proposed REM framework illustrating different processes for a sample microblog.

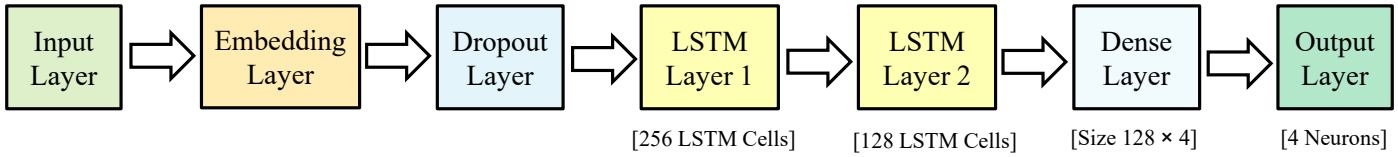


Figure 2: LSTM network architecture used in the proposed REM.

3.1. Microblogs Processing

Twitter is a popular microblog platform and it allows emoticons with texts. Thus, Twitter microblogs are collected and processed for REM in this study. Social media data contains noisy information that needs to be cleaned up to use in the system. Then the clean microblogs with emoticons and texts go through the Processes 1 to 3 (Fig. 1). In Process 1, each emoticon is replaced with corresponding text (i.e., equivalent word for the emoticon) using a function, called *Emoticon.meaning(),* with the help of a lookup table with equivalent words and emoticons. Process 2 is the Tokenization step: it removes unnecessary information and then, generates an integer vector sequence of words (*IW)* through integer encoding. Finally, Process 3 transforms *IW* to a defined fixed length size (say *S*) of words with zero initial paddings. If *IW* contains *W* integer values, the padding outcome *P* vector will contain zeros (i.e., 0) in initial *S*-*L* positions and the rest are *L* values from *IW*.

3.2. Emotion Recognition Using LSTM

The proposed structure of the LSTM network for the REM is shown in Figure 2. The network consists of an input layer, an embedding layer, a dropout layer, two LSTM layers, a dense layer, and finally, the output layer. The input in the LSTM network comprises a sequence of an integer number (defined fixed length *S*) with zero paddings in initial positions. The embedding layer simply transforms each integer word into a particular embedding vector. In the proposed architecture, the sizes of input integer words and embedding vectors are 78 and 128, respectively. Therefore, the output of the embedding for a microblog text is a 78×128 sized 2D vector. A dropout layer is placed just after the input layer, which randomly selects input features during training. The purpose of the dropout layer is to reduce overfitting and improve the generalization of the system.

There are two LSTM layers in the proposed architecture which are the main functional elements of the system. The first and second LSTM layers contain 256 and 128 hidden LSTM cells, respectively. Each LSTM cell in the first layer processes 128 sized embedding vectors and generates single output; therefore, the first LSTM layer produces 256 values which propagate to the input of each LSTM cell of the second layer. The second LSTM layer produces 128 values and the dense layer generates the emotional response from the values. Emotion recognition of this study is a multiclass (i.e., 4-class) classification problem to classify microblogs into four different emotion categories. Thus, it requires the dense layer to be of size 128×4. The output layer has to yield one of the four classes and therefore, the output layer comprises four neurons where each neuron represents a particular emotional state.

An LSTM cell is the heart of the LSTM network architecture illustrated in Fig. 3, which shows the basic building block of an

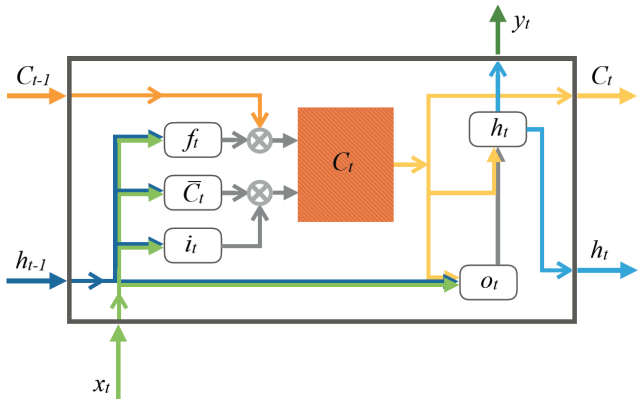


Figure 3: Basic building block of an LSTM cell.

LSTM cell. The LSTM cell consists of a forget gate (*f*), a memory cell (*C*), an input gate (*i*), and an output gate (*o*), At any state *t*, the memory block uses both the current input (x_t) and the previous hidden layer output (h_{t-1}) as inputs and generates new output (h_t) of the hidden layer. This memory block enables the LSTM network in forgetting and memorizing information as required. Hyperbolic tangent or *tanh (*symbol ϕ) and *sigmoidal* (symbol σ) functions are used as the gates. The memory unit calculates the candidate memory $\overline{C_t}$, f_t and input gate i_t at state *t* according to Eqs. (1-3).

$$\overline{C_t} = \phi\,(W_C * h_{t-1} + U_C * x_t + b_C) \quad (1)$$

$$f_t = \sigma\,(W_f * h_{t-1} + U_f * x_t + b_f) \quad (2)$$

$$i_t = \sigma\,(W_i * h_{t-1} + U_i * x_t + b_i) \quad (3)$$

Then, memory C_t is calculated using Eq. (4).

$$C_t = f_t * C_{t-1} + i_t + \overline{C_t} \quad (4)$$

h_t comes from C_t through the output gate o_t following Eqs. (5-6).

$$o_t = \sigma\,(W_o * h_{t-1} + U_o * x_t + b_o) \quad (5)$$

$$h_t = o_i * \phi(C_t) \quad (6)$$

In Eqs. (5-6), *W* and *U* denote the respective shared weights, and *b* denotes the bias vector. Finally, the output of an LSTM cell comes from h_t through the weight vector *V* defined as

$$y_t = V * h_t\,. \quad (7)$$

The LSTM is suitable for modeling complex time-series data since it can classify from any given sequence upon training. Therefore, the LSTM is chosen in the proposed REM to classify the processed microblogs. A detailed description of the LSTM is available in [34].

4. Experimental Studies

This section describes Twitter data preparation, experimental settings and experimental results of this study.

4.1. Dataset Preparation

To make the processed microblog data compatible with LSTM network, tokenization is performed using Keras open-source neural-network library [36] to convert the words to numerical values, and then padded with zeros for a fixed-sized vector. The size of the padded numeral blog was 78, whereas, it was 33 while emoticons were discarded.

Table 1: Emoticons and corresponding word meanings

Sl.	Emoticon	Unicode	Textual Meaning	Emotion
1	😀	U+1F600	Grinning face	Happy
2	😄	U+1F604	Grinning face with smiling eyes	
3	😁	U+1F601	Beaming face with smiling eyes	
4	☺	U+263A	Smiling face	
5	😭	U+1F62D	Loudly crying face	Sad
6	😢	U+1F622	Crying face	
7	🥺	U+1F97A	Pleading face	
8	☹	U+2639	Frowning face	
9	😠	U+1F620	Angry face	Angry
10	😡	U+1F621	Pouting face	
11	😤	U+1F624	Face with steam from nose	
12	🤬	U+1F62C	Face with symbols on mouth	
13	😍	U+1F60D	Smiling face with heart-eyes	Love
14	🥰	U+1F970	Smiling face with hearts	
15	😘	U+1F618	Face blowing a kiss	
16	😚	U+1F617	Kissing face with closed eyes	

Table 2: Samples Twitter microblogs and emotion calabel

Microblog	Emotion Category
Yay! Very well deserved 😀	Happy
I'm so sorry 😭	Sad
😭 I just want the pain to go away	Sad
Apologies to him right now! 😡	Angry
😡 go follow right now	Angry
😍😍😍 Good night love	Love

4.2. Experimental Settings

Adam algorithm [37], a popular optimization algorithm in computer vision and natural language processing applications, is used to train LSTM. Softmax and categorical-cross entropy are considered as activation function and loss function, respectively. The dropout rate of the dropout layer is set to 0.3, and each LSTM layer contains 30% dropout and 20% recurrent dropout while training the model. Batch-wise training is common nowadays and LSTM training was performed for batch sizes 32, 64, and 128 which are commonly used in related studies. Among the collected 16012 tweets, 75% (i.e., 12009) were used to train LSTM, and the remaining 25% (i.e., 4003) were reserved for the test set to check the generalization ability of the system.

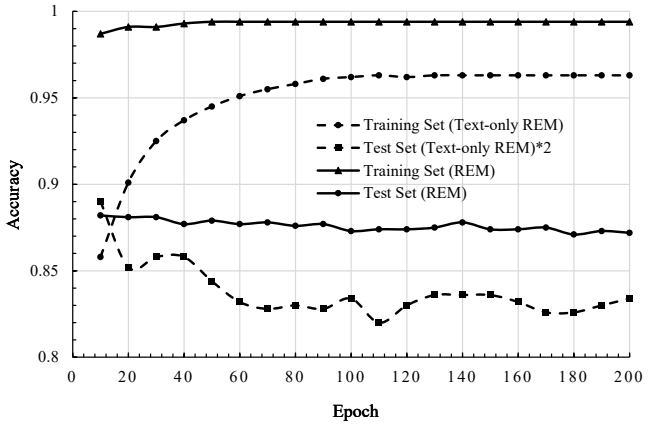
The proposed model was implemented in Python programming language. The experiments were performed using a jupyter notebook as well as Kaggle online environment. A PC with the following configuration is used for conducting the experiment model: HP ProBook 450 G4, processor: Intel(R) Core (TM) i5-5200U, CPU: 2.20 GHz, RAM: 8 GB, OS: Windows 10.

4.3. Experimental Results and Analysis

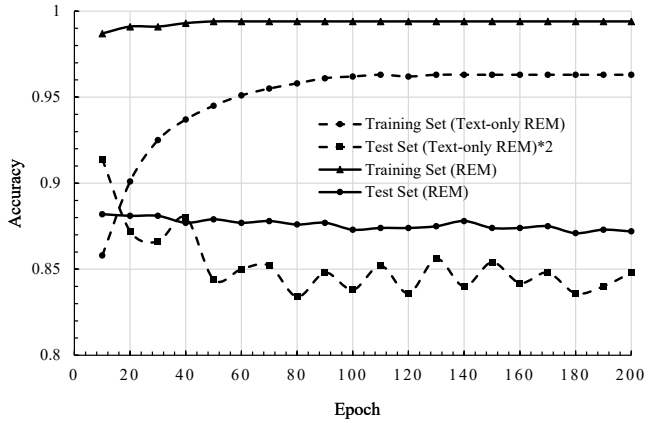
The emoticon consideration with text is the core significance of the proposed REM (with emoticons embedded in texts) from real-life Tweeter data. An experiment discarding emoticons (i.e., using texts only) is also carried out, it may be called REM without emoticon or text-only REM. The outcomes of text-only REM are compared with the proposed REM to observe the effect of emoticon.

In Figure 4, the accuracies of the LSTM for both the training and test sets are evaluated by varying the training epochs up to 200 for the different batch sizes (BSs). It is very clear from the graph that the accuracy of the proposed REM (as shown by the solid curves) is always better than the accuracy of the text-only REM (as shown by the dashed curve). Remarkably, the accuracy with the text-only case is compatible with the proposed REM while training. For example, at 100 epochs for BS=32 (in Fig. 4(a)), the achieved accuracies on the training set of the proposed REM and text-only REM are 0.994 and 0.957, respectively. However, regarding the test data, the accuracy of the proposed REM is much better than that of the method without emoticon (i.e., text-only REM). It is remarkable that the text-only REM test set accuracy is placed in a graph doubling its achieved value to make the graph better visualization. At a glance, the test set accuracy of the proposed REM is almost double that of the text-only REM with any BS values. As an example, at 100 epochs for BS=64 (in Fig. 4(b)), the achieved test set accuracy for proposed REM is 0.873; whereas, the value is only 0.424 for text-only REM. A similar observation is also visible for BS=128 in Fig. 4(c).

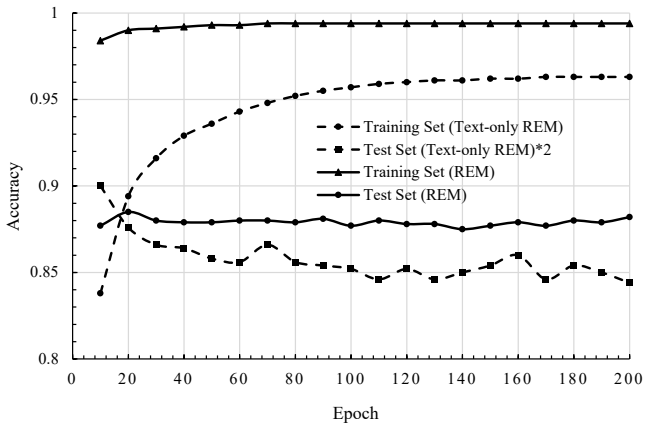
Figure 5 shows the best accuracies from individual experiments with different BS values. The training set accuracy for the text-only REM and the proposed REM are 0.963 and 0.994 for any BS value, respectively. Although performance on the



(a) BS= 32



(b) BS= 64



(c) BS= 128

Figure 4: Recognition accuracies of proposed REM (using both text and emoticon) and text-only REM varying training epochs for three different batch sizes (BSs) 32, 64 and 128.

training set is competitive with and without emoticon, it is remarkable on test set accuracy with emoticons. The best test set accuracies for the text-only REM are 0.445, 0.457, and 0.45 at BS values 32, 64, and 128, respectively. On the other hand, considering both emoticon and text, the proposed REM achieved much better test set accuracies and the values are 0.884, 0.882, and 0.885 at BS values 32, 64, and 128, respectively. Notably, training set accuracy indicates the memorization ability and test set accuracy indicates the generalization ability of a system to

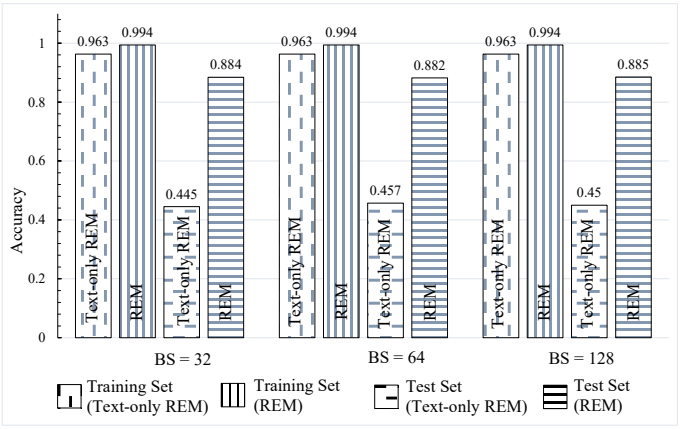


Figure 5: Best accuracies for training and test sets for proposed REM (text and emoticon) and text-only REM for training batch sizes (BSs) 32, 64 and 128.

work on unseen data. Test set accuracy is the key performance indicator for any machine learning system and it is a better score of proposed REM over text-only REM (i.e., without emoticon), which revealed that the use of emoticons enhances the ability of the proposed method in learning the emotion properly. However, the reason behind the worse performance with text-only REM is that the texts are limited in the selected tweet data and, in many cases, the text becomes meaningless without emotion, which has been explained in the data preparation section. Moreover, people do not care about meaning with text only when they use emoticon within it.

Table 3 and Table 4 show the emotion category-wise performance matrices of the proposed REM and text-only REM, respectively, for the best test accuracy cases shown in Fig. 4. The table shows the variation in actual and predicted emotions labeled for the individual emotion category. In the test set, 'Happy', 'Sad', 'Angry' and 'Love' emotion categories hold 1024, 956, 949, and 1074 tweet data consecutively. For the 'Happy' case in Table 3, for example, the proposed REM truly classified 914 cases, and the remaining 110 cases were misclassified into 'Sad', 'Angry' and 'Love' categories as 41, 33, and 36 cases, respectively. On the other hand, text-only REM truly classified only 442 cases as from Table 4. The proposed REM showed the best performance for the 'Angry' category by truly classifying 862 cases out of 949 cases

Table 3: Emotion category wise test set performance matrix for proposed REM method considering both emoticon and text.

Emotion Category	Total Sample	Detected Category			
		Happy	***Sad***	***Angry***	***Love***
Happy	1024	**914**	41	33	36
Sad	956	21	**857**	42	36
Angry	949	29	31	**862**	27
Love	1074	43	68	53	**910**

Table 4. Emotion category wise test set performance matrix for text-only REM.

Emotion Category	Total Sample	Detected Category			
		Happy	***Sad***	***Angry***	***Love***
Happy	1024	**442**	175	173	234
Sad	956	177	**419**	194	166
Angry	949	195	209	**434**	111
Love	1074	205	190	145	**534**

(i.e., 90.83% accuracy). On the other hand, the text-only method showed the best performance for the 'Love' category, truly classifying 534 cases out of 1074 cases (i.e., 49.72% accuracy). The performance comparison in the individual emotion category visualizes the proficiency of the proposed REM with emoticon and text.

On Twitter data, Table 5 compares the classification accuracy of the proposed method with other existing methods. The table also includes the methods used by various studies with a variety of dataset sizes. The existing methods considered Naive Bayes, CNN, and BiLSTM. The self-processed 16012 Twitter data used in the present study. The dataset used in [21] is larger than this study, but the authors did not mention how training and test sets are partitioned. Due to varying dataset sizes, the comparison with other methods may not be completely fair. However, the proposed method has outperformed any other methods showing a test set accuracy of 88.5%. The achieved accuracy is much better than the traditional machine learning with Naive Bayes [31] and deep learning methods with CNN and BiLSTM [21] [22]. It is already mentioned that study in [22] used emoticons but processed by separating them from the text. The main reason behind the outperforming ability of the proposed method is its emoticon management with text which is not appropriately handled in the existing methods. Finally, managing emoticons and texts simultaneously and classification with LSTM have been revealed as a promising emotion recognition method for microblogs.

Table 5: Performance comparison with existing methods on Twitter data.

Sl.	Author [Work Ref], Year	Sample Size (Training + Test)	Method Used	Test Set Accuracy
1	Wikarsa and Thahir [31], 2015	268 (116+152)	Naive Bayes	71.3%
2	Yang et al. [22], 2018	7200 (4600+2600)	CNN	72.6%
3	Batbaatar et al. [21], 2019	19678 (Not Mentioned)	BiLSTM + CNN	61.3%
4	Proposed Method	16012 (12009 + 4003)	LSTM	88.5%

5. Conclusions

Nowadays, people are very active on social media and frequently express their emotions using both texts and emoticons in microblogs. Emotion recognition from social media microblogs (i.e., REM) emerges as a promising and challenging research issue. It is essential to consider all necessary microblog information for comprehensive REM. Unlike many existing methods that only view the textual expressions for simplicity, this study has investigated REM utilizing both emoticons and texts simultaneously. Using the underlying LSTM technique, the proposed REM could interpret the emoticons in the context of text expressions in Twitter data to precisely classify the user emotions and outperformed the existing methods. The proposed REM method is expected to be an effective tool in emerging emotion recognition-based applications and play a vital role in social communication.

This study has revealed the proficiency of REM managing emoticons, and at the same time, several research directions are opened from its motivational outcomes and gaps. REM is developed collecting Twitter data for only 16 selected emoticons related to the four emotions (Happy, Sad, Angry, and Love); and system including other emotional states (e.g., Disgust, Surprise) and more emoticons might be interesting. Another thing, the texts were limited in the selected blogs, information degraded due to emoticon removal in text-only REM, and finally, recognition performance with LSTM was poor without emoticons. It might be interesting research to investigate text and emoticon trade-off effects on REM performance. In addition, instead of LSTM, any other deep learning method (e.g., CNN) might also be investigated owing to achieve better classification performance. We wish to work in such directions for developing a more comprehensive REM in the future study.

References

[1] J. Islam, S. Ahmed, M.A.H. Akhand, N. Siddique, "Improved Emotion Recognition from Microblog Focusing on Both Emoticon and Text," in 2020 IEEE Region 10 Symposium (TENSYMP), IEEE: 778–782, 2020, doi:10.1109/TENSYMP50017.2020.9230725.

[2] L. Al-Shawaf, D. Conroy-Beam, K. Asao, D.M. Buss, "Human Emotions: An Evolutionary Psychological Perspective," Emotion Review, **8**(2), 173–186, 2016, doi:10.1177/1754073914565518.

[3] A.R. Avila, Z. Akhtar, J.F. Santos, D. OShaughnessy, T.H. Falk, "Feature Pooling of Modulation Spectrum Features for Improved Speech Emotion Recognition in the Wild," IEEE Transactions on Affective Computing, **12**(1), 177–188, 2021, doi:10.1109/TAFFC.2018.2858255.

[4] M.A.H. Akhand, S. Roy, N. Siddique, M.A.S. Kamal, T. Shimamura, "Facial Emotion Recognition Using Transfer Learning in the Deep CNN," Electronics, **10**(9), 1036, 2021, doi:10.3390/electronics10091036.

[5] G. Castellano, L. Kessous, G. Caridakis, Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech, Springer Berlin Heidelberg, Berlin, Heidelberg, Heidelberg: 92–103, 2008, doi:10.1007/978-3-540-85099-1_8.

[6] N. Alswaidan, M.E.B. Menai, "A survey of state-of-the-art approaches for emotion recognition in text," Knowledge and Information Systems, **62**(8), 2937–2987, 2020, doi:10.1007/s10115-020-01449-0.

[7] A. Java, X. Song, T. Finin, B. Tseng, "Why we twitter: understanding microblogging usage and communities," in Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07, ACM Press, New York, New York, USA: 56–65, 2007, doi:10.1145/1348549.1348556.

[8] A. Ortigosa, J.M. Martín, R.M. Carro, "Sentiment analysis in Facebook and its application to e-learning," Computers in Human Behavior, **31**(1), 527–541, 2014, doi:10.1016/j.chb.2013.05.024.

[9] S. Gangrade, N. Shrivastava, J. Gangrade, "Instagram Sentiment Analysis: Opinion Mining," SSRN Electronic Journal, 2019, doi:10.2139/ssrn.3372757.

[10] M. Nam, E. Lee, J. Shin, "A Method for User Sentiment Classification using Instagram Hashtags," **18**(11), 1391–1399, 2015.

[11] M.Z. Naf'an, A.A. Bimantara, A. Larasati, E.M. Risondang, N.A.S. Nugraha, "Sentiment Analysis of Cyberbullying on Instagram User Comments," Journal of Data Science and Its Applications, **2**(1), 88–98, 2019, doi:10.21108/jdsa.2019.2.20.

[12] L. Huang, S. Li, G. Zhou, Emotion Corpus Construction on Microblog Text, 204–212, 2015, doi:10.1007/978-3-319-27194-1_21.

[13] R. Poonguzhali, V. Waldiya, S. Vinothini, K. Livisha, "Sentiment Analysis on LinkedIn Comments," International Journal of Engineering Research & Technology (IJERT), **6**(07), 2018.

[14] Z. Peng, Q. Hu, J. Dang, "Multi-kernel SVM based depression recognition using social media data," International Journal of Machine Learning and Cybernetics, **10**(1), 43–57, 2019, doi:10.1007/s13042-017-0697-1.

[15] S.R. Braithwaite, C. Giraud-Carrier, J. West, M.D. Barnes, C.L. Hanson, "Validating Machine Learning Algorithms for Twitter Data Against

Established Measures of Suicidality," JMIR Mental Health, **3**(2), e21, 2016, doi:10.2196/mental.4822.
[16] D. Gräbner, M. Zanker, G. Fliedl, M. Fuchs, Classification of Customer Reviews based on Sentiment Analysis, Springer Vienna, Vienna: 460–470, 2012, doi:10.1007/978-3-7091-1142-0_40.
[17] S. Shayaa, N.I. Jaafar, S. Bahri, A. Sulaiman, P. Seuk Wai, Y. Wai Chung, A.Z. Piprani, M.A. Al-Garadi, "Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges," IEEE Access, **6**, 37807–37827, 2018, doi:10.1109/ACCESS.2018.2851311.
[18] S. Chaffar, D. Inkpen, Using a Heterogeneous Dataset for Emotion Analysis in Text, 62–67, 2011, doi:10.1007/978-3-642-21043-3_8.
[19] S. Nirenburg, K. Mahesh, Knowledge-Based Systems for Natural Language Processing., 1997.
[20] S.B. Kotsiantis, I.D. Zaharakis, P.E. Pintelas, "Machine learning: a review of classification and combining techniques," Artificial Intelligence Review, **26**(3), 159–190, 2006, doi:10.1007/s10462-007-9052-3.
[21] E. Batbaatar, M. Li, K.H. Ryu, "Semantic-Emotion Neural Network for Emotion Recognition from Text," IEEE Access, **7**, 111866–111878, 2019, doi:10.1109/ACCESS.2019.2934529.
[22] G. Yang, H. He, Q. Chen, "Emotion-semantic-enhanced neural network," IEEE/ACM Transactions on Audio Speech and Language Processing, **27**(3), 531–543, 2019, doi:10.1109/TASLP.2018.2885775.
[23] N. Colneric, J. Demsar, "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model," IEEE Transactions on Affective Computing, **11**(3), 433–446, 2020, doi:10.1109/TAFFC.2018.2807817.
[24] A. Yousaf, M. Umer, S. Sadiq, S. Ullah, S. Mirjalili, V. Rupapara, M. Nappi, "Emotion Recognition by Textual Tweets Classification Using Voting Classifier (LR-SGD)," IEEE Access, **9**, 6286–6295, 2021, doi:10.1109/ACCESS.2020.3047831.
[25] K. Sailunaz, R. Alhajj, "Emotion and sentiment analysis from Twitter text," Journal of Computational Science, **36**, 101003, 2019, doi:10.1016/j.jocs.2019.05.009.
[26] M.S. Schlichtkrull, "Learning affective projections for emoticons on Twitter," in 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), IEEE: 539–543, 2015, doi:10.1109/CogInfoCom.2015.7390651.
[27] J. De Silva, P.S. Haddela, "A term weighting method for identifying emotions from text content," in 2013 IEEE 8th International Conference on Industrial and Information Systems, IEEE: 381–386, 2013, doi:10.1109/ICIInfS.2013.6732014.
[28] N. Chirawichitchai, "Emotion classification of Thai text based using term weighting and machine learning techniques," in 2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE: 91–96, 2014, doi:10.1109/JCSSE.2014.6841848.
[29] Z. Yuan, M. Purver, "Predicting emotion labels for Chinese microblog texts," Studies in Computational Intelligence, **602**, 129–149, 2015, doi:10.1007/978-3-319-18458-6_7.
[30] F. Jiang, Y.-Q. Liu, H.-B. Luan, J.-S. Sun, X. Zhu, M. Zhang, S.-P. Ma, "Microblog Sentiment Analysis with Emoticon Space Model," Journal of Computer Science and Technology, **30**(5), 1120–1129, 2015, doi:10.1007/s11390-015-1587-1.
[31] L. Wikarsa, S.N. Thahir, "A text mining application of emotion classifications of Twitter's users using Naïve Bayes method," in 2015 1st International Conference on Wireless and Telematics (ICWT), IEEE: 1–6, 2015, doi:10.1109/ICWT.2015.7449218.
[32] J.K. Rout, K.K.R. Choo, A.K. Dash, S. Bakshi, S.K. Jena, K.L. Williams, "A model for sentiment and emotion analysis of unstructured social media text," Electronic Commerce Research, **18**(1), 181–199, 2018, doi:10.1007/s10660-017-9257-8.
[33] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, U. Kaymak, "Exploiting emoticons in sentiment analysis," in Proceedings of the 28th Annual ACM Symposium on Applied Computing - SAC '13, ACM Press, New York, New York, USA: 703, 2013, doi:10.1145/2480362.2480498.
[34] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural Computation, **9**(8), 1735–1780, 1997, doi:10.1162/neco.1997.9.8.1735.
[35] Full Emoji List, Available: https://unicode.org/emoji/charts/full-emoji-list.html, Accessed: March 31, 2021.
[36] Chollet François, Keras: The Python Deep Learning Library, 2015.
[37] D.P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 356-361 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Designing a Model of Consciousness Based on the Findings of Jungian Psychology

Toshiki Watanabe[*,1], Hiroyuki Kameda[2]

[1]*Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology, Tokyo, 192-0982, Japan*

[2]*School of Computer Sciense, Tokyo University of Technology, Tokyo, 192-0982, Japan*

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 10 October, 2021
Online: 23 October, 2021

Keywords:
Model of consciousness
Jungian psychology
Artificial Intelligence

ABSTRACT

As artificial intelligence (AI) develops, it is expected that humans and AI will become more closely related than now. At the same time, however, the more closely humans and AI are related to each other, the more clearly they will face a moral dilemma, i.e., artificial intelligence will face a moral dilemma. To solve the moral dilemma problem, AI should understand and take into account human values and ethics. From this point of view. we designed a consciousness model based on Jungian post-psychological notes. As a result, we found that in order to implement the model of consciousness on a computer, it is necessary to design it with a structure similar to that of human beings, referring to human structures in various fields.

1. Introduction

This paper is an extension of work originally presented in SCIS-ISIS 2020 [1]

1.1. The social value of artificial intelligence technology

Artificial intelligence (AI) is currently being used in a variety of fields, including medicine, weather, and robotics. It has become an indispensable part of our lives, and its use is still being actively researched.

It is expected to evolve further in the future and will be applied and used in even wider fields than at present. According to an announcement by Japan's Ministry of Internal Affairs and Communications(https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd113230.html), the use of AI is expected to lead to more efficient operations and the development of new products and business models. It is also expected to have the effect of creating innovation by freeing humans from cumbersome tasks and enabling them to focus on creative tasks that only humans can perform. Several research organizations have analyzed the effect of AI on productivity. For example, according to an analysis by Accenture, the labor productivity,which is defined as (results produced or amount of value-added) / (quantity of labor), e.g, number of workers or workers * working hours, in Japan is expected to increase by 34% in 2035 compared to the baseline. The value of AI will become even higher in the future.

1.2. Current Status and Trends of Artificial Intelligence

In recent years, AI has become so prevalent in our society that we can ask so-called voice assistants such as Alexa (Amazon), Siri (Apple), and Google Assistant to tell us today's weather, or help input a destination into car navigation or show us the route on a map. In addition, there is a lot of research that focuses on interaction with humans, such as mentioned in [2] and[3], where robots and androids communicate with humans to heal and support them. In the future, humans and AI will be more closely related to each other.

As humans and AI become more closely involved, some of the decisions that have traditionally been made by humans will be delegated to them. For example, in the area of automated driving, Waymo currently commercializes Level 4 automated driving in the U.S., and Japan has also released a roadmap for fully automated driving by 2030 (https://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20200715/2020_roadmap.pdf), while Honda LEGEND (https://www.honda.co.jp/LEGEND/), which is currently Level 3 automated driving, is currently being sold to consumers.

However, with the development of these technologies, AI is facing the problem of moral dilemma. A thought experiment of Massachusetts Institute of Technology's Moral Machine (https://www.moralmachine.net/) exposes problems like the trolley problem to gather human perspectives on how self-driving cars should behave. We believe that this kind of problem is an important issue that must be considered in the future of AI.

*Corresponding Author: Toshiki Watanabe, g212003578@edu.teu.ac.jp

1.3. Position of this study

In this study, we hypothesized that if humans were to delegate decision-making to AI in the future, they would be able to make better and more flexible decisions by understanding and taking into account human values, ethics, and local culture and customs than now. We hypothesized that if the conscious and unconscious minds that control self-control could be realized on a computer, it would be able to understand and take into account human values, ethics, and local culture and customs. Self-control is considered to be important for this, and we thought that the relationship between the conscious and unconscious mind, which controls behavior and judgment, makes this possible. Therefore, we decided to model the conscious and unconscious as a decision-making system for AI and reproduce it on a computer for verification.

Therefore, to realize the conscious and unconscious on a computer, we designed a model of consciousness based on the findings of Jungian psychology.

2. Related works

2.1. The treatment of consciousness in the field of psychology

Freud is one of the first persons to discover the unconscious. He likened the human mind to the tip of an iceberg and divides conscious into three layers: the consciousness, the pre-conscious which can be made conscious, and the unconscious which cannot be made conscious, He also stating that repressed desires are pushed into the unconscious.

In his work, the author described the structure of the mind as consisting of the *es*, which is a mass of instincts, the *superego*, which is responsible for the protection of ethics, and the *ego*, which regulates them and guides conclusions to action [4].

The authors did not break down the human being structurally, as his study was centered on the theory of the purpose, which states that the individual mind cannot be divided and acts according to its present purpose [5].

The authors divided human conscious into the *consciousness*, the *personal unconscious*, and the *collective unconscious*.

Consciousness is that in the sense of being aware, and so on., while unconsciousness refers to other things.

> "Many of the difficult perceptual computations, such as invariant face recognition or speaker-invariant speech recognition. that ware recently addressed by AI correspond to nonconscious computations in the human brain. [...] Many of these C0 computations have now been captured by AI, particularly using feedforward convolutional neural networks (CNNs)."

The personal unconscious is the consciousness that has been suppressed by experiences and can be viewed as things that we are no longer conscious of, but which we can become aware of if we want to.

The collective unconscious is the unconscious that is common to all human beings and cannot be conscious.[6]

In Jungian psychology, the unconscious accounts for about 90% of the total, with the collective unconscious accounting for a large percentage of the total.

He was also the first person to use the term *complex*, as it is commonly referred to use in an even broader sense than its general scope.

In a recent study on the role of the conscious and unconscious mind the authors in [7],

> "(1) Recent findings in social psychology research indicate that many human behaviors, including self-control, depend on automatic processes that do not require consciousness (2) We examine the relationship between conscious and unconscious processes in self-control and propose effective self-control in which the two processes work together. The purpose of this study is to examine the relationship between conscious and unconscious processes in self-control and to propose an effective self-control system in which the two processes work together. The role of conscious processes seems to be to form new coalitions through abstract mental editing. […]One of the important functions of conscious processes was also to make unconscious self-control more effective. Thus, it is only by assuming the collaboration of the two that the essential functions of unconscious and conscious processes can be seen."

Through experiments, he came closer to elucidating the roles and functions of the conscious and unconscious.

2.2. Treatment of consciousness in the medical field

In the field of medicine, consciousness is often treated as a level of conscious or arousal. According to the website(https://bsd.neuroinf.jp/wiki/%E6%84%8F%E8%AD%98) ,it says

> "The level of consciousness or arousal is the state of awareness. The level of consciousness is the highest when you are awake and clear-headed, lower when you are sleepy and fuzzy-headed, or when you are asleep and not dreaming, or when you are under deep anesthesia."

Typical scales for the level of consciousness include the Grady Coma Scale. In [8] the authors, the authors say

> "The grade I patient is only slightly confused. The grade II patient requires a light pain stimulus (such as a sharp pin tapped lightly over the chest wall) for appropriate arousal, or may be combative or belligerent. The grade III patient is comatose but will ward off deeply painful stimuli such as sternal pressure or nipple twist with an appropriate response. The grade IV patient reacts inappropriately with either decorticate or decerebrate posturing to such deeply painful stimuli, and the grade V patient remains flaccid when similarly stimulated."

Consciousness in neuroscience and neurology deals with both psychological and medical consciousness. In [9] the authors discussed how to make computers conscious from the perspective of neuroscience. They argue that

He also claims that In other words, much of what AI does today, such as face and voice recognition, corresponds to unconscious computations in the human brain. They also claim

In other words, he claimed that if the Unconscious processing (C0) can be equipped with Global availability (C1) and Self-monitoring (C2), the machine will behave as if it has conscious.

2.3. Related studies in the field of engineering

In related research in the field of engineering, models of the mind and emotion have been actively studied.

" We contend that a machine endowed with C1 and C2 would behave as if it were conscious—for instance,"

In [10], the author builds a model that is aware of the whole picture of the mind and is hierarchical at five levels, from raw data to complex concepts, and can integrate low-level signal processing with high-level symbolic processing, which AI is good at. They built a model called the "Aesopworld" and implemented it in the Aesopworld project.

In [11], the author built a model of the mind based on the passive consciousness hypothesis, dividing it into two major parts: the conscious system and the unconscious system, assuming the passivity and illusion of consciousness. Thereby he proposed a method of conceiving the mind that has clear boundaries and is feasible, unlike conventional cognitive models.

In [2], the robot used cold reading and other techniques in communication to read the minds of customers and use them in sales. In [3] the authors adopt a model of personality to "heal" people.

These can be said to be equipped with unique models to communicate with humans.

In [12] the authors analyzed the trends in incorporating emotions into AI and what people need to be aware of when building models. People would like to create a foothold for clarifying the nature of emotions, which have been considered ambiguous, by incorporating knowledge about emotions from different fields such as psychology, neurology, and cognitive science, and other emotion research using a constructive approach. The reason why people can't reach the essence of emotions is probably that we can't abolish naive psychology and that the emotion research in various fields is conducted in different ways. He pointed out that it is important to abolish naive psychology to build a model of emotions and the mind.

2.4. Position of this study

In this research, we focus on being able to reproduce human consciousness on a computer. This is because if we can reproduce human consciousness on a computer, we will be able to systematically explain what people think and how they behave. And we believe that this will be useful not only in the field of AI but also in various other fields such as psychology and brain science.

3. Consciousness model in this study

In this study, we modeled consciousness in psychology, while including ethics, culture, and the human thinking and judgment, and aimed to construct a model of consciousness as a decision-making system, so that AI with the model can understand and take into account their values, ethics, and local culture and customs.

Psychology remains at a conceptual level and is difficult to implement directly in a computer, but it most clearly represents the role of consciousness and human behavior. On the other hand, research in the field of engineering has concrete models and roles to be implemented in computers, but they do not represent human behavior. In addition, the role of consciousness and human behavior was clear in the field of brain science, although the level of conscious in the medical field is different in meaning from the consciousness dealt with in this study. However, we still could not find a model that could be immediately implemented in a computer.

Compared to previous studies, my research is based on psychology, which allows me to design a model that eliminates naive psychology as much as possible, and by dividing the functions of consciousness, I can create a more detailed and clear model for designing programs.

The psychology we refer to when modeling is Jungian psychology, which divides the flow of thought into the consciousness, personal unconscious, and the collective unconscious, and explains the role of them in particular detail.

In [9] the authors found that it is effective to divide the functions of consciousness, so our model in this study is also designed by dividing the functions to some extent to clarify the roles of each process.

3.1. Definition of our Consciousness Model

We define the consciousness model as a decision-making function that selects action based on ethics, culture, and values accumulated through personal experience.

3.2. Requirements for an Consciousness model

The model of consciousness will be set to meet the following requirements.

- To fulfill the roles of consciousness, personal unconscious, and collective unconscious in Jungian psychology.
- Input and output are based on the human body.
- The model has memory retention and storage functions.
- The model can be reproduced as a program on a computer.

3.3. Pre-design of Consciousness Model

Consciousness in Jungian psychology is regarded as a system, and the functions of "consciousness," "personal unconscious," and "collective unconscious" are assumed to be interrelated and to make decisions by processing some data.

The procedure for constructing a model of consciousness is as follows: First, we clarify the functions of conscious in Jungian psychology. Next, we consider the connections among the functions of consciousness, and the roles these connections play in each other. We define the functions of consciousness as shown in Table1.

Table 1: Functions of Consciousness

Main components	Functions
Consciousness	Able to retrieve information from the unconscious that is related to the recognized information.
	Able to elicit needs and desires
	Able to repress needs and desires
Personal Unconscious	Complex (a desire or want that has been strongly suppressed by experience, education, and so on.)
	Suppressed needs and desires (currently unconscious needs and desires)
Collective unconscious	Human instincts (maternal instincts, sexual desires, and other archetypes shared by all human beings)

Assuming that the consciousness, personal unconscious, and collective unconscious have functions as shown in Table 1, there should be the interaction of data transfer and processing amang main components.

Therefore, as a result of considering the data transfer and processing between the functions of consciousness as a decision-making system, Figure 1 was obtained.

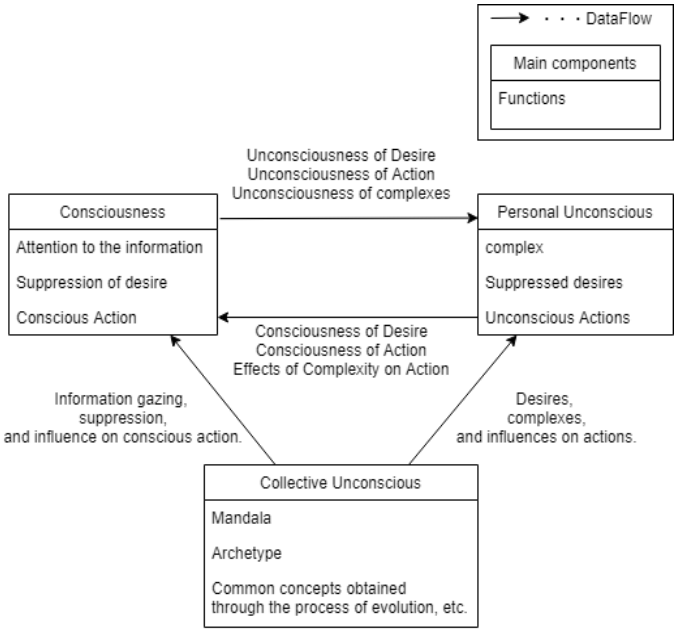


Figure 1: Conceptual diagram of consciousness

Figure 1 shows the connections between consciousness, the personal unconscious, and the collective unconscious, but this is too abstract to be reproduced as a program on a computer.

There must be connections between processes in the consciousness mind, connections between processes in the unconscious mind, and connections between small processes between the conscious mind and the unconscious mind.

Therefore, we further extended Figure 1 to consider the connections between each process and created Figure 2.

3.4. Overview of the Consciousness Model

By extending the conceptual diagram of consciousness in Figure 1, we obtained the following Figure 2. Using this figure as a model of consciousness, we can consider the operation of each process.

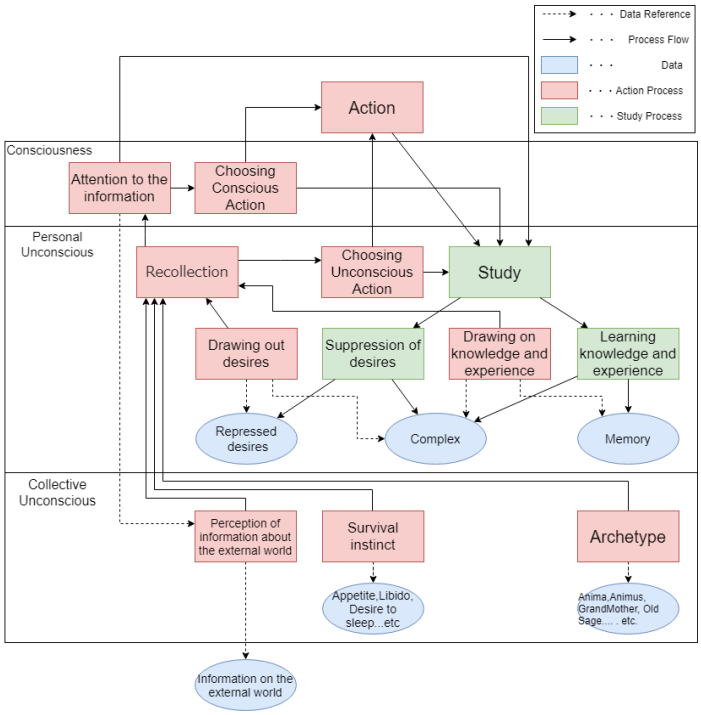


Figure 2: Consciousness Model

The process shown in Figure 2 can be divided into three major parts: "data," "Action process," and "Study process". For data, these include values obtained from sensors and episodes. The arrows between data and process represent data referrence.

The action process is the process of taking action based on the data, and the study process is the process of retaining and storing the data as experiences and desires in the individual's memory after taking action.

For example, when a sensor detects hunger, the survival instinct perceives hunger and recalls associates information about food and how to eat it. Then, priming memory and other memories are sent to unconscious behavioral processes to select actions. Episodic memory and other memories are sent to information gazing to link the perception of external information with the information in episodic memory. After that, the conscious action selection process selects the action, and the action process selects the final action from the conscious and unconscious actions.

The focused information and selected behaviors are sent to the learning process, where the information is divided into the desire suppression process and knowledge-experience learning, and is prepared in the form of each memory before being stored.

The followings are the explanations of the components in Figure 2.

- Action

The actual action is taken by the system or the signal that directs it. The process of acting is selected by conscious and unconscious action selection.

- Information on the external world

Information that is not internal to the system, such as the five senses in humans, or sensors in machines. And it is unprocessed one, such as what you see with your camera or eyes, or what you feel.

- Attention to the information

To pay attention to objects and phenomena that are highly related to the objects recognized through the recognition of information about the external world, and to perform a deeper recall of these objects and phenomena.

- Choosing Conscious Action

To select the higher rewarding action required by the information you have been paying attention to.

- Choosing Unconscious Action.

To select highly learned behaviors picked up by recall.

- Recollection

The process of retrieving retained information in the right place and at the right time. Memory recall is the process of recalling a specific memory from among the memories stored in the brain. It induces memory destabilization, refixation, extinction learning, and memory association. In terms of the movement of the process, it refers to the action of searching for knowledge and experience to achieve desires.

- Drawing out desires

To search for and elicit relevant desires.

- Drawing on knowledge and experience

To search for and elicit relevant knowledge.

- Study

The act of filling in the blanks of knowledge. It is close to the meaning of the word in psychology and pedagogy and is distinguished from learning. It refers to the overall act of learning, such as extracting, categorizing, and storing features from the information obtained at that time as a result of an action, and adjusting each parameter. Knowledge is all that is gained through experience and includes all information stored in the unconscious. The rewards received from actions and related information are parameterized and recorded into knowledge, experiences, wants desires, and complexes.

- Suppression of desires

To repress what could not be done or failed in the drawn-out desire-lust.

- Learning knowledge and experience

To record the results and rewards of actions taken, and to store a series of actions as episodic memory.

- Repressed desires

A set of data repressed by desire.

- Memory

It is classified into sensory memory, short-term memory, and long-term memory, and it records information about what happens to us and what we feel, both temporarily and long-term. Sensory memory has the shortest storage time and is present in each sense. Only the information to which we pay attention becomes short-term memory. Short-term memory lasts only a few tens of seconds and is limited in the amount of information that can be stored at one time. Long-term memory is a part of short-term memory that is retained and becomes long-term memory through memory fixation. In this study, data is stored by learning, and organized by study.

- Perception of information about the external world

To recognize the type of object and distance from the external world information.

- Survival instinct

It elicits life-sustaining desire parameters such as sexual desire, appetite, and desire to sleep. The survival instinct is considered to be part of the collective unconscious.

It is the minimum means of survival stored as data.

- Archetype

To draws information about repression related to external world information and desires from stored data, such as anima and animus. Instincts and concepts cultivated in the group are individually shared data, not necessarily shared things in common.

Of the collective unconscious, the history and common sense cultivated by humankind fall into this category.

- Complex

In this study, we believe that complexes can be treated as a cohesive group of closely related memories and impulses, and as a collection of related memories. Complexes are said to be complexes of impulses, memories, and ideas that are intertwined by some emotion. By using these complexes, we will be able to handle memories that are highly relevant to recall. The parameters stored in the memory are treated as data that is associated with a distribution or set.

- Appetite, Libido, Desire to sleep...and so on.

Internal information such as a sensation of human internal organs, mental state, drowsiness, and so on., and remaining battery power and continuous operation time of the robot.

- Anima, Animus, Great Mother, Old Wise Man...and so on.

Conceptual things that humanity has in common, such as a sense of fear.

4. Discussion and further works

4.1. Discussion

The assumed environment and data are built on the premise that the input is received from the sensors installed in the computer, and the action is derived from the recall from this model.

In the future, when we implement this consciousness model, we need to consider various methods for each process.

Therefore, to clarify the data flow of this model and to make it easier to implement, we created a DFD(Data Flow Diagram) that is supposed to be installed in a robot or a simulator, and as a result, we obtained the following Figure 3.

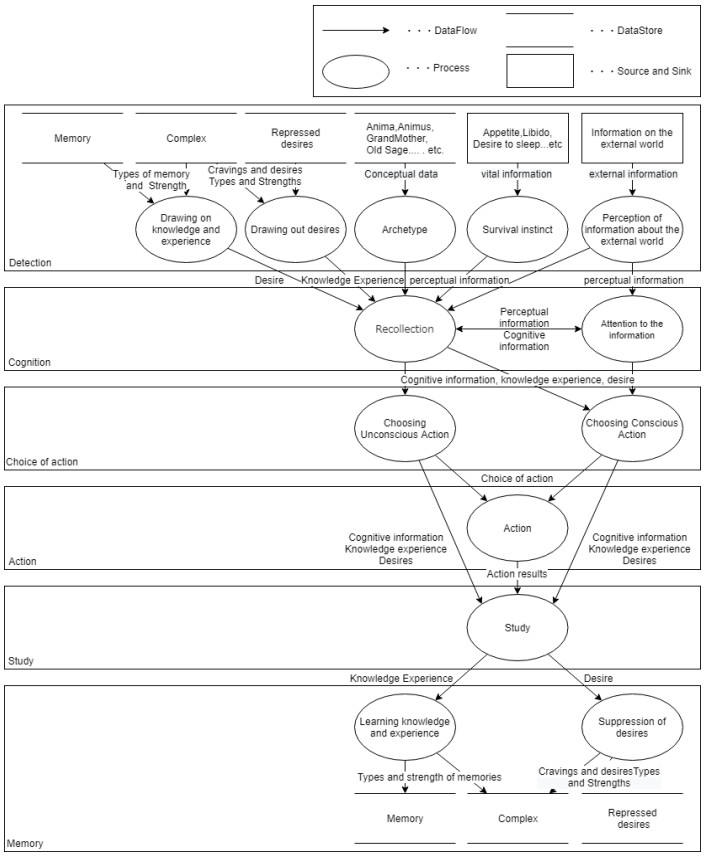


Figure 3: Diagram of consciousness DFD

Figure 3 shows the DFD version of the consciousness model in Figure 2. This DFD indicates that data will be processed from the top and passed to the next process. The arrows represent the direction in which the data is passed, and the names between the arrows represent the data.

Figure 3 shows that the flow of data is roughly in the following order: detection, data retrieval, action selection, action, study, and memory. This is roughly in line with the order of the process of detection-cognition-judgment-action, which is called the basic human thought process (https://ipsj-is.jp/isdic/1070/). We believe that this model can reproduce the basic thought process of human beings and still meet the requirements of the consciousness model defined in this study.

To implement the model on the agent in Unity for model validation, we will create a class diagram from Figure 2 for object-oriented implementation. Based on the obtained diagram, we plan to code and implement the model with the help of other researchers.

4.2. Further works

It will be necessary to implement the consciousness model in computers and robots to prove whether the model can function as a decision-making program and make decisions based on an understanding of human ethics and values.

When we implement the consciousness model in the future, we need to be especially careful about the methods we use for each process, because some methods may cause unintended behavior. For example, the data structure of memory should be based on the structure of memory used in psychology and brain science, and the implementation should be based on human evidence.

In the learning process, it is also necessary to consider specific implementation methods on computers, such as the learning method and how to process the data to be handled.

5. Conclusion

In this study, we have developed a model of consciousness as a decision-making system based on Jungian psychology for implementation on computers and robots.

In the future, various issues remain to be solved in implementing this model, but we believe that solving them one by one will not only solve the ethical problems of AI but also lead to a better understanding of human thinking and brain mechanisms.

References

[1] T. Watanabe, H. Kameda, “Proposal of a computer-reproducible model of consciousness,” in 2020 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS), IEEE: 1–4, 2020, doi:10.1109/SCISISIS50064.2020.9322763.

[2] M. Watanabe, K. Ogawa, H. Ishiguro, “Minami-chan: Application and Verification of Androids to Society Through Selling,” **57**(4), 1251–1261, 2016.

[3] T. Harada, S. Yamamoto, K. Yui, H. Kanda, Y. Chiba, H. Kameda, “Healing Robot PDDIN – Design and Implementation of prototype system -,” The 65th National Convention Lecture and Paper Series, **2003**(1), 71–72.

[4] H. Myouki, Introduction to Freud, ChikumaShinsho, 2000.

[5] I. Kishimi, Introduction to Adlerian Psychology: For Better Relationships, bestsellers, 1999.

[6] H. Kawai, Introduction to Jungian Psychology, baifukan, 1967.

[7] M. Oikawa, H. Oikawa, “Consciousness and the unconscious in self-regulation: The effects of conscious compilation on goal priming,” The Japanese Journal of Psychology, **81**(5), 485–491, 2010, doi:10.4992/jjpsy.81.485.

[8] H.K. Walker, W.D. Hall, J.W. Hurst, Clinical Methods, 3rd edition, Butterworth, 1990.

[9] S. Dehaene, H. Lau, S. Kouider, “What is consciousness, and could machines have it?,” Science, **358**(6362), 486–492, 2017, doi:10.1126/science.aan8871.

[10] N. Okada, “A Computational Model of the Mind,” IEICE ESS Fundamentals Review, **1**(4), 30–45, 2008, doi:10.1587/essfr.1.4_30.

[11] T. Maeno, “How to Make a Conscious Robot-Fundamental Idea based on Passive Consciousness Model-,” Journal of the Robotics Society of Japan, **23**(1), 51–62, 2005, doi:10.7210/jrsj.23.51.

[12] H. Chie, H. Takato, N. Takayuki, “OS-18 Emotion and AI,” Artificial Intelligence, **34**(6), 881–887, 2019, doi:10.11517/jjsai.34.6_881.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 5, 362-368 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Impact Assessment of the Renewable Energy Sources Implementation in Bulgarian Single-Family Houses on the Greenhouse Gas by HOMER Pro Software

Dilyana Gospodinova*, Peter Dineff

Department of Electrical Apparatus, Technical University of Sofia, Sofia, 1000, Bulgaria

ARTICLE INFO

Article history:
Received: 14 July, 2021
Accepted: 04 October, 2021
Online: 23 October, 2021

Keywords:
Greenhouse gases
HOMER Pro
Renewable energy sources

ABSTRACT

It is widely known that cities now house more than half of the world's population. Within this framework, this study presents the possibilities for real-world application of renewable energy sources (RES) in urban areas, as well as their contribution to the urban deployment of the new energy paradigm. A comparison is made between hybrid power systems, operating on Bulgarian territory and conventional (traditional) power systems. Proposed hybrid system is designed and actually implemented to power a single-family house and consists of wind turbine (WG) photovoltaics (PV), lithium-ion batteries for energy storage and suitable converter. The HOMER Pro software was used to model and explore the long-term continuous implementation of a hybrid power system and greenhouse (GHG) gas emissions investigation. The article discusses the amount of carbon dioxide (CO_2) and nitrogen oxides (NO_X), that can be reduced by using a hybrid power system (solar and wind) in conjunction with a battery storage system (BSS - lithium-ion batteries) in single-family houses. Renewable energy sources combined with energy storage, according to this report, result in a 50% reduction in dangerous carbon dioxide and nitrogen oxide emissions. The proposed system is optimized based on the lower cost of energy (COE) and proper dispatch strategies (load following and cycle charging), resulting in greenhouse gas emissions distributed by regional Bulgarian cities.

1. Introduction

This paper is a continuation of work that was originally presented at the 2020 12th Electrical Engineering Faculty Conference (BulEF), [1]. The main task of this extended article is to expand knowledge and demonstrate a modern ability to share computing and modern technologies to assess important parameters such as the generation of harmful emissions. Kumar discuss [2] the opportunity to assess the generated harmful emissions is of particular importance and is a societal challenge for all of us. The rapid development of renewable technologies, as well as energy storage technologies, as well as their ever-increasing disposal, could to some extent provide a solution to the existing problem – environmental pollution generated at electricity produced for domestic use [3].

Many literary sources define the different types of green energy; some say there are five, while others say there are six or even seven [4–9]. Electricity production using fossil fuels has historically provided adequate costs but emits GHG.

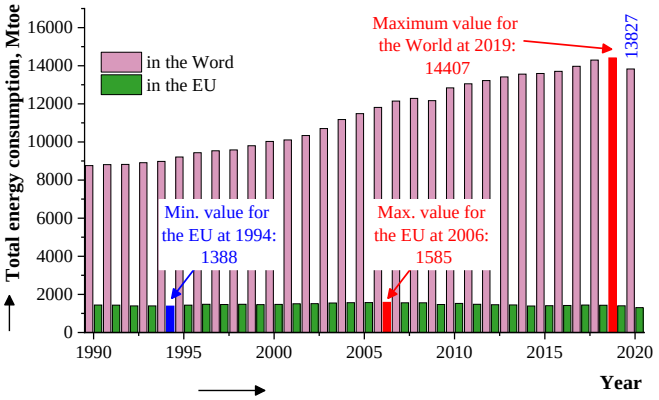


Figure 1: Total energy consumption over 1990 – 2020 in the Word and the EU.

According to the European Environment Agency, the European Union ranks third in the world in terms of greenhouse gas emissions, trailing only China and the United States [10]. As reported by Enerdata, contrary to expectations, in the face of a global pandemic, global energy consumption growth slowed by

*Corresponding Author: Dilyana Gospodinova, dilianang@abv.bg

4% in 2020, compared to an average of 2% per year from 2000 to 2018 and a 0.8% slowdown in 2019, Figure 1, [11].

Except for China, which is the world's largest energy consumer, most countries have reduced their energy consumption.

The difference in total power consumption between Europe and the rest of the world is between 6 and 10.6 times, and the trend is upward Figure 1. The data show that the global trend of increasing energy consumption continues. The year 2020 is an exception for both the world and Europe, due to the global pandemic caused by COVID-19 and its spread. Globally, energy consumption is steadily increasing, whereas in Europe, consumption is decreasing, most likely as a result of the European Union's policy to increase efficiency. Lockdown measures and limited economic activity had a significant influence on energy consumption in 2020. It dropped by about 7% in the European Union. Energy consumption fell across Africa and the Middle East as well, [11]. Under the Paris Agreement, the European Union (EU) has pledged to reduce GHG emissions by at least 40% by 2030 compared to 1990 (2016). The European Union is taking it a step further, aiming for nearly zero emissions by 2050 under the European Green Pact, [12]. The European Union is dedicated to achieving lofty goals in the fight against hazardous climate change. The European Commission announced a strategy for long term and achieving a climate friendly economy by 2050 in November 2018. By the second half of this decade, the EU hopes to be carbon neutral, [13]. The European Union has passed regulations in a variety of areas to attain the goal of reducing GHG emissions. The emissions trading scheme has undergone changes, for example, which oversees the progressive reduction of emissions from energy and industry while requiring polluters to pay a penalty. Various industries, such as agriculture and construction, have established national emission reduction targets. Because energy is at the heart of the emissions problem, the EU is working toward a strategy of cleaner sources and more efficient use. In 2018, Parliament passed laws requiring renewable energy sources (sun, wind, and water) to account for 32% of overall consumption and allowing households on their own to generate their own "green" energy, [14]. Furthermore, the EU has implemented rules on buildings and household appliances to increase energy efficiency by 32.5% by 2030, [15]. The “Clean Air For Europe programme (CAFE)” 6 of 2013 maintained the goal of achieving full compliance with current air quality regulations throughout the EU as soon as possible, with 2020 and 2030 targets specified.

The policy initiatives themselves are: the *first basis* consists of ground-level ozone, particulate matter, and nitrogen oxides ambient air quality standards, hazardous heavy metals, and a variety of other pollutants set out in the directives on ambient air quality; the *second basis* according to [16] is the Directive on national emission ceilings for the most major transboundary air pollutants, which establishes national emission reduction targets.

It is widely known that more than half of the world's population now lives in cities. Human and monetary activities both require energy. Fossil resources are decreasing in tandem with rising energy demand. In terms of global development, energy supply and security are major issues. The hunt for new energy sources based on RES has been prompted by the scarcity of energy and the demand for sustainable systems. In comparison to traditional fuel-powered generators, numerous renewable energy (RE) solutions are now well developed, cost-effective, and reliable, [17].

The true potential of solar energy is found in the numerous applications it offers at the municipal level, as it can be used to generate both heat and power in public and residential buildings and facilities. Solar energy has grown to be one of the most popular RES in Europe and Bulgaria as a result of its low cost, Figure 2. Solar energy is becoming a more common source of energy and a viable alternative to overcome energy shortages. According to [18], the main benefits of solar energy are accessibility, lack of pollution i.e. obtaining "green" energy, ease of operation, availability, and tranquillity. Furthermore in [19], the author point out, solar panels are known for their low maintenance, ease of installation and disassembly, and lack of moving components. A solar panel, as a RES technology, is simple to install and safe in urban environments. PV panels can be distributed and put in any sun- location exposure to the sun [20].

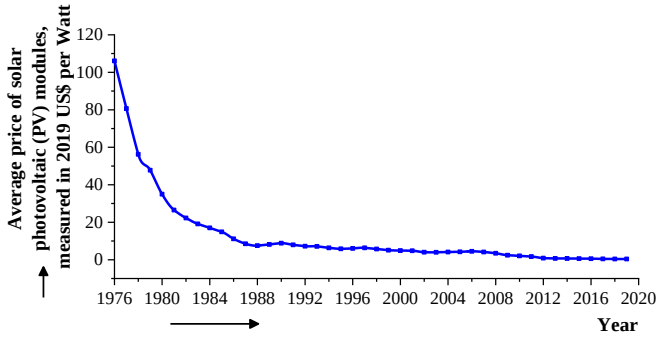


Figure 2: The average price of solar photovoltaic modules around the world, measured in US dollars per watt in 2019, [21].

PV systems and generators powered by the wind (WT) can be combined to generate power in a hybrid system. This system produces no greenhouse emissions. A system like this has the advantages of being practicable, dependable, simple, almost no maintenance and low cost. The economics of a solar-wind hybrid system is always rising due to constant reductions in component pricing. Because the energy obtained from the sun and wind fluctuates over time, a battery storage device must be included in the hybrid system, [21].

This extended article presents also new comparative data for Bulgaria, the EU and the World in connection with the use of RES and their share in the overall energy mix, Figure 3.

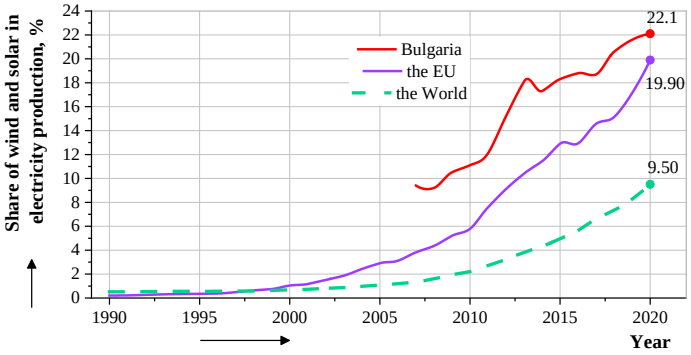


Figure 3: The percentage share of wind and solar in electricity production over 1990 – 2020 in the World, European Union and Bulgaria.

Increasing 0the amount of renewable energy in Bulgaria provides further economic, scientific, and political benefits: decreases the state's reliance on fossil fuels. Bulgaria first began promoting renewable energy in 2007, well after the formation and

execution of the institutional and legislative framework in the other old EU member states. In two periods (2007÷2012 and 2012÷2016), the state saw a significant increase in its percentage of renewable energy. Bulgaria now has approximately 1700 MW of installed renewable energy capacity (mostly wind and photovoltaic). Bulgaria is one of 11 EU countries that has already met its 2020 renewable energy goal. Energy from renewable sources accounted for 20.5% of gross final energy consumption in 2018, according to the Republic of Bulgaria National Statistical Institute [22-24].

A hybrid power station that combines PV and WT is promising future energy generation technology. Furthermore, Prasad proposes PV and wind power have an effect of synergy, with PV and wind power dominating during daytime and nocturnal hours, respectively [25]. Several studies have suggested that battery energy storage systems (BESS) be used to provide a variety of services for performance in order to deal with the variability of PV/wind energy generation [26-28]. The operation of a hybrid PV-WT system requires a battery storage system. The installation of a battery storage system enhances dependability, efficiency, and load smoothing. The hybrid power system is balanced thanks to the battery storage system [18,29].

The current computer-based study provides new valuable systematic price information connected to the so important chemical source of electricity - lithium-ion batteries as shown in Figure 4.

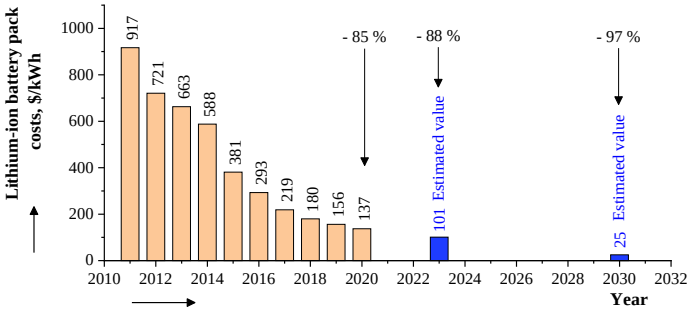


Figure 4: Costs of lithium-ion battery packs worldwide between 2011 and 2030 [30].

Energy storage is critical for integrating renewable energy sources into the power system. Furthermore, the combination of rooftop solar panels and battery storage is becoming increasingly popular in the hose-hold market. The cost of manufacturing lithium-ion batteries is steadily decreasing [30, 31]. These facts support the argument made in this article for the use of battery energy storage devices.

The complicated architecture of the power grid makes calculating the actual amount of emission reductions from energy efficiency and renewable energy measures difficult. Because air hazard pollutants like ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen oxides (NO_X) can be harmful to human health, it is critical to understand how to limit their emissions. Increased energy efficiency is one of the most well-known technologies. The usage of renewable energy sources is the second option.

The quantity of harmful emissions that can be decreased by employing a hybrid power system (WT and PV) in conjunction with a BSS is estimated in this article.

Numerous experiments have been conducted to study how to reduce costs by dispatching a battery without using forecasting techniques. There are also numerous battery optimization studies that use only one forecasting technique to achieve the least costs or greatest revenues. Although the complexity of the models was minimized in these investigations, the impact of different forecasting approaches on the optimization results could not be observed. As a result, the forecasting findings from several forecasting methodologies are integrated into the optimization model in this study [32-36].

The impact of a hybrid energy system with integrated energy storage technology on a single-family house in terms of harmful CO_2 and NO_X emissions has been investigated using an appropriate software product. The available literature does not present results on the distribution of harmful gas emissions after the implementation of RES on Bulgarian territory. The preliminary quantitative assessment yields result indicating a reduction in harmful CO_2 and NO_X emissions.

This article explains the contributions of the study in contrast to the article [1]. This paper's major contributions include the following: 1) The presentation of a dispatch strategy proposed by HOEMR Pro. 2) The demonstration of applying HOMER Pro software product to efficiently solve a hybrid system complex optimization problem, from a green gas emissions point of view, a single-family house power system that includes different renewable energy sources and battery storage system. 3) The use of a Load following (LF) dispatch strategy resulted in significantly lower hybrid power system operation costs. 4) The impact of the RES implementation and energy storage on the amount of generated emissions of CO_2 and NO_X, as well as the economic effect on the territory of Bulgaria, which is a member of the European Union, was assessed.

2. Design specification for the proposed hybrid system

This article takes a computer-assisted method. The simulation compares greenhouse gas emissions from hybrid energy supply systems in several Bulgarian regions. The proposed system configuration is simulated, optimized, and analysed using HOMER. The simulation assesses the system's technical feasibility and performs system optimization based on several system configurations to identify which one is the most suitable.

Forecasting PV and/or wind power generation is critical for developing a battery optimization plan for battery energy management. To meet the load, the HOMER software provides alternative dispatch algorithms to manage the operation of the generator and storage bank.

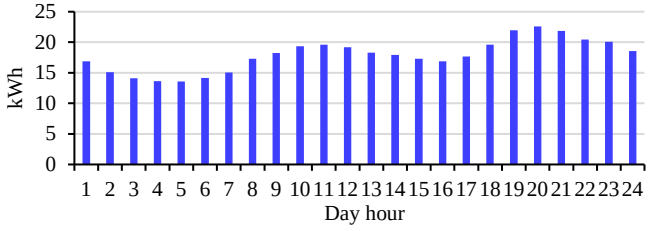


Figure 5: Daily profile for a Sofia-area (Bulgaria) urban household.

The shown standardized load profile is standard and pertains to residential customers in the Sofia metropolitan area, as recognized by Bulgaria's Energy and Water Regulatory Commission

(EWRC). The distribution company in the Sofia area of Bulgaria offered the developed load profile, Figure 5.

The HOMER Pro software is capable of simulating tree dispatch strategies - cycle charging, load following, and combined [37]. Which is ideal relies upon numerous variables, including the spans of the generators and battery bank, the fuel cost, the operation and maintenance (O&M) cost of the generators, etc.

When used in a load-following strategy (LF), a generator only produces enough power to meet the demand. LF works best in systems that use a lot of renewable energy because the renewable energy output occasionally exceeds the load. A generator is used in the cycle charging (CC) approach, and excess power is used to charge the battery bank. In systems with little or no renewable energy, cycle charging is preferred [37].

One of the objectives of this article is to compare alternative working strategies (LF- and CC-) in terms of greenhouse gas emissions in order to determine the best dispatch technique. It has been chosen to model both dispatch strategies, HOMER simulates each system separately.

Figure 4 depicts a schematic diagram of the microgrid designed by HOMER. A grid-connected hybrid system generates power from the sun and wind, which is then converted into grid-compliant electricity.

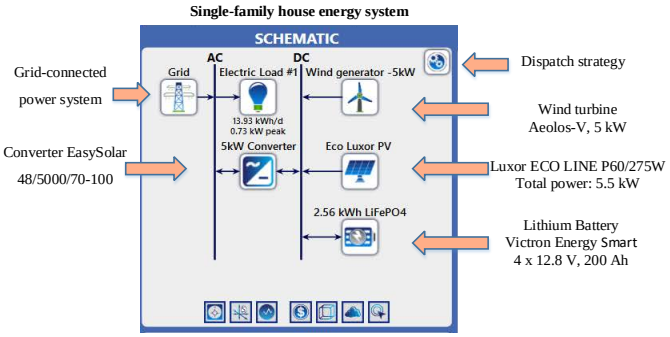


Figure 6: A schematic diagram of the hybrid power system that was designed to power a single-family home in Bulgaria.

2.1. Hybrid system components

The solar panels (Luxor ECO Line P60/275W) chosen had an initial and replacement cost of 3050 EUR, as well as an annual operation and maintenance cost of 20 EUR. It has a service life of 25 years. Without a tracking system, the selected panel's external current is DC, and the ground reflection coefficient is assumed to be 20%. The average efficiency coefficient was assumed to be 90% in this study. The coefficient of performance was used to represent the performance degradation caused by factors such as snow, shadows, high temperature, dust, wiring loss, and so on. It is usually 90%, but in hot weather it drops to 70%.

Four (4) identical series-connected 12V, 200Ah Li-ion Victron Energy batteries with an initial and replacement cost of 9600 EUR and an annual operating and maintenance cost of 50 EUR were used. The HOMER software accepts that the battery's highlights remain constant throughout its useful life and are not influenced by external factors such as temperature.

A converter is required for any system that includes both AC and DC components. For the simulation, a 5 kW converter with an initial and replacement cost of 4000 EUR is used. For grid compatibility in Bulgaria, the output voltage is set to 230V at 50Hz. The inverter has a 25-year lifespan and an efficiency of 95% as an inverter and 95% as a rectifier at a relative capacity of 100%.

A 5kW direct current wind generator was used. It had an initial and replacement cost of 800 EUR, as well as an annual operating and maintenance cost of 50 EUR. It has a 25-year life expectancy and was installed at a height of 10 meters above ground.

Reference [38] provides a detailed description of the peak electric load and standardized load profile.

2.2. HOMER's Greenhouse Gas Emission Assessment Capabilities

In [39], the author stated that there is a plethora of energy modelling software for renewable energy systems simulation available on the web. However, only five of them are appropriate for simulation procedures, and their environmental impact approach is appropriate.

The software needed to input technical and economic data, as well as energy component specs and resource data for renewable and non-renewable resources. Once all of the data has been entered into the software, HOMER calculates the environmental impact based on direct CO_2, CO, NO_X, and SO_2 emissions.

HOMER calculates emissions for each pollutant by multiplying the emissions factor (in g/kWh) by the annual fuel usage (in kWh) [37].

3. Discussion of simulation results

A 25-year prediction was used in the simulation. HOMER gave simulation results for a hybrid power system. All of the results reported here are for a hybrid system in Sofia, Bulgaria - 42°39.4'N, 23°21.2'E.

HOMER ran 114 simulation alternatives for the proposed hybrid power system, Figure 6, and all of them were feasible. This microgrid uses 27 kWh per day and has a max power output of 5 kW. The electrical load was served by the generation sources in the proposed system. This optimization has a net present cost of 20 800 EUR, while the cost of generating RE is 0.16 EUR per kilowatt-hour.

PV penetration of 136% for energy supply increases net grid purchase, whereas wind penetration of 11.4% decreases it. A renewable portion of 72.6% appears to be an appropriate load sharing between the grid, wind generator and PV panels.

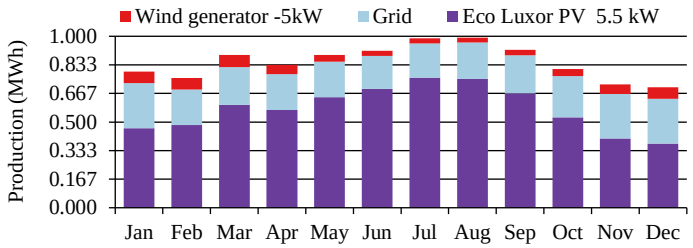


Figure 7: Average monthly electric production of a grid-connected hybrid power system in Sofia, Bulgaria.

Figure 7 indicates that the hybrid power system generates a total of 7516 kWh/year to sustain the load, lowering the grid purchase to 2697 kWh/year. It is clear that the solar component of the energy system has the biggest share, followed by power from the conventional grid. Despite a clear trend of increasing

renewable energy generation from wind turbine during the winter months of the year, the share of this type of source remains low.

In [1], it clearly shows the economics and operational statistics for the solar system, wind generator, and converter separately. At 136% penetration, the solar system's reported value is 5.50 kW. The Aeolos Wind Energy Ltd. wind turbine installation, rated at 5.00 kW, produces 581 kWh/yr of electricity.

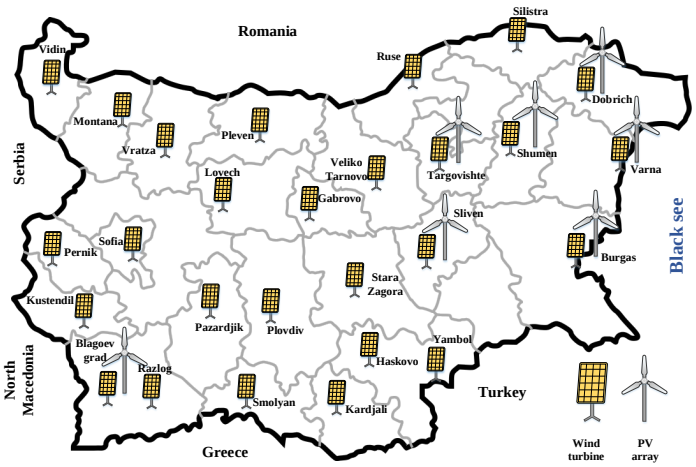


Figure 8: Preferred RES implementation by Bulgarian geographical area in accordance with [32], as determined by HOMER software.

HOMER calculates and displays the total amount of each pollutant generated yearly by the power system in kilograms per year of the six pollutants listed below: carbon dioxide (CO_2), carbon monoxide (CO), unburned hydrocarbons (UHC), particulate matter (PM), sulfur dioxide (SO_2), and nitrogen oxides (NOx). Pollutants are produced through the use of fuel and biomass in generators, boilers, and reformers, as well as by the use of grid electricity [29]. The HOMER Pro software application allows for the presentation of Bulgaria's current carbon dioxide (CO_2) emissions situation. HOMER adds 8.93 EUR/ton CO_2 emissions to the total annual cost of the hybrid power system. Grid emissions are lowered as a result of power sales to the grid. According to our previous article [28], the recommended distribution of used RES corresponds to that shown in Figure 8.

The proposed distribution map was based on the demand for the cheapest per-kWh energy. The HOMER software package allows to estimate the cost of energy for each proposed system, and this parameter, as the most crucial, was tracked throughout the simulation process. The presented hybrid power supply system was multiplied and applied to each of Bulgaria's provinces. One of the goals of this research is to develop a roadmap for the use of renewable energy sources in single-family houses across the country, resulting in significant reductions in GHG emissions and the cost of energy, [40].

Figure 9 shows the results and comparison of electricity costs with renewable energy sources in a single-family home, as well as the distribution by major Bulgarian district cities. Two power systems were compared: one powered by the local electricity distribution company, and the other combining conventional electricity and RES (PV and WT) with a battery storage system.

In general, the cost of energy produced by a hybrid power system is higher than that produced by the conventional power system.

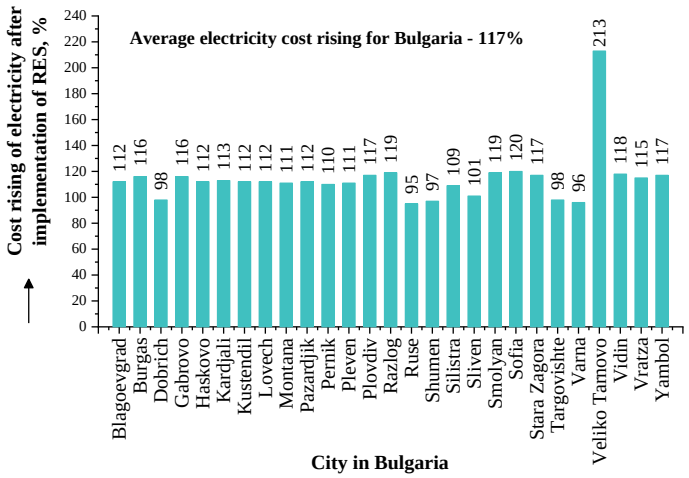


Figure 9: Costs of energy (COE) comparison between hybrid power system and conventional grid-connected system based on preferred energy sources shown in Figure 8.

When powering a single-family home, the proportion of renewable energy in the overall energy mix is an important factor in the modelling and analysis of the results. Figure 10 shows a clear trend toward a high percentage of energy generated from renewable sources in the total energy mix. The share of renewable energy accounts for between 67.7% and 79.0% of total energy consumption. This percentage stays high enough to power Bulgaria's entire territory with renewable energy.

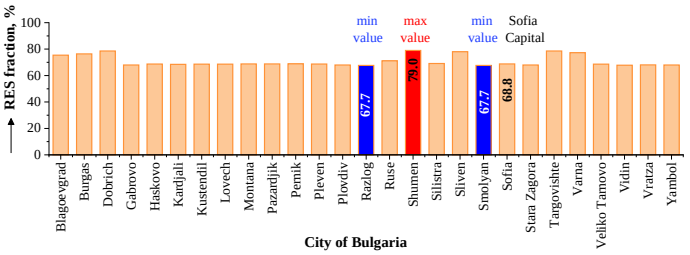


Figure 10: Distribution of RE fractions (%) in Bulgaria's major cities using a hybrid system, calculated by HOMER Pro.

Harmful CO_2 emissions for a single-family house with a hybrid power supply system are between 1.318 and 1.905 kg per year. Figure 11 shows that the baseline level of CO_2 emissions when using solely a conventional electrical system is 3.213 kg/year.

When a hybrid energy system is combined with battery storage, the percentage change with a negative sign goes from 41% to 59%, indicating a considerable reduction in harmful CO_2 emissions. Several locations of the country, including Shumen, Varna, Targovishte, Sliven, and Dobrich, have the potential to reduce CO_2 emissions by twofold.

Nitrogen oxides (NO_X) are a group of seven chemical molecules that pollute the air. The results for nitrogen oxides are also presented in that article in a different manner. The situation with nitrogen oxides, or NO_X, is identical to CO_2 emissions. Nitrogen oxides emissions from a single-family hose powered by the traditional energy system are 6.81 kg per year. In a single-family house, using a hybrid power system reduces hazardous nitrogen oxides by 41% to 57%, depending on the location where the home is located.

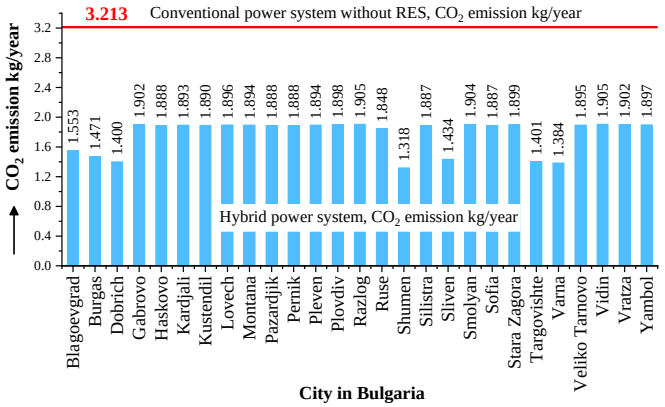


Figure 11: Comparison of CO_2 (kg/year) emissions by Bulgarian regions for conventional grid-connected and hybrid power systems.

4. Conclusions

The renewable energy fractions in the proposed hybrid power system for a single-family dwelling in Bulgaria are high, reaching 79 percent in some regions of the country.

Despite the limited percentage of wind energy, there are some areas of the country where a greater number of wind turbines would result in a further reduction in electricity prices.

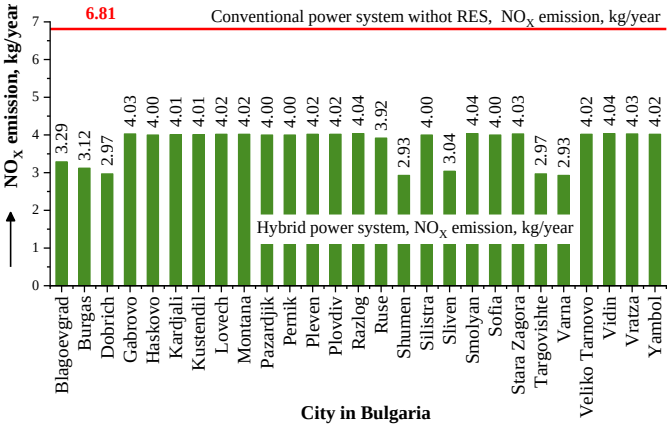


Figure 12: Comparison of NOx (kg/year) emissions by Bulgarian regions for conventional grid-connected and hybrid power systems.

The HOMER Pro software has been successfully used to anticipate greenhouse gas emissions from grid-connected single-family houses in Bulgaria.

When a hybrid power system is used in conjunction with proper energy source and storage management, harmful CO_2 emissions can be reduced by up to 59% in some parts of the country. Similarly, NO_X generation reductions in some sections of the country might reach 57%, which is an excessively high percentage.

To summarize, the software package Homer Pro may be used effectively for quickly assess harmful greenhouse gas emissions, both in Bulgaria and throughout the European Union's territory, which is particularly important for transboundary air pollution.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

The authors would like to thank the National Scientific Program NNP "*Low Carbon Energy for Transport and Households - Eplus*", D01-214/28.11.2018 for the financial support.

References

[1] D. Gospodinova, P. Dineff, K. Milanov, "Greenhouse Gas Emissions Assessment after Renewable Energy Sources Implementation in Bulgarian Grid-Connected Single-Family Houses by HOMER Pro Software," 2020 12th Electrical Engineering Faculty Conference, BulEF 2020, 2020, doi:10.1109/BULEF51036.2020.9326082.
[2] M. Kumar, Social, Economic, and Environmental Impacts of Renewable Energy Resources, IntechOpen, 2020, doi:10.5772/INTECHOPEN.89494.
[3] K.P. Tsagarakis, A. Mavragani, A. Jurelionis, I. Prodan, T. Andrian, D. Bajare, A. Korjakins, S. Magelinskaite-Legkauskiene, V. Razvan, L. Stasiuliene, "Clean vs. Green: Redefining renewable energy. Evidence from Latvia, Lithuania, and Romania," Renewable Energy, **121**, 412–419, 2018, doi:10.1016/J.RENENE.2018.01.020.
[4] N.K.M.A. Alrikabi, "Renewable Energy Types," Journal of Clean Energy Technologies, 61–64, 2014, doi:10.7763/JOCET.2014.V2.92.
[5] IRENA, Overview of renewable energy, Overview of Renewable Energy, 27, Sep. 2021.
[6] U.S.E.I.A. (EIA), Renewable energy explained - types and usage, 2021.
[7] J. Kantenbacher, R. Shirley, Renewable Energy: Scaling Deployment in the United States and in Developing Economies, Butterworth-Heinemann: 89–109, 2018, doi:10.1016/B978-0-12-813964-6.00005-7.
[8] I. Yahyaoui, Renewable Energies and Irrigation, Elsevier: 1–14, 2017, doi:10.1016/B978-0-12-812039-2.00001-6.
[9] N. El Bassam, "Energy basics," Distributed Renewable Energies for Off-Grid Communities, 113–122, 2021, doi:10.1016/B978-0-12-821605-7.00005-2.
[10] European Environment Agency, Air pollution, 2020.
[11] Enerdata, World Energy Consumption Statistics | Enerdata, Global Energy Consumption Fell in 2020 (-4%), Due to Lockdown Measures and Transport Restrictions, 2021.
[12] United Nations, Paris agreement, 2015.
[13] European Commission, 2050 long-term strategy, 2018.
[14] European Commission, Fourth Biennial Report from the European Union Required under the United Nations Framework Convention on Climate Change, Brussels.
[15] European Commission, Energy label and ecodesign, About the Energy Label and Ecodesign, 2020.
[16] European Commission, A Europe that protects: Clean air for all.
[17] Y. Yongchang, Y. Chuanan, "Implementation of a MPPT Controller Based on AVR Mega16 for Photovoltaic Systems," Energy Procedia, **17**, 241–248, 2012, doi:10.1016/j.egypro.2012.02.090.
[18] A. Mohanty, M. Viswavandya, D.K. Mishra, P.K. Ray, S. Pragyan, "Modelling & Simulation of a PV Based Micro Grid for Enhanced Stability," Energy Procedia, **109**, 94–101, 2017, doi:10.1016/j.egypro.2017.03.060.
[19] K. Adam, V. Hoolohan, J. Gooding, T. Knowland, C.S.E. Bale, A.S. Tomlin, "Methodologies for city-scale assessment of renewable energy generation potential to inform strategic energy infrastructure investment," Cities, **54**, 45–56, 2016, doi:10.1016/J.CITIES.2015.10.015.
[20] A. Gasparatos, C.N.H. Doll, M. Esteban, A. Ahmed, T.A. Olang, "Renewable energy and biodiversity: Implications for transitioning to a Green Economy," Renewable and Sustainable Energy Reviews, **70**, 161–184, 2017, doi:10.1016/J.RSER.2016.08.030.
[21] H. Ritchie, M. Roser, Energy, 2020.
[22] Statista, Electricity: share of renewable energy 2007-2018 | Statista, Share of Energy from Renewable Sources in Electricity Generation in Bulgaria from 2007 to 2018, 2021.
[23] European Commission, Europe 2020 targets: statistics and indicators for Bulgaria | European Commission, Europe 2020 Targets: Statistics and Indicators for Bulgaria, 2020.
[24] National Statistical Institute, Share of renewable energy in gross final energy consumption | National statistical institute, 2021.
[25] A.A. Prasad, R.A. Taylor, M. Kay, "Assessment of solar and wind resource synergy in Australia," Applied Energy, **190**, 354–367, 2017, doi:10.1016/J.APENERGY.2016.12.135.
[26] H. Zhao, Q. Wu, S. Hu, H. Xu, C.N. Rasmussen, "Review of energy storage system for wind power integration support," Applied Energy, **137**, 545–553, 2015, doi:10.1016/J.APENERGY.2014.04.103.

[27] A. Berrada, K. Loudiyi, "Operation, sizing, and economic evaluation of storage for solar and wind power plants," Renewable and Sustainable Energy Reviews, **59**, 1117–1129, 2016, doi:10.1016/J.RSER.2016.01.048.

[28] O.M. Toledo, D. Oliveira Filho, A.S.A.C. Diniz, "Distributed photovoltaic generation and energy storage systems: A review," Renewable and Sustainable Energy Reviews, **14**(1), 506–511, 2010, doi:10.1016/J.RSER.2009.08.007.

[29] F. Baghdadi, K. Mohammedi, S. Diaf, O. Behar, "Feasibility study and energy conversion analysis of stand-alone hybrid renewable energy system," Energy Conversion and Management, 2015, doi:10.1016/j.enconman.2015.07.051.

[30] Statista, Worldwide - lithium ion battery pack costs | Statista, Lithium-Ion Battery Pack Costs Worldwide between 2011 and 2030, 2021.

[31] M.S. Ziegler, J.E. Trancik, "Re-examining rates of lithium-ion battery technology improvement and cost decline," Energy & Environmental Science, **14**(4), 1635–1651, 2021, doi:10.1039/D0EE02681F.

[32] I. Miranda, N. Silva, H. Leite, "A Holistic Approach to the Integration of Battery Energy Storage Systems in Island Electric Grids with High Wind Penetration," IEEE Transactions on Sustainable Energy, **7**(2), 775–785, 2016, doi:10.1109/TSTE.2015.2497003.

[33] S. Mazzola, C. Vergara, M. Astolfi, V. Li, I. Perez-Arriaga, E. Macchi, "Assessing the value of forecast-based dispatch in the operation of off-grid rural microgrids," Renewable Energy, **108**, 116–125, 2017, doi:10.1016/J.RENENE.2017.02.040.

[34] C. Zhang, Y. Xu, Z.Y. Dong, J. Ma, "Robust Operation of Microgrids via Two-Stage Coordinated Energy Storage and Direct Load Control," IEEE Transactions on Power Systems, **32**(4), 2858–2868, 2017, doi:10.1109/TPWRS.2016.2627583.

[35] H.H. Abdeltawab, Y.A.R.I. Mohamed, "Market-oriented energy management of a hybrid wind-battery energy storage system via model predictive control with constraint optimizer," IEEE Transactions on Industrial Electronics, **62**(11), 6658–6670, 2015, doi:10.1109/TIE.2015.2435694.

[36] M.R. Sandgani, S. Sirouspour, "Coordinated Optimal Dispatch of Energy Storage in a Network of Grid-Connected Microgrids," IEEE Transactions on Sustainable Energy, **8**(3), 1166–1176, 2017, doi:10.1109/TSTE.2017.2664666.

[37] H.P. Software, HOMER Pro, 2020.

[38] D. Gospodinova, K. Milanov, M. Minchev, P. Dineff, "Techno-Economic Feasibility Analysis of Nearly-Zero Hybrid Energy System for the City of Sofia in Bulgaria," in 2019 11th Electrical Engineering Faculty Conference, BulEF 2019, 2019, doi:10.1109/BulEF48056.2019.9030697.

[39] H. Sharma, É. Monnier, G. Mandil, P. Zwolinski, S. Colasson, "Comparison of environmental assessment methodology in hybrid energy system simulation software," Procedia CIRP, **80**, 221–227, 2019, doi:10.1016/j.procir.2019.01.007.

[40] D. Gospodinova, K. Milanov, P. Dineff, "Route map for renewable energy sources implementation for household in Bulgaria," in IOP Conference Series: Materials Science and Engineering, IOP Publishing: 012016, 2020, doi:10.1088/1757-899X/878/1/012016.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 368-375 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Personalized Serious Games for Improving Attention Skills among Palestinian Adolescents

Malak Amro[1] , Stephanny VicunaPolo[2], Rashid Jayousi[1], Radwan Qasrawi[1,*]

[1]Department of Computer Science, Al Quds University, Jerusalem, 9103401, Palestine

[2]The Center of Innovation Technology, Al Quds University, Jerusalem, 9103401, Palestine

ARTICLE INFO

Article history:
Received: 15 July, 2021
Accepted: 09 August, 2021
Online: 26 August, 2021

Keywords:
Attention
Understanding
Cognitive skills
Serious Games
Game-based learning
Digital Games
Education outcomes

ABSTRACT

Serious games (SGs) are interactive and entertaining digital games with a special educational purpose. Studies have shown that SGs are effective in enhancing educational skills. Cognitive skills training through serious games have been used in improving students learning outcomes. In this article, we introduce the 'plants kingdom' serious game for improving adolescents' cognitive skills, mainly attention (Focus, selection, and sustained attention) and understanding skills. The game used the grade 8 Science book in designing the game content. The plant kingdom lesson was used for developing the game story and objects, its methods and tools were designed for the purpose of attention and understanding skills improvement. The game was evaluated on 43 students from public schools between the ages of 13-15 years, the study selected data from the students who had completed 5 playing sessions. The attention and understanding skills were assessed using the automatic recording and analysis of the game player's data. The variables utilized from the players' data included player ID, session number, gender, number of trials, level, drag and drop time, distance, reason for failure, position, speed, status, time, and playing tool. Results showed that the game improved the attention and understanding skills of students by 27% and 25 % respectively. The study showed the significant effect of serious games in enhancing students' cognition; thus, integrating serious games into the education system can potentially improve learning objectives and outcomes.

1. Introduction

In a rapidly changing digital environment, and the ever-growing number of persons interacting with digital technologies daily, Digital Games Based Learning (DGBL) have gained momentum in their application as learning aids for informal and formal education. DGBL provides users with the ability to improve decision-making, memory, mathematical, and spatial skills [1]. Digital Games that are used with the aim of educating, instructing, or training have come to be known as serious games (SG) [2]. Serious games were first defined by [1] as activities that "unite the seriousness of thought and problems that require it with the experimental and emotional freedom of active play"[1].

SGs differ from computer games in purpose, while the computer or digital games aim to entertain, serious games seek to achieve a certain educational, medical, or social outcome for the players. SGs have been used in a variety of fields, such as to aid children with Autism Spectrum Disorders (ASD) with Social Emotional Learning (SEL) [3], training the eyes of children with oculomotor dysfunction (OMD) [4], and cognitive training for chronic stroke survivors[5].

Several researchers have studied a further application of serious games whereby games are used to enhance educational skills [6], [7]. Serious games are directly correlated with an increase in the engagement, attraction, and interest of players in learning new skills[8]. Likewise, SGs may also enhance cognitive skills, such as selective attention, an important skill for students' academic performance [9]. Serious games show to increase players' motivation and adaptability to new educational content as the games can create scenarios that aren't easily accessible to them[10].

The ability to focus mental resources on the information most relevant at a given moment is referred to as attention. Students' attention skills are volatile and fluctuate subject to the environment they are in, which may hinder their learning skills or overall academic performance. Nonetheless, serious games allow for the creation of personalized training experiences that match

*Corresponding Author: Radwan Qasrawi, Email: radwan@staff.alquds.edu

the student's abilities and training needs and thus aid in the development of skills that enhance their cognitive functions, namely attention [11], [12].

2. Literature Review

Given that serious games aim to utilize new gaming technologies to enhance the skills of users in a particular field [12], this paper aims to study the application of serious games in adolescents' learning skills, particularly attention skills, to improve their educational outcomes. Several studies have explored the correlation of serious gaming and learning skills, and most recently, researchers have begun to focus on specific skills, such as attention and memory in a variety of sectors and applications [13], [14].

2.1. Serious Games and Learning Skills

In the article "Serious Play: Literacy, Learning and Digital Games" [15], the author discussed the Serious Play Project implemented by three Australian universities across two Australian states. The project aimed to learn about how digital games may be incorporated into education, namely the opportunities that games provide for creativity and innovation, and how learning through games challenges multimodal literacy learning. The project introduced games to classrooms between years 1 through 10 in a wide range of subjects, such as Information Technology, English, Literacy, and Social Studies among others. The findings show that serious games support learning skills in all subject areas, particularly in formal curriculum areas like Mathematics and Geography. The games enhanced students' interpersonal collaboration, negotiation, and autonomy.

In a previous study, a new model for the automatic collection of players' data and analysis of their skills was introduced by [8]. The model used game-learning analytics and robotic process automation for data collection and analysis, the model was tested on a sample of Palestinian public-school students between the ages of 13 and 14 (eighth grade). An automatic analysis model was created within a game focused on improving adolescent's attention and understanding of a given topic. The study found that serious games have the potential to be used as educational tools to ultimately aid in students' learning process and attention skills. Furthermore, the study found that the merging of game analytic tools and robotic process automation can be replicated in serious game development in order to measure players' enhanced skills. Thus, this approach to serious games could replace the traditional pre-post testing methodology currently used in such interventions.

The influence of serious games on formal education by examining their design and deployment was explored by [16]. Although the majority of studies on the topic of SGs have focused on their application in informal learning. In [16] researchers sought to develop a framework for introducing SGs into the formal pedagogical curriculum which pays special attention to the role of the educator. The review finds that serious games have the potential to significantly contribute to students' learning skills, yet, taking into account that game-based environments are rapidly evolving, educators and practitioners must be trained and prepared for the incorporation of new game-based teaching methodologies.

2.2. Serious Games and Attention (Cognitive skills)

In [17], researchers introduced a design for a computerized serious game to increase cognitive skills among the elderly "Smart Thinker." The game aimed to increase cognitive skills, such as memory and attention in elderly adults to fortify their cognitive performance. Their study consisted of 59 older adults playing Smart Thinker under the supervision of social workers. The findings show that the game, "Smart Thinker," had a significant impact in regards to the enhancement of cognitive and attention skills of the intervention group. Thus, serious games show to improve memory and attention skills among the elderly.

Similarly, in [5], the authors aimed to assess the value of "Neuro-World," a serious game, in cognitive training for stroke survivors who, as a result, present mild or moderate cognitive disabilities. The study utilized a hybrid model between ANOVA and Tukey's posthoc tests and found that all outcomes presented significant advancements except for language. Therefore, Neuro-World showed to enhance the cognitive function and decrease depression symptoms among the subjects in the intervention group [5].

In [11], researchers sought to study the assessment step of an SG for attention enhancement. A serious game for enhancing attention skills based on an Open Learned Model (OLM) was developed (Keep Attention), and a study was conducted to evaluate to what extent the OLM influences users' decision-making in attention training [11]. The study concluded that the Open Learner Model is effective in personalizing the user's experience, guaranteeing transparency, and helping users self-regulate their skills.

Furthermore, in [10], the authors took it one step further by approaching SGs for improving cognitive functions through the lens of Virtual Reality (VR). Thus, their study explored Virtual Reality Serious Games (VRSGs) to optimize the cognitive performance of users. To accomplish this, they utilize a combined 'Learning Mechanics –Game Mechanics' (LM-GM) model and further add VR characteristics. The study found that VRSGs improve the presence, immersion, and cognitive performance of users as the VR component reinforces players' embodied cognition through a purposeful and interactive design[10]., VR offers additional gaming experiences that don't translate to traditional computerized games.

In [18], the authors reviewed the use of educational serious games in all educational levels with a focus on knowledge, learning, memory, and attention. The review of the literature shows that a large number of researchers support the use of ICTs [19], in particular serious games, for enhancing attention skills and overall improving academic performance. Moreover, in [1], the authors showed that players demonstrate better selective attention over space, and can focus on one specific object at a given time with less effort than non-players as a result of their gaming activity. Thus, the study concludes that video games, particularly serious games, have the potential of significantly enhancing spatial skills, promote communication, improve memory and attention skills in students from all grade levels, but in particular in children with Attention Deficit Disorders (ADD).

The literature reviewed shows a consensus among researchers about the positive correlation of serious games with learning skills,

in particular skills that relate to cognition, such as memory and attention [20]. The impact of SGs in attention skills is increasingly applicable in school-aged children, particularly children with cognitive or intellectual disabilities.

3. Research Methodology

3.1 Serious Game Design

In this research, we developed an educational serious game for improving the basic cognitive processes of memory and attention, such as focus, orientation, recall, and selection, among school children. The game's design provides students with a series of training activities through game playing that would lead to the enhancement of their cognitive abilities, particularly attention skills. The game used a mixture of instructional and interactive video-game-play methods to achieve the proposed objectives.

The "Plant Kingdom" SG was developed based on Bloom's hierarchy of cognitive learning [21]. Therefore, the game seeks to aid students in navigating several levels to achieve increased attention skills. The research focused on improving the attention and understanding cognitive skills of adolescents, including logical thinking, intellectual, awareness, observation, knowledge, interaction, intuition, and decision-making. The game content was planned in a way to stimulate the player's recall, concentration, and memory skills through visual and audio activities, and motivation and rewards actions.

The grade eight science book was used in developing the serious game's playing materials and objects. The course learning outcomes were considered as a reference for the game objectives' design. The SG includes eight playing levels with ascending difficulties designed based on the desired skills with the help of gaming tools such as mazes, mix-match, basket-filling, and maps. The levels' difficulty was adjusted by altering game variables, including touch sensitivity, length and complexity of the game pattern, game tool, response time, success and failure, and reasons for failure.



Figure 1: The left side indicates lesson subtopics in Arabic (4 titles: plants characteristics, physical materials, structures of plants, and plants classification. The right side shows plant pictures as described in the students' textbook.

The serious game application utilized the learning contents of the grade 8 science course (specifically: the plant kingdom lesson in the students' science course curriculum) and furthered modeled them in the framework of skills development. To achieve this, objects' shapes, colors, motion, size, location, and other related human-computer interaction standards were considered in the game application development. In addition, the development team considered brain processing measurement features during game design, such as response, response time, and decision making. A group of experts (serious game experts, teachers, and education supervisors) worked on the creation of the game story, scenarios, objects, and characters.



Figure 2. Two game levels (simple and complex). Simple level (left-side) used a standard plan classification tree, and the complex level (right-side) used a recall method of the classification order of plants.



Figure 3. Two game levels with different playing tools. The Maze (left-side) drags plant properties from the upper list onto the corresponding baskets through the maze playing tool. The Tree (right-side) drags the plant properties to the correct location using finger touch movements

Figures 1, 2, and 3 show a sample of game-playing tools that have been used in game design to enhance the development of cognitive skills among adolescents.

3.2 Game Development

The "Plant Kingdom" game was developed following the 'normal serious game' design methodology that achieved the game objectives in enhancing adolescents' cognitive skills. Figure 4 illustrates the methodological steps undertaken during game design, from the definition of game objectives until the game's evaluation. Steps 1-5 include the development of the game's story and scenarios, which reflect the plant kingdom textbook material. The objects and pictures were selected to match those shown in the student textbook.

The scenarios followed the teaching material's learning outcomes and were interactively designed to enhance students' motivation within the learning environment. Step 6 concerns the game tools design, in which Mazes, drag and drop, and mix and were used for the study's purposes. These tools are commonly found in games and are considered efficient tools for enhancing players' concentration, planning, and visual-motor skills integration. Touch sensitivity and movement speed have also been considered in the game design. Sensitivity and speed complexity become increasingly difficult with higher game levels. The final step concerns game evaluation, whereby game player data were automatically collected and managed using robotic process automation and automatic analytical tools for evaluating the serious game objectives [8].

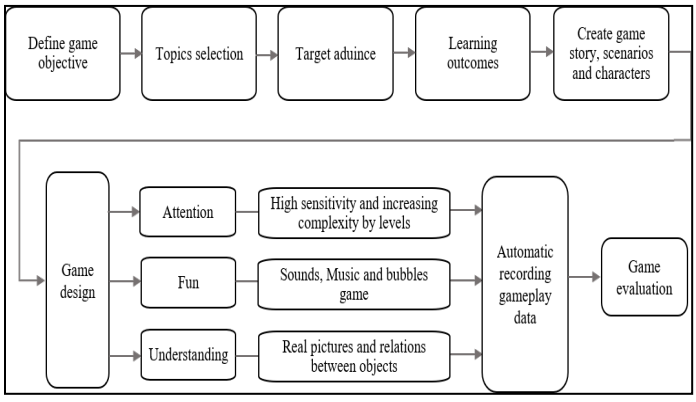


Figure 4: Methodological steps of game design

Utilizing the standard game design procedure, the game development stage for this study included the following aspects: storytelling, goal, feedback, time, rewards, instructional methods, and rules. Storytelling was employed in different ways, such as drag and drop and Mix and match. The maze itself was designed as an aid for players' attention as users were forced to concentrate on the element while moving it across the screen. The elements were shuffled after a certain period of time and players had to withdraw specific images before and during shuffling. Figures 1-3 illustrate the design of game levels using the maze, drag, and drop, and mix and match.

The software development incorporated the use of the Unity 3D mobile App development environment, in addition to free software development tools, such as PHP and MySQL to maintain user records and feedback database. Figures 5 and 6 show screenshots of the game generation stage using the Unity 3D video-game engine (version 2019.3.11), PHP MyAdmin 5.1.1, and MySQL development tools. Microsoft Visual Studio 2019 was used to edit the program code in C#.

In addition, a Web-based platform was created to manage users' game data, increase visibility, and deliver information about the SG. Programming and video-game experts were responsible for game design and implementation, as per international guidelines. The prototype was validated by end-users (stakeholders from across several domains) and tested in real-life scenarios with real data in the early development stages.

3.3 Game Testing and Validation

Prior to the release of the final prototype, the game was pilot-tested twice, firstly by 3 teachers and 4 eighth-grade students, and secondly by two groups. The groups comprised the original testing sample and an additional group made up of 2 teachers and 5 eighth-grade students. Test groups were asked to play the game 'as many times as desired' for approximately 30-40 minutes dedicated to each level in order to complete all 8 levels. Pre-post data were collected upon finalization of the pilot-test, analyzed, and considered in the next version of the game. Finally, the game's final version was launched and reviewed by experts and teachers before being given to the student sample.

3.4 Data Collection

The game's target audience was eighth-grade students between the ages of 13 and 14. The study's target population was selected following the recommendations of the Palestinian Ministry of Education (MoE) supervisors. The target age was chosen given that 8th grade proves critical as students transition to adolescence and are increasingly exposed to smart technologies.

A sample of 60 students weighted by gender was selected, however, only 43 students completed the study. The rest of the students failed to complete more than 3 sessions and were



Figure 5: Sample of game programing using Unity

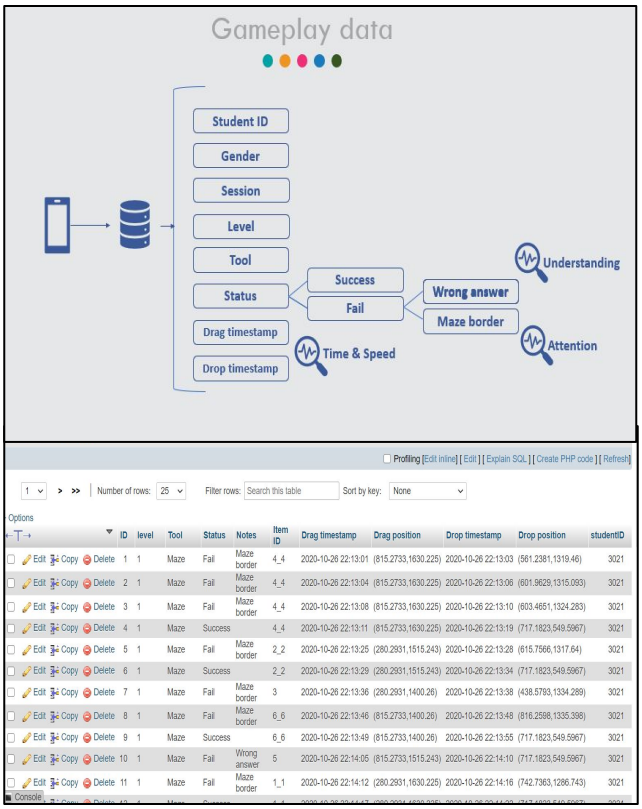


Figure 6: Sample of game programing using PHP and MySQL

therefore excluded from the analysis. The student's parents signed consent forms before gameplay. Participants were required to have access to a mobile phone with an active internet connection.

The game's application was accessible for download on the Google Play store. A research assistant guided the student in downloading the game into their smartphones. Students were asked to play the game in their homes due to COVID19 pandemic protection procedures. The students' school teachers and science teachers supervised and mentored students throughout game-playing. The player's game data was automatically stored in the game platform. The data included the variables: player ID, gender, number of trials, playing tool, session number, level number, location, drag and drop time, status (Success or Fail), the reason for failure (bumped into the maze border or wrong answer), distance, time, and speed. Moreover, attention and understanding variables were created in the following way:

a) The attention variable: Interpreted as the relationship between the amount of time it takes to complete a task, the player's status (Success or Failure), and the cause of failure.
b) The understanding variable: Interpreted as the ratio between correct and incorrect responses, the number of trials per level, the rise in correct answers, and the decline of incorrect answers.

3.5. Experiment Setting

The research took place under direct supervision and partnership with the Palestinian Ministry of Education (MoE). The MoE approved the testing of the game prototype and was responsible for the nomination of 2 science supervisors and 2 science teachers to oversee the study's implementation. The Ministry was also responsible for identifying and selecting participants. Following parents' approval upon communication with the supervisors, students attended an in-person orientation session to explain the research objectives and aid participants in downloading and installing the SG onto their mobile phones. Additionally, supervisors provided students with detailed instructions for gameplay as participants were to use the game at home. Figure 7 shows participants making use of the "Plant Kingdom" serious game.

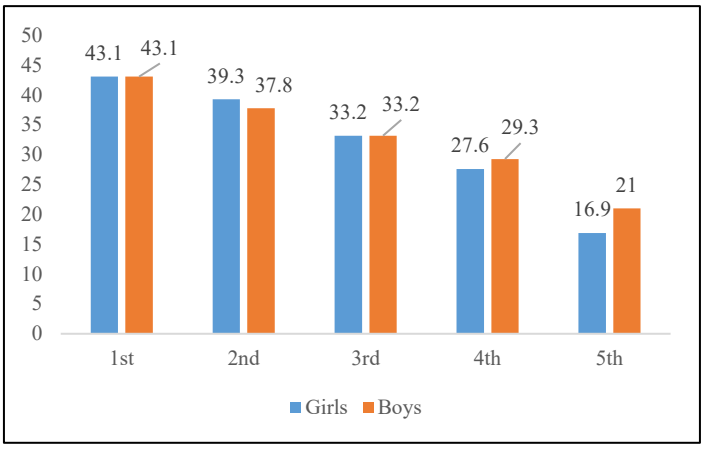




Figure 7: A sample of students playing the game.

The students were asked to play one level per day. To achieve the game objectives, the minimum required number of playing sessions was five. Out of a sample of 60 participants, 43 students successfully completed the 5 sessions required.

4. Results

4.1 Population Studied

The game was tested on 43 grade 8 students (31 girls and 13 boys) aged 13-15 years. The results are shown in Table 1 detail the percentage of participants' distribution by gender and playing sessions. Overall, the number of playing trials for the 43 players was 25,494. The results evidenced that the first session had a higher number of trials than the last session among girls (69.7% and 67.5% respectively). On the other hand, boys reported a higher number of trials in the last session than in the first session (32.5% and 30.3% respectively. However, girls presented a higher overall number of trials than boys (69% and 31% respectively);

this has been attributed to the fact that the girls to boys ratio in the sample is 31:13.

Table 1: Participants distribution of playing trials by gender and playing sessions.

		Gender		
		Girls	Boys	Total
		n (%)	n (%)	N
Session	1	4730 (69.7)	2059 (30.3)	6789
	2	3982 (69.4)	1756 (30.6)	5738
	3	3366 (69.2)	1498 (30.8)	4864
	4	3039 (68.1)	1422 (31.9)	4461
	5	2458 (67.5)	1184 (32.5)	3642
	Total	17575(69)	7919(31)	25494

4.2 Outcomes

Results in Table 2 show the average playing time (minutes) by gender and playing session. Results indicated that the average playing time decreased as the playing sessions increased for both girls and boys. The average playing time for Session 1 compared to Session 5 by gender (in minutes) was 10.7 to 5.9; 3.7 to 2.0 for both girls and boys respectively. Overall, the playing time deviation decreased by 2 minutes while the number of playing sessions increased.

Table 2: Average playing time (minutes) by gender and playing session.

		Player Gender		
		Girls	Boys	Total
Session	1	10.7	3.7	14.4
	2	8.9	3.1	12.0
	3	7.6	2.4	10.0
	4	6.7	2.5	9.2
	5	5.9	2.0	7.9

Figure 8 shows the percentage of player status (success or fail) by gender and playing sessions. The failures decreased as playing sessions increased among both girls and boys. 43.1% of both boys and girls failed in session 1, while only 16.9% of girls and 21% of boys failed in Session 5. The girls reported a failure enhancement of 26.2%, while boys reported an enhancement of 24.8%. The results indicated that girls reported an overall 83.1% success rate in the last session, while boys reported a success rate of 79.0% in the last session.

The status of participants was analyzed through the collection of individual data on the reason for failure at each level. The results shown in Table 3 indicate participants' reported reason for failure as either a wrong movement or simply a wrong answer.

Table 3 compares players by gender. In session 1 (pre-test), girls showed a higher percentage of wrong movements than boys (35.8% and 32.3% respectively). In the fifth session (post-test), girls had a greater decrease in the amount of border touch and wrong movements than boys (7% and 8.4% respectively).

Table 3: The game-play status by gender and playing session.

		Player Gender					
		Female		Male		Total	
		Wrong Touch	Wrong Answer	Wrong Touch	Wrong Answer	Wrong Touch	Wrong Answer
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Session	1	980 (35.8)	1057 (32.6)	455 (32.3)	433 (33.1)	1435 (34.6)	1490 (32.8)
	2	706 (25.8)	859 (26.5)	358 (25.4)	306 (23.4)	1064 (25.7)	1165 (25.6)
	3	494 (18.1)	623 (19.2)	256 (18.2)	242 (18.5)	750 (18.1)	865 (19.0)
	4	363 (13.3)	476 (14.7)	220 (15.6)	197 (15.0)	583 (14.1)	673 (14.8)
	5	192 (7.0)	224 (6.9)	118 (8.4)	131 (10.0)	310 (7.5)	355 (7.8)

Table 4: Average time difference of player fails per gender and playing session (minutes).

		Player Gender					
		Female		Male		Total	
		Wrong Touch	Wrong Answer	Wrong Touch	Wrong Answer	Wrong Touch	Wrong Answer
		X^2	X^2	X^2	X^2	X^2	X^2
Session	1	.15	.13	.12	.11	.14	.13
	2	.12	.13	.12	.09	.12	.12
	3	.13	.13	.10	.08	.12	.12
	4	.16	.13	.10	.10	.14	.12
	5	.18	.17	.10	.10	.15	.14

Table 5: Deviation percentage in attention and understanding using pre-posttest analysis by gender and player status.

	Girls	Boys	Total	P-Value
Attention	27.1	23.5	26	0.0001
Understanding	25.7	23.1	25	0.0001

The average playing time was reported for each level and playing session. The average time of game completion was 2.5 minutes, while the average time per session was 34 minutes. Table 4 shows the players' fail average time difference, the results show an increase in the average time difference between sessions one and five. The girls' fail time increased from 0.15 minutes to 0.18 minutes from sessions 1 to 5. Therefore, girls spent more time making use of their focus and attention skills while game-playing, which explains the decrease of fails from 35.8% to 7% as shown in Table 3. For boys, the time difference decreased from 0.12 minutes to 0.1 minutes.

The serious game effectiveness was assessed by evaluating the changes in-game players' results. The percentage of deviation in fails rate, average playing time, and fails between sessions one and five indicates the effectiveness of the proposed game in improving students' attention and understanding skills. Results in Table 5 show the deviation percentage for attention and understanding skills by gender. Attention skills improved by 26% (27.1% girls and 23.5% Boys) and understanding skills improved

by 25%. Girls reported a higher improvement than boys for both skills.

5. Discussions and Recommendations

Towards our goal of enhancing cognitive skills training among school children, we have created a serious game titled "plant-kingdom" and measured its impact on attention skills improvement through a random sample of grade 8 students. We used the game players' data for analyzing the game results and evaluating the outcomes. The average playing time per session leading to improvements in attention skills was 34 minutes a day, two times a week. The analysis of data considered every student who completed a minimum of five playing sessions. The results obtained were positive, we observed a statistically significant improvement in attention skills through the increase of average playing time, decrease in playing fails, and increase in sessions' success rate. Our results were found to be consistent with other similar studies [22, 23].

The study also yielded a statistically relevant reduction of 'fails' as playing sessions increased. Hence, the data analyzed indicated that the total number of trials decreased as the number of sessions increased. Likewise, the success rate increased while the failure rate and average playing time per level decreased. The above indicators reflect that students attained higher levels of attention and understanding of the science lesson as they continued to play the 'plant kingdom's serious game.

In the same manner, a comparison of players' outcomes indicates that the skills of sustained and selective attention grew among players as sessions increased. This is evidenced in the fact that as game-playing continued, students, girls to a larger extent, developed the ability to stay within the playing borders on the touch screen while dragging and dropping the game object.

Moreover, a notable advancement in understanding skills was observed during the answer activities. The students obtained a lesser amount of wrong answers as the game sessions progressed. Thus, the understanding of the game's content, and of the lesson plan by extension, has remarkably improved after game-play. Self-correction skills, understanding, and attention are all skills that the students had to improve on during game-play in order to decrease the 'fail' rate.

Given that the game was designed based on the students' 8th-grade Science material, the 'Plant-kingdom' serious game proved to be useful in the achievement of the course's educational outcomes. Therefore, our study's results evidence the effectiveness of serious games in improving students' cognitive skills. The skills enhancement reported in other studies [24], [25] is consistent with the results obtained. The reported improvement in players' outcomes reflects the value of serious games in educational contexts, particularly attention skills (focus, selection, and sustained attention). Furthermore, the current study, in resonance with similar previous studies [16], [26], [27], demonstrates that serious games are able to promote the enhancement of attention and understanding skills among adolescents.

Nonetheless, it is worthy to note that the study had several limitations. Firstly, with only 43 students studied, the research possessed a small sample size. Although the preliminary results presented here were enough to examine the effect of the Plant-Kingdom game in improving attention and understanding skills, we recommend the replication of this study with a larger and more significant population.

Secondly, the experiment was conducted under COVID-19 restrictions and Ministry of Health guidelines, therefore challenging the safe and correct implementation of the research. For this reason, a virtual platform was used for game supervision, thus limiting the game's evaluation. We recommend future researchers deliver their proposed intervention in person to facilitate the evaluation process, and supervise the correct use of the serious game, especially for children.

Finally, the game used the automatic recording and analysis of game players' data instead of the traditional pre-post testing measure. Thus, the number of variables used in the evaluation of game effectiveness was limited. We recommend that future studies consider the evaluation of a larger number of variables

6. Conclusions

Firstly, the preliminary results presented in this study could be used in larger-scale research to further investigate the impact of SGs in achieving learning outcomes and enhancing student's cognitive skills development. The expansion of this research could bring about innovative change to the field of education and technology.

The unification of serious games in formal education is just at its onset, but as evidenced in this research, could be greatly beneficial for students' learning outcomes. Thus, this study concludes that utilizing serious games as a learning tool informal education settings (i.e. schools, academies, universities), would enhance students' cognitive skills and improve their learning outcomes by extension, all the while providing a fun, innovative, and interactive learning methodology.

Furthermore, the serious game methodology in formal education could prove even more beneficial than evidenced in this paper for students diagnosed with Attention Deficit Disorders, Autism, or other learning disabilities.

7. Abbreviations

Digital Games Base Learning (DGBL), Serious Games (SG), Autism Spectrum Disorder (ASD), Social Emotional Learning (SEL), Oculomotor Dysfunction (OMD), Open Learner Model (OLM), Virtual Reality (VR), Virtual Reality Serious Games (VRSG), Learning Mechanics- Game Mechanics (LM-GM), Attention Deficit Disorders (ADD), Hypertext Pre-Processor (PHP), Ministry of Education (MoE).

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to thank and acknowledge all participants, their families, and the schools that contributed to the success of this

study. The authors acknowledge the role of the l thanks to the Palestinian Ministry of Education, the INTETECH company, and Al Quds University for aiding in the implementation of this research project.

References

[1] G. Papanastasiou, A. Drigas, C. Skianis, M.D. Lytras, "Serious games in K-12 education: Benefits and impacts on students with attention, memory and developmental disabilities," Program, **51**(4), 424–440, 2017, doi:10.1108/PROG-02-2016-0020.

[2] P. Wouters, H. van Oostendorp, Overview of Instructional Techniques to Facilitate Learning and Motivation of Serious Games, 2017, doi:10.1007/978-3-319-39298-1_1.

[3] C. Wu, Q. Zheng, Simulation and Serious Games for Education, 2016.

[4] I. Heldal, C. Helgesen, Q. Ali, D. Patel, A.B. Geitung, H. Pettersen, "Supporting school aged children to train their vision by using serious games," Computers, **10**(4), 2021, doi:10.3390/computers10040053.

[5] H.T. Jung, J.F. Daneault, T. Nanglo, H. Lee, B. Kim, Y. Kim, S.I. Lee, "Effectiveness of a serious game for cognitive training in chronic stroke survivors with mild-to-moderate cognitive impairment: A pilot randomized controlled trial," Applied Sciences (Switzerland), **10**(19), 2020, doi:10.3390/APP10196703.

[6] J.I. Navarro, E. Marchena, C. Alcalde, G. Ruiz, I. Llorens, M. Aguilar, "Improving attention behaviour in primary and secondary school children with a Computer Assisted Instruction Procedure," International Journal of Psychology, **38**(6), 359–365, 2003, doi:10.1080/00207590244000042.

[7] D.K. Ramos, H.M. Melo, "Can digital games in school improve attention? A study of Brazilian elementary school students," Journal of Computers in Education, **6**(1), 5–19, 2019, doi:10.1007/s40692-018-0111-3.

[8] R. Qasrawi, M. Amro, R. Jayousi, "Automatic analytics model for learning skills analysis using game player data and robotic process automation in a serious game for education," Proceedings - 2020 International Conference on Promising Electronic Technologies, ICPET 2020, 94–98, 2020, doi:10.1109/ICPET51420.2020.00026.

[9] T. Moore, M. Zirnsak, "Neural Mechanisms of Selective Visual Attention," Annual Review of Psychology, **68**, 47–72, 2017, doi:10.1146/annurev-psych-122414-033400.

[10] M. Alcañiz, S. Göbel, M. Ma, M. Fradinho, J. Baalsrud, H. Tim, M. Eds, D. Hutchison, Serious Games, 2017, doi:10.1007/978-3-319-70111-0.

[11] N. Hocine, "Personalized Serious Games for Self-regulated Attention Training," ACM UMAP 2019 Adjunct - Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization, (ExHUM), 251–255, 2019, doi:10.1145/3314183.3323458.

[12] J. Mishra, D. Bavelier, A. Gazzaley, "How to Assess Gaming-Induced Benefits on Attention and Working Memory," Games for Health Journal, **1**(3), 192–198, 2012, doi:10.1089/g4h.2011.0033.

[13] H.-L. Chan, W. V. Giannobile, R.M. Eber, J.P. Simmer, J.C. Hu, " Characterization of Periodontal Structures of Enamelin -Null Mice ," Journal of Periodontology, **85**(1), 195–203, 2014, doi:10.1902/jop.2013.120651.

[14] N. Akshoomoff, "Selective attention and active engagement in young children," Developmental Neuropsychology, **22**(3), 625–642, 2002, doi:10.1207/S15326942DN2203_4.

[15] سینا ا., No Title1386 ,قانون در طب.

[16] S. Arnab, R. Berta, J. Earp, F. de Sara, M. Popescu, M. Romero, I. Stanescu, M. Usart, "Framing the adoption of serious games in formal education," Electronic Journal of E-Learning, **10**(2), 159–171, 2012.

[17] H. Chi, E. Agama, Z.G. Prodanoff, "Developing serious games to promote cognitive abilities for the elderly," 2017 IEEE 5th International Conference on Serious Games and Applications for Health, SeGAH 2017, 2017, doi:10.1109/SeGAH.2017.7939279.

[18] G. Papanastasiou, A. Drigas, C. Skianis, M.D. Lytras, "Serious games in K-12 education," Program, **51**(4), 424–440, 2017, doi:10.1108/PROG-02-2016-0020.

[19] I.D. Danilov, "Geological and paleoclimatic evolution of the Arctic during Late Cenozoic time," The Arctic Seas, **10**, 759–760, 1989, doi:10.1007/978-1-4613-0677-1_27.

[20] V. Bolón-Canedo, A. Alonso-Betanzos, Intelligent Systems Reference Library 147 Recent Advances in Ensembles for Feature Selection.

[21] H.A. Efe, R. Efe, "Evaluating the effect of computer simulations on secondary biology instruction: An application of Bloom's taxonomy," Scientific Research and Essays, **6**(10), 2137–2146, 2011, doi:10.5897/sre10.1025.

[22] C. Alonso-Fernández, I. Martínez-Ortiz, R. Caballero, M. Freire, B. Fernández-Manjón, "Predicting students' knowledge after playing a serious game based on learning analytics data: A case study," Journal of Computer Assisted Learning, **36**(3), 350–358, 2020, doi:10.1111/jcal.12405.

[23] Y. Wang, P. Rajan, C.S. Sankar, P.K. Raju, "Let Them Play: The Impact of Mechanics and Dynamics of a Serious Game on Student Perceptions of Learning Engagement," IEEE Transactions on Learning Technologies, **10**(4), 514–525, 2016, doi:10.1109/tlt.2016.2639019.

[24] C. Boletsis, S. McCallum, "Smartkuber: A Serious Game for Cognitive Health Screening of Elderly Players," Games for Health Journal, **5**(4), 241–251, 2016, doi:10.1089/g4h.2015.0107.

[25] M.H. Chen, W.T. Tseng, T.Y. Hsiao, "The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis," British Journal of Educational Technology, **49**(1), 69–77, 2018, doi:10.1111/bjet.12526.

[26] P. Cardoso-Leite, D. Bavelier, "Video game play, attention, and learning: How to shape the development of attention and influence learning?," Current Opinion in Neurology, **27**(2), 185–191, 2014, doi:10.1097/WCO.0000000000000077.

[27] J.M. Halperin, D.J. Marks, A.C. V. Bedard, A. Chacko, J.T. Curchack, C.A. Yoon, D.M. Healey, "Training Executive, Attention, and Motor Skills: A Proof-of-Concept Study in Preschool Children With ADHD," Journal of Attention Disorders, **17**(8), 711–721, 2013, doi:10.1177/1087054711435681.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 376-384 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques

Zhiyuan Chen*, Howe Seng Goh, Kai Ling Sin, Kelly Lim, Nicole Ka Hei Chung, Xin Yu Liew

School of Computer Science, University of Nottingham Malaysia, Semenyih, 43500, Malaysia

ARTICLE INFO

Article history:
Received: 24 June, 2021
Accepted: 10 August, 2021
Online: 26 August, 2021

Keywords:
Agriculture Commodity Price Prediction
Machine Learning
Long Short-Term Memory Model
Mean Square Error
ARIMA
SVR
Prophet
XGBoost

ABSTRACT

The intention of this research is to study and design an automated agriculture commodity price prediction system with novel machine learning techniques. Due to the increasing large amounts historical data of agricultural commodity prices and the need of performing accurate prediction of price fluctuations, the solution has largely shifted from statistical methods to machine learning area. However, the selection of proper machine learning techniques for automated agriculture commodity price prediction still has limited consideration. On the other hand, when implementing machine learning techniques, finding a suitable model with optimal parameters for global solution, nonlinearity and avoiding curse of dimensionality are still biggest challenges. In this research, we address these problems by conducting a machine learning strategy study and propose a web-based automated system to predict agriculture commodity price. In the two series experiments, five popular machine learning algorithms, ARIMA, SVR, Prophet, XGBoost and LSTM have been compared with large historical datasets in Malaysia. The results validate the efficiency of the proposed Long Short-Term Memory Model (LSTM) to serve as the prediction engine for the proposed system. Particularly in the long-term experiment testing, the average performance of LSTM with MSE has improved 45.5% while ARIMA has dropped 74.1% and the average MSE of LSTM is 0.304 which outperformed all other four algorithms.

1. Introduction

The increasing availability of large amounts of agricultural commodity prices historical data and the need of performing accurate predicting of price fluctuations in agricultural economy demands the definition of robust and efficient techniques, which are able to infer from current observations. The traditional way to solve the prediction problem lies in linear statistical methods (such as ARIMA models), and more recently with the emergence of machine learning techniques, the solution has largely shifted from statistical methods to machine learning area. However, the selection of proper machine learning techniques for automated agriculture commodity price prediction still has limited consideration. On the other hand, when implementing machine learning techniques, finding optimal parameters of learning algorithm for global solution, nonlinearity and avoiding curse of dimensionality are still biggest challenges, therefore machine learning strategies studies are needed.

In practice, volatility in price of agricultural commodities is often unpredictable as they are affected by eventualities for example fluctuations of oil price, greenhouse effects and natural disasters such as flood or attacks by disease. Uncertainties of agricultural commodities price endanger the accessibility of food by consumers which leads to food insecurity and causes starvation and malnutrition. Instability of agricultural commodities price due to oversupply or lack of demand causes unnecessary food wastage. In recent years, the price of some agricultural commodities (such as palm oil and rubber) has been steadily decreased. This downward trend is making an impact to Malaysia's economy and can contribute to a slower economic growth in various investments in Malaysia. There are many exogenous factors that could cause such a trend and time-series data analysis is required to forecast this trend to improve Malaysia's agricultural plantation plan for better country development. Being able to anticipate the fluctuations and patterns in agricultural commodities price will enable the government to propose new policies that can help prevent the country into worse economy state. Further, the agricultural commodities providers are able to control their supply based on the time series analysis in order to prevent a bad plantation plan.

Most of the time series analysis on agricultural commodities are based on the US and China commodities market and there is no concrete research done about the Malaysia commodities

*Corresponding Author: Zhiyuan Chen,Email: zhiyuan.chen@nottingham.edu.my

market. Therefore, thorough agricultural commodities prices analysis should be performed based on the Malaysia prices.

In this research, we address these problems by conducting a machine learning strategy study and propose a web-based automated system to predict agriculture commodity price.

The main contribution of our work is summarized as follows:

- We formulate prediction of agriculture commodity price as a machine learning problem, which is more accurate, robust and efficient with the increasing availability of large amounts of agricultural commodity prices historical data than the traditional statistical methods.
- The proposed web-based automated agriculture commodity price prediction system with the best performance machine learning model has been presented as an advanced solution to fulfill the need of performing accurate predicting of price fluctuations in agricultural economy demands.
- Five popular machine learning algorithms, ARIMA, SVR, Prophet, XGBoost and LSTM have been compared with large historical datasets in Malaysia, which could serve as the foundation for other Malaysia commodities market research.

This paper is an extension of work originally presented in ITCC 2020 [1].

2. Background

In the earliest studies, the focus of analysis is on commodity futures markets. William G et al. discovered that futures prices were good price predictors in the corn, soybean and potato market established from empirical forecast assessment [2]. Many other subsequent studies on different agricultural futures markets in 1970's was the result of high dependency used by the researchers, who were interested in specific market conditions and traditional econometric models [3]. However, there are two obvious problems when implementing the traditional econometric models in the earliest days. First is the limited analysis due to the capacity of the model and the high computational cost [4], [5]. Second is those models are estimated through assumption that variables are independent, normal distribution, which is unrealistic in the real world [6].

Later on, several research studies have been proposed to implement agriculture price prediction scheme using different machine learning algorithms [7]-[10]. However, the performance of machine learning in agricultural commodity futures prediction is rarely explored. Future forecasting is usually done by analysing past price data of each commodity, climate, location, planting area, and several other conditions. Therefore, it is challenging because of its inherent complexity and dynamism. According to the lowest error percentage, many researchers have selected ANN and PLS as prediction algorithms.

Table 1: Summary of Related Work

Article	Agriculture Commodities	Technique
Tomek et al. [2]	Corn, Soybean, Potato	Empirical Forecast Assessment: Linear Regression
Kofi et al. [3]	Wheat, Potato	Traditional Econometric Model: Linear Regression
Sariannidis et al. [4]	Rice	Statistical Models: Autoregressive Conditional Heteroskedasticity (ARC), Generalized Autoregressive Conditional Heteroskedasticity (GARC)
Zulauf et al. [5]	Corn, Soybean	Econometric models: Price-level and Percent-change Models
Onour et al. [6]	Wheat, Rice, Beef, Groundnut, Sugar, coffee	Statistical Models: ARC/GARC models, Stochastic Volatility (SV) models.
Xiong et al. [7]	Cotton, Corn	Vector Error Correction model (VECM) – multi-output support vector regression (MSVR)
Peng et al. [8]	Cabbage, Bok choy, Watermelon, Cauliflower	Autoregressive Integrated Moving Average (ARIMA), Partial Least Square (PLS), and Artificial Neural Network (ANN)
Kumar et al. [9]	Sugarcane	K-Nearest Neighbor, Support Vector Machine (SVM), Least Squared Support Vector Machine
Manjula et al. [10]	--	Optimal Neural Network classifier (ONN)
Cao et al. [11]	--	SVM, RBF neural network
Connor et al. [12]	--	Recurrent Neural Networks
Dasgupta et al. [13]	--	Gaussian Dynamic Boltzmann Machine (DyBM) model with a recurrent neural network (RNN)
Tang et al. [14]	--	Feed Forward, Backpropagation Neural Network models, Standard Box-Jenkins model
Namaki et al. [15]	--	Deep Neural Network
Siami-Namini et al. [16]	--	Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA)
Hochreiter et al. [17]	--	Long Short-Term Memory (LSTM)
Chen [1]	Chicken, Chili, Tomato	ARIMA, SVR, Prophet, XGBoost and LSTM

In [9], the author proposed a system to apply prediction by analyzing past soil and rainfall datasets. In another research paper, Askunuri Manjula [10] has done crop prediction using multiple features, such as weather forecasting, pesticides and fertilizers and past revenue. Both of these two research works have been done the prediction with implementing pre-processing and feature reduction functions.

A few research studies have focused on implementing neural networks for prediction [11]-[13]. However most of these works are mainly about interval prediction and the point forecasting of agricultural commodity prediction has been taken less notice.

Further on, because of gradient vanishing, these existing methods fail to capture very long-term information [11]. Next, the dynamic dependencies among multiple variables are not being taken into consideration [12]. Besides that, these studies fall short in distinguishing a mixture of short-term and long-term repeating patterns explicitly [13].

During the recent years, research works of the promising deep neural networks algorithms in time series forecasting can be classified into three categories [14]. First category is to identify statistically significant events; the second is to find and predict inherent structure and the third is to do accurate prediction on numerical value [15]. Looking into time series prediction through deep neural networks, the most popular approach is Long Short Term Memory (LSTM). This approach has been highly discussed due to its promising result and the capability of not only modelling nonlinear patterns, realizing complex causal relationships, as well as the learning rate on huge historical datasets. The LSTM model is said to be more accurate than the prediction of the ARIMA model by 85% on average from the results obtained in the research by Namini, S. S, Tavakoli, N. and Namin, A. S [16]. Furthermore, LSTM were introduced and aimed for a better performance by tackling the vanishing gradient issue that recurrent networks would suffer when dealing with long data sequences [17].

Other advance technologies, such as Box–Jenkins approach [18] and irrigation monitoring systems based on IoT technologies, GPS and LoRaWAN [19] have been presented by researchers recently. However due to the application in different domain or system complexity, we will not consider these techniques in this paper but it might be useful for our future work.

A summary of all these related works can be found in Table 1. Other background information and technical review of ARIMA, SVR, Prophet, XGBoost and LSTM have been discussed in our conference paper [1].

3. Proposed System

The main objective of the proposed automated agriculture commodity price prediction system is to assist government or farmers for a better agricultural plantation plan. To achieve the objective, this research utilizes machine learning techniques to provide an agriculture price forecasting feature into the web system.

The project studies the needs of the farmers in doing agricultural activities, and these studies have been adapted into the system and be delivered in a simpler, more comprehensive way to suffice the farmers' knowledge in doing agricultural activities.

3.1. System Requirement Specification

3.1.1. Functional Requirements

The system should have a web app which consists of:

- A sign up and login page
- A forecast page to show all graph and data
- A commodity information page
- A user profile page
- An enquiry page

The system should forecast the future prices of agriculture commodities

- The forecast should be shown in a graph form
- User should be able to choose duration of prediction
- User should be able to rescale the x-axis of the graph
- The forecast should be based on previously available data
- User shall be able to access different type of model in price forecast (Univariate / multivariate)
- Graph should be updated upon type of commodity, duration of view, forecast period, and type of model

Visualization of forecast result should be clear for users

- Past prices and forecast results shall be separated with different color in one graph
- When the cursor hovers above the graph line, price information of the x-axis value shall be shown to users

Users shall able to select interested commodity

- Commodity chosen can only be pre-existing in the database
- System should update graph to show new data point based on new data added
- User shall be able to access different model type for each commodity (univariate, multivariate)

System should store and retrieve data in a database

- Developers should be able to update the database with new data
- Developers should be able to edit previous data stored in the database

Commodities data shall be downloadable

- User shall be able to download complete past prices of selected commodities in .csv file

A user profile shall be created automatically upon registration

- User shall be able to edit the user profile and save it

System shall be able to authenticate existing user

- User account shall be saved in database during registration

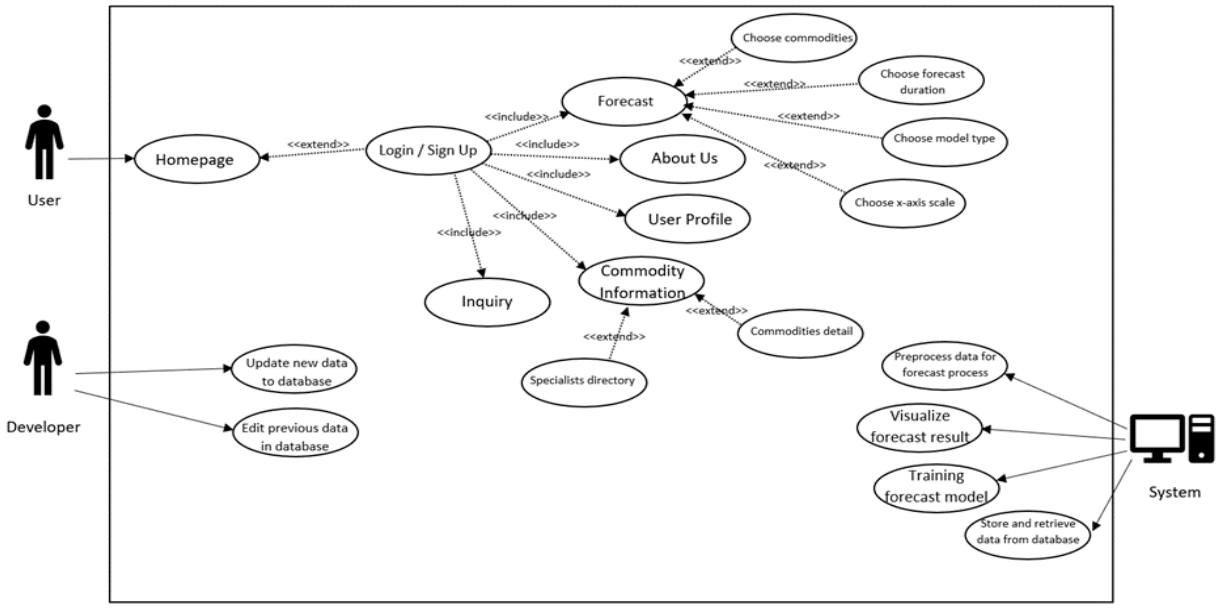
3.1.2. Non-Functional Requirements

Reliability

- The system shall be able to access the system anywhere with an internet connection
- The system shall be able to handle user queries not exceed 30 seconds when Tensorflow backend is running

Usability

- The system shall be simple to access for a first-time user
- The redirecting page amount for any feature of the system shall not exceed 5 pages



Performance

- The forecasting result shall be available for at least 7 commodities
- For each commodity, univariate model and multivariate model for forecast process shall be able to be accessed
- The meantime of download a file in csv format from a software shall not exceed 10 seconds

Security

- The system would require user account to access to all features
- User shall register their account with their email address one time only, system will reject account registration of existing account

Commodity products

- Commodities listed in the software shall be in English
- All commodities prices in the software shall be in Ringgit Malaysia, and local pricing for every commodities

Interface

- The system shall be portable in devices including personal computers, iPad and any mobile devices with Google Chrome software installed.
- The system layout shall be responsive to both full-size window and minimized window

3.2. Use Case Diagram

The use case diagram is shown in Figure 1. It explains the interactions that occur between the users and the system itself.

3.3. System Overview

Figure 2 describes the system Flow. The web system follows the Model-View-Controller (MVC) architecture. When a user enters a URL in their browser, the browser sends an http request to the web server, then the web server forwards the request to the application server and the URL setting contained in the urls.py file selects the view according to the url specified in the request, the view communicates with the database via models.py, renders the html or other format using templates (loads the static files) and returns the http response to the web server and finally, the web server provides the desired page to the browser.

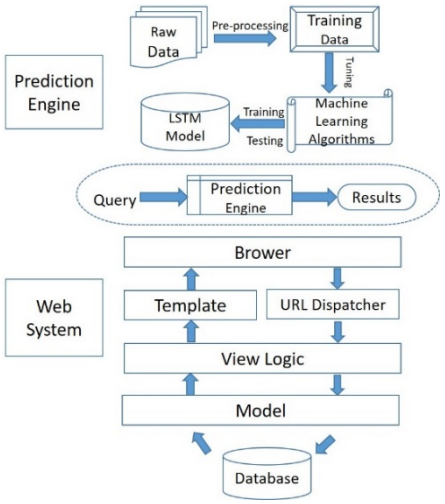


Figure 2: System Flow of the Automated Agriculture Commodity Price Prediction System

There are four processes in the prediction engine design of the proposed system: data pre-processing process, tuning process, training and testing process and decision-making process. Process 1 involves cleaning and reformatting the dataset; while process 2 focuses on defining optimal parameter value of each machine learning techniques. All machine learning models will be trained and tested in process 3. Five machine learning methods were compared in this research, which are ARIMA [20], SVR [21],

Prophet [22], XGBoost [23] and LSTM [17]. Technical details of these methods have been presented in our conference paper [1]. Finally, process 4 selects the best model out of all models to serve as the prediction engine. An algorithm snippet of the prediction engine is shown in Figure 3.

```
# Part 3 - Make the predictions and visualise the results
data_test = data.tail(52)
data_test = data_test[:]['Harga Ladang']
X_test = []
X_test.append(data_train_scaled[(len(data_train_scaled)-52):])


X_test = np.array(X_test)
X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)

predicted_price = regressor.predict(X_test)
predicted_price = sc.inverse_transform(predicted_price)
predicted_price = predicted_price.reshape(-1 ,1)
```

3.5. Implementation Practice

The main implementation of the application for machine learning algorithms to predict is Python programming language. It is an interpreted, high-level, general-purpose programming language used widely for machine learning. For the frontend design implementation of the application, Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript have been used. These are high-level languages widely used to build a design and features of web applications.

3.5.1. Collaborative Software and Version Control

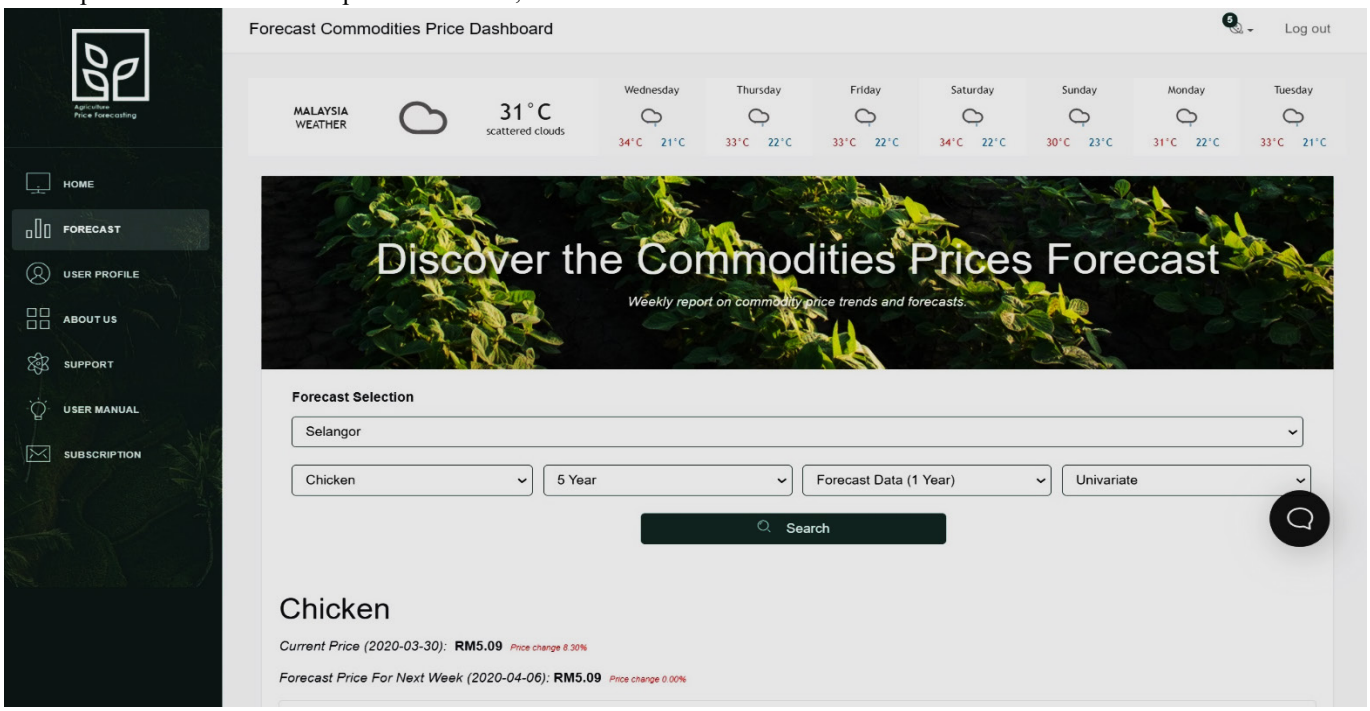
Django is the main framework used to build our web application, which is Python-based software that follows the model-template-view architectural pattern. Besides, a free cloud service which supports free GPU named Google Colab is in service to assist while implementing Long short-term memory (LSTM) algorithms in python programming language and it also supports the development for deep learning applications using popular libraries such as TensorFlow, Keras, PyTorch and so on.

Github, an open source version control website, is used as a centralized repository to keep, share, update codes and most importantly is to do version control. This platform is used as our service-oriented architecture for the application, also to manage the codebased, deliver the latest code and manage configuration of our web application.

3.5.2. Graphical User Interface (GUI)

During the initial implementation to achieve our prototype user interface (UI), we have used Hypertext Markup Language (HTML) to build the main structure of the homepage, login page and the forecasting dashboard. Thus, we modify using Cascading Style Sheet (CSS) and JavaScript to implement our proposed design accordingly.

In the latest design, there are 8 pages that build up the website, particularly, register, login, dashboard, user profile, support, weather forecast, about us and subscription page. Firstly, the user will be required to have an account in order to login the page, this will be handled in the register page for the user to create an account. Secondly, after getting an account, the user will be redirected to the login page with a simple input form for the user to log in with their account. All of the pages will require the user to login in order to view it. After logging in, there will be a side navigation bar containing each Uniform Resource Locator (URL) for relevant pages. In the forecast page as shown in Figure 4, users will also be able to select relevant commodities, data types, and duration of forecast they would prefer to see in the graph. The user profile will handle all the information that belongs to the user and



users will be able to edit and update their profiles. In the support page, users may submit their enquiries via the form provided. The about us page is a brief introduction regarding the website and its purpose and usage. In the subscription page, the user will be able to subscribe to different plans to unlock different features. Lastly, each page contains the identical top navigation bar for them to log out anytime.

4. Experiments and Results

The forecasting feature of the prediction system is the main focus of the project. In the decision of technique implemented to serve as the prediction engine, a series of experiments have been designed to discover the most optimal algorithm which should perform with highest accuracy and best capability of handling increasing data.

Based on our literature review analysis, five algorithms namely, ARIMA, SVR, Prophet, XGBoost and Long short-term memory have been selected as our potential prediction engine and their performance will be compared to learn the best model of interest from large datasets.

4.1. Datasets

Our experimental agriculture datasets are extracted from the report made by FAMA official government website (https://sdvi.fama.gov.my/analisahargamingguan/DownloadPassword.asp). This website consists of weekly data analysis reports from 2007-2020. The price datasets of different categories (December 2008 to March 2020) are selected and migrated into a raw dataset. Two series of experiments have been conducted, one is to consider only time-series dataset with 11 years' datasets (December 2008 to March 2019) and the other is with multivariable (Temperature, Humidity, Precipitation and Crude Oil Price) with the whole dataset.

Table 2 shows the statistics of the price with three selected commodities from two categories, chicken for poultry, chili and tomato for fruits. All three commodities datasets with 2% missing values.

Table 2: Price Statistics of the Raw Data

Commodities	Mean	Minimum	Maximum	StdDev
Chicken	4.84	3.50	6.25	0.52
Chili	5.92	2.90	12	1.55
Tomato	2.19	0.50	6.35	0.83

4.2. Experimental Setting

4.2.1. Configuration of ARIMA

ARIMA is a combination of two models that is Auto regressive (AR) and moving average (MA). ARIMA consists of three parameters (p,d,q), where: p is the number of autoregressive terms, d is the number of nonseasonal differences needed for stationarity, and q is the number of lagged forecast errors in the prediction equation [20]. A series of actions have been performed to preprocess the data. Firstly, an Augmented Dickey Fuller Test (ADCF) is performed to determine whether the time series is stationary and perform data transformation to assure the stationarity of the data [22]. And then a differencing method is used to transform a non-stationary time series into a stationary one, a lag of 1 has been given as the result shows that the mean and variance are more constant over time in comparison to before transformation, thus resulting in ARIMA (d,1).

AR demonstrates the correlation between the previous time period with the current. Thus, in order to predict the value of P, a partial auto-correlation function (PACF) graph will plot. This plot provides a summary of the relationship between an observation in a time series with observations at prior time steps with the relationships of intervening observations removed [25].

Inevitably, there is always noise or irregularity attached in a time series. Hence, MA has been implemented and the main purpose is to figure and average out the noises that could potentially affect the model. This has been achieved by plotting an Auto Correlation Function (ACF) autocorrelation plot. An ACF plot is an (complete) auto-correlation function which gives values of auto-correlation of any series with its lagged values, it describes how well the present value of the series is related with its past values [24][26]. From our experimental results, it is observed that the diagram shows that the graph cuts off and drops to zero for the first time on x-axis in 1 on ACF plot and 1.5 on ACF plot, thus, giving q=1 and p=1.5.

4.2.2. Configuration of SVR

To capture the non-linear pattern of the data prediction, the implemented SVR model has used Radial Basis Function as the kernel function. As a popular function used in most kernelling algorithm, it was claimed of having well performance under general smoothness assumptions [27] and this suits the situation in our study. All parameter settings remain default, however, some parameter values have been tuned in building the model, such as:

- The kernel coefficient (gamma) setting value is set as 1/total number of features to deal with various amounts of features.
- Regulation parameter(C) is set by fine-tuning to find the most suitable value for the model.
- The epsilon setting is used with default value.

To resolve the missing values, two different approaches have been implemented to the model:

- The missing value in the initial raw data - fill the missing value with valid, but not influence value (e.g. -99999) to avoid clearance of missing data affecting the prediction process.
- The missing value in the post-process raw data - clear the missing data to maintain the same amount of row in the data frame.

Training of the model has utilized the python sklearn function called train_test_split. The dataset is randomly split into the corresponding train set and test set, in this case, the ratio of training set and test set is 9:1.

4.2.3. Configuration of Prophet

The default setting of Prophet:

- growth ('linear')
- n.changepoints (n)
- changepoints.range (0.8)

- changepoint_prior_scale (m)

The number of change points n has been set as one per month. The changepoint_prior_scale is to decide how flexible the changepoints are allowed. To avoid the overfitting problem, m has to set to [10,30]. Prophet is relatively robust to missing data, so no resampling methods have been implemented. Cross-validation method has been implemented to split the date into training and testing sets.

4.2.4. Configuration of XGBoost

The default setting of XGBoost:

- objective='reg:linear',
- colsample_bytree=0.8,
- learning_rate=0.1,
- max_depth=8,
- n_estimators=1000,
- silent=1,
- subsample=0.8,
- scale_pos_weight=1,
- seed=27
- early_stopping_rounds=50

4.2.5. Configuration of LSTM

The default setting of LSTM:

- Number of epochs: 100
- Batch size: 10
- Layers: 4 LSTM layer (including input layer) each of size 50 unit and subsequent by a dropout layer of +/-0.2 based on the dataset after each layer. 1 Output dense layer of size 52.
- Optimizer and loss function: adam, mse.
- Activation: hyperbolic tangent
- Data transform: data are scaled by a minmaxscaler

4.3. Results

The final mean squared error score (MSE) for the prediction in the first series of experiments has been shown in Figure 5. The mean squared error score gives a better justification of the optimal algorithm that could be chosen as the prediction engine. From the results, ARIMA has shown as the most stable model amongst the three commodities, as it has the lowest mean square error value (0.251) on average, particularly with Chili, it gets as low as 0.027 for MSE.

From the literature review, LSTM is more superior than other machine learning models such as SVR and ARIMA in terms of accuracy and difficulties in handling the data. However, in the first series of experiments, due to the small sample size, it was not able to perform outstanding predictions to fit the system. We can conclude ARIMA has outstanding ability in handling small amounts of data in performing predictions.

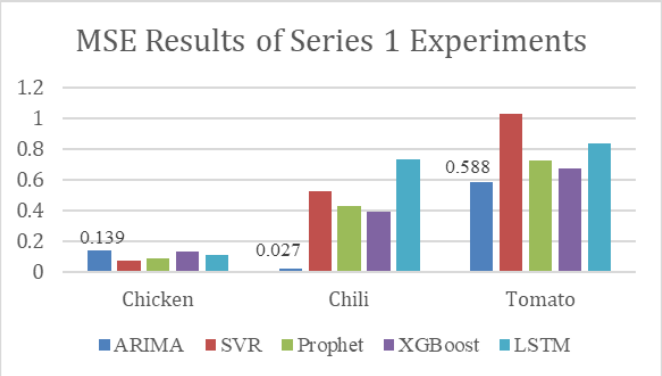


Figure 5: MSE Results of Series 1 Experiments

In order to test the longer term of agriculture commodity price prediction, which is more approximate to the real world situation, a second series of experiments have been conducted. In the second attempt of the experiment, the data has been enlarged to March 2020, and the feature space has been increased from one dimension to multi-dimensions with Temperature, Humidity, Precipitation and Crude Oil Price features being added to the model training process. Figure 6 shows the latest results of the second series of experiments.

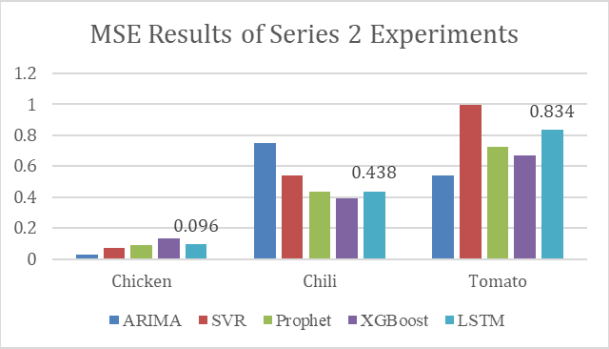
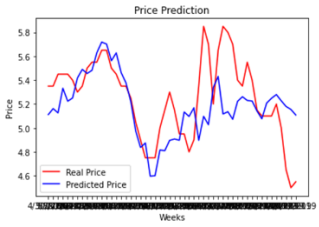


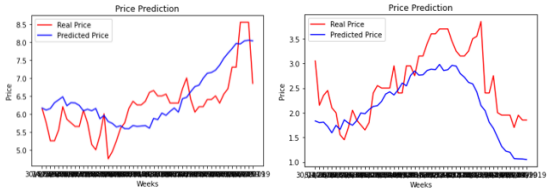
Figure 6: MSE Results of Series 2 Experiments

From the results, the average performance of LSTM with MSE has improved 45.5% while ARIMA has dropped 74.1%. The average MSE of LSTM is 0.304 which outperformed all other four algorithms. The improvement of the LSTM model shows its potential in handling increasing data and higher complexity of data compared to other models and in a long-term overview. Considering the fact that the collected data will continuously increase in terms of size and complexity, hence, LSTM is a better choice comparing to ARIMA (which shows the best result in the first series of experiment). Meanwhile LSTM's outperform capability of handling bigger dataset, as well as in the terms of system maintainability and scalability have fully justified the suitability to serve as the prediction engine for the proposed automated agriculture commodity price prediction system.

Figure 7 (a) to (c) are the graphs that illustrate the original prices, as well as the prediction prices using LSTM algorithm for 3 experimental commodities (Chicken, Chilli and tomato). The graph can help to visualise the outcome of predicted price in order to make a visual presentation to users.



(a) Chicken Price Prediction



(b) Chili Price Prediction (c) Tomato Price Prediction

Figure 7: LSTM Prediction Results

5. Conclusion

In this research, we study the challenges of prediction on price fluctuations in agricultural economy with increasing large amounts of agricultural commodity prices historical data. An accurate, robust and efficient agriculture commodity price prediction system with novel machine learning techniques have been proposed, which is able to automatically infer from current observations. The traditional way to solve the prediction problem lies in linear statistical methods (such as ARIMA models), and in this paper we present the novel and complete solution in machine learning area.

Firstly, the prediction of agriculture commodity price has been formulated as a machine learning problem. It is more accurate, robust and efficient for the increasing availability of large amounts of agricultural commodity prices historical data than the traditional statistical methods. Secondly the proposed web-based automated agriculture commodity price prediction system with the best performance machine learning model engine has been presented as an advanced solution for the need of performing accurate predicting of price fluctuations in agricultural economy demands. Thirdly five popular machine learning algorithms, ARIMA, SVR, Prophet, XGBoost and LSTM have been compared with large historical agriculture commodity price datasets in Malaysia, which could serve as the foundation for other Malaysia commodities market research.

In future, correlation analysis between different supporting factors and the historical price will be investigated in depth and the optimization of the proposed prediction engine model will be further improved. The proposed system will also enhance its trading aspect in near future, by reporting more in-depth location-specific price analysis to realize trading among farmers.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is part of a research funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2018/ICT02/UNIM/02/1).

References

[1] Z.Y. Chen, K.L.Sin, "Long Short-Term Memory Model Based Agriculture Commodity Price Prediction Application," in Proceedings of the 2020 2nd International Conference on Information Technology and Computer Communications, Association for Computing Machinery, New York, NY, USA: 43–49, 2020, doi: https://doi.org/10.1145/3417473.3417481.

[2] W.G. Tomek, R.W. Gray, "Temporal Relationships Among Prices on Commodity Futures Markets: Their Allocative and Stabilizing Roles," American Journal of Agricultural Economics, **52**(3), 372–380, 1970, doi:https://doi.org/10.2307/1237388.

[3] T.A. Kofi, "A Framework for Comparing the Efficiency of Futures Markets," American Journal of Agricultural Economics, **55**(4_Part_1), 584–594, 1973, doi:https://doi.org/10.2307/1238343.

[4] N. Sariannidis, E. Zafeiriou, "The spillover effect of financial factors on the inferior rice market," Journal of Food, Agriculture & Environment, **9**(1), 336–341, 2011, doi:https://doi.org/10.1234/4.2011.1962.

[5] C.R. Zulauf, S.H. Irwin, J.E. Ropp, A.J. Sberna, "A reappraisal of the forecasting performance of corn and soybean new crop futures," Journal of Futures Markets, **19**(5), 603618, 1999, doi:https://doi.org/10.1002/(SICI)1096-9934(199908)19:5<603::AID-FUT6>3.0.CO;2-U.

[6] I. Onour, B.S. Sergi, "Modeling and Forecasting Volatility in the Global Food Commodity Prices (Modelování a Prognózování Volatility Globálních cen Potravinářských Komodit)," Agricultural Economics-Czech, **57**(3), 132–139, 2011, doi:https://doi.org/10.17221/28/2010-AGRICECON.

[7] T. Xiong, C. Li, Y. Bao, Z. Hu, L. Zhang, "A combination method for interval forecasting of agricultural commodity futures prices," Knowledge-Based Systems,77,92–102, 2015, doi:https://doi.org/10.1016/j.knosys.2015.01.002.

[8] Y.-H. Peng, C.-S. Hsu, P.-C. Huang, "Developing crop price forecasting service using open data from Taiwan markets," in 2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI), 172–175, 2015, doi: https://doi.org/10.1109/TAAI.2015.7407108.

[9] A. Kumar, N. Kumar, V. Vats, "Efficient crop yield prediction using machine learning algorithms," International Research Journal of Engineering and Technology,5(06),3151–3159,2018. Online Access.

[10] A. Manjula, G. Narsimha, "Crop Yield prediction with aid of optimal neural network in spatial data mining: New approaches," International Journal of Information and Computation Technology, **1**(6), 25–33, 2016. Online Access.

[11] L.J. Cao, F.E.H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," IEEE Transactions on Neural Networks, **14**(6), 1506–1518, 2003, doi:https://doi.org/10.1109/TNN.2003.820556.

[12] J. Connor, L.E. Atlas, D.R. Martin, "Recurrent networks and NARMA modeling," in Advances in neural information processing systems, 301–308, 1992, doi:https://dl.acm.org/doi/10.5555/2986916.2986953.

[13] S. Dasgupta, T. Osogami, "Nonlinear Dynamic Boltzmann Machines for Time-Series Prediction," Proceedings of the AAAI Conference on Artificial Intelligence,**31**(1),2017,doi:https://dl.acm.org/doi/10.5555/3298483.3298506.

[14] Z. Tang, C. de Almeida, P.A. Fishwick, "Time series forecasting using neural networks vs. Box- Jenkins methodology," SIMULATION, **57**(5), 303–310, 1991, doi:https://doi.org/10.1177/003754979105700508.

[15] M.H. Namaki, P. Lin, Y. Wu, "Event pattern discovery by keywords in graph streams," in 2017 IEEE International Conference on Big Data (Big Data), 982–987, 2017, doi: https://doi.org/10.1109/BigData.2017.8258019.

[16] S. Siami-Namini, N. Tavakoli, A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1394–1401, 2018, doi: https://doi.org/10.1109/ICMLA.2018.00227.

[17] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural Computation,**9**(8),1735-1780,1997, doi:https://doi.org/10.1162/neco.1997.9.8.1735.

[18] A.M. Ashik, K.S.Kannan, "Time Series Model for Stock Price Forecasting in India", in Logistics, Supply Chain and Financial Predictive Analytics. Asset Analytics (Performance and Safety Management). Springer, Singapore, 2018, https://doi.org/10.1007/978-981-13-0872-7_17.

[19] D. M. Matilla, Á. L. Murciego, D. M. Jiménez Bravo, A. Sales Mendes and V. R. Quietinho Leithardt, "Low cost center pivot irrigation monitoring systems based on IoT and LoRaWAN technologies," 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), 2020, 262-267, doi: 10.1109/MetroAgriFor50201.2020.9277548.

[20] G.E.P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, Inc., USA, 1990, doi:https://dl.acm.org/doi/10.5555/574978.
[21] M. Awad, R. Khanna, Support vector regression, Springer: 67–80, 2015, doi:https://doi.org/10.1007/978-1-4302-5990-9_4.
[22] S.J. Taylor, B. Letham, "Forecasting at Scale," The American Statistician, 72(1), 37–45, 2018, doi:https://doi.org/10.1080/00031305.2017.1380080.
[23] T. Chen, C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA: 785–794, 2016, doi: https://doi.org/10.1145/2939672.2939785.
[24] Y. W. Cheung, K. S. Lai, "Lag order and critical values of the augmented Dickey–Fuller test", Journal of Business & Economic Statistics, **13**(3), 277-280, 1995,doi: https://doi.org/10.1080/07350015.1995.10524601.
[25] R. H. Shumway, D. S. Stoffer, Time series analysis and its applications: With R , Principles of fuel cells, Springer, 2006.
[26] G. E. P. Box, W. G. Hunter, J. S. Hunter, Statistics for experimenters: an introduction to design, data analysis, and model building, John Wiley and Sons, 1978.
[27] A. J. Smola, B. Schölkopf, K. R. Müller, "The connection between regularization operators and support vector kernels", Neural networks, **11**(4), 637-649, 1998, doi: 10.1016/S0893-6080(98)00032-X.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 385-394 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Scheduling Algorithm with RTiK+ for MIL-STD-1553B Based on Windows for Real-Time Operation System

Jong-Jin Kim, Sang-Gil Lee, Cheol-Hoon Lee*

Department of Computer Engineering, Chungnam National University, Daejeon, 305764, Korea

ARTICLE INFO

Article history:
Received: 01 July, 2021
Accepted: 09 August, 2021
Online: 26 August, 2021

Keywords:
RTX
RTiK+
APIC
Tablet PC
Inspection Equipment
Real-Time Operating System
Windows
MIL-STD-1553B
RM Scheduling Algorithm

ABSTRACT

In devices using Windows operating system based on x86 system, the real-time performance is not guaranteed by Windows. It is because Windows is not a real-time operating system. Users who develop applications in such a Windows environment generally use commercial solutions such as the RTX or the INtime to provide real-time performance to the system. However, when using functions and API for simple real-time processing, an issue of high development cost occurs in terms of cost-effectiveness.

In this paper, the RTiK+ was implemented in the type of a device driver by controlling the MSR_FSB_FREQ register to generate a timer interrupt independent of Windows in the Windows 8 and providing a real-time functionality to the user mode by re-setting the local APIC count register. And the operating frequency of the CPU is changed to minimize power consumption for battery life in a mobile device such as a Tablet PC.

In particular, the weapon system uses highly reliable MIL-STD-1553B communication and performs BC and RT functions of MIL-STD-1553B to transmit and/or receive data in communication between component and component. It is significantly importance to guarantee integrity of data without loss data during communication. For this purpose, it is proposed to implant the Scheduling algorithm with the RTiK+ for MIL-STD-1553B communication for Windows 8 to support a real-time processing in the Windows operating system on the embedded system, and to use the periods of 2ms (max), 5ms and 10ms provided by the RTiK+ for real-time processing when performing the BC and RT functions of MIL-STD-1553B communication. In this paper, the method of the scheduling algorithm with RTiK+ for MIL-STD-1553B to provide real-time processing is proposed for the Windows based on x86 system.

1. Introduction

1.1. Research Background

Recently, with the development of hardware and embedded systems, various mobile devices such as Tablet PC and smart phone used in daily life are being used instead of dedicated equipment for special purposes according to the user environment. Therefore, it is necessary to support an accurate real-time processing function in order to transmit/receive accurate commands or data to a mobile device without data loss of information from the target equipment. And this paper is an extension of work originally presented in Real-time Processing Method for Windows OS Using MSR_FSB_FREQ Control [1-5].

In general, a real-time system should be predictable in any situation, which means that time deterministic should be guaranteed. In this paper, to ensure time determinism, the MSR_FSB_FREQ register, which determines the operating frequency of the CPU to provide real-time processing performance in the Windows 8 environment, is controlled and the local APIC timer count register value is reset to operate a window independent timer. A study was conducted to guarantee the periodicity set by the user. Also, when checking the function and performance of a guided weapon system using a window-based inspection equipment in a guided weapon that uses fast MIL-STD-1553B communication such as 2ms between components, The RTiK+ is first installed to provide real-time performance to the Windows 8, and then it use the period provided by RTiK+ for scheduling algorithm to meet deadline in MIL-STD-1553B communication in guided weapons system, and also a study for miss rates to

*Corresponding Author: Prof. Cheol-Hoon Lee, Department of Computer Engineering, Chungnam National University, Korea, Email: clee@cnu.ac.kr

guarantee data integrity without loss is conducted through experimental tests to be proofed.

1.2. Research Purpose

In the case of a tablet PC, The Windows as operating software is mainly used as an operating system to provide compatibility with libraries that have been developed and operated, dependency on the device's operating system, and extensibility for various applications. To support a real-time processing function in the tablet PC, a commercial solution that is installed in addition to the Windows operating system installed in the tablet PC and provides a real-time functionality must be used. Such products include IntervalZero's RTX, which is currently widely used. The purchase price is very high, and when installed in equipment for mass production, royalty and maintenance costs are high, which causes a high development cost burden on developers. To improve this matter, The RTiK+ was designed to provide a real-time performance in Windows 8 based on x86 system.

In addition, the Windows operating system used for inspection equipment prioritizes compatibility, stability and reliability due to the characteristics of the weapon system, so older versions such as Windows 7 and Windows 8 are used rather than the latest versions such as Windows 10 [6]. For example, the flight body inspection equipment that recently developed an unmanned aerial vehicle developed in 2017 uses the Windows 7 released in 2009 [7].

In this paper, the RTiK+ is designed as a method by controlling the MSR_FSB_FREQ register to generate a window-independent timer interrupt in Windows 8 and providing a real-time period to the user mode by resetting the counter register of the local APIC, and the CPU operating frequency. A study how to provide a real-time performance by accessing the MSR_PKG_CST_CONFIG_ CONTROL register that determines the C-States so that the frequency controlling the C-States does not change.

A MIL-STD-1553B communication with high reliability is used mainly in the precision guided weapon system [8], and the proposed RTiK+ to provide a real-time processing function in the Windows operating system of the equipment is implanted in Windows 8, and it supports the BC and RT function of the MIL-STD-1553B to provide a real-time processing. The scheduling algorithm with RTiK+ supports a real-time period to the BC and RT function when performing communication.

Finally, an evaluation for experiment results is conducted to proof the miss rate that complies with the deadline in the set period in the MIL-STD-1553B communication which RTiK+ for a real-time processing is applied. The RTiK+ to provide a real-time processing is implanted in the Windows 8 based on x86 system, and the scheduling algorithm is applied to perform the MIL-STD-1553B communication. The integrity is verified using the RDTSC for measurement of the period of a real-time function and PASS 3200 for data monitoring of the MIL-STD-1553B, and the miss rate of the scheduling algorithm for the MIL-STD-1553B is analyzed. And to see if the RTiK+ can be used as a replacement item for the RTX, a real-time performance is measured and analyzed using the same experimental configuration for measuring performance of the RTX, showing that the RTiK+ can be used as a substitute for a third party.

The paper consists of chapters 1 to 6. Chapter 2 studies RTX, window scheduling, RM algorithm, and MIL-STD-1553B. Chapter 3 describes the real-time processing in the tablet PC proposed in this paper, and the real-time performance implementation using the MIL-STD-1553B communication scheduling algorithm. Chapter 4 describes the experimental environment, composition, and results. Chapter 5 analyzes the algorithm and results, Chapter 6 concludes by describing the summary of this paper and future research.

2. Related Research

2.1. RTX

IntervalZero released the RTX operating in kernel mode (ring 0) based on x86 computers in 1995. The RTX is software that provides a real-time performance to support a real-time function in Windows XP or Windows 7 but there is no solution for Windows 8 currently. The RTX provides a development environment that is easy to access and familiar to developers who use Windows based on a real-time operating systems, which are the advantages of Windows [9].

2.2. Local APIC

An APIC is an interrupt controller provided by the x86 architecture. It implements the function of delivering a hardware interrupt to the interrupt descriptor table with the address of the interrupt handler. A local APIC is an interrupt controller provided by the x86 architecture as an extension of the programmable interrupt controller [10-13].

2.3. Scheduling of Process and Thread in Windows

The Windows operating systems based on x86 system have processes and threads to process specific tasks. In general, the threads are divided into user-level threads and kernel-level threads. A kernel-level thread is used in Windows, and in Windows, a thread is assigned in a processor by a kernel scheduler and implements a scheduling. It is a scheduling algorithm that is scheduled according to the priority of all threads created in the operating system regardless of the priority of the process. The Windows scheduler uses a round-robin scheduling method. A process has six priorities, and a thread has seven priorities, and consists of a total of thirty-one priorities from 0 to 31 [14].

2.4. RM Scheduling Algorithm

The RM Scheduling, The RM means Rate-Monotonic, for a real-time system is a scheduling algorithm that gives the highest priority to the shortest period. The operating system to which this algorithm is applied is generally called preemptive and has a decisive characteristic for response time [15]. When there are n processes, the upper limit of total utilization on the system can calculate using the equation (1) [16]. If the total system utilization is less than or equal to the utilization U_{RM} calculated by the equation (1), it proves that scheduling is possible with the RM scheduling algorithm.

$$U_{RM}(n) = n\left(2^{\frac{1}{n}} - 1\right) \quad (1)$$

The RM scheduling algorithm is a method of fixed priority scheduling in which a fixed priority is given to each task, and when a task having a higher priority arrives, the currently executing task is preempted. The RM algorithm determines the priority based on the period of the task, and it is an algorithm that gives a high priority to a task with a short execution period [17-19].

2.5. C-States

It can command the CPU to enter a low-power mode to save battery when the CPU is idle. Each CPU has multiple power modes, collectively known as the C-States. In general, in the case of a mobile device such as tablet PC in the x86 architecture, the low-power technique provides five states of the CPU called the C-States as shown in Table 1 [20].

Table 1: C-States mode in Intel's CPU

Mode	Status of CPU
C0	Processor state: Full On
C1	Processor state: Auto-Halt
C1E	Processor state: Auto-Halt
C3	Processor state: Deep Sleep
C6	Processor state: Deep Power Down

2.6. MIL-STD-1553B

The MIL-STD-1553B, which is mainly applied to aircraft weapon systems and guided weapon systems, has a maximum transmission speed of 1Mbps, and it is a half-duplex and a serial communication method.

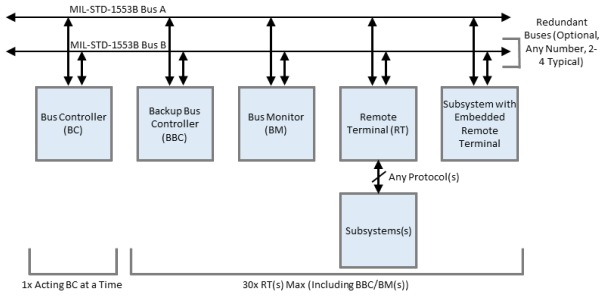


Figure 1: Interface Diagram of MIL-STD-1553B bus

Also, there is a restriction that the length of data that can transmit or receive data at a time is limited to a maximum of 32 words by a protocol. As shown in Figure 1, a bus controller (BC), a remote terminal (RT) and a monitoring (MT) are interfaced to the MIL-STD-1553B bus. The main function of the BC is to control the transmission and reception of all RTs on the bus. The RT has a function of transmitting or receiving data by responding to commands from the BC, and the MT has a function of monitoring all data moving on the bus [21, 22].

3. Methodology

3.1. Overview

In Windows 8, the initial count register (FFF0 0380h), as a local APIC timer register, is initialized to 0 after the operating system is booted. Therefore, since it is impossible to utilize a timer interrupt independent of the Window, there is a matter in providing a real-time processing function corresponding to the computer's system clock. In addition, The C-States are used to increase battery operation time by reducing power consumption of mobile device. The C-States function has the problem that the system clock is changed frequently flexibly, and an accurate period calculation for guarantying a real-time is not allowed [23, 24].

In this paper, in order to solve this issue, The RTiK+ is designed as shown in Figure 2, which uses a timer interrupt independent of the Windows to provide a real-time processing functions in the Windows 8 to support real-time performance for threads in the user mode. And the scheduling algorithm to support a real-time performance for MIL-STD-1553B communication is implemented based on RTiK+.

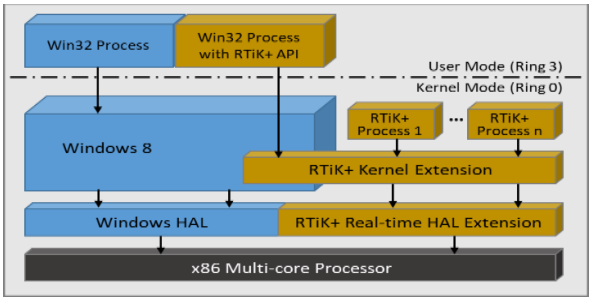


Figure 2: Architecture of RTiK+ for Real-Time Operating System

3.2. Local APIC Timer Control

As shown in Figure 2, The RTiK+ supplies a timer interrupt independent of the window through Kernel Resources access. And as shown in Figure 3, The RTiK+ is implemented to provide a real-time functionality in the Windows using the local APIC and kernel resources of the application processor [25, 26].

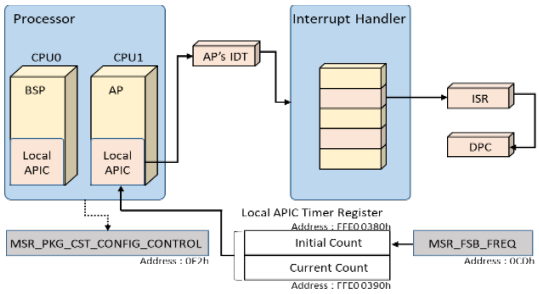


Figure 3: A Real-Time Processing Operating Process of the RTiK+

Before the timer interrupt is occurred, the initial counter register (FEE0 0380h) of the local APIC of the application processor sets the initial counter register value according to the value set in the MSR_FSB_FREQ register (0CDh) to generate a timer interrupt independent of the Windows. For such interrupt processing, a real-time task is created by registering an interrupt object in the interrupt descriptor table (IDT). In addition, The RTiK+ guarantees the periodicity of the real-time thread by using the MSR_PKG_CST_CONFIG_CONTROL register and minimizes the period error tolerance of the set task by controlling the C-States operation mode of the CPU to ensure the periodicity of the interrupt for thread processing [27-29].

3.3. Time Interrupt Control through FSB

As shown in Figure 4, x86-based system has two chipsets called the Northbridge and the Southbridge, the Northbridge is

between memory controller hub and the memory/graphic card slots, and the Southbridge is between I/O controller hub and PCI slots.

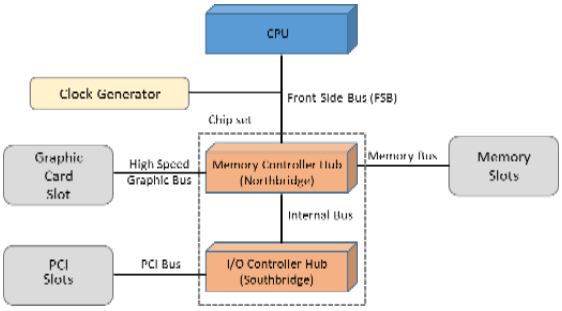


Figure 4: A Block Diagram of Chipset based on x86 Architecture [30]

A FSB also provides the ability to use a clock multiplier to determine the operating speed of the CPU and it provides synchronization by affecting the memory clock speed. The chipset block diagram of the x86 architecture is as shown in Figure 4, and the frequency generated by the clock generator is connected to the FSB and provided by the CPU and set as the initial counter register (FFF0 0360h) value of the timer register of the local APIC and then involves in determining the generation period of the timer interrupt. And in the case of the Windows 8, the operating frequency of the FSB according to the CPU workloads is provided through the lower three bits of D2:D0 of the MSR_FSB_FREQ register (0CDh), which is a hardware resource.

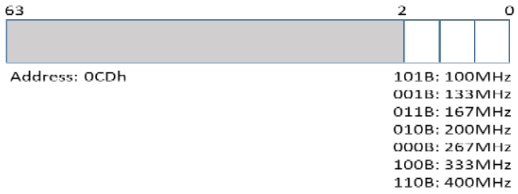


Figure 5: A MSR_FSB_FREQ Register

Therefore, as shown in Figure 5, to set the timer interrupt generation period of the local APIC to a fixed value, check the operating frequency of the currently executing FSB in the MSR_FSB_FREQ register, and set the local APIC timer to the value of the initial counter to generate an independent timer interrupt [31].

3.4. Time Deterministic through controlling C-States

Unlike the existing Windows 7, The Windows 8 provides a power saving function that reduces power consumption in real time by turning off power to some devices by adjusting the CPU clock low through access to the C-States for low-power control of the CPU. The following equation (2) is applied as the equation used to determine the CPU clock in Windows 8 [24, 32].

$$\text{CPU clock} = \text{FSB frequency} \times \text{Clock multiplier} \quad (2)$$

The FSB can control the power consumption by controlling the value of the clock multiplier by the Windows according to the hardware specifications. However, the method of adjusting the value of the clock multiplier to lower the clock of the CPU has a matter in that it is difficult to guarantee the time determinism of the timer interrupt of the local APIC. In addition, a problem occurs in that the aforementioned time determinism cannot be guaranteed because a delay time occurs due to the time when the CPU is activated from the idle state or sleep state.

In this paper, The C-States are controlled through the MSR_PKG_CST_CONFIG_CONTROL register (0E2h) provided by x86 system to guarantee the periodicity of the RTiK+. And each speed of the core clock according to the setting of the C-States is seemed as shown in Table 2. When the C-States are enabled, the clock is automatically adjusted based on the utilization rate the CPU is processing tasks. Also, when the C-States are disabled, the clock is kept at its maximum value regardless of the CPU utilization [33].

Table 2: A CPU Clock Status depending on the C-States

Type of CPU		C-States Enable	C-States Disable
Core 1	Standard	3199.86 MHz	3199.86 MHz
	Current	3199.86 MHz	3199.86 MHz
Core 2	Standard	1599.93 MHz	3199.86 MHz
	Current	1599.93 MHz	3199.86 MHz

3.5. Process in User Mode of the RTiK+

A study on the real-time processing function of the user mode provided in the existing Windows such as the Windows XP and the Windows 7 before the Windows 8 has a matter that the periodicity, data accuracy, and data integrity are not guaranteed because event signals are lost in an environment of multi-core processor. To solve this problem, the RTiK+ provides a real-time processing function in the user mode as shown in Figure 6.

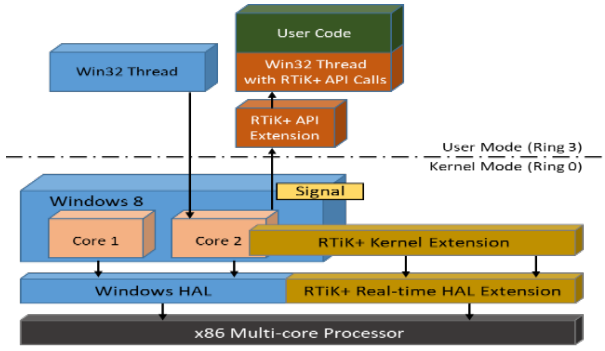


Figure 6: An Operation Process of the RTiK+ for a Real-Time Processing

The algorithm coded in the RTiK+ to set the priority of the process to the highest to guarantee data accuracy and data integrity is shown in Figure 7 [34]. It is implemented to call the Process_Affinity() function inside the real-time thread, and the real-time thread inside Process_Affinity(), by calling the SetProcessAffinityMask() function, an API provided by the Windows, the CPU is specially designated so that the real-time thread operates only on the core where RTiK+ operates.

Algorithm 1:

Result:

```
void Process_Affinity()
{
   SetPriorityClass(GetCurrentProcess(), REALTIME_PRIORITY_CLASS);
   SetProcessAffinityMask(GetCurrentProcess(), 0x02);
}
```

Figure 7: An Algorithm to set for Priority of Process

In addition, by calling the SetPriorityClass() function to set the priority of the real-time thread and process of RTiK+, the current real-time thread is designated as the highest level as a REALTIME_PRIORITY_CLASS, so that the CPU is not preempted by other Window's threads [35].

When the RTiK+ process is created, the process must be created by setting it to the REALTIME_PRIORITY_CLASS priority level, which is the highest priority. Figure 8 is the algorithm coded to provide a real-time processing function in the user mode of the RTiK+.

Algorithm 2:

Result:

```
void main()
{
    hThread = (HANDLE)_beginThreadex(NULL,0,ThreadFunction,NULL,0,(unsigned*)&dwThreadID);
    SetThreadPriority(hThread,THREAD_PRIORITY_TIME_CRITICAL);
}
```

Figure 8: An Algorithm to set for Priority of Thread

3.6. Priority-based Scheduling Algorithm

In the priority-based scheduling algorithm, a scheduling is implemented online, and when an event such as the creation or termination of a task occurs, the scheduling is performed first in the high-priority task. For this reason, the scheduling is performed while tasks are being performed, and there is an advantage of being able to respond more flexibly according to the state of the system, but on the other hand, it sometimes causes overhead in the operating system [36].

In this paper, the RM scheduling algorithm, a fixed-priority scheduling algorithm proposed by Liu and Layland, is applied to the RTiK+. The RM is an algorithm of giving a high priority to a task with a short period when all tasks are periodic tasks, the deadline coincides with the period, and the tasks are independent from each other. In this paper, the total utilization rate of the RM scheduling algorithm is U_{RM} (n) = n (2^(1/n)-1). If tasks are maximum 100, U_{RM} (100) = 0.69555. And no matter how many tasks are, the total utilization U_{RM} is within 0.69314.

3.7. A Process in User Mode of RTiK+

3.7.1. The BC Algorithm for a Real-Time Performance

In the MIL-STD-1553B, the BC function controls the direction of data when transmitting data through the bus, and transmits data to the RT by sending transmitter (Tx) and receiver (Rx) commands from the BC of the MIL-STD-1553B communication. In order to guarantee a real-time processing function of MIL-STD-1553B communication, it is implemented to operate the Tx and Rx commands in the real-time thread provided by the RTiK+. Figure 9 is the BC algorithm of MIL-STD-1553B communication that implements the Tx and Rx commands in the periodic thread of the RTiK+.

In this algorithm, Tx command releases the Rx link with the m1553_delete_link() function, connects the Tx link with the m1553_add_link_to_chain() function, sends the Tx command from the BC with the m1553_delete_link() function, and checks the data transmitted from the RT with m1553_read_link_data() function. And the Rx command releases the Tx Link with the m1553_delete_link() function, connects the RX Link with the m1553_load_chain() function, and transmits data to the RT with the m1553_load_chain() function.

Algorithm 3:

Result:

```
m1553_delete_link(device_number, RxLink.link_id, RxLink.chain_id);
status = m1553_add_link_to_chain( device_number, &TxLink );
if ( status == FAILURE )
{
    printf( "FAILURE\n" );
    printf( "%s\n", sbs_read_error() );
}

status = m1553_load_chain( device_number, 1 );
if ( status == FAILURE )
{
    printf( "FAILURE\n" );
    printf( "%s\n", sbs_read_error() );
}

uSleep(600);
m11553_read_link_data(device_number, TxLink.link_id, TxLink.buff_id[0], r_buffer);

m1553_delete_link(device_number, TxLink.link_id, TxLink.chain_id);
status = m1553_add_link_to_chain( device_number, &RxLink );
if ( status == FAILURE )
{
    printf( "FAILURE\n" );
    printf( "%s\n", sbs_read_error() );
}

m11553_write_link_data(device_number, RxLink.link_id, RxLink.buff_id[0], r_buffer);
status = m1553_load_chain( device_number, 1 );
if ( status == FAILURE )
{
    printf( "FAILURE\n" );
    printf( "%s\n", sbs_read_error() );
}
uSleep(600);
```

Figure 9: BC Algorithm for MIL-STD-1553B based on Real-Time Processing

3.7.2. The RT Algorithm for a Real-Time Performance

In the MIL-STD-1553B, the RT function transmits or receives data through the bus according to the BC commands. In order to check a real-time performance of the RTiK+, it is designed to store data in the buffer at the same period as the period set in the BC of the inspection equipment by the RTiK+ to send 32 words of the value increased by one (1) from the previous data. However, the matter of data loss due to the gap time for enabling the RTiK+ implanted in the BC and the RT function also occurs in the RT same as the BC. In order to solve this matter, a double buffer was used in the same way as the BC. Figure 10 shows the RT algorithm of the MIL-STD-1553B communication using double buffer.

Algorithm 4:

Result:

```
before_actptr_TX = actprt_TX;
actprt_TX = LL_get_ram( device_number, btrptr_TX + rt1sa2t.sa + 32);

If (before_actprt_TX != actptr_TX)
{
    if ( rt1sa2t.buff_id[0] == actprt_TX )
    {
        for (i=0; i<32, i++)
        {
            w_buffer[i] = (SBS16)( cnt );
        }

        m1553_write_sa_buffer( device_number, 32, rt1sa2t.buff_id[1], w_buffer);
        cnt++;
    }
    else if (rt1sa2t.buff_id[1] == actprt_TX
    {
        for (i=0, i<32, i++)
        {
            w_buffer[i] = (SBS16)( cnt );
        }

        m1553_write_sa_buffer( device_nmber, 32, rt1sa2t.buff_id[0], w_buffer);
        cnt++;
}
```

Figure 10: RT Algorithm for MIL-STD-1553B based on Real-Time Processing

3.7.3. Scheduling Algorithm of MIL-STD-1553B for a Real-Time Performance

When performing the MIL-STD-1553B communication, it takes time to activate the RTiK+ each time to utilize the RTiK+. A period error occurs in communication. Also, when a data transmission occurs in the BC due to the application of polling method applied in the paper and a data is received at the RT. In a state in which all data transmitted from the BC are not correctly received, the next data to be transmitted updates the buffer, resulting in a matter that correct data cannot be received at the RT. Therefore, in order to improve this problem, it is necessary to set the period of the RTiK+ for MIL-STD-1553B communication by considering the time it takes to transmit a data with the MIL-STD-1553B's BC command and the time it takes to receive data with the RT command. And a scheduling with RTiK+ for MIL-STD-1553B is required. Also, in the MIL-STD-1553 communication proposed in this paper, there is a problem in that it does not guarantee the periodic operation of data transmission from the BC to the RT because the Rx command is immediately sent after the Tx command is finished. In addition, since the RTiK+ operates at the period initially set, the period of the Tx command and the Rx command cannot be set respectively, so there can be a difference in the operating period of the RT and data loss can occur in this case. To solve this matter, when the Rx function is finished, the Rx Link is released, and the Tx Link is connected immediately to execute Tx function. The scheduling algorithm using the RTiK+ to guarantee a real-time processing performance when transmitting and receiving data in the MIL-STD-1553B communication was proposed.

This scheduling algorithm applied when performing the MIL-STD-1553B communication between the host PC and the target PC is assumed as follows.

- The RTiK+ uses the RM algorithm and preempts the CPU.
- The MIL-STD-1553B communication occurs periodically.
- Task period (Period, pi) and deadline (Deadline, Di) are the same. (pi = Di)
- The task start time (Release Time, Øi) is 0.
- The RTiK+ is implanted to both the host PC and target PC, and is activated only when the RTiK+ is utilized.
- The BC and the RT function are for MIL-STD-1553B, and interrupt cannot be used in an asynchronous method.
- The Tx and Rx command occur alternately. The data received by the Tx command is read from the buffer and transferred to the Rx command as it is.
- Transmits up to 32 words of data as in the Tx and Rx commands, and the transmission time is up to 0.80ms for all tasks [37].
- The Tx and Rx command cannot occur within a half of a period, therefore disconnect and create each time Tx to Rx transitions. Thus Tx(ei) > Rx(ei) or Tx(ei) < Rx(ei).
- The BC function and the RT function each use a double buffer.

Figure 11 shows the scheduling algorithm for the MIL-STD-1553B communication. The MIL-STD-1553B occurs when the Tx and Rx commands are alternated between the host PC and the target PC, and each task Ti (T0, T1, T2, T3, ..., Tn) has a maximum of 32 words of data. When transmitted, the execution time ei of each task is 0.80ms. At this time, since the deadline is the same as the period, it is scheduled so that transmission or reception of data within the period is normally completed without overhead. Also, when the MIL-STD-1553B communication is performed, RTiK+(pi), Tx(pi), Rx(pi), Buffer1(pi) and Buffer2(pi) are operated with an independent period each to match the period provided by the RTiK+.

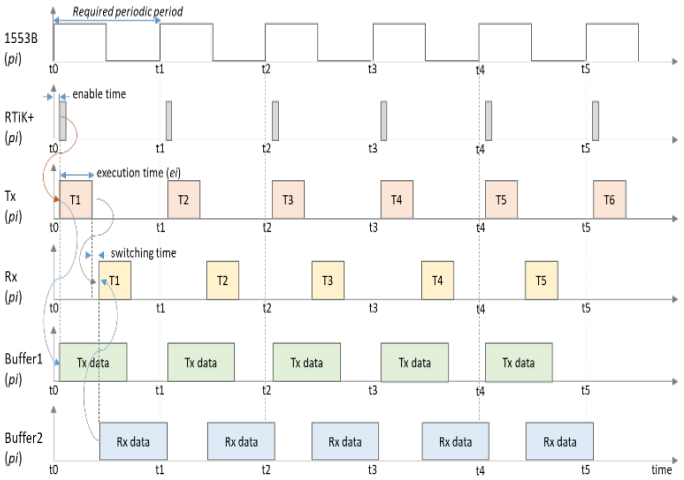


Figure 11: The Scheduling Algorithm with RTiK+ for MIL-STD-1553B

And if the task for the MIL-STD-1553B communication scheduling algorithm is as shown in Table 3, there are three Ti = (Øi, pi, ei, Di) tasks in the Tx and Rx commands linked to the RTiK+ pi period, verified if scheduling is possible through simulation that when the same 32 words of data transmission is occurred in the MIL-STD-1553B communication. That is, it was assumes that when this simulation is satisfied, all other communication conditions are also satisfied. At this time, when the Tx and Rx commands are executed, each task needs a maximum execution time of 0.80ms to transmit 32 words.

Table 3: A related Tasks of Tx/Rx for the MIL-STD-1553B Scheduling

Tasks	Release time (Øi)	Period (pi)	Execution time (ei)	Deadline (Di)
RTiK+ pi	0ms	2ms	0.01ms	2ms
Tx pi	0ms	2ms	0.80ms	2ms
Rx pi	0ms	2ms	0.80ms	2ms
T1	0ms	2ms	0.80ms	2ms
T2	0ms	2ms	0.80ms	2ms
T3	0ms	2ms	0.80ms	2ms

Using this method of the scheduling algorithm, each T1, T2, T3, …, Tn while Tx and Rx commands are being executed. It shows that the BC function of the MIL-STD-1553B can be implemented without data loss within the period set in the RTiK+ by the tasks. If the equation (3) is applied to calculate the system utilization rate for Tx(pi) and Rx(pi), the system utilization rate is 0.80/2 + 0.08/2 = 0.80, and the total utilization rate U_{RM} is 0.82 by applying the equation (1). It can be seen theoretically that the MIL-STD-1553B scheduling algorithm using the RM algorithm

operates normally with the system utilization rate (0.80) < U_{RM} (0.82).

Therefore, in the scheduling algorithm proposed in this paper, even if a Tx command executes in the MIL-STD-1553B communication and an Rx command executes immediately, the MIL-STD-1553B communication in real time meets the deadline without data loss within the period set in the RTiK+.

4. Results

4.1. Experimental Environment

To verify the real-time performance of the scheduling algorithm with the RTiK+ for MIL-STD-1553B communication, the experimental specifications for the host PC and the target PC is shown in Table 4.

Table 4: An Experimental Environment between host PC and target PC

Item	Host PC	Target PC
CPU	Intel Pentium Dual-Core 2117U @1.80 GHz	Intel Core i5-2500 @3.30 GHz
OS	Windows 8	Windows 7

4.2. Experimental Method

A verification is a measurement method that calculates the period by storing data using the RDTSC command to get a timestamp value to proof a real-time processing performance of the RTiK+. The algorithm coded in RDTSC for this experiment is shown in Figure 12.

The current timestamp stored in 64-bit values of EAX and EDX is read, store the timestamp value in the timearray[i].QuadPart array at i index, and calculate the difference between adjacent index values in the array. The f file is designed to be saved. In this way, a timer interrupt set with a period of 2ms occurs in the RTiK+, and whenever a real-time thread is executed in the user mode, the value of the timestamp is stored in the array.

Algorithm 5:

Result:

```
Hthread()
{
   _asm
   {
      RDTSC
      MOV lowval, EAX
      MOV highval, EDX
   }
   clockcal2_LowPart = lowval;
   clockcal2_HighPart = highval;
   timearray[i].QuadPart = clockval2.QuadPart;
   if(i==32000)
   {
      for(j=1, j<32000, j++)
      {
         fprint(f, "%.10f\n", (double)(timearry[i].QuadPart-timearray[j-1].QuadPart))/CPU);
      }
      i=0;
      break;
   }
}
```

Figure 12: An Algorithm with RDTSC for Period Measurement of the RTiK+

4.3. Measurement and Results

4.3.1. A Measurement for MIL-STD-1553B Communication

To verify a real-time performance support of the RTiK+ in the MIL-STD-1553B communication, the experimental environment was configured as shown in Figure 13. The RTiK+ is implanted to the host PC equipped with the MIL-STD-1553B communication, and the Period of the MIL-STD-1553B communication is set to 2ms with the RTiK+. The experimental data were measured with the RDTSC command as previously explained.

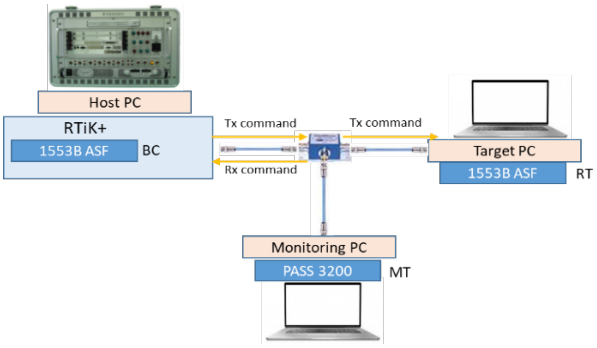


Figure 13: A Configuration to measure Period of MIL-STD-1553B

And to check the integrity of the data that there is no loss of transmitted and received data when performing the MIL-STD-1553B communication, The SBS Technologies' PASS 3200, a special-purpose MIL-STD-1553B monitoring equipment, is as shown in Figure 13, it connects between the host PC operating with the MIL-STD-1553B BC function and the target PC operating with the MIL-STD-1553B RT function to operate the MIL-STD-1553B MT function.

4.3.2. Result of Measuring Period during MIL-STD-1553B

In this paper, the maximum experiment period of 2ms was measured by RDTSC command during the MIL-STD-1553B communication. In addition, in order to compare a real-time performance for the MIL-STD-1553B between RTiK+ and RTX, the Period of the MIL-STD-1553B communication was measured as shown in Table 5 by applying the same experimental environment and experimental method.

Table 5: Comparison with Period of MIL-STD-1553B between RTiK+ and RTX

Period	Period of RTiK+			Period of RTX		
	Min.	Max.	Tolerance	Min.	Max.	Tolerance
2ms	1.993	2.031	1.57%	1.901	2.015	4.91%
5ms	4.955	5.003	0.90%	4.998	5.252	5.04%
10ms	9.993	10.005	0.06%	9.996	10.003	0.04%

4.3.3. A Result of Monitoring Data during MIL-STD-1553B

During the MIL-STD-1553B communication, in order to check the accuracy and integrity of the transmitted data, all data passing through the MIL-STD-1553B communication bus is stored with the PASS 3200 using the MIL-STD-1553B MT function as Figure 13. As shown in Figure 14, the first data transmitted by performing Tx command from the MIL-STD-1553B communication BC function to RT function is 32 words, all of which are 0xC912, and 32 words received by Rx command from RT function are all 0xC912. It seems that there is no loss or wrong data.

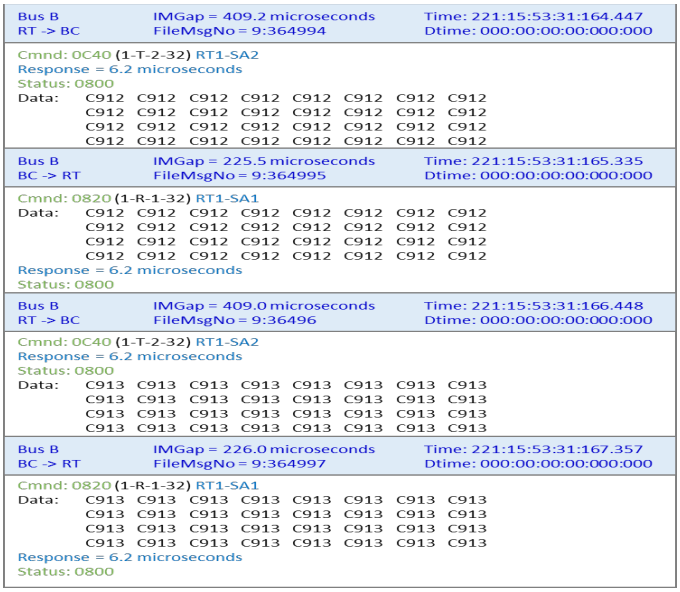


Figure 14: A Result of the MIL-STD-1553B monitoring by PASS 3200

5. Discussion

5.1. Process in User Mode of RTiK+

If the scheduling algorithm for the MIL-STD-1553B is analyzed based on the experimental results for period of the MIL-STD-1553B communication, Figure 15 and 16 is shown. As can be seen from this result, there was no case of overload the deadline in the 2ms period of the MIL-STD-1553B communication to which the scheduling algorithm with RTiK+ was applied, and it was proofed that the MIL-STD-1553B communication operates normally using the 2ms period of the RTiK+. As shown in Table 5, when the tolerance of the RTiK+ used in the MIL-STD-1553B communication is 1.57%, it is ±0.03ms from 2ms, but when the period for each task of Tx(pi) and Rx(pi) transmits 32 words, the maximum transmission time is 0.80ms, which is less than the half period of 2ms. The system utilization rate is 0.32 from the equation (3), and U_{RM} is calculated as 0.82 from the equation (1). It means that the scheduling algorithm for the MIL-STD-1553B operates normally without any communication errors.

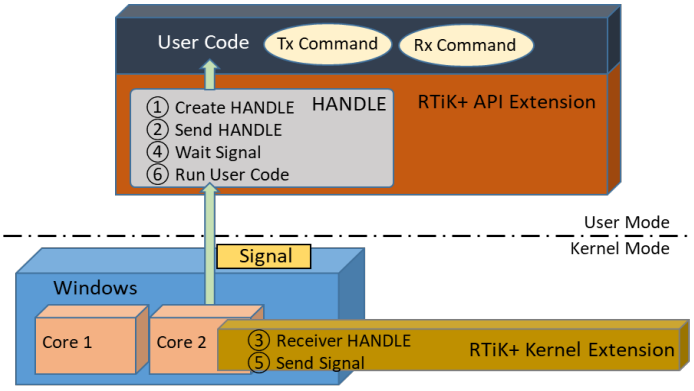


Figure 15: An Operation Diagram of Tx and Rx command for MIL-STD-1553B Scheduling Algorithm

In addition, this is the result of proving that the miss rate is 0 in the 2ms period provided by the RTiK+ for real-time processing function in the MIL-STD-1553B communication.

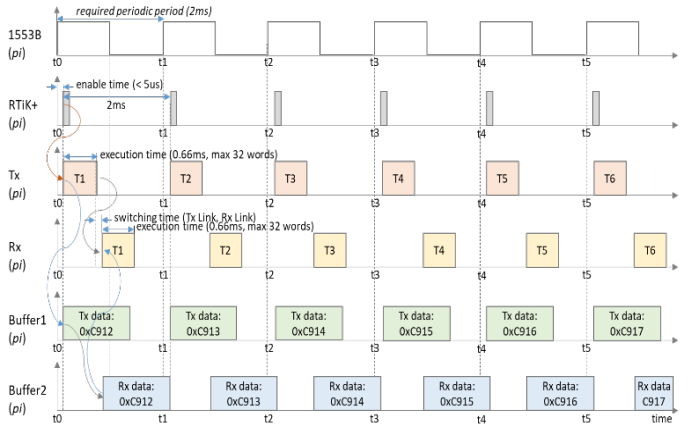


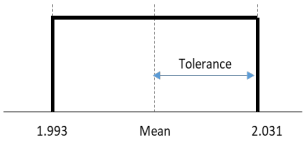
Figure 16: An Analysis Diagram for MIL-STD-1553B Scheduling Algorithm

5.2. Evaluation for Real-Time Performance of RTiK+

In order to evaluate a real-time performance of the MIL-STD-1553B communication with the result of Table 5, the 2ms period is schematically shown in Figure 17. In Figure 17(b), it can be seen that all the periods occurring in the scheduling algorithm with a period of 2ms of the RTiK+ for the MIL-STD-1553B are between 1.993ms and 2.031ms.

	Period of Scheduling algorithm with RTiK+ for MIL-STD-1553B		
	Min	Max	Tolerance
2ms	1.993	2.031	1.57%

(a) Result of Period of 2ms for MIL-STD-1553B



(b) Graph of Period of 2ms for MIL-STD-1553B

Figure 17: A Performance of 2ms Period for MIL-STD-1553B

For a performance measurement of a real-time system, 32 words of data are transmitted according to the period provided by the scheduling algorithm with the RTiK+ in Windows for the MIL-STD-1553B communication, and the miss rate is applied to verify the jobs performed after the deadline. The figure 18 is shown for miss rate during the MIL-STD-1553B communication to decide success or failure in real-time system.

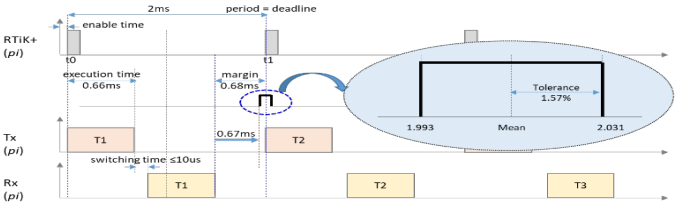


Figure 18: A Miss Rate of the MIL-STD-1553B Communication

The system based on x86 with Windows 8 applied in this paper is a hard-real-time system that must meet a deadline. Therefore, if all deadlines are met, the MIL-STD-1553B communication is a success, otherwise it is a failure.

$$\text{Margin of MIL-STD-1553B} = (pi\text{-}ei)\text{-Tolerance (\%)} \quad (5)$$

Substituting the measured result into Equation (5), actually the margin of the MIL-STD-1553B is 0.39ms that (2ms-1.60ms) - (2ms-1.9938ms). And the miss rate of the MIL-STD-1553B is defined as in equation (6).

Miss rate of MIL-STD-1553B = Number of Error / Number of Communications (6)

The MIL-STD-1553B using PASS 3200 analyzed all data as shown in Figure 14 in the period of 2ms, there was no communication error and all data were normal during the MIL-STD-1553B.

Table 6: Performance of Scheduling Algorithm by RTiK+ for MIL-STD-1553B

Period	Min	Max	Miss Rate
2ms	1.997ms	2.006ms	0%
5ms	4.961ms	5.038ms	0%
10ms	9.995ms	10.007ms	0%

If the MIL-STD-1553B communication period is 1ms and the algorithm proposed in this paper is applied, the maximum transmitted data is Tx(ei) = Rx(ei)) = 0.80ms. This is 0.60ms from the 1ms period of MIL-STD-1553B communication, so even if a double buffer is used, overhead may occur. RTiK+ itself can provide a period within 3% error when there are no workloads on the CPU in a 1ms period.

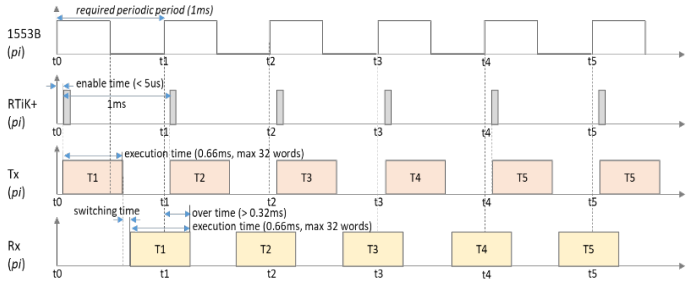


Figure 19: An Analysis Result of 1ms Period for Scheduling Algorithm

It is predicted that MIL-STD-1553B communication will not operate normally due to miss rate because of failure to comply with the deadline due to the lengthy duration as shown in Figure 19.

6. Conclusion

In recent year, with the development of high-speed communication technology and computer technology, the scope of a real-time information sharing has been expanded, and the performance of electronic equipment has been improved remarkably. The accuracy of data collection and real-time data communication are required in industry and defense system.

The RTiK+ presented in this paper controls the MSR_FSB_FREQ register to support a real-time performance on Windows in an environment where x86-based Windows 8 is installed, and calculates the CPU clock tick value required for the operation of the local APIC timer to be a Windows independent timer interrupt was made to occur, and in this way, the matter that local APIC is initialized after booting in Windows 8 is solved.

The experimental verification was checked by measuring the period by calculating a timestamp in the PC's internal memory using RDTSC, and measuring a real-time processing performance of the RTiK+.

And to support a real-time performance in the MIL-STD-1553B communication, The RTiK+ was implanted in the Windows based on x86 system, and a real-time processing functions were added to the BC and RT functions. Here, the real-time performance was provided when BC and RT function of the MIL-STD-1553B communication are implemented, and by applying the scheduling algorithm with RTiK+ for a high-speed MIL-STD-1553B communication such as 2ms of period is implemented to prevent data loss due to error of the period. To proof this scheduling algorithm experimentally, The RTiK+ was implanted in the Window 8 and the MIL-STD-1553B communication of the period of 2ms, 5ms, and 10ms was set to verify whether a real-time processing function of the RTiK+ normally while MIL-STD-1553B communication is performing.

In addition, in order to compare a real-time performance between the RTiK+ and the RTX as a third party for the MIL-STD-1553B communication, the experiment result for the period of the MIL-STD-1553B by applying the experimental environment and experimental method is from 1.993ms to 2.031ms in the period of 2ms. The RTiK+ had a tolerance of 1.5%, and the results of this experiment proved that RTiK+ can be applied to systems by replacing third party such as RTX. Finally, a real-time processing performance of the scheduling algorithm with RTiK+ proposed in this paper for MIL-STD-1553B implemented has a miss rate of 0 during the MIL-STD-1553B communication with a period of 2ms for system that requires high reliability. It also proved that the real-time performance and data integrity for scheduling algorithm with RTiK+ are guaranteed.

In the future research, the methods proposed in this paper is needed in the RS-232C, RS-422 and Ethernet used mainly in industry and defense system, and the research for the Windows 10 based on x86 system is needed as well.

Acknowledgment

I would like to thank my dissertation advisor Professor Lee Cheol-Hoon from CNU for his supervise and expert advice, as well as Mr. Muammar Abdulla Abushehab for extraordinary support.

References

[1] C.M. Krishna, Gang G. Shin, Real-Time Systems, McGrow-Hill, 1997.
[2] C. Koo et al, "Distributed Simulator design by using of SimNetwork to overcome speed limit on GenSim", Recent Advances in Space Technologies (RAST), Proceedings of 5th International Conference on RAST 2011, 430-435, 2011, doi: 10.1109/RAST.2011.5966872.
[3] H. Kim et al., "The Design and Performance Verification of Real-Time Inspection Equipment Software based on WindowsOperating System", Journal of the Korea Contents Association, **17**(10), 1-8, 2018, doi: 10.5392/JKCA.2017.17.10.001.
[4] S. Koh et al., "The method of development for enhancing reliability of missile assembly test set", The Journel of The Korea Academia-Industrial Cooperation Society, **19**(8), 37-43, 2018, doi: 10.5762/KAIS.2018.19.8.37.
[5] J. Kim et al., "Real-time Processing Method for Windows OS Using MSR_FSB_FREQ Control", Journal of Korea Multimedia Society, **24**(1), 95-105, 2021, doi: 10.9717/KMMS.2020.24.1.095.
[6] J. Lee et al., "Timer Implementation and Performance Measurement for Providing Real-time Performance to Windows 10", Journal of the Korea Contents Association, **20**(10), 14-24, 2020, doi: 10.5392/JKCA.2020.20.10.014.
[7] S. Kwon , "Design and Implementation of Air Vehicle Test Equipment for Unmanned Aerial Vehicle", The Journal of Advanced Navigation Technology", **24**(5), 251-260, 2020, doi: 10.12673/JANT.2020.24.4.251.
[8] J. Jung et al., "Analysis for Next-generation High-Speed MIL-STD-1553B Bus Technology", Journal of Aerospace Ssytem Engineering, **11**(6), 76-83, 2017, doi: 10.20910/JASE.2017.11.6.76.
[9] http://www.intervalzero.com
[10] Intel, Intel® 64 and IA-32 Architectures Software Developer's Manual, **2**

(2A, 2B, 2C & 2D), Volume 3A and Volume 3B, Intel, 2016.
[11] Intel, Intel® 64 Architectures x2APIC Specifiation, Intel, 2010.
[12] Intel, Intel Core i7-600, i5-500, i5-400 and i3-300 Mobile Processor Series 2010.
[13] A. Jo, "Integrated Middleware for Real-Time Device Driver on Windows", Journal of Korea Contents Association, **13**(3), 22-31, 2013, doi: 10.5392/JKCA.2013.13.03.022.
[14] H. Lihua, Analysis of Fuel Cell Generation System Application, Ph. D Thesis, Chongqing University, 2005.
[15] https://en.wikipedia.org/Rate-monotonic_scheduling
[16] J. Liu, "Real-Time Systems", Prentice Hall, 2000.
[17] M. Gong et al., "Dynamic Voltage Scaling for Scheduling Mixed Real-Time Tasks", Proceeding ICESS 2007 Third International Conference, 488-497, doi: 10.1007/978-3-540-72685-2.
[18] M. Lee, C. Lee, "Low Power EccEDF Algorithm for Real-Time Operating Systems", The Journal of the Korea Contents Association, **15**(1), 31-43, 2015, uci: 1410-ECN-0101-2016-004-001099668.
[19] M. Jung et al., "Optimal RM scheduling for simply periodic tasks on uniform multiprocessors", Proceedings of the 2009 International Conference on Hybrid Information Technology, 383-389, 2009, doi: 10.1145/1644993.1645064.
[20] M.E. Russinovich, "Windows Internals, Part 2 (6th Edition)", Microsoft, 2012.
[21] https://en.wikipedia.org/wiki/MIL-STD-1553
[22] ILC DDC, MIL-STD-1553 Designer's Guide (5th Edition), ILC Data Device Corporation, 1998.
[23] J. Kim et al., "Real-Time Supports on Tablet PC Platforms", Proceeding IEEE 2020 15th International Conference for Internet Technology and Secured Transactions(ICITST), 111-117, 2021, doi: 10.23919/ICITST51030.2020.935.1322.
[24] J. Kim, "Real-time Processing Method for Windows OS Using MSR_FSB_FREQ Control", Journal of Korea Multimedia Society, **24**(1), 95-105, 2021.
[25] D.A. Solomon, "Inside Microsoft Windows 2000, Third Edition", Microsoft, 2000.
[26] http://msdn.microsoft.com/en-us/library/ms810029.aspx
[27] D.A. Godse, A.P. Godse, Microprocessors, Technical Publications Pune, 432-472, 2007.
[28] O. Bailey, Embedded systems: desktop integration, Wordware Publishing Inc, 34-56, 2005.
[29] Intel, "MultiProcessor Specification (Version 1.4)", Intel, 1997.
[30] J. Choi, "A Study on Developing the Missile System Test Set Software with DDS", Proceeding of Korea Institute of Communication Sciences Conference, **2018**(11), 614-615, 2018, uci: 1410-ECN-0101-2019-567-000045708.
[31] https://en.wikipedia.org/wiki/Front-side_bus
[32] https://en.wikipedia.org/wiki/CPU_multiplier
[33] https://docs.microsoft.com/en-us/windows-hardware/drivers/ddi/wdm/nf-wdm-mmmapiospace
[34] https://docs.microsoft.com/en-us/windows/win32/procthread/scheduling-priorities
[35] http://msdn.microsoft.com/en-us/library/ms685100 (VS.85).aspx
[36] C. Lee, Real-Time Systems, CNU, 2012.
[37] K. Kim et al., "An Indirect Data Transfer Technique based on MIL-STD-1553B for the Software Upgrade of Embedded Equipments on a Missile Assembly with Booster", Journal of the Korea Institute of Military Science and Technology, **19**(1), 43-49, 2016, doi: 10.9766/KIMST.2016.19.1.043.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 385-408 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Survey of FPGA Robotics Applications in the Period 2010 – 2019

Dimitris Ziouzios[1], Pavlos Kilintzis[2], Nikolaos Baras[1,*], Minas Dasygenis[1]

[1]*Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani, 50100, Greece*

[2]*Department of Mechanical Engineering, University of Western Macedonia, Kozani, 50100, Greece*

ARTICLE INFO

Article history:
Received: 08 January, 2021
Accepted: 03 April, 2021
Online: 27 June, 2021

Keywords:
FPGA
Robotics
Hardware Accelerators

ABSTRACT

FPGAs constitute a flexible and increasingly popular controlling solution for robotic applications. Their core advantages regarding high computational performance and software-like flexibility make them suitable controller platforms for robots. These robotic applications include localization / navigation, image processing, industrial or even more complex procedures such as operating on medical or human assistant tasks. This paper provides an overview of the publications regarding different robotic FPGA application fields as well as the most commonly-used robot types used for those applications for the 10-year period of 2010-2019. A short description of each paper reviewed is also included, providing a total view of FPGA technology trends in robotic applications over the last decade.

1. Introduction

Field programmable gate array (FPGA) devices are reconfigurable circuits composed of numerous interconnected logic cells. FPGAs contain an array of programmable logic blocks, and a hierarchy of "reconfigurable interconnects" that allow the blocks to be wired together. They are still used as a high-performance computing platform in multiple applications nowadays, despite their overshadowing by CPUs and GPUs. The general core advantages of FPGAs are that they can be easily connected to most types of circuits due to their protocol-independent serial links and they offer hard real-time computations and highly parallel effective programming with low power consumption, using algorithmic high-level synthesis [1]. FPGAs can be reconfigured and reprogrammed according to the user's requirements, unlike fixed hardware structures that CPUs have. This allows users to utilize them in a vast variety of applications.

Robots today are used for multiple tasks, for both industrial and consumer purposes that require high performance and low power consumption. They can be used for manufacturing or logistics purposes, while offering medical, educational, home, defense and other activities in the services' sector. So, the applications implemented on FPGAs in the robotics sector contain several multi-factorial tasks from multiple domains. Primarily those tasks include vision, sound or touch response systems and object tracing through processing of a high data flow coming from sensor inputs. Furthermore, increased real-time requirements for controlling robotics' functionalities such as the movement of a robot, make FPGAs to be a more attractive processing solution since they can be reconfigured as many times as it is necessary, even during runtime, while they are utilized, until the required functionality is properly achieved. FPGAs can actually rewire their internal circuitry in order to allow configuration, after the control system is installed onto the field. Moreover, the high-power consumption of robotics applications indicates FPGAs as suitable computational systems in comparison to other solutions. In fact, the advantages of using reconfigurable dynamic systems do not exclusively lie on the increased adaptability to changing conditions, but also on reducing power consumption without affecting the robot's capabilities. Overall, the size and speed of an FPGA is similar to the application-specific integrated circuit (ASIC), but their shorter design cycle and reconfigurability capabilities make them more flexible for use in robotic applications [2].

FPGAs are a mature implementation technology, spanning over 35 years of technology improvements. At the same time, after 70 years of evolution in robotics the controlling complexity of them has reached a level that only high-end processing cores can handle. Investigating the blending between the robotics and the FPGAs in the related literature, we located a gap. In this paper, we aim to answer the research question: Given all the advantages in CPUs and GPUs, is it viable to use FPGAs in robotics applications, and if yes which are the areas that FPGAs excel in the robotics domain and what challenges loom ahead? Towards

*Corresponding Author: Nikolaos Baras, Email: nbaras@uowm.gr

this, we investigated IEEE listed scientific publications in the 10-year period between 2010 and 2019 in order to get an overall view of the typical application fields of FPGAs in robotic applications. Figure 1 shows the absolute quantity of the total publications regarding journals, conferences, early-access articles and magazines in the IEEE Xplore and ACM digital library depicting the actual publication tendency over the last 27 years, since the library includes publication data since 1993. The total publications count from 1993 to 2019 regarding FPGAs in robotics is 1174 publications, including the aforementioned four types of publications. The publication count for the testing period of 2010-2019 counts 815 publications in total, where 762 are conference papers, 49 journal papers, 2 early-access articles and 2 magazine articles. The average publication number over the 10 year investigation period is almost 80 publications per year. As accrued from Figure 1, from 1993 to 2001 the publication count remains relatively steady with less than 10 publications / year, while from 2002 to 2009 there is a yearly significant increase, reaching the top value of 81 publications in 2009. Since then, a fluctuating count is observed in the testing time period of 2010-2019, with the maximum value accomplished in 2019 with 102 publications. Figure 1 imprints the technological boom in the FPGA capabilities that started in the millennium (2000), allowing wide FPGA usage in robotic applications

This tendency lasted for almost a decade, until 2009 and since then the robotic FPGA applications - according to the publications count - remain relatively steady with fluctuating tendencies. Since the total publications for the tested time period are almost 800, the researchers choose to analyze 20 popular publications for each year, in order to imprint the latest scientific tendencies. The total number of publications tested for the 10-year period is 200.

The paper is structured as follows. First, a brief theoretical analysis of the modern robot types regarding their application fields and interaction mechanism is given, according to the latest literature trends. The description of the three main robot categories as classified in this survey, follows. The six main FPGA application fields, as uncovered in the survey analysis, are then quoted.

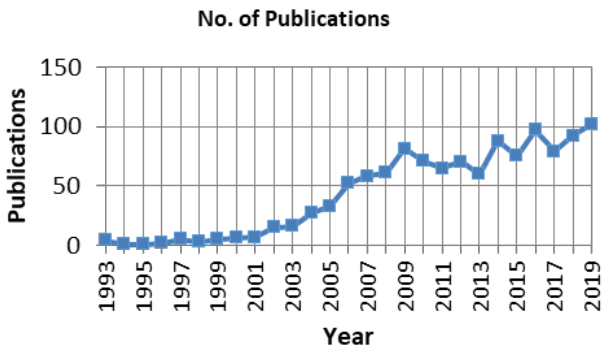


Figure 1: FPGA Robotics-related publications in IEEE Xplore and ACM digital library per year

The conclusion section sums up our findings, as indicated by the survey findings. All reviewed literature is quoted with application field and robot type categorization, a short description of each paper and the year of the publication.

2. Related Work

A shortlist of FPGA-related studies uncovering related application fields, architectures and technology trends has been revealed in the survey. The most detailed view for survey applications for a distinct period of time is included in [1]. 22 distinct technology applications of FPGAs (+ "other" quoted as distinct category) for a 15-year period are documented, investigating 815 publications in the IEEE Xplore and ACM digital library for the aforementioned duration. This survey offers a relatively indicative view of the FPGA technology trends until 2015, however it does not investigate the FPGA architectures designed specifically for robotics applications. Paper [3] highlights the availability of reconfigurable applications as platforms for digital controllers and quotes the most important factors for the technology trend of FPGAs utilization in comparison to other architectures, primarily due to the acceleration and flexibility of FPGAs, reduction of costs and reduced energy consumption. Study [4] quotes a general description of FPGAs, regarding their use as accelerators and simulator platforms, the usage of their integrated memory and their fault tolerance as well as the programming languages and computation methodologies. However, it does not quote applications per year or include a specific distinction of application fields The FPGA virtualization techniques are included in [5], separated into three categories - resource level, node level and multi-node level. The first level investigates whether FPGA's resources are reconfigurable or not. The node level explores the applications needed in both software and hardware terms regarding virtualization at a single FPGA. The existing architectures of how to connect multiple FPGAs are described in the multi-node level. The survey [6] investigates FPGA's routing architectures, dividing them into island-style routing and hierarchical routing. Furthermore, it quotes a brief description of the recent technology applications of FPGAs, distinguishing them into eight distinct categories. In conclusion, there are no specific surveys like this one and for this reason we believe our paper will play a pivotal role as a reference for researchers in the field of implementing robotic applications on FPGAs.

3. Application Fields of FPGAs in Robotics

In our survey, we categorized the papers into 6 major categories, according to the application. In order for an application field to be categorized as a distinct category, at least 5 publications had to be identified in the databases. Some application fields can be mixed, for instance, a robotic controller may use a regular FPGA algorithm and Fuzzy logic architecture or a navigation application can use SoC FPGA image processing systems. Distinct categories as well as mixed ones are referred to the Application Field column of Table 1, where the research overview is depicted. The Robot type column includes the type of robot which this publication's application is designed for. A short description of the publication's application regarding the algorithms used, the software or the type of the FPGA controller chips is included in the next column. The publication year and the reference number are depicted in the last two columns of Table 1.

3.1. Classification of Robots

While there is no strict classification of robots, most robotic applications can be either categorized according to their mechanism of interaction or according to their application field [7]. Figure 2 depicts the robot classification according to their mechanism of interaction [7]:

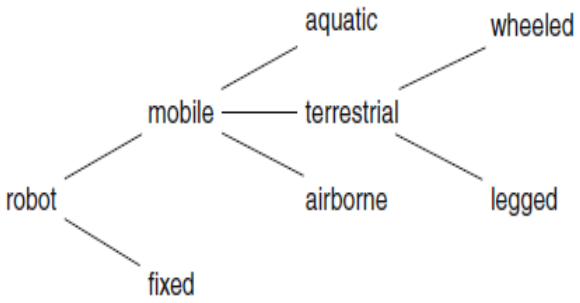


Figure 2: Robot classification by interaction mechanism [7]

The basic classification of robots according to their application field is shown on Figure 3. It has to be noted that the six sub-categories of Fig. 3 can also include variations. For instance, educational robots can either be wheeled or humanoid (legged) and industrial robots can be either articulated robots, cartesian robots, parallel, spherical, or SCARA, etc. [8]. The aforementioned categories are based on the primary means of navigation (for educational robots) or the robot's basic configuration (for industrial robots).

However, this survey includes the robotic applications that are FPGA-applied, so the classification of robots is done according to those categories that can use the application fields found in the survey publications and analyzed in the next chapter. The main robot categories of the FPGA-based robotic applications included in this survey are the following:

- Industrial Robots (A1)
- Autonomous Mobile Robots (A2)
- Humanoid Robots (A3)

For a robot type to be classified as a distinct category, at least 5 publications had to be found for the selected period. In other cases, this column is classified as "other", implying a different type of robot, described in the "description" column such as medical, defense, educational, etc.

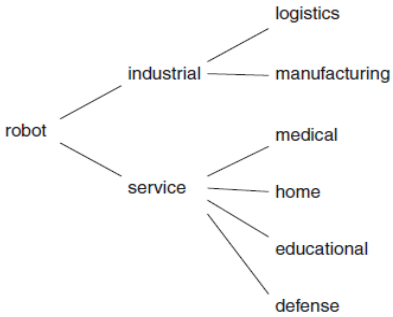


Figure 3: Classification of robots according to their application field [7]

3.1.1. Industrial Robots

Industrial robots constitute one of the major of the FPGA-related publications in this survey, as depicted in Fig. 4. The main subcategories of this main robot category are [10]:

- Articulated robots
- Cartesian robots
- Cylindrical robots
- SCARA robots
- Delta robots
- Polar robots

The articulated robot type is amongst the most common types of industrial robotics. It includes a mechanical arm conjoined to a stable base. Rotary joints connecting the arm associations may vary from two to ten joints although the most common robots are the ones with six joints. The joints give the robot additional freedom to its movement while the joints can be either orthogonal or parallel. Cartesian or Gantry robots are also amongst the most common robot types. They have a rectangular configuration with three prismatic joints (X, Y, Z) allowing linear motion, while some Gantry robots have an additional wrist allowing rotational motion. Cylindrical robots use a main arm making an up-and-down motion. This arm is internally embedded with a cylinder allowing the extension/reaction motion. Cylindrical robots use a 3-dimensional coordinated architecture using a specifically selected reference axis for selecting point positions. Cylindrical robots are mostly used in picking-placing materials operations. SCARA (Selective Compliance Assembly Robot Arm) robots use two parallel joints providing movement at a specifically selected track. The rotary tracks are placed vertically and the operating tooling attached to the arm moves horizontally, making SCARA robots most suitable for vertical assembly operations where efficient and quick pick and place operation is critical. The primary characteristic of the Delta robots is their parallel link configuration consisting of parallel joint connections linked to a common stable base. The positioning of the effector is controlled with the robotic arm making the robot type suitable for speedy pick-and-place applications. Polar (or spherical) robots use stationary arms with spherical work envelopes which can be placed in a polar-coordinated architecture. They use one linear and two rotary joints conjoining the links and a twisting joint conjoined to the arm.

3.1.2. Autonomous Mobile Robots

Autonomous Mobile Robots (AMR) have been classified as a distinct category because they can be used for industrial as well as domestic, medical, defense or educational purposes. An AMR is designed to move on wheels or with legs (human walking simulation is excluded from AMR category) across the ground floor to specified locations operating as a flexible assistant performing numerous tasks. The movement of an AMR is accomplished by using a total of advanced on-board sensors, FPGA-architectures and processors that allow it to comprehend and interpret its external environment. The aforementioned means offer the AMR the capability of combining the information from

multiple different sources such as video or GPS in order to navigate when there is a malfunction or full absence of data from one or more data sources. The core operation of an AMR is its control system, where FPGAs may constitute an implementation solution in many cases. While FPGA-based control systems are analytically explained in the next section, the design requirements of an effective AMR control system typically include:

- **Multiple sensor and processor integration:** As mentioned above, the possible malfunction, unreliability or inapplicability of any individual sensor or set of sensors must be coped with the integration and effective operation of the rest of AMR sensors. Furthermore, AMR must be able to accomplish all assigned tasks despite concurrent and often conflicting actions that need to be performed.
- **Reactivity to environment:** AMR should be capable of reacting to unscheduled sudden changes in the external environment by executing fast and efficient decisions in order to fulfill the assigned tasks. Software and architecture efficiency play a crucial role at that requirement, including the accuracy and usability of FPGA software algorithms.
- **Reprogrammability and Expandability:** An expandable architecture of an AMR allows the incremental build-up of the system and its continuous improvement and evolution. FPGA chips offer a significant advantage over other processing solutions in this case since they are reconfigurable and can be programmed to adopt a different behavior every time is required, changing the configuration of FPGA's internal circuitry.

3.1.3. Humanoid Robots

Humanoid robots are robotic applications that mimic human body structures in both behavioral and motion terms. They are considered a relatively complex robot type since it is necessary to consider their mechanical (musculoskeletal) structure, the computational and sensory nervous system, the applied methods of information processing in the robot's primary decision-making units and the applied software algorithms as a whole and for any given application that they have to perform. FPGA chips can be used for multiple humanoid robotic applications such as their motion control, communication modules between robots and users, visual sensing and processing that constitute the robot's total behavior in mimicking human functionality. However, the primary distinguishing element of humanoids in comparison to other robot types is their human mimetic motion and human-like robot reaction. The first element involves the robot's skeleton structure, the body proportions, the muscle arrangement and the joint performance [10]. Human-like robot reaction mainly represents the robot's responses for safety. This can be applied actively by using a feedback loop or passively by applying a suitable robot's mechanical design. Humanoids with high gains actuators tend to have a low performance on estimating correctly the forces applied by the human for specific functions, while the usage of integrated early-stage additional sensors shows better performance on that specific task [11]. FPGA chips constitute an efficient choice as a fused sensor architecture.

3.2. Application Fields

The sections analyzed below, will briefly introduce the primary application domains of FPGA-related technology components on robotics applications. These applications can be either mathematical and software algorithms, hardware FPGA-oriented architectures, mixed systems with FPGA components or combinations of all the above.

3.2.1. Robotic controllers

A robot constitutes an autonomous closed-loop control application because it is affected by the inputs from its sensors. Due to the highly effective integration of the intellectual properties of the embedded processors in the FPGA, the mathematical algorithms executing the demanding computations required by the robot controllers can be implemented by software in the FPGA. This application field mainly includes the coding of FPGA motion control systems of immobile robotic applications and industrial robots. Industrial immobile robots are widely used in order to solve tasks such as assembly, handling or welding.

Servo control is the most crucial part of robotic controller architecture since a possible failure to maintain a closed-loop behavior may delay the production process, damage the robot itself or even cause accidents and danger for human life. A servo-control solution based on FPGAs is considered amongst the most effective ones. Specifically, an FPGA-based control interface is responsible for multiple tasks in the robotic operation such as decoding feedback position data, processing communication signals and control outputs and transmitting the data to the actuators. The instructions provided to the robot used to perform specific actions are processed by the robotics control system. The programming language is actually used to describe all the variables that affect the robotics operation, primarily regarding movements on specifically set paths. There are several path types that a robot can operate under, depending on its purpose. The simplest type of robot motion is the movement from one point to the other, without having to perform any other parallel simultaneous actions. This motion can be executed as joint displacements and this type of control is called joint control [12]. The joint control motion expresses the interactions of position, acceleration and speed in the robot's language describing the necessary movements to perform the requested action. The main limitation of this control method is that the user needs joint dimensional information since he must comprehend the robot's joint displacements in order to drive the robot into the desired position. This limitation can be overcome by using the Cartesian space, which helps to interpret robot's movements by using Cartesian coordinates, based on the Jacobian Transpose matrix. This type of control is called Cartesian control. Due to this reason, the mathematical algorithms representing the kinematics of Cartesian control motion are significantly important in order to propose the most suitable control structures. The FPGA-based controller is then used to transform the trajectory created in the Cartesian space into a parallel robot's trajectory in joint space. This application field includes research undertaken for kinematics algorithms and software applications regarding the robotic controllers, mainly representing immobile robots executing industrial tasks.

3.2.2. Mapping – Navigation

This application field includes research regarding autonomous mobile robots, analyzing the usage of FPGA controllers in crucial operations, such as robot's localization, mapping and navigation. In general, the navigation of robots is divided into two parts: The 1st part is about the robot being capable of selecting the most suitable path between a starting and a finishing (goal) point. This path can be computed by either making calculations based on pre-loaded existing data or by processing incoming real-time data by GPS and other applications. The 2nd part includes the robot's real-time effectiveness on adapting its behavior to the external environment characteristics, such as avoiding obstacles, making turns and so on. FPGA applications can be used for both parts. Robot navigation is performed by using line following and obstacle avoiding algorithms that can operate with sensor data and probabilistic algorithms that study traffic and external environment patterns. The FPGA control unit is receiving and processing all incoming data from sensors and modules in order to choose the most suitable path.

Localization is another crucial process of mobile robotics as it is used in order to determine a robot's position in relation to an external reference point, called landmark. There are two localization methods: The first method calculates the distance from the landmark and its angle relative to the north, called the azimuth. The second method measures the angles to the object from two different positions, called triangulation. FPGA control units can process data and by applying the proper algorithms are capable of determining the right course. In recent years both localization techniques have been replaced by GPS location and navigation. FPGA receivers are often used to process GPS data and apply the proper navigation algorithms.

Simultaneous Localization and Mapping (SLAM) is the next-in-line step in robot mapping and navigation. A primary task of a mobile robot in an unknown environment is to create and frequently update a map of the environment and compute its current location within this map at the same time. This operation is called the SLAM operation. At normal conditions, the robot can use sensor or odometry measurements to gather information over its environment in order to perform localization and mapping. SLAM is used when these conditions are altered such as the localization at an underwater or space mission. Some researchers have proposed the use of probabilistic methods, such as Extended Kalman Filters [13]. The Extended Kaman Filters (EKF) algorithm simultaneously computes a model map of the environment and the current robot position based on odometric and exteroceptive sensor data. FPGA-based architectures usually implement the EKG algorithm. Alternative solutions avoiding the errors produced by odometry rely on scan-matching techniques that usually rely on range measurements acquired from robot sensors. Such an approach is the SMG algorithm [14]. This algorithm can perform SLAM by applying highly improved parameters which in turn considerably increase the algorithm's computational requirements. A considerable acceleration over software can be managed, by using the programmable resources of FPGA circuits and the parallelization properties of the algorithm.

3.2.3. Image Processing

FPGAs are widely used for the execution of digital image processing algorithms. They are used even more intensively, specifically in the case of real-time embedded applications, where latency and power are considered important issues. Typical topics in digital image processing are the evaluation of single images or numerous frames of a video signal for specifically set criteria such as movement information extraction or object tracking. For example, an FPGA embedded in a camera can execute image processing directly via the transmission of the image through the sensor. The camera provides a fully processed output stream that way, rather than just an image sequence.

The sector of robotics significantly benefits from the advantages provided by digital image processing on FPGAs. The image processing system of a robot is crucial for the enhancement of processing performance, since it includes the ability for spatial recognition and multiple detective algorithms to carry out the mission tasks. These tasks primarily include localization/navigation in the case of autonomous mobile robots. The primary reason for applying algorithms in FPGA robotics is the parallelism option which allows advantageous capabilities such as real-time image filtering for pre-processing purposes. Moreover, the local memory embedded in current FPGA architectures allows the buffering of pertinent image information to limit the communication with external memories, a situation that can potentially form a bottleneck [1].

Furthermore, FPGAs are more effective in comparison to CPUs in several aspects, regarding the ideal image processing system. CPUs execute their operations sequentially, so the first operation must process the entire image before the second operation starts. Since FPGAs are operating in parallel, every operation of the algorithm can operate on different pixels of the image simultaneously. This denotes that the time needed to receive the first processed pixel and process the whole image is much shorter, in comparison to an exclusive CPU application. Even if a combined co-processing FPGA-CPU architecture is used to transfer or receive images from the CPU, the total processing time is still shorter than an architecture that exclusively uses the CPU. Moreover, image processing architectures for robotics based both on USB and PC interface, are usually applied by multithread software solutions, so hard real-time response cannot be achieved for the proper integration of image and behavioral usage. Therefore, a robotic vision system that can perform image processing through an FPGA, which is directly connected to the CMOS sensor is more efficient.

3.2.3. Neural Networks

Artificial Neural Networks (ANNs) can be used in a variety of tasks performed by robots such as trajectory and image recognition, vehicle control, image processing, pattern and sequence recognition and medical diagnostics. They are statistical models destined to adapt and self-program, by applying learning algorithms that make them execute the aforementioned tasks. ANNs operate as parallel distributed computing networks and they operate correspondingly to the biological neural system. According to this, there are multiple input signals directed to neurons. Every input is assigned to a relative weight value which correspondingly influences the impact of that specific input. In

modelling, weight values are the adaptive coefficients within the network that calculate the magnitude of the input signal. The neuron output signal is then created by adding all the weighted inputs algebraically, simulating the biological cell body [15].

In order to perform properly, developers arrange ANN architectures in specific layers that are implemented in parallel. Since ANNs require significant computational power, they are better implemented using parallel hardware architecture [16]. ANNs can be implemented by using either software or hardware implementation. Software implementation is the favorite choice, supporting ANN training and simulation on sequential computers for imitating a large variety of neural network models, with significant flexibility. That is so since the developers do not need to know the internal operations of the neural network elements, but can be exclusively focused on the software application of the ANN. On the contrary, hardware FPGA-based architectures offer high applicability and the significant advantage of parallelism. The two basic characteristics of the FPGA-based reconfigurable computing architecture are its dynamic adoption and parallelism. These architectures are most suitable to apply ANNs as a developer can achieve concurrency and easily reconfigure in order to adjust the topology and the weights of an ANN. In total, FPGA-based architectures offer numerous advantages for neural implementations such as:

- Prototyping through reprogrammability: The neural FPGA-based architectures can be directly tested in order to find the most effective one, without any additional costs. Furthermore, an effective application architecture already developed and implemented can be later replaced by another one, without the need to design a new chip.
- Embedment: FPGA architectures can be used for all kinds of embedded applications, in order to take advantage of the effectiveness of neural calculations.
- On-chip leaning: Usually on-chip learning is avoided as it is considered a complicated process that can easily lead to loss of efficiency due to a number of reasons. On the contrary, off-chip learning is a more preferable choice when there are no requirements for dynamic learning. In the case of FPGAs, on-chip learning can be executed before the chosen implementation of the learned ANN on the same chip.
- Mapping: FPGA-based ANN architectures can be mapped onto more modern and effective FPGAs, whenever technology advancements indicate so. The basic prerequisite for ANN implementation is the scalability level of the implementation method.

3.4. Fuzzy Logic Control

Control algorithms for detailed dynamic models often require the exact specification of accurate target values. These equations may be complex and nonlinear, requiring longer "run-time" in order to get processed. An alternative approach to the usage of classical logic controllers described in section III subsection B.1 is the usage of Fuzzy logic control algorithms, which use imprecise specifications in order to reach the same goal. Fuzzy logic controllers are implemented to applications that are difficult by nature to model precisely. They are often considered more robust algorithms in planning parameter value changes, in comparison to regular control algorithms, since they do not require high precision and high computational power.

In the sector of robotics, fuzzy logic controllers are often used for automatic navigation purposes. This is due to their capacity to process a high number of imprecise and incomplete input signals. In such a case, signal input processing may allow robots to navigate in uncertain environments. A robot is equipped with multiple sensors that allow it to sense and process the environmental information. Fuzzy controllers receive sensor outputs as inputs for processing according to the fuzzy rule database that dictates the robot's possible behaviors. The robot decision making mechanism implemented by the fuzzy controller consists of three basic steps in order to reach the final decision. First, during fuzzification a crisp input value is converted to a fuzzy value that is performed by the use of the information in the fuzzy rule database. Next, fuzzy inference (rule evaluation) formulates the mapping from a given input to an output using fuzzy logic, applying the rules to the fuzzy inputs. The mapping includes the rule framework under which robot's decision making will take place. Then, during the defuzzification step the values from the output variables are calculated from the fuzzy values resulting in the controller's final decision [17]. In total, fuzzy algorithms allow the robot to choose amongst various combinations of decisions the most effective one, rather than being forced to make a pre-designed single decision.

A fuzzy logic controller is capable of processing inputs from a wide range of sensors. For instance, ultrasonic navigation sensors can calculate the distance between a robot and the obstacles of the external environment. GPS applications can be used for localization purposes, detecting the robot's current position. Fuzzy algorithms constitute a suitable controller solution for navigation in 2D and 3D environments, industrial robot manipulators [18], image processing algorithms [19] and several other robotic domains that require high computational power.

3.2.5. Swarm Robotics

Swarm robotics is the task of using multiple robots to perform different activities simultaneously, rather than using one robotic application with enhanced capabilities. It is considered a distributed approach that aims at coordinating robot behaviour by copying natural decision mechanisms of the biological entities (humans, animals). These algorithmic mechanisms allow a group of robots to achieve performance levels that could never be achieved by a single or a few robotic applications. The usage of multiple robots may significantly shorten manufacturing time, performing different tasks with simpler, scalable and more robust robotic applications. The major issue in swarm robotics is the effectiveness and precision of the coordination algorithms, as they are usually performing on the same physical object or environment.

There are two main approaches of designing swarm robotics. The first approach is the centralized system, according to which a primary robot acts as a leader-coordinator for the rest of the robots of the swarm [20]. In FPGA robotic controllers, this is a relatively simple solution for navigation tasks. In this case, a master (leader) robot plans the path planning to traverse to the direction of a certain goal point, with the rest of the robots to follow. The

behavioral control mechanism creates a constant mutual communication between the leader and the followers, whose behaviour is checked and corrected at specific time periods, according to the leader's motion. The core advantage of this approach is that all follower or slave robots can be simple robotic applications compared to the master robotic unit. The main disadvantage is that if the leader robotic unit fails to perform as planned or properly communicate with the followers, then the whole swarm fails to complete the task.

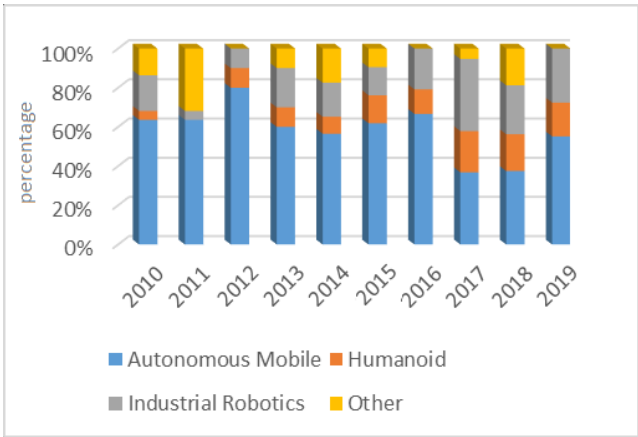


Figure 4: Robot types

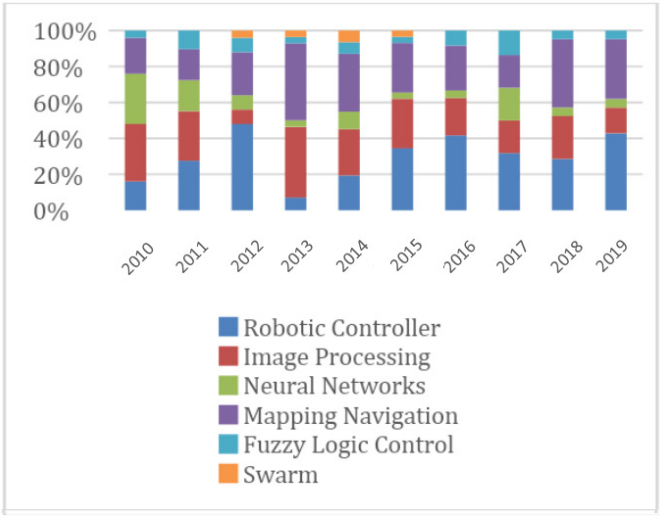


Figure 5: Distribution of FPGA applications for the period of 2010-2019

The second approach is the decentralized system that operates without a leader robotic unit and it is not affected by the cases of a single swarm unit failure or miscommunication amongst swarm members (agents). According to this approach, the motion of each robot is coordinated using local information exchange such as speed, obstacle existence or geographical coordinates, by the closest neighbor agents at a given time period. These motions are usually developed according to animal aggregations. Also, there is a physical decentralized system based on the physical interaction among agents. The communication amongst robots is either directional, by exerting forces to each other or in directional by manipulating or placing a common object. Directional communication may be applied by transmitting/receiving digital signals (WiFi, Bluetooth, etc), sound or light. In directional communication may be applied by either leaving optical markings in the ground or placing/manipulating an object in the environment that will be used as a marking point by other agents [20]. In total, the centralized architectures are simpler and used primarily for navigation purposes while decentralized architectures are more complicated and can be used for performing activities such as simulations, biotechnology or military tasks.

In Figure 4 and Figure 5 the robot types and the distribution of FPGA applications found in this survey are depicted, according to the total count of 200 publications tested.

4. Reviewed Literature

In this section, we present and analyze the examined research publications of the IEEE & ACM database. For each application field, we present the publications by giving a short description for each publication. Only selected publications that utilized high performance, throughput and efficiency were selected. This description for each publication includes a summary of the publication, the key points and its specifications.

4.1. Application Field: Image Processing

In [161] we see an approach about a real time image processing patrol service robot. The robot consists of a CMOS camera module that sends data to the FPGA board. The genetic algorithm optimizes threshold values for the filtering operations according to various lighting environments. The visual navigation system has been developed for the linkage between flexible hardware circuits and real-time software applications for robot vision purposes. Experimental results show the feasibility of the proposed methodology.

In 2013, an FPGA algorithm used to detect corners in colored stereo images by implementing Harris corner detection technique [155]. The FPGA used is the XC5VLX50 and the architecture consists of 3 pipeline modules consisting of OmniVision image sensors in RGB555 format. Data is stored into a single DDR2 memory of Micron. By using less than 75% of FPGA resources and a 100MHz system clock, the authors achieved a corner detection rate of 0.33 pixels per clock cycle (ppcc) corresponding to a corner detection frequency of 54Hz for the stereo images. They highlighted the performance benefits of their methodology by conducting experiments.

A novel human-robot communication algorithm for both voice recognition and image processing embedded in an autonomous mobile robot [149] is proposed. The main controller is built on a Spartan FPGA processor, processing visual and voice information. The result of the research is a very flexible and powerful development tool for human-robot multimodal communication research and implementation. The hardware resources of this robot are wisely integrated by using the LabView graphical software toolkit, enabling a large amount of voice signals and image processing algorithms testing. The designed software drivers also allow various sensor modality including vision and speech signals analyzing. Endowed with these novel capabilities, the mobile robot displays intelligence and cooperativeness in its behavior, and becomes a versatile hardware system for human-robot multimodal communication implementation.

In 2014, researchers proposed a bio-inspired event-driven vision system using pseudo-simultaneous sensing [137]. The control system consists of a Spartan6 FPGA implementing a Convolutional Neural Network and a 3D construction algorithm. Experimental results demonstrate the application of this paradigm to implement Gabor filters and 3D stereo reconstruction systems. This architecture can be applied to real systems which need efficient and high-speed visual perception, like vehicle automatic driving, robotic applications in unstructured environments, or intelligent surveillance in security systems. Experimental results highlight the performance gains of the proposed methodology.

The authors of [147] published a research paper about a hexapod (6-legged spider like) robot using an FPGA-based robotic platform (Spartan 6). The system architecture involves an image sensor controller, a servo-motor controller and an AVR processor. Overall, the implementation of the servo motor controller in an FPGA has offered several advantages in terms of circuit design flexibility and simultaneous command executions when compared to conventional microcontroller-based system.

In 2014, researchers proposed an FPGA-based OV driver for VGA cameras [132]. For their implementation, they used an FPGA board (Cyclone EP2C5T144C8N) and optimized it using VHDL. The results from the conducted experiments suggest that the devised solution outperforms when compared to the coexisting implementations in terms of resources utilization.

Researchers published a research paper implementing a visual sensor unit for mapping and localization using a six-dimensional Micro Aerial Vehicle (MAR - drone) [126]. The sensor design is based on a XILINX FPGA System-on-Chip combining FPGA resources with an ARM A9 Dual-core processor

One year later, a group of researchers proposed a real-time machine vision system for an industrial robot that grasps items from an assembly line [119]. The items are in the same class, meaning that they have the same general shape but different details. The conducted experiments suggest that the machine parts can be located accurately under arbitrary in-plane rotations and can be classified correctly according to the details in their shapes. Guo's system can run with an industrial camera at a resolution of 640×480 and a speed of 50 fps (frames per second) or higher.

In 2015, the authors of an interesting research paper proposed a pragmatic incremental algorithm for achieving gain control and white balancing on an FPGA [112]. The proposed methodology significantly reduces the effects of the variations in lighting on the thresholds. In their paper, they propose a new color space which maximises the hue separation of the different coloured regions to improve the colour discrimination. All of the algorithms are designed to operate on data directly streamed from the camera, enabling a low latency FPGA implementation. The authors conducted real world experiments to evaluate the stability and performance of the proposed prototype.

In 2015, a group of researchers proposed a humanoid robot head with a vision system which has 22 degrees of freedom, including eyebrows sliding mechanism, the eye mechanism, jaw and neck mechanism [108]. The control system of the proposed prototype is an FPGA board. Facial movements are realized through driving traction mechanisms by controlling servos.The vision system of the robot S HFR-II Ihas a human-computer interaction interface with four kinds of work patterns,which includes facial expression recognitions and representations, face recognitions, gender recognitions and expression representations. The authors conducted experiments and the results suggest that the robot can recognize six basic expressions and vividly reproduce them.

Similarly to the aforementioned paper, in the same year, a group of researchers published an implementation of a face recognition system that captures facial images and recognizes a face and detects slight expressional changes of the test image [104]. In this paper, the authors read various facial images and store them. Then, images are compared bit by bit and any mismatch is detected. If the image is matched then a "Match found" message is displayed.

The author of [97] proposed a hardware accelerator to reduce the computation time of the scale-space analysis and the feature detection of the KAZE algorithm on Zynq-SoC family of FPGAs.

An interesting paper is proposed by [91] demonstrates an implementation of image mosaic algorithm on the Xilinx Zynq-7020 FPGA board. The system is able to process more than 60 frames per second (fps) and is capable of maintaining low power consumption. The authors verified the proposed methodology in a synthetic environment simulated using MATLAB.

A real time target tracking and positioning system based on FPGA is proposed by the authors of [89]. The hardware platform for the design is the development board by Xilinx´The system uses monocular camera to measure the orientation of target in real time. It acquires the video stream from a CMOS camera. The system realizes target tracking by extracting the color of the target as the feature, filtering detecting feature of the target and then displaying on the screen. The system realizes target positioning according to the centroid of the tracked target, then computing the orientation which can be used for the intelligent robot vision.

A group of researchers proposed a system that integrates different techniques to perform dimensional measurement of thick steel plate bevels using a DE0-Nano FPGA board [88]. The groove modeling is achieved through the use of line detection algorithms that look for straight lines and segments. The system level modeling and synthesis and the integration to the whole digital linear welding system, using only 28% of the FPGA's slices. Results Show the approach is a viable option to automate the process that is currently widely adopted in industry.

An interesting work about a high speed ball target tracking system based on FPGA (ZYNQ) [85] is proposed. With a CMOS sensor, the original frame can be transferred to the FPGA port. After the video streams go through the parallel image processing core, the color and corner features can be extracted. When compared to the predefined target's pattern, the target can be detected. The authors conducted real-world experiments to verify the feasibility of the proposed methodology.

In 2017, a group of researchers published their paper about an FPGA based real-time virtual blind cane system for the visually impaired, detecting obstacles in front of the user in real time [70]. Described three novel architectures of image segmentation, implemented on Virtex UltraScale XCVU190-2FLGC2104EES9847 FPGA. Working as a blind cane, the proposed system helps a blind person find the type of obstacle and the distance to it. The distance from the user to the obstacle is estimated by extracting the laser coordinate points on the obstacle, as well as tracking the system pointing angle. The paper provides a simple method to classify the obstacle's type by analyzing the laser intersection histogram. Real experimental results were also presented in the paper to highlight the validity and accuracy of the proposed system.

The same year, a group of researchers presented three novel architectures of image segmentation, implemented on Virtex UltraScale XCVU190-2FLGC2104EES9847 FPGA [69]. The sequential design which allows control over minute details of arithmetic design, The hybrid architecture provides a full fledged parallel and pipelined implementation of the algorithm and the hybrid architecture overall provides at least two times gain in acceleration.

One year later, a group of researchers propose a real time tracking hardware based on FPGA [60]. They converted a simulated and verified MATLAB code directly into HDL. The key point of their research is that the simulated and hardware outputs were consistent. Three types tracking algorithms (Kalman filter, extended Kalman filter and αβγ filter) were implemented to validate the proposed methodology.

In 2018, a research paper about a design approach of accelerating the execution of a robust algorithm based on Horn's quaternion solution [59] is proposed. The authors presented a HW architecture as well as its FPGA implementation. In their research paper, they also included results of the experiments they conducted.

Figure 6 depicts the total number of research papers that implemented FPGA Image Processing and the total Image Processing papers in robotics, from 2010 to 2019. It is evident that the popularity of FPGA and Image processing is steadily growing.

4.2. Application Field: Robotic Controller

A group of researchers propose an inverse kinematic algorithm for a parallel milling robot. The proposed robot consists of three robotic arms connected to a moving platform inside which there is a rotating spindle along a trajectory in the cartesian x, y, z space. The system is implemented using Altera Cyclone IV FPGA chip and it is evaluated by experimental results.

A group of researchers propose a robotic platform based on Lego Mindstorms in [162]. In their research, they included hardware, software and mechanical modules. A mathematical model for the motor and the PI controller are included. Experimental results highlight the performance gains of their implementation.

A system architecture controlling and monitoring household appliances through GSM. The system consists of an FPGA Xilinx SP605 implemented in VHDL, the GSM circuit, the PIR sensor and the interface circuit that monitors the appliances

A connecting algorithm for interfacing the FPGA-based controllers to high level robotic software frameworks. The Unity-Link algorithm conjoins PC-processed architectures with nodes that provide demanding real-time control of distributed robotic systems

A group of researchers published their paper about an inverse kinematics FPGA algorithm for a biped robot, based on CORDIC operators of hyperbolic and circular vectoring. The speed of the proposed hardware FPGA application is compared to that of the Nios II CPU. The authors evaluated the proposed methodology by conducting experiments in a synthetic environment.

In 2019, the authors of [25] proposed an FPGA implementation of an adaptive neuro-fuzzy control for a teleoperation system. The controllers were developed in the MATLAB-Simulink environment and were implemented on FPGA using Simulink's fixed point tool and HDL Coder. The experimental results were very promising, as the latencies of these filters on the Field Programmable Gate Array (FPGA) platform were below 300ns with an average latency reduction of 188% (maximum of 570%) over the software versions running on a desktop PC CPU. This open-source event-based filter IP library for FPGA was tested on two different platforms and scenarios using different synthesis and implementation tools for Lattice and Xilinxvendors.

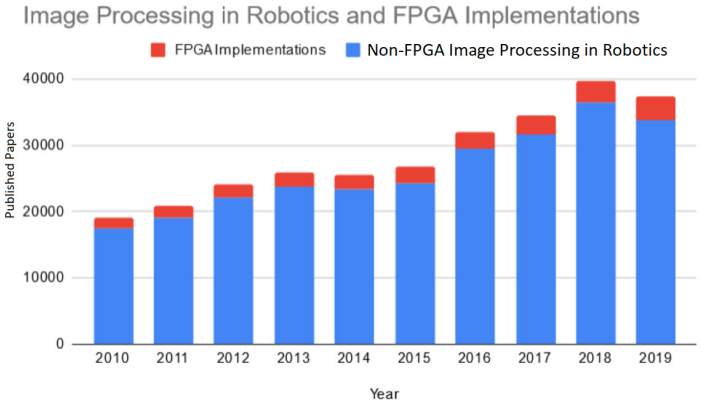


Figure 6. Research papers published about Image Processing using FPGA in Robotics compared to non-FPGA Image Processing papers in Robotics

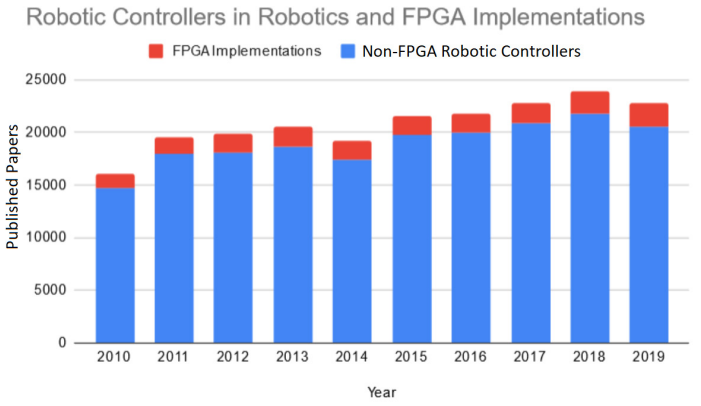


Figure 7. Research papers published about Robotic Controllers using FPGA compared to non-FPGA Robotic Controller papers

Figure 7 depicts the total number of research papers that implemented FPGA Robotic Controllers and the total Robotic Controllers papers in robotics, from 2010 to 2019. It is evident that the popularity of FPGA and Robotic Controllers is increasing.

4.3. Application Field: Neural Network

In 2012, the authors proposed an FPGA-based Neural Network algorithm for lane following of autonomous mobile robots [164]. The neural function consists of two main parts, the fixed points coding and the Sigmoïde function coding. The VHDL simulation is compiled in A Quartus II and implemented in Cyclone EP2C35F484C6 FPGA. The experimental results are promising and suggest high performance gains.

In [163], the authors publish their research paper about a stereo vision landing system architecture for an Unmanned Aerial Vehicle (UAV). The proposed FPGA-based Artificial Neural Network (ANN) algorithm is designed to apply real-time object tracking, 3D position computation via Visual Odometry methodology, Euclidean distance estimation from the landing target and horizontal displacement. The algorithm implementation was done on a Xilinx XC2V100 FPGA chip

In the same year, the authors proposed a Neural Processing Unit implementation [120] which reached a range from 6k to 90k iterations per second for the targeted neural map size. The authors propose a Neural Processing Unit designed as a programmable accelerator implementing recent equations close to self-organizing maps and neural fields. The proposed architecture is validated on FPGAdevices and compared to state of the art solutions. The trade-off proposed by this dedicated but programmable neural processing unit allows us to achieve significant improvements and makes our architecture adapted to many embedded systems.

An alternative architecture that exploits stochastic computation for doing classification with deep belief networks was proposed by the authors of [110] in 2015. The network was trained offline in MATLAB on the MNIST training dataset, and then implemented on an FPGA for classification at variable stochastic computation precision. Experimental results highlight the performance gains of the proposed methodology.

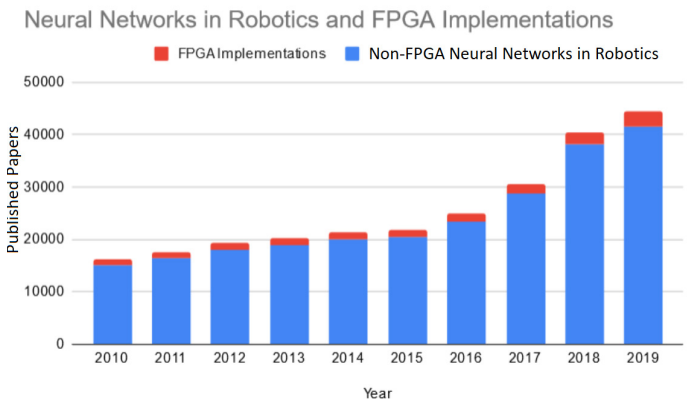


Figure 8. Research papers published about Neural Networks using FPGA in Robotics compared to non-FPGA Neural Network papers in Robotics

Figure 8 depicts the total number of research papers that implemented FPGA Neural Networks and the total Neural Network papers in robotics, from 2010 to 2019. The popularity of FPGA and Networks is increasing steadily. Compared to the early 2010s, the Neural Networks using FPGA and the Neural Networks research papers in general have increased by a factor of 2.92.

4.4. Application Field: Mapping & Navigation

In [146], a color tracking algorithm was proposed for autonomous mobile robots, using a CMOS camera and an ultrasonic sensor. The design implementation was done on a Digilent Nexys2 FPGA. The video is captured and further processed by the FPGA board and displayed on a 3.2 inch screen. They conducted real world experiments to evaluate the stability and performance of their prototype.

A group of researchers proposed a processor embodying a Radio Frequency Identification Detection (RFID) algorithm and a CORDIC algorithm on a Kintex-7 FPGA platform [140]. The proposed system is used for navigation and anti-collision purposes. Experimental results highlight the performance gains of the proposed methodology.

In [135], a Micro Aerial Vehicle (MAV) architecture was proposed consisting of two subsystems: a) An image processing subsystem consisting of two VGA cameras and an Inertial Measurement Unit (IMU) and b) A control system using an Altera SoC FPGA. The interesting aspect of this research is that motion estimation and outlier rejection algorithms were included.

In the same year, an FPGA-based in-door flight localization/navigation system with a pan-tilt camera was proposed [129]. The FPGA (XC6SLX100) was used for image analysis and quad rotor control. The authors validated the proposed methodology by conducting real world experiments.

The authors of [119] proposed an improved ant colony system algorithm for path planning by establishing two new mechanisms for pheromone updating. They included partial pheromone updating and opposite pheromone updating. The authors evaluated the proposed methodology by conducting experiments in a synthetic environment.

[114] presents an effective data cache configuration where the number of active ways and hence its total capacity can be configured at run-time. When using the best data cache configuration, a reduction of 7% the power consumed at a 1% execution time penalty is suggested and it is the same result for the SLAM-EKF algorithm.

The authors of [113] present a novel methodology based entirely on the on-board approach. The obstacle avoidance system that is proposed is suitable for running on low-power embedded devices, leveraging a light-weight low power stereo vision system on FPGA used on Micro Aerial Vehicles (MAVs).

A group of researchers proposed a robot navigation system using a combination of sophisticated image processing algorithms [111]. Their system is used to generate a map and an extended A* algorithm to plan the path. A turtlebot localizes itself in an environment containing several obstacles and a destination. The authors evaluated the proposed methodology by conducting experiments in a synthetic environment.

[109] presents an FPGA-based hardware architecture of EKF Prediction Stage applied to the Mobile Robotics Localization Problem. It runs on an Altera Cyclone IV FPGA with a Nios II processor, adapted to the mobile platform Pioneer 3AT. The experimental results look very promising, as the execution time results point out that the hardware solution has a speed up factor greater than 2 in comparison with the embedded software implementations.

In 2015, a hardware structure for a vision-based autonomous vehicle navigation system was proposed [107]. Results showed that it is possible to implement the structure without consuming a high number of FPGA resources and achieve acceptable execution times. Analyzing the final hardware structure this goal was achieved with a considerably high frame-rate.

In the same year, an investigation was conducted [100]on the amount and sophistication of sensing and processing hardware needed by a mobile robot for performing high-level tasks in real time, such as azimuth calculation and dynamic obstacle tracing. The paper suggests that one ultrasonic sensor along with a coordinate rotation digital computer processor are adequate for azimuth calculation while one more ultrasonic sensor is required for dynamic obstacle tracing.

In [98], a hardware architecture is proposed to accelerate Rapidly-Exploring Random Trees (RRT) path planning algorithm using multiple Kd-trees to conduct the parallel search. The authors evaluated the proposed methodology by conducting experiments in a synthetic environment.

One year later, the authors of [78] propose a technique to evaluate sparse equations on an FPGA by restricting the maximum amount of items in the system. The implementation is done using CλaSH which allows a transformation from mathematical descriptions to a hardware design. The system uses a sparse data notation and puts constraints on the number of loop closures to restrict the amount of computations. Experimental results highlight the performance gains of the proposed methodology.

The first multilevel implementation of the Harris-Stephens corner detector and the Oriented FAST and rotated BRIEF (ORB) feature extractor running on FPGA hardware was published [74] in 2017. The Tarsier device performs excellently as it successfully generates descriptors for real images with sufficiently low latency for integration into a 1080p resolution 60 fps loop.

In the same year, the authors of [62] present their research about the development of a neuromorphic system on a Spartan 6 FPGA generating locomotion patterns (gaits) for three different legged robots (biped, quadruped, and hexapod). The authors validated the proposed methodology by conducting real world experiments. Moreover, the proposed spiking neuron model consumes less hardware resources and power than those based on coupled.

In 2019, an angular position tracking control was implemented using modified D channel PID algorithm [40]. It was developed in such a way that the reading of data, the calculation of the control signals, and the controlling of the motors are all implemented on FPGA. Experimental results highlight the performance gains of the proposed methodology.

Figure 9 depicts the total number of research papers that implemented FPGA Mapping & Navigation and the total Mapping & Navigation papers in robotics, from 2010 to 2019. It is evident that the popularity of FPGA and Mapping is increasing. This could possibly be caused by the general increased popularity of autonomous vehicles (Unmanned Aerial Vehicles, Unmanned Ground Vehicles and Unmanned Underwater Vehicles).

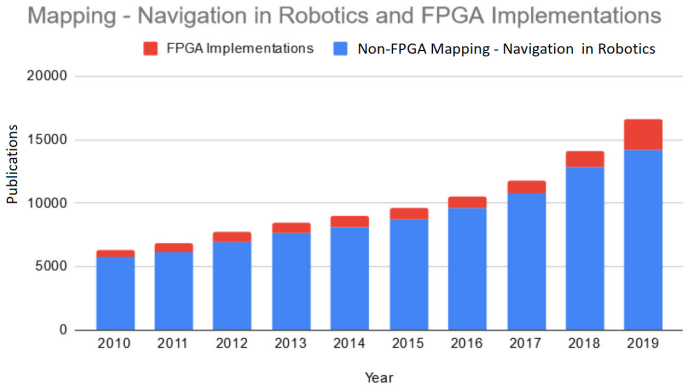


Figure 9. Research papers published about Mapping & Navigation using FPGA in Robotics compared to the non-FPGA Mapping & Navigation papers in Robotics

4.5. Application Field: Fuzzy Logic Control

In 2012 a fuzzy-logic controller [168] for an autonomous navigation robot was proposed. The controller is implemented using an Atmega16 microcontroller and tested using the whole robotic application. The FPGA implementation consists of the Fuzzy logic controller, the UART receiver (it translates data between serial to parallel form), the LCD controller (for controlling the LCD operations) and the ASCII to binary converter and binary to ASCII controllers.

In the same year, an interesting work is proposed about an autonomous parking FPGA-based system for a mobile robot [167]. The system scans the area for available parking space and performs the necessary maneuvers for parking. The whole system consists of three major parts, the car-like robot design, the development of a simulation model using MATLAB and Simulink and the fuzzy logic controller design. The evaluation of the paper was performed using a synthetic realistic environment.

A group of researchers published their research paper about an autonomous mobile robotic platform [165]. Their prototype is designed within the context of a robotic course. The system uses two separate FPGA-based controllers: A Proportional integral derivative (PID) controller and a fuzzy logic-based controller. Apart from the FPGA card, the robot's hardware design consists of two motors, eight sonar sensors, a radio frequency identification reader and a Basic Stamp microcontroller. For the modeling and implementation, they utilized MATLAB and Simulink.

Two fuzzy logic algorithms are developed: one for avoiding obstacles and another for avoiding obstacles / reaching a specific point. For the obstacle avoidance an ultrasonic sensor is used, while software is implemented using LabVIEW modules

In 2013, an interesting work is proposed about a fuzzy logic algorithm that is applied on an autonomous mobile robot for line tracking and obstacle avoidance tasks [145]. The software is

implemented by an Altera Quartus platform while Eclipse software is used to simulate the fuzzy rules. Several practical tests that were conducted by the authors reveal the effectiveness of the proposed fuzzy rules for which MRTQ was able to track the path more precisely and softly at different speeds and avoid the detected obstacle through the path readily.

In 2013, a paper was presented about a fuzzy-logic color extractor used on NI myRIO embedded system package [133]. In total, the fuzzy logic module, the real-time module, and the FPGA LabVIEW module are required for the implementation of the extractor on myRIO embedded device. The authors conducted multiple experiments and evaluated its effectiveness and performance gains.

In 2015, an interesting work is proposed about an omni-directional mobile platform control using the artificial intelligence technique of Fuzzy Logic; the control allows a practical and reliable driving control of 4 omni-directional wheels implemented on FPGA. The experimental results show increased performance and low power consumption.

Figure 10 depicts the total number of research papers that implemented FPGA Fuzzy Logic Control and the total Fuzzy Logic Control papers in robotics, from 2010 to 2019.

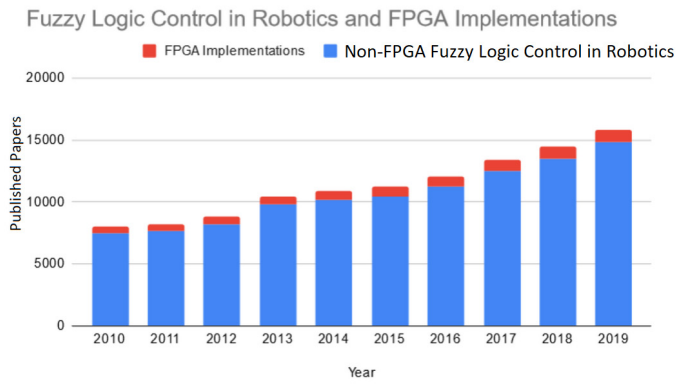


Figure 10. Research papers published about Fuzzy Logic Control using FPGA in Robotics compared to non-FPGA Fuzzy Logic Control papers in Robotics

4.6. Application Field: SWARM

In 2015, an interesting work is proposed about a behavioural control algorithm of a centralized swarm consisting of two robots [101]. In this paper, the authors present the challenges of multi robots and how the leader and follower robots can change their mode of behavioural control by switching between centralization and decentralization methods with a novel approach using implicit communication. According to the authors and their experimental results, the proposed implementation is very energy efficient. Robots are developed with minimal sensors (ultrasonic sensors, digital compass, RF Transceiver (PMOD-RF2)) and are implemented using a Spartan 3e FPGA board

Another interesting implementation was proposed by the authors of [84] in 2014. The heuristic approach based leader follower method is incorporated between robots for behavioural control. The robots were able to navigate in both static and dynamic conditions by using the modified M* based obstacle avoidance mechanism with CORDIC module. The proposed methodology was validated by the implementation of the algorithm on Spartan 6 FPGAs (XC6SLX45 CSG324C).

In 2015, an interesting work is proposed about a pragmatic incremental algorithm for achieving gain control and white balancing on an FPGA [112]. The proposed methodology significantly reduces the effects of the variations in lighting on the thresholds. In their paper, they propose a new color space which maximises the hue separation of the different coloured regions to improve the colour discrimination. All of the algorithms are designed to operate on data directly streamed from the camera, enabling a low latency FPGA implementation.

Figure 11 depicts the total number of research papers that implemented FPGA SWARM and the total SWARM papers in robotics, from 2010 to 2019. Swarm and FPGA applications in robotics are not as popular as other fields, like mapping, image processing and controllers. However, SWARM and FPGA are still rising in popularity over the years.

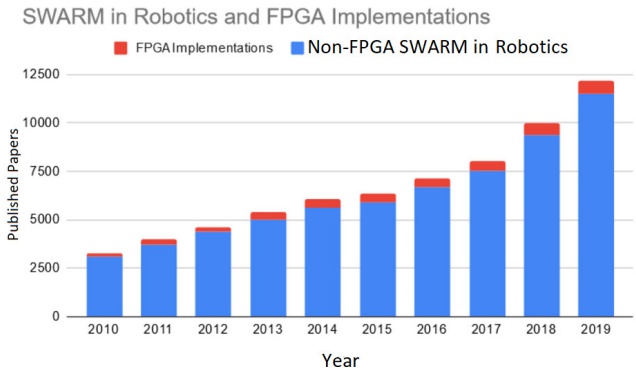


Figure 11. Research papers published about SWARM using FPGA in Robotics compared to non-FPGA SWARM papers in Robotics

5. Discussion

In the 2000s, FPGAs were seen as an inexpensive alternative to the traditional ASICs (Application Specific Integrated Circuit) for heavy computation tasks. Along with the decreased production cost and the technological advancements, this led to the increased popularity of FPGAs. This time period had significant improvements. Most importantly, they were software as well as hardware programmable. In boards released in the late 2000s, for the first time, the microprocessor was not the simple hardware block dropped into the FPGA as was done in older days, but included a full environment with caches, busses, Network-on-Chip and peripherals. Bundled software included operating systems, compilers and middleware: an entire ecosystem, rather than an integrated function block. The increased capabilities of FPGA boards of this era led to the increased usage in robotic applications, such as image processing, robotic controllers, mapping and navigation. In these fields, 10% of the published research papers utilize FPGA boards, presumably for the performance benefits they offer.

In the early and mid 2010s, FPGAs started to gain prominence in another field: neural networks, fuzzy logic control and swarm. Running neural network inference models takes significant

processing power. Graphics processing units (GPUs) are often used to accelerate inference processing, but in some cases, high-performance FPGAs might actually outperform GPUs in analyzing large amounts of data for machine learning.

Just as customers seeking custom integrated circuits 30 years ago were attracted to FPGAs over the complexity of ASICs, many are now attracted to multicore processors, graphic processing units (GPU) and software-programmable Application Specific Standard Products (ASSPs). These alternative solutions provide pre engineered systems with software to simplify mapping problems onto them. They sacrifice some of the flexibility, the performance and the power efficiency of programmable logic for ease-of-use. It is clear that, while there are many FPGA users who need to exploit the limits of FPGA technology, there are many others for whom the technological capability is adequate, but who are daunted by the complexity of using that technology.

The complexity and capability of devices have driven an increase in capability of design tools. Modern FPGA tool-sets include high-level synthesis compilation from C, Cuda and OpenCL to logic or to embedded microprocessors. Vendor-provided libraries of logic and processing functions meet design costs. Working operating systems and hypervisors control FPGA SoC operation. Team modules, including control, are built into FPGA design systems. Some capabilities are built by the vendors themselves, others are part of the growing FPGA ecosystem. Clearly, usability is critical to this next age of FPGAs and their adoption in robotics. If FPGA manufacturers focus on ease of use and solve the steep learning curve, FPGAs will be able to compete with GPUs and become dominant in high performance applications.

6. Conclusions

This survey revealed popular technological, research and industry trends of the past 10 years, regarding robot types used and processed in the research as well as application fields of the FPGAs. As depicted in Figure 4, the vast majority of robots for which applications are being developed are autonomous mobile robots especially when the algorithms developed for these robots are related to localization, mapping and navigation purposes.

In order to further increase the execution speed and efficiency of algorithms, developers utilize FPGA boards, as they offer an innovative solution. The popularity of FPGA boards has been steadily increasing over the last decades, as we can see in Figure 1. FPGAs offer significant advantages over software running on a CPU or GPU, including execution speed and energy efficiency. In a traditional local system, the FPGA board is typically located on-site, along with the other components of the device. However, in modern systems, FPGA boards can be found in the cloud, along with a high performance traditional server. As the server receives data from modern embedded IoT systems, images the server offloads the necessary information to be processed on the FPGA boards. Many boards support PCIe connectivity, with a maximum theoretical bandwidth of 15.75 GB/s in version 3.0. The main disadvantage of FPGA boards, however, is the increased engineering cost and complexity. The development process on FPGA boards includes the design of custom hardware circuits. Traditionally, these hardware circuits are described via Hardware Description Languages (HDL), such as VHDL and Verilog, whereas software is programmed via one of a plethora of programming languages, such as Java, C and Python. An upcoming trend on FPGA is High Level Synthesis (HLS). HLS is allows the programming of FPGAs using regular programming languages such as OpenCL or C++, allowing for a much higher level of abstraction. However, even when using high level languages, programming FPGAs is still an order of magnitude more difficult than programming instruction based systems. Accelerating the computation process via FPGA boards is only efficient at larger scales, where the performance gain of the system outweighs the increased complexity and engineering cost.

The parallel architectures offered by the FPGAs easily offer the high computational power that continuous real-time localization and navigation tasks demand at a low cost, while mixed FPGA/DSP controllers offer flexibility over the process of complex image data that robotic visual systems include. This particular finding is in accordance with the high number of mapping / navigation application fields shown in Figure 5, for each consecutive year. As shown, the mapping and navigation application field is between half and one third of the total application fields for the 10-years tested period.

As depicted in Figure 4, the field of industrial robotics is the second most usually appearing robot field in the 2010-2019 publication search. FPGAs constitute a flexible solution in the implementation of kinematics algorithms that parallel multi-axis industrial robots often use, in comparison to DSP controllers. The example of the commonly used CORDIC algorithm implemented for rotation tasks in industrial robotics reveals the flexibility of FPGAs. That specific algorithm includes a significant extent of mathematical functions, so the development of a VLSI-based FPGA algorithm is often preferred as the implementation architecture of building a controller for multi-axis motion of robotic limbs. Figure 5 shows the high number of robotic controllers found in the survey, representing the flexibility of FPGAs in industrial robotic control.

Humanoid robots are the third most commonly appeared robot type in the survey. The multi-sensory humanoid robotic applications often demand multi-processor controllers based on mixed DSP/FPGA circuitry. The FPGA connectivity with DSP processors constitutes an effective solution of robotic applications that require multi-processing. The FPGA application fields applied to humanoid robots include regular robotic controllers for motion purposes, image processing applications for the process of visual data for pure image processing or mapping and navigation tasks, or neural network architectures for navigation or else. Swarm robotics also constitutes a wide FPGA application field. Complex behavioural algorithms which demand high computational power and high-levels of real-time communication between swarm members, can be easily implemented through FPGA architectures.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research has been implemented in the framework of the Action “Business Enhancement for Research projects” and co-

financed by the European Union and national resources through the B.P. WESTERN GREECE 2014-2020 (project code: DEP5 - 0019433). Table 1 contains other significant works that were reviewed but not fully addressed in section 4.

Table 1: Other significant works that were reviewed

Field	Description	Year	Ref
4.1	A fast frame based on Block RAM reuse	2019	9
4.1	4.1 of a SoC FPGA using RGB-D sensor	2018	21
4.5 / 4.2	A hardware system for on-board support vector machine (SVM) model training for robotic transtibial prostheses locomotion recognition based on SoC-FPGA	2018	22
4.1	FPGA implement for multi-scale pedestrian detection using 4 different size histogram of oriented gradients (HOG) features	2018	23
4.2	An exemplary system for the calculation of the inverse kinematics of a robot arm, which is tailored for the implementation on a FPGA	2018	24
4.5 / 4.2	FPGA-based platform for integrated multi-axis motion control and motor drive	2018	26
4.2	A high-speed, high-accuracy fully pipelined architecture to solve inverse kinematics of a 3-DOF robot leg	2018	27
4.1 / 4.3 / 4.4	SqueezeNet DCNN is accelerated using an SoC FPGA in order for the offered object recognition resource to be employed in a robotic application	2018	28
4.3	A low cost and robust deep-learning based on gaze estimator is proposed to control surgical robots	2018	29
4.5 / 4.2	A novel robot control system based on EtherCAT In order to increase control frequency and expansibility for 7-degree-of-freedom (DoF) light-weight	2018	30
4.2	A control system architecture for knee joint surgery robot based on computer and motion (FPGA) control card	2018	31
4.4 - 4.3	An implementation of mechanical, electronics and software for a man-portable AUV with hybrid architecture	2018	32
4.1	A novel heterogeneous computer-on-module, integrating a Xilinx Zynq SoC and an Adapteva Epiphany multi-core processor	2018	33
4.3	A framework to integrate Deep Learning algorithms on the PYNQ-Z1 at the embedded system level	2018	34
4.2	An open-architecture controller with modularity	2018	35
4.1 - 4.4	A wireless panorama camera development specially for snake robot sensing	2018	36
4.2	A novel distributed SoC-based (PS + PL) impedance controller solution implemented in a robotic hand	2018	37
4.3	Implementation of the You only look once (YOLO) object detector on an FPGA	2018	38
4.1	A design approach of accelerating the execution of a robust algorithm based on Horn's quaternion solution and presents a HW architecture as well as its FPGA implementation	2018	39
4.4	Development of a Triangulation based Navigation Algorithm implemented on FPGA	2017	41

4.2	A mixed FPGA-microprocessor control architecture for a multi-axis industrial robot with vision-based behavior	2017	42
4.2	Robotic application performing auto-detection for high-voltage insulator inspection. The FPGA controller is used for step motor control, communication and signal acquisition	2017	43
4.2	A mixed FPGA-Microcontroller control architecture is included, using a Kawasaki RA10N robot body	2017	44
4.2	A high precision model of minimally invasive surgical robot	2017	45
4.2	A robotic system using a hard processor plus FPGA solution where the FPGA performs all real-time tasks, freeing the processor to run lower-frequency high level control and interface to other devices	2017	46
4.3	An FPGA implementation of Q-learning with Artificial 4.3 (ANN) reducing processing time	2017	47
4.2	A robust method based on a Dual-Core ARM Cortex-A9 System-on-Chip FPGA board that controls the motion of a Cartesian robot	2017	48
4.1 - 4.4	Three novel architectures of image segmentation, implemented on Virtex UltraScale XCVU190-2FLGC2104EES9847 FPGA. The sequential design which allowing control over minute details of arithmetic design, The hybrid architecture provides a full fledged parallel and pipelined implementation of the algorithm and the hybrid architecture overall provides at least 2X gain in acceleration	2017	49
4.1 - 4.4	An FPGA based real-time virtual blind cane system for the visually impaired, detecting obstacles in front of the user in real time. Described three novel architectures of image segmentation, implemented on Virtex UltraScale XCVU190-2FLGC2104EES9847 FPGA	2017	50
4.4	A novel approach of a robot navigation in smart indoor environment developed with the triangulation techniques using Coordinate Rotation Digital Computer (CORDIC) method	2017	51
4.2	Using inverse kinematics and zero moment point algorithmic program (IK-ZMP) to command robotic legs in parallel manner which algorithms were programmed in VHDL	2017	52
4.2	An overall design of the triple inverted pendulum on a cart model which can be used to illustrate various control design problems related to underactuated nonlinear mechanical systems	2017	53
4.4	The first multilevel implementation of the Harris-Stephens corner detector and the Oriented FAST and rotated BRIEF (ORB) feature extractor running on FPGA hardware, for computer vision and robotics applications. The Tarsier device performs excellently-it successfully generates descriptors for real images with sufficiently low latency for integration into a 1080p resolution 60 fps loop	2017	54
4.2	An open and high-performance industrial robot controller which consists of dual MCU + FPGA	2017	55
4.5 - 4.3	A swarm-based algorithm using spiking 4.3 (SNN) in order to simulate biological neuronal assemblies. The FPGA SNN system uses an FPGA Artix-7 chip executing mapping tasks	2017	56
4.3 - 4.2	A cerebellum-like spiking neural network model to be used as the short-range timing function for the talking robot's control system	2017	57
4.4	A technique to evaluate sparse equations on an FPGA by restricting the maximum amount of items in the system. The implementation is done using CλaSH which allows a transformation from mathematical descriptions to a hardware design. The system uses a sparse data notation and puts constraints on the number of loop closures to restrict the amount of computations	2017	58
4.5	A walking robot developed with the FPGA controller for investigating the robot's performance for moving in any direction and avoid obstacles. An FPGA Altera Max II microcontroller has been used, programmed with VHDL and fuzzy logic functions to control the walking robot	2016	61

4.4	An FPGA-based control unit is designed for a patrol robot, using a Spartan 3E board. Depth-first search (DFS) algorithm is used for visiting all the nodes in the patrolling path and this is implemented using black tags (RFID) and IR sensors	2016	63
4.4-4.6	A multi robots path planning like shortest path, behavioral control and obstacle avoidance using a Spartan 6 FPGA with modified M algorithm has resulted with less distance by that it consumes less power	2016	64
4.1	A high speed ball target tracking system based on FPGA (ZYNQ)	2016	65
4.4	A prototype for an autonomous robot using the FPGA DE0-Nano platform to implement. The system designed for having easier maintenance, control and, environmentally friendly	2016	66
4.2	Analyzing and accelerating the inverse Jacobian, pseudo inverse Jacobian, gradient descent, and the Newton method through an FPGA design using high-level synthesis (HLS). Implemented four numerical methods for solving the inverse kinematic problem of a 4 Degrees Of Freedom (DOF) robotic arm	2016	67
4.1	A system that integrates different techniques to perform dimensional measurement of thick steel plate bevels using a DE0-Nano FPGA board. The groove modeling is achieved through the use of line detection algorithms that look for straight lines and segments	2016	68
4.1	An implementation of image mosaic algorithm which is comprised of image registration and image fusion for better effect on the Xilinx Zynq-7020 FPGA board. The system is able to process more than 60 frames per second (fps) and maintaining a low power	2016	71
4.4 - 4.1	A processor-centric FPGA-based architecture for a latency reduction in the vision-based robotic navigation. Also, it can be used to other tasks like detection and tracking of image salient points	2016	72
4.4 - 4.1	A 6 Degree of Freedom (DoF) camera pose estimation algorithm using a monocular camera and non co-planar markers, design for robotics applications, and an implementation on a low cost CPU and the FPGA that can conduct precise 6-DoF estimations at 100Hz	2016	73
4.2	An effective method for calculating the onboard real-time path planning of large-scale space redundant manipulators	2016	75
4.4 - 4.1	A vision-based lane detection algorithm which requires no knowledge of any physical parameters like position and orientation of the camera and is hence very flexible. The algorithm is tested with 14 videos and demonstrate super fast speed, high accuracy, and super energy-efficient	2016	76
4.1	A hardware accelerator to reduces the computation time of the scale-space analysis and the feature detection of the KAZE algorithm on Zynq-SoC family of FPGAs	2016	77
4.3	A design automation tool, named Deep Burning, allowing the application developers to build from scratch learning accelerators that target their specific 4.3 models with custom configurations and optimized performance	2016	79
4.4-4.1	An investigation on the amount and sophistication of sensing and processing hardware needed by a mobile robot for performing high-level tasks in real time such as azimuth calculation and dynamic obstacle tracing which shown that one ultrasonic sensor along with a coordinate rotation digital computer processor are adequate for azimuth calculation while one more ultrasonic sensor is required for dynamic obstacle tracing	2016	80
4.4 - 4.6	Behavioural navigation control algorithm of a centralized swarm consisting of two robots, using FPGA Spartan 3e board	2015	81
4.4 - 4.1	A practical analysis of a SoC architecture that incorporates an FPGA with an embedded ARM processor for Visual SLAM applications	2015	82
4.2	A design and implementation of a novel system for robot-assisted fracture manipulation. The system was evaluated through positioning trials to measure its accuracy and repeatability	2015	83

4.4 - 4.1	A real-time image de-blurring by dynamics-based approach and parallel system architecture	2015	86
4.4 - 4.1	A hardware structure for a vision-based autonomous vehicle navigation system. Results showed that it is possible to implement the structure without consuming a high number of FPGA resources and achieve acceptable execution times. Analyzing the final hardware structure this goal was achieved with a considerably high frame-rate	2015	87
4.3	An alternative architecture that exploits stochastic computation for doing classification with deep belief networks. The network was trained offline in MATLAB on the MNIST training dataset, and then implemented on an FPGA for classification at variable stochastic computation precision	2015	90
4.1 – 4.6	Pragmatic incremental algorithms for achieving gain control and white balancing have been described and implemented on an FPGA	2015	92
4.4 - 4.1	A novel entirely on-board approach, obstacle avoidance system that is suitable for running on low-power embedded devices, leveraging a light-weight low power stereo vision system on FPGA used on Micro Aerial Vehicles (MAVs)	2015	93
4.4	An effective data cache configuration where the number of active ways and hence its total capacity can be configured at run-time. When using the best data cache configuration reveal a reduction of 7% the power consumed at a 1% execution time penalty and it is the same results for the SLAM-EKF algorithm	2015	94
4.2	A finger control for an anthropomorphic piano robot. The robot can play the right note and right rhythm, the punching force of the fingers are controlled just right while the robot is playing	2015	95
4.4	An improved ant colony system algorithm for path planning by establishing two new mechanisms for pheromone updating, including partial pheromone updating and opposite pheromone updating	2015	96
4.1	A real-time machine vision system for industrial robot to grasp from an assembly line a class of machine parts which are similar in the general shape but different in details using FPGA for acceleration	2015	99
4.4 - 4.1	Image-processing navigation algorithm using C language	2014	102
4.2	A dynamic (mathematical) model of a robot finger simulating human motion. An FPGA-based EtherCAT master-slave platform is developed in order to validate the model	2014	105
4.1 – 4.4	A visual sensor unit for mapping/localization used by a six-dimensional Micro Aerial Vehicle (MAR - drone). The sensor design is based on a XILINX FPGA System-on-Chip combining FPGA resources with an ARM A9	2014	106
4.4 - 4.1	A Micro Aerial Vehicle (MAV) architecture consisting of two subsystems: a) An 4.1 subsystem consisting of two VGA cameras and an Inertial Measurement Unit (IMU) and b) A control system using an Altera SoC FPGA. Motion estimation and outlier rejection algorithms are included	2014	115
4.2 - 4.1	A high-speed image visual servoing (VS) system is developed for a SCARA robot. VS is the procedure of the right positioning of the manipulator in order to mitigate the effects of disturbances. The proposed system uses a Cyclone III FPGA controller with a camera and a Position Sensitive Detector (PSD) for position computation	2014	116
4.1 - 4.3	A bio-inspired event-driven vision system not represented by frames, using pseudo-simultaneous sensing. The control system consists of a Spartan6 FPGA implementing a Convolutional Neural Network and a 3D construction algorithm	2014	117

4.2	An FPGA smart sensor interface remotely communicating the main controller with the sensors. An EP3CI6F484 FPGA controller interfaces with SDRAM controller, vibration sensors and a camera. The robot can be used for disaster affected areas and other fields where remote communication is demanded	2014	118
4.2	A Kinematics architecture is applied on an Altera FPGA-based parallel industrial 3-axis milling robot, using the CORDIC algorithm	2013	121
4.2	A Kinematics algorithm is applied on an Altera FPGA-based parallel industrial milling robot, using C language.	2013	122
4.4 - 4.1	An FPGA Based Platform for localization/navigation through a software/hardware solution	2013	123
4.2	A modular robot manipulator consisting of three modules with embedded FPGA controllers, for increased precision. The manipulator uses the Virtual Decomposition Control (VDC) technique where the host computer executes only kinematics calculations, and all the other control calculations are done by local decentralized embedded slave computers	2013	124
4.5	A fuzzy logic algorithm is applied on an autonomous mobile robot for line tracking and obstacle avoidance tasks. The software is implemented by an Altera QuantusII platform while Eclipse software is used to simulate the fuzzy rules	2013	125
4.2	A Spartan-3 FPGA chip is used as a controller for a servomotor control technique implemented for a hexapod robotic platform. The robot consists of 18 servomotors. The controller is programmed with VHDL and the simulation is done with a Xilinx ISE Simulator	2013	127
4.4 - 4.3	A multi-layer neural network perceptron application using a Spartan III FPGA chip designed to drive two servomotors of an autonomous mobile robot. FPGA has been described in VHDL	2013	128
4.2	An ARM + FPGA multi-axis motion controller is implemented on an industrial robotic platform (SCARA robot). The controller is adapted to independent real-time network communication in order to achieve real time servo control	2013	130
4.2	A smoothing robust control algorithm using the Manifold Deformation Design Scheme (MDDS) is implemented for a Delta robot. The controller is a combined Kernel FPGA and DSP. PFGA is used for computation of the feedback signals and synchronously control each joint while DSP performs MDDS calculations	2013	131
Other	A median filter based on Xilinx Spartan-3 FPGA. It is used to filter the output of the measurement module of a line sensor IP core by smoothing the signals, suppress impulse voice and preserve edge	2013	134
4.5	Two fuzzy logic algorithms are developed: one for avoiding obstacles and another for avoiding obstacles / reaching a specific point. For the obstacle avoidance an ultrasonic sensor is used, while software is implemented using LabVIEW modules	2013	136
4.1 - 4.4	A stereovision measurement algorithm based on (Scale Invariant Feature Transform) SIFT key points, implemented on an embedded 4.1 board based on DSP and FPGA. SIFT algorithm looks for key points in the image scale space and extracts invariant locations. The system uses two CCD cameras and an 4.1 board	2013	138
4.2	A connecting algorithm for interfacing the FPGA-based controllers to high level robotic software frameworks. The Unity-Link algorithm conjoins PC-processed architectures with nodes that provide demanding real-time control of distributed robotic systems	2013	139
4.1	A real time 4.1 patrol service robot consisting of a CMOS camera module sending data to the FPGA board	2012	141
4.2 - 4.4	A robotic platform - based on Lego Mindstorms - is presented, including hardware, software and mechanical modules. A mathematical model for the motor and the PI controller are included	2012	142

4.3	A stereo vision landing system architecture for an Unmanned Aerial Vehicle (UAV). The proposed FPGA-based Artificial Neural Network (ANN) algorithm is designed to apply real-time object tracking, 3D position computation via Visual Odomotery methodology, Euclidian distance estimation from the landing target and horizontal displacement. The algorithm implementation was done on a Xilinx XC2V100 FPGA chip	2012	143
4.3	An FPGA-based Neural Network algorithm for lane following of autonomous mobile robots. The neural function consists of two main parts, the fixed points coding and the Sigmoïde function coding. The VHDL simulation is compiled in A Quartus II and implemented in Cyclone EP2C35F484C6 FPGA	2012	144
4.5	A fuzzy-logic controller for an autonomous navigation robot. The controller is implemented using an Atmega16 microcontroller and tested using the whole robotic application. The FPGA implementation consists of the 4.5ler, the UART receiver (it translates data between serial to parallel form), the LCD controller (for controlling the LCD operations) and the ASCII to binary converter and binary to ASCII controllers	2012	148
4.3	A hardware application of an FPGA-based controller using an artificial neural approach. The controller is capable of landscape learning, path planning, obstacle avoidance and sensory motor controlling. The controller is implemented as a SoC embedded on the robot, using a Virtex 6 VLX240T FPGA. The basic architecture of the controller consists of three basic layers: the pre-processing layer, the Self Organizing Map layers based on Kohonen maps and the computing layer	2012	150
4.4	A path planning / navigation algorithm for an autonomous mobile robot, using CORDIC architecture for implementing rotation. The robot's external interface includes an FPGA board consisting of Spartan XC2S50 FPGA and three ultrasonic sensors. The coding has been performed using VHDL and synthesized using Xilinx ISE	2012	151
4.1	A multi-resolution real-time dense stereo 4.1 algorithm is implemented. The algorithm's implementation is done on an Altera Stratix IV FPGA board with GigE vision cameras, using six levels of Gaussian Pyramids	2012	152
4.2	A hardware / software robotic application consisting of an operating software system that drives the FPGA controllers. The robotic (5 axe) system is an MK-2 industrial robotic arm and its controllers are operated by a XilinxVirtex II FPGA board and two PowerPC processors	2012	153
4.1	Digital image compression and decompression architecture applying M-JPEG2000 image compression standard, used for an UAV data link system. The two primary components of the system are an FPGA and an ADV212	2012	154
4.2	An FPGA-based speed controller for a DC motor designed for an autonomous mobile robot. The proportional integral (PI) controller used to achieve acceleration realized by the FPGA, while for decelerating control a braked deceleration circuit module is proposed. An additional motor protection circuit to avoid the motor driver IC damage is also included	2012	156
4.3	A 4.2 consisting of a multi-processor System on Chip implemented on a Spartan 3A FPGA, performing bio-inspired walking on a six-ledged robot (hexapod). The whole system consists of one soft processor used for high level decisions in motion decisions and six soft processors running independent control loops for the six legs respectively	2012	157
4.2	A hybrid architecture consisting of parallel FPGA (Spartan 6) and Advance RISC Machine microcontroller for visual trajectory-tracking. The system's framework consists of 3 main layers: The data layer where the FPGA performs input/output extension and hardware acceleration, the real-time layer where the microcontroller performs real-time control and the high performance layer where a computer executes complex tasks like navigation, map generating and artificial intelligence	2012	158
4.4	An FPGA-based algorithm for path planning and obstacle avoidance of autonomous mobile robots. The system consists of ultrasonic sensors and is implemented in a Xilinx FPGA Spartan II	2012	159
	A dexterous six-DOF, five-joint robotic arm for ping-pong is proposed. Each joint contains a motor position sensor, a current sensor and a temperature sensor and an FPGA for communication and control. More		

4.2	specifically, each joint contains four modules, the sensor signal processing, the joint servo control, the motor current control and the communication module	2012	160
4.4	An in-door navigation architecture implemented on FPGA based on impulse-based ultra-wideband (UWB) technology. The system consists of an UWB receiver installed on an autonomous mobile robot that utilizes time-difference on arrival (TDOA) between pairs of anchor nodes for localization purposes. These nodes transmit a coded repeated stream of ultra-short pulses	2011	166
4.5	A locomotion / visual information common function based on fuzzy logic is implemented on an FPGA.	2011	169

References

[1] R. Johannes, M. Porrmann, and U. Rückert. "Survey of FPGA applications in the period 2000–2015" Technical Report, Center of Exclellence Cognitive Internaction Technology, Bielefeld University, Germany, 2017.

[2] M. Prado, "An FPGA-based open architecture industrial robot controller" IEEE Access 6, 13407--13417, 2018. .

[3] P. Carlos and M. Porrmann. 2007. The utilization of reconfigurable hardware to implement digital controllers: a review. In 2007 IEEE International Symposium on Industrial Electronics (ISIE'19). 2380--2385.

[4] T. Russell, K. Pocek, and A. DeHon "Reconfigurable computing architectures". In Proceedings of the IEEE. IEEE 103, **3** (2015), 332--354, 2015

[5] V. Anuj, K. Dang Pham, and D. Koch, "A survey on FPGA virtualization", 28th International Conference on Field Programmable Logic and Applications, 1310--1317, 2018..

[6] S. Gandhare and B. Karthikeyan "Survey on FPGA Architecture and Recent Applications" International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN'19), 1--4, 2019.

[7] B. Ari and F. Mondada, "Elements of robotics" Springer Nature, 2018.

[8] W. Mike. 2014. "Implementation of robot systems: an introduction to robotics, automation, and successful systems integration in manufacturing" Butterworth-Heinemann.

[9] H. Martin, "Industrial robotics", Springer handbook of robotics. Springer Nature, 2016

[10] A. Yuki, K. Okada, "Design principles of a human mimetic humanoid: Humanoid platform to study human intelligence and internal body systems". Science Robotics **2**, 13, 2017.

[11] S. Olivier and Thomas Flayols. 2019. An overview of humanoid robots technologies. In Biomechanics of Anthropomorphic Systems. Springer Cham, (2019), 281--310.

[12] S. Sánchez, R. Cortés, "Cartesian control for robot manipulators". In Robot manipulators trends and development, 165--212, 2010

[13] B. Vanderlei, E. Marques, and G. A. Constantinides, "A floating-point extended kalman filter implementation for autonomous mobile robots". Journal of Signal Processing Systems **56**, 1, 41--50, 2019.

[14] G. Mingas, Emmanouil Tsardoulias, and Loukas Petrou. 2012. An FPGA implementation of the SMG-SLAM algorithm. Microprocessors and Microsystems 36, 3, 190--204, 2012

[15] P. Young-seuk, S. Lek. "Artificial Neural Networks: Multilayer Perceptron for Ecological Modeling". Developments in environmental modelling, **28**, 123--140, 2016.

[16] Poliac, M. O., J. M. Zanetti, and D. M. Salerno. 1991. Performance measurements of seismocardiogram interpretation using neural networks. In 1991 Proceedings Computers in Cardiology (IEEE'91). 573--576.

[17] K. Erdal, M. Ahmadieh Khanesar "Fuzzy neural networks for real time control applications: concepts, modeling and algorithms for fast learning", Butterworth-Heinemann, 2016.

[18] P. Yanik, George Ford, and William McDaniel. 2010. An introduction and literature review of fuzzy logic applications for robot motion planning. In Proceedings of ASEE Southeast Section Conference (ASEE'10).

[19] I. Mario and M. Chacón. 2006. Fuzzy logic for image processing: definition and applications of a fuzzy image processing scheme. In Bai Y., Zhuang H., Wang D. (eds) Advanced Fuzzy Logic Technologies in Industrial Applications. 101--113.

[20] S. Giandomenico "Swarm Robotics". MDPI-Multidisciplinary Digital Publishing Institute, 2019

[21] A. Podlubne and D. Göhringer. FPGA-ROS: Methodology to Augment the Robot Operating System with FPGA Designs. In 2019 International Conference on ReConFigurable Computing and FPGAs (ReConFig'19). 1--5, 2019

[22] C. Fu and Y. Yu "FPGA-based Power Efficient Face Detection for Mobile Robots". In 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO'19). 467--473, 2019

[23] T. Sun, Y. Liu, and Y. Wang "Design and implementation of a High-speed Lidar Data Reading System based on FPGA", In 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR'19). 322--327, 2019

[24] K. Hocine, R. Mellah, and H. Talem, "Neuro-fuzzy Control of a Position-Position Teleoperation System Using FPGA". In 2019 24th International Conference on Methods and Models in Automation and Robotics (MMAR'19). 64--69., 2019

[25] L. Alejandro, "Low latency event-based filtering and feature extraction for dynamic vision sensors in real-time FPGA applications", IEEE Access 7, 134926--134942, 2019.

[26] C. Chien, and C. Hsu. HW/SW Co-design and FPGA Acceleration of a Feature-Based Visual Odometry. In 2019 4th International Conference on Robotics and Automation Engineering (ICRAE'19), 148--152, 2019.

[27] R. Martínez Peralta, E. Zamora Gómez, and J. Humberto Sossa Azuela "Efficient FPGA hardware implementation for robot manipulator kinematic modeling using rational trigonometry" IEEE Latin America Transactions, 1524--1536, 2019

[28] P. Christos, T. Nikos, and B. Michael. Design and Implementation of an APSoC-Based Robotic System with Motion Tracking Teleoperation. In 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST'19). 1--4, doi: 10.1109/MOCAST.2019.8742037, 2019

[29] C. Zhengdong et al. 2019. Design of HEVC Intra Model Decision Based on Zynq. In 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR'19). 311--315.

[30] A. Kyriakos et al. High Performance Accelerator for CNN Applications. In 2019 29th International Symposium on Power and Timing Modeling, Optimization and Simulation (PATMOS'19). 135--140. doi: 10.1109/PATMOS.2019.8862166.

[31] G. Ce et al. Customisable control policy learning for robotics. In 2019 IEEE 30th International Conference on Application-specific Systems, Architectures and Processors (ASAP'19). 91--98.

[32] B. G. Oliveira and Jorge Lobo. Interactive Demonstration of an Energy Efficient YOLOv3 Implementation in Reconfigurable Logic. In 2019 5th Experiment International Conference (exp. at'19). 235--236.

[33] Z. Chen et al. Eye-to-hand robotic visual tracking based on template matching on FPGAs. IEEE Access **7** (2019), 88870--88880.

[34] H. Chih-Lyang, H. Wu, and W. Hung. Software/Hardware-based hierarchical finite-time sliding-mode control with input saturation for an omnidirectional autonomous mobile robot. IEEE Access **7** (2019), 90254--90267, 2019

[35] D. , Zeyang and L Jing. Real-Time Attitude Estimation of Sigma-Point Kalman Filter via Matrix Operation Accelerator. In 2019 IEEE 13th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC). IEEE, 342--346, 2019

[36] D. Lee, C. Cheung, and Dan Pritsker. Radar-based Object Classification Using An Artificial Neural Network. In 2019 IEEE National Aerospace and Electronics Conference (NAECON'19). 305--310, 2019

[37] A. Kamaleldin et al. Modular memory system for RISC-V based MPSoCs on Xilinx FPGAs. In 2019 IEEE 13th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC'19). 68--73, 2019

[38] L. Ning et al. When Deep Learning Meets the Edge: Auto-Masking Deep Neural Networks for Efficient Machine Learning on Edge Devices. In 2019 IEEE 37th International Conference on Computer Design (ICCD'19). 506--514, 2019

[39] R. Tapiador Morales et al. "Spiking Row-by-Row FPGA Multi-Kernel and Multi-Layer Convolution Processor". In 29th International Conference on Field Programmable Logic and Applications (FPL'19). 248--249, 2019

[40] S. Zsolt, and S. Tihamér, Quadcopter control implemented on FPGA. In 2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo'19). 000049--000054, 2019

[41] C. Pedro-F et al. A proposal for a SoC FPGA-based image processing in RGB-D sensors for robotics applications. In 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA'18). 1--6, 2018

[42] M. Jingeng et al. Implementing a SoC-FPGA Based Acceleration System for On-Board SVM Training for Robotic Transtibial Prostheses. In 2018 IEEE International Conference on Real-time Computing and Robotics (RCAR'18). 150--155., 2018

[43] W. Ming-Shi and Zhe-Rong Zhang. FPGA implementation of HOG based multi-scale pedestrian detection. In 2018 IEEE International Conference on Applied System Invention (ICASI'18). 1099--1102.

[44] Koepper Alexander and Karsten Berns. Behavior-Based Approach for Calculation of a Robot Arm's Inverse Kinematics on a FPGA. In ISR 2018 50th International Symposium on Robotics (ISR'18). 1--8, 2018

[45] O. Rahnama, D. Frost, O. Miksik and P. Torr.. Real-Time Dense Stereo Matching With ELAS on FPGA-Accelerated Embedded Devices. In IEEE Robotics and Automation Letters **3**, 3 2018

[46] Y. Sun et al. "An SoC-based platform for integrated multi-axis motion control and motor drive" In 2018 International Power Electronics Conference (IPEC-Niigata 2018-ECCE Asia). 560--564

[47] G. Evangelista, Carlos Olaya, and Erick Rodríguez. Fully-pipelined CORDIC-based FPGA Realization for a 3-DOF Hexapod-Leg Inverse Kinematics Calculation. In 2018 WRC Symposium on Advanced Robotics and Automation (WRC SARA'18). 237--242.

[48] P. G. Mousouliotis et al. Expanding a robot's life: Low power object recognition via FPGA-based DCNN deployment. In 2018 7th International Conference on Modern Circuits and Systems Technologies (MOCAST'18). 1--4, 2018

[49] P. Li et al. Efficient and low-cost Deep-Learning based gaze estimator for surgical robot control. In 2018 IEEE International Conference on Real-time Computing and Robotics (RCAR'18). 58--63, 2018

[50] G. Zhang et al. A Control System Design for 7-DoF Light-weight Robot based on EtherCAT Bus. In 2018 IEEE International Conference on Mechatronics and Automation (ICMA'18). 2169--2174, 2018

[51] H. Sheng et al. Design of robot control system for minimally invasive surgery of knee joint. In 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC'18). 2583--2586.

[52] A. Cadena et al. Development of a hybrid autonomous underwater vehicle for benthic monitoring. In 2018 4th International Conference on Control, Automation and Robotics (ICCAR'18). 437--440.

[53] D. Klimeck et al. Resource-efficient Reconfigurable Computer-on-Module for Embedded Vision Applications. In 2018 IEEE 29th International Conference on Application-specific Systems, Architectures and Processors (ASAP'18). 1--4.

[54] L. Stornaiuolo, M. Santambrogio, and Donatella Sciuto. On how to efficiently implement deep learning algorithms on pynq platform. In 2018 IEEE Computer Society Annual Symposium on VLSI (ISVLSI'18). 587--590.

[55] M. Martínez-Prado et al. An FPGA-based open architecture industrial robot controller. IEEE Access 6 (2018), 13407--13417.

[56] L. Chen et al. Teleoperation Snake Robots with Panorama Vision. In 2018 11th International Workshop on Human Friendly Robotics (HFR'18). 54--59.

[57] S. A. Pertuz et al. A Modular and Distributed Impedance Control Architecture on a Chip for a Robotic Hand. In 2018 31st Symposium on Integrated Circuits and Systems Design (SBCCI'18). 1--6.

[58] H. Nakahara, A demonstration of FPGA-based you only look once version2 (YOLOv2). In 2018 28th International Conference on Field Programmable Logic and Applications (FPL'18). 4570--4571.

[59] N. Dimou et al. Parallel Robust Absolute Orientation on FPGA for Vision and Robotics. In 2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS'18). 665--668.

[60] G. Haritha et al. Analysis of real-time tracking filters implementation in FPGA. In 2018 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER). 158--162.

[61] S. , Yuhei, Acceleration of Publish/Subscribe Messaging in ROS-Compliant FPGA Component. In Proceedings of the 8th International Symposium on Highly Efficient Accelerators and Reconfigurable Technologies (HEART2017). Association for Computing Machinery, New York, NY, USA, Article 13, 1--6.

[62] Y. Z. Jing et al. Design of auto detection system for high-voltage insulator inspection robot. In 2017 International Conference on Robotics and Automation Sciences (ICRAS'17). 16--19, 2017

[63] L. Chen et al.. Development of an industrial robot controller with open architecture. In 2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM'17). 754--756, 2017

[64] Z. Jiang et al. Kinematics calculation of minimally invasive surgical robot based on FPGA. In 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO'17). 1726--1730, 2017

[65] Ben P. Jeppesen, et al. An FPGA-based controller for collaborative robotics. In 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE'17). 1067--1072, 2017

[66] P. Reddy Gankidi and Jekan Thangavelautham. FPGA architecture for deep learning and its application to planetary robotics." In 2017 IEEE Aerospace Conference (AeroCon'17). 1--9.

[67] E. S. Fiestas and S. G. Prado, Motion control of a cartesian robot using a dual- core ARM cortex-A9 system-on-chip FPGA. In 2017 Latin American Robotics Symposium (LARS'17) and 2017 Brazilian Symposium on Robotics (SBR'17). 1--6, 2017

[68] R. Nahar, FPGA based parallelized architecture of efficient graph based image segmentation algorithm. In 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO'17). 98--103, 2017

[69] K. Fan et al. Hardware implementation of a virtual blind cane on FPGA. In 2017 IEEE International Conference on Real-time Computing and Robotics (RCAR'17). 344--348, 2017

[70] M. C. Chinnaaiah et al. A versatile autonomous navigation algorithm for smart indoor environments using FPGA based robots. In 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT'17). 1196--1200.

[71] M. Yaqoob Javed et al. Implementation of FPGA based efficient gait controller for a biped robot. In 2017 2nd International Conference on Robotics and Automation Engineering (ICRAE'17). 42--47, 2017

[72] V. Šetka, R. Čečil, and M. Schlegel. Triple inverted pendulum system implementation using a new ARM/FPGA control platform. In 2017 18th International Carpathian Control Conference (ICCC'17). 321--326, 2017

[73] J. Weberruss et al. FPGA acceleration of multilevel ORB feature extraction for computer vision. In 2017 27th International Conference on Field Programmable Logic and Applications (FPL'17). 1--8.

[74] L. Chen, et al. Development of an industrial robot controller with open architecture. In 2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM'17). 754--757.

[75] F. Palumbo et al. Feasibility study of real-time spiking neural network simulations on a swarm intelligence based digital architecture. In 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW'17). 247--250.

[76] T. Vo Nhu, and Hideyuki Sawada. Cerebellum-like neural network for short-range timing function of a robotic speaking system. In 2017 3rd

International Conference on Control, Automation and Robotics (ICCAR'17). 184--187.

[77] R. Appel et al. Design-time improvement using a functional approach to specify GraphSLAM with deterministic performance on an FPGA. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'17). 797--803.

[78] Z. Zhou et al. Mechatronic design of an ankle-foot rehabilitation robot for children with cerebral palsy and preliminary clinical trial. In 2017 IEEE International Conference on Industrial Technology (ICIT'17). 825--830.

[79] S. Chiang et al. 2017. The design of intelligent interactive service robots. In 2017 International Conference on Advanced Robotics and Intelligent Systems (ARIS'17). 100--103.

[80] S. Pullteap. Development of a walking robot by using FPGA controller. In 2016 11th France-Japan & 9th Europe-Asia Congress on Mechatronics (MECATRONICS)/17th International Conference on Research and Education in Mechatronics (REM'16). 054--057.

[81] J. Savage et al. Configurable Mobile Robot Behaviors Implemented on FPGA Based Architectures. In 2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC'16). 317--322.

[82] M. C. Chinnaaiah, et al. A new approach: An FPGA based robot navigation for patrolling in a service environment. In 2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS'16). 1--4.

[83] M. C. Chinnaaiah, et al. A versatile path planning algorithm with behavioural control using FPGA based robots. In 2016 10th International Conference on Intelligent Systems and Control (ISCO'16). 1--6.

[84] C. Lyu et al. High-speed target tracking base on FPGA. In 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR'16). 272--276.

[85] T. Kho et al. Enhance implementation of embedded robot auto-navigation system using fpga for better performance. In 2016 3rd International Conference on Electronic Design (ICED'16). 309--314.

[86] F. Schwiegelshohn, Florian Kästner, and Michael Hübner. FPGA design of numerical methods for the robotic motion control task exploiting high-level synthesis. In 2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE'16). 1--5.

[87] C. Rafael Steffens et al. FPGA Based Sensor Integration and Communication Protocols for Automated Robot Control in Linear Welding. In 2016 XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR'16). 115--120.

[88] C., Chen, et al. Real-time target tracking and positioning on FPGA. In 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR'16). 448--453.

[89] Y. San Woo et al. 2016. Enhance implementation of flying robot auto-navigation system on FPGA for better performance. In 2016 3rd International Conference on Electronic Design (ICED'16). 315--320.

[90] W. Zhou et al. Real-time implementation of panoramic mosaic camera based on FPGA. In 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR'16). 204--209.

[91] P. Čížek, J. Faigl, Low-latency image processing for vision-based navigation systems. In 2016 IEEE International Conference on Robotics and Automation (ICRA'16). 781--786.

[92] R. Konomura and K. Hori. FPGA-based 6-DoF pose estimation with a monocular camera using non co-planer marker and application on micro quadcopter. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16). 4250--4257.

[93] K. Boikos. Semi-dense SLAM on an FPGA SoC. In 2016 26th International Conference on Field Programmable Logic and Applications (FPL'16). 1—4, 2016.

[94] J. Yu et al. On-board real-time path planning design for large-scale 7-DOF space manipulator. In 2016 IEEE International Conference on Mechatronics and Automation, Harbin (ICMA'16). 1501--1506.

[95] K. Huang et al. 2016. A scalable lane detection algorithm on COTSs with OpenCL. In 2016 Design, Automation & Test in Europe Conference & Exhibition (DATE'16). 229--232.

[96] L. Kalms, A. Elhossini, and B. Juurlink. FPGA based hardware accelerator for KAZE feature extraction algorithm. In 2016 International Conference on Field-Programmable Technology (FPT'16). 281--284.

[97] S. Xiao, A. Postula, and N. Bergmann. Optimal random sampling based path planning on FPGAs. In 2016 26th International Conference on Field Programmable Logic and Applications (FPL'16). 1--2.

[98] Y. Wang et al. DeepBurning: automatic generation of FPGA-based learning accelerators for the neural network family. In 2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC'16).1--6.

[99] P. Vyas et al. Cordic-based azimuth calculation and obstacle tracing via optimal sensor placement on a mobile robot. In IEEE/ASME Transactions on Mechatronics 21, 5 (2015), 2317--2329.

[100] M. C. Chinnaiah, A novel approach in navigation of FPGA robots in robust indoor environment. In 2015 International Conference on Advanced Robotics and Intelligent Systems (ARIS'15). 1--6.

[101] V. Hugo Schulz, A SoC with FPGA Landmark Acquisition System for Binocular Visual SLAM. In 2015 12th Latin American Robotics Symposium and 2015 3rd Brazilian Symposium on Robotics (LARS-SBR'15). 336--341.

[102] G. Dagnino et al. Design and real-time control of a robotic system for fracture manipulation. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'15). 4865--4868.

[103] R. Saumyarup et al. Face detection system using FPGA. In 2015 Online International Conference on Green Engineering and Technologies (IC-GET'15). 12--15.

[104] G. Quintal et al. Real-time FPGA decentralized inverse optimal neural control for a shrimp robot. In 2015 10th System of Systems Engineering Conference (SoSE'15). 250--255.

[105] M. D. Kim and Jun Ueda. 2015. Real-time image de-blurring and image processing for a robotic vision system. In 2015 IEEE International Conference on Robotics and Automation (ICRA'15). 1899--1904.

[106] M. A. Dias and Fernando S. Osório. 2015. Reconfigurable hardware architecture for vision-based driving systems. In 2015 12th Latin American Robotics Symposium and 2015 3rd Brazilian Symposium on Robotics (LARS-SBR'15). 97--102.

[107] X. Ke, Vision development of humanoid head robot SHFR-III. In 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO'15). 1590--1595.

[108] L. Contreras et al.. Hardware architecture of the EKF prediction stage applied to mobile robot localization. In 2015 IEEE 6th Latin American Symposium on Circuits & Systems (LASCAS'15). 1--4.

[109] K. Sanni et al. Fpga implementation of a deep belief network architecture for character recognition using stochastic computation. In 2015 49th Annual Conference on Information Sciences and Systems (CISS'15). 1--5.

[110] J. Rettkowski, Robot navigation based on an efficient combination of an extended A∗ algorithm, bird's eye view and image stitching. In 2015 Conference on Design and Architectures for Signal and Image Processing (DASIP'15). 1--8.

[111] D. Bailey, Miguel Contreras, and Gourab Sen Gupta. Towards automatic colour segmentation for robot soccer. In 2015 6th International Conference on Automation, Robotics and Applications (ICARA'15). 478--483.

[112] H. Oleynikova, D. Honegger, and M. Pollefeys. Reactive avoidance using embedded stereo vision for MAV flight. In 2015 IEEE International Conference on Robotics and Automation (ICRA'15). 50--56.

[113] B. A. Silva et al. Run-time cache configuration for the LEON-3 embedded processor. In Proceedings of the 28th Symposium on Integrated Circuits and Systems Design (SBCCI'15). 1--6.

[114] Y. Li and C. Huang. Force control for the fingers of the piano playing robot—A gain switched approach. In 2015 IEEE 11th International Conference on Power Electronics and Drive Systems (PEDS'15) 265--270.

[115] C. Hsu et al. FPGA-based path planning using improved Ant Colony Optimization Algorithm. In 2015 IEEE 5th International Conference on Consumer Electronics-Berlin (ICCE-Berlin'15). 443--444.

[116] M. Pena et al . Fuzzy logic for omni directional mobile platform control displacement using FPGA and bluetooth. In IEEE Latin America Transactions 13, 6, 2015, 1907--1914.

[117] L. Zouari, M. Ben Ayed, and M. Abid. A hardware in the loop simulation for electrically driven robot manipulator. In 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE'15). 1495--1501.

[118] H. Guo et al. Real-time detection and classification of machine parts with embedded system for industrial robot grasping. In 2015 IEEE International Conference on Mechatronics and Automation (ICMA'15). 1691--1696.

[119] L. Fiack, L. Rodriguez, and B. Miramond. 2015. Hardware design of a neural processing unit for bio-inspired computing. In 2015 IEEE 13th International New Circuits and Systems Conference (NEWCAS'15). 1--4.

[120] M. C. Chinnaaiah et al. 2014. A novel approach in implementation of rendezvous behavioral control between FPGA robots. In 2014 International Conference on Advanced Robotics and Intelligent Systems (ARIS'14). 97--102

[121] J. Heshmatpanah, S. Boroumand, and Mehdi Tale Masouleh. 2014. FPGA design and implementation for real time vision applications on NTACO mobile robot. In 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM'14). 172--177.

[122] A. M. Ashir, Atef A. Ata, and M. Shukri Salman. 2014. FPGA-based image processing system for Quality Control and Palletization applications. In 2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC'14). 285--290.

[123] S. Agrawal and Ravi N. Prakash.. Implementation of WSN which can simultaneously monitor Temperature conditions and control robot for positional accuracy. In 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE'14). 1--6.

[124] M. Hou et al. Analysis of the multi-finger dynamics for robot hand system based on EtherCAT. In 2014 10th International Conference on Natural Computation (ICNC'14). 1061--1065.

[125] J. Nikolic et al. A synchronized visual-inertial sensor system with FPGA pre-processing for accurate real-time SLAM. In 2014 IEEE international conference on robotics and automation (ICRA'14). 431--437.

[126] S. Seok et al. A highly parallelized control system platform architecture using multicore CPU and FPGA for multi-DoF robots. In 2014 IEEE International Conference on Robotics and Automation (ICRA'14). 5414--5419.

[127] F. Mota Muñoz, S. Martínez de la Piedra, and Karla C. Gómez. Development of a multi-DOF robotic controller for academic purposes. In 2014 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE'14). 104--109.

[128] R. Ladig and K. Shimonomura.. Fpga-based fast response image analysis for autonomous or semi-autonomous indoor flight. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'14). 668--673.

[129] C. Guo et al. Reconfigurable fault-tolerant joint control system for space robot. In Proceeding of the 11th World Congress on Intelligent Control and Automation (WCICA'14). 1899--1904.

[130] T. Zhang et al. Design and control of a multisensory five-finger prosthetic hand. In Proceeding of the 11th World Congress on Intelligent Control and Automation (WCICA'14). 3327--3332.

[131] E. Yusifli et al. An FPGA Based Resources Efficient Solution for the OmniVision Digital VGA Cameras Family. 2014 IEEE Intl Conf on High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC'14, CSS'14, ICESS'14). 558--561.

[132] D. Oswald et al. Implementation of fuzzy color extractor on ni myrio embedded device. In 2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems (MFI'14). 1--6.

[133] C. Andres Lara-Nino, Cesar Torres-Huitzil, and Jose Hugo Barron-Zambrano. Versatile educational and research robotic platform based on reconfigurable hardware. In 2014 International Conference on ReConFigurable Computing and FPGAs (ReConFig14). 1--6.

[134] G. Zhou et al. 2014. On-board inertial-assisted visual odometer on an embedded system. In 2014 IEEE International Conference on Robotics and Automation (ICRA'14). 2602--2608.

[135] M. H. Liyanage and N. Krouglicof . A single time scale visual servoing system for a high speed SCARA type robotic arm. In 2014 IEEE International Conference on Robotics and Automation (ICRA'14). 4153--4160.

[136] L. Alejandro, Teresa B. Linares-Barranco. Event-driven sensing and processing for high-speed robotic vision. In 2014 IEEE Biomedical Circuits and Systems Conference (BioCAS'14) Proceedings. 516--519.

[137] B. Y. Kumar and S. Ilaiyaraja. 2014. Smart interface for sensors in robot. In 2014 Sixth International Conference on Advanced Computing (ICoAC'14). 145--149.

[138] H. B. et al. Embedded FPGA controller for robot arm in material handling using reconfigurable nicompact RIO. In Second International Conference on Current Trends In Engineering and Technology (ICCTET'14). 125--131.

[139] S. Bhashis Roy, J. Bag, and S. Kumar Sarkar. Design and VLSI implementation of a robot navigation processor deploying CORDIC based anti-collision algorithm with RFID technology. In 2014 Annual IEEE India Conference (INDICON'14). 1--6.

[140] M. Petko et al. Acceleration of parallel robot trajectory generation in FPGA. 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA'13). 1123--1128.

[141] M. Petko et al. Acceleration of parallel robot trajectory generation in FPGA. In 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA'13). 1123--1128.

[142] A. Ghorbel et al. An FPGA based platform for real time robot localization. In 2013 International Conference on Individual and Collective Behaviors in Robotics (ICBR'13). 56--61.

[143] W. Zhu et al. Precision control of modular robot manipulators: The VDC approach with embedded FPGA. In IEEE Transactions on Robotics 29, 5 (2013), 1162--1179.

[144] S. Boroumand et al. Path tracking and obstacle avoidance of a FPGA-based mobile robot (MRTQ) via fuzzy algorithm. In 2013 13th Iranian Conference on Fuzzy Systems (IFSC'13). 1--5.

[145] S. Bold, Min A. Jeong, S. Min Jeon, and S. Ro Lee. FPGA based real time embedded color tracking mobile robot. In 2013 International Conference on IT Convergence and Security (ICITCS'13). 1--4.

[146] N. Alisa Ali et al. PWM controller design of a hexapod robot using FPGA. In 2013 IEEE International Conference on Control System, Computing and Engineering (ICCSCE'13). 310--314.

[147] B. Velasquez et al. Navigation control in FPGA for a differential-drive mobile robot. In 2013 16th International Conference on Advanced Robotics (ICAR'13). 1--6.

[148] C. Szász and G. Husi. Novel multimodal communication skills implementation on the NI-9631 robot. In IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society (IECON'13). 7837--7842.

[149] R. Li, Shiyao Lin, and Jian Chen. Embedded Motion Controller Design Based on RTEX Network. In 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC'13). 326--329.

[150] H. Lin "The implementation of smoothing robust control for a delta robot" In 2013 Second International Conference on Robot, Vision and Signal Processing (RVSP'13). 208--213.

[151] Yen-Fang Li and L. Chuang. Controller design for music playing robot—Applied to the anthropomorphic piano robot. In 2013 IEEE 10th International Conference on Power Electronics and Drive Systems (PEDS'13). 968--973.

[152] C. C. Wong and C. C. Liu. FPGA realisation of inverse kinematics for biped robot based on CORDIC. Electronics letters 49, **5** (2013), 332--334.

[153] L. Mutauranwa and Magripa Nleya. 2013. An efficient median filter in a robot sensor soft ip-core. In 2013 Africon Pointe-Aux-Piments (Africon'13). 1--5.

[154] M. Fatih Aydogdu, M. Fatih Demirci, and C. Kasnakoglu. Pipelining Harris corner detection with a tiny FPGA for a mobile robot. In 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO'13). 2177--2184.

[155] M. Mysorewala, et al. Design and implementation of fuzzy-logic based obstacle-avoidance and target-reaching algorithms on NI's embedded-FPGA robotic platform. In 2013 IEEE Symposium on Computational Intelligence in Control and Automation (CICA'13). 168--173.

[156] S. Cruz et al. FPGA implementation of a sequential extended Kalman filter algorithm applied to mobile robotics localization problems. In 2013 IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS'13). 1--4.

[157] H. Song et al. A fast stereovision measurement algorithm based on SIFT keypoints for mobile robots. In 2013 IEEE International Conference on Mechatronics and Automation (ICMA'13). 1743--1748.

[158] A. Blaabjerg Lange, Ulrik Pagh Schultz, and Anders Stengaard Soerensen. Unity-link: A software-gateware interface for rapid prototyping of experimental robot controllers on fpgas. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'13). 3899--3906.

[159] B. Thanh Trung and N. Van Cuong. Monitoring and controlling devices system by GPRS on FPGA platform. In 2013 International Conference on Advanced Technologies for Communications (ATC'13). 713--717.

[160] K. Asami, H. Hagiwara, and M. Komori. Visual navigation system based on evolutionary computation on FPGA for patrol service robot. In The 1st IEEE Global Conference on Consumer Electronics 2012 (GCCE'12). 295--298.

[161] M. D.A. Siqueira, J. Conforto, and Marcos B.R. Vallim. An FPGA-Based Robotics Platform for Educational and Research Applications. In 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium (SBR'12, LARS'12). 313--318.

[162] A. Din "Embedded low power controller for autonomous landing of UAV using artificial neural network" In 2012 10th International Conference on Frontiers of Information Technology (FIT'12). 196--203.

[163] A. Benjemmaa. Implementations approches of neural networks lane following system. In 2012 16th IEEE Mediterranean Electrotechnical Conference (MELECON'12). 515--518.

[164] R. Alba-Flores. Project-based learning approach in a senior level course in autonomous mobile robots. In 2012 Proceedings of IEEE Southeastcon (Southeastcon'12). 1--8.

[165] K. Gac. FPGA based hardware accelerator for calculations of the parallel robot inverse kinematics. In Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA'12). 1--4.

[166] N. Scicluna et al. FPGA-based autonomous parking of a car-like robot using Fuzzy Logic Control. In 2012 19th IEEE International Conference on Electronics, Circuits, and Systems (ICECS'12). 229--232.

[167] T. Muthu et al. 2012. Fuzzy logic controller for autonomous navigation. In 2012 International Conference on Communication and Signal Processing (ICCSP'12). 81--92.

[168] H. Hagiwara, K. Asami, and Mochimitsu Komori. Real-time image processing system by using FPGA for service robots. In The 1st IEEE Global Conference on Consumer Electronics 2012 (GCCE'12). 720--723.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 395-400 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Devices and Methods for Microclimate Research in Closed Areas – Underground Mining

Mila Ilieva-Obretenova*

Electrical Engineering Department, The University of Mining and Geology, "St. Ivan Rilski", Studentski Grad, "prof. Boyan Kamenov" Street, Sofia 1700, Bulgaria

ARTICLE INFO

Article history:
Received: 29 April, 2021
Accepted: 14 August, 2021
Online: 28 August, 2021

Keywords:
Sensors
Measurement Devices
Microcontroller
Data transfer
Underground Mining

ABSTRACT

Technical safety and health are especially important for mining-extracting industry. Even though the respective lows and good engineering practices exist, technologies develop and could address even better security for humans and equipment. The research question is to survey microclimate sensors in underground mining and to find whether they are ready for automation. The article is inspired from the research in "Computer System for Microclimate Management in Closed Areas of the Post-Mining Galleries and Greenhouses". The author offers a short review of state and trends in development of sensors and devices for monitoring and reporting of environmental parameters in underground mining. All environmental parameters: air temperature; temperature of surrounding building constructions, heated surfaces of technological machines and equipment; heat flow, heat irradiation; relative humidity; velocity of air flow; noise; illumination; gas concentration; dust level; ionizing radiations; radon concentration in air are represented with relevant measurement devices and measurement units. The next step is representing of fast-developing sensors using scientific references. Author performs quality assessment of their suitability for automated data transfer and management. The assessment criteria are: Analogue measurement devices, Digital measurement devices, Availability of microcontroller. Findings are proposed for discussion: recent used devices in underground mining are not suitable for automatization, because they miss a controller. The availability of controller also presumes availability of management services. On the base of articles about management of sensors and controllers future work is proposed: 1. Integration of existing elements sensor and controller and defining of its management; 2. Moving of new elements and synthesis of algorithms for its management. This will lead to more precise assessment of industrial risk and improvement of safety activities.

1. Introduction

Mining-extraction sector is especially important for the economics in Bulgaria. Therefore, the technical safety and professional health must embrace different activities on protection of employees and assets by reducing to minimum of dangers, hazards, failures, and oversights. Even though the respective lows and good engineering practices exist technologies develop and could address even better security of humans and equipment.

The article is inspired from research in [1] which considers a computer system for microclimate management in closed areas of post-mining galleries and greenhouses. Other authors describe this problem too. In [2], the authors represent Pictograms for admissible norms of harmful substances, but they miss devices and measurement methods. In [3], the organizers display a standard for microclimate design, but they miss details for modern devices for monitoring and reporting. In [4], the authors consider safety technics in underground mining, but without details in trends for devices and sensors. In [5], the author describes methods and devices of illumination technology in underground mining and tunnels, but without including other parameter of environment and the appropriate measurement devices. In [6] and [7], the authors represent rules for radioactivity (radon) measurement, while [8] portrays harmless work with cyanides, and [9] displays design of industrial ventilation, but they all miss automated data transfer. In [10], the authors describe a wireless permanent monitoring of coal

*Corresponding Author: Mila Ilieva-Obretenova, Bulgaria, Sofia 1505, Tcherkovna Str. 21, App. 8, milailieva@abv.bg, mila.ilieva@mgu.bg

strata, which could be applied to wireless permanent monitoring of microclimate. In [11], the author explains levels of IoT environment in underground mining as the emphasis is on the security of data transfer, but different sensor types and their management are not detailed.

The purpose of the article is to make a short review of state and trends of sensors and devices for monitoring and reporting of environmental parameters in underground mining and to accomplish a quality assessment of their suitability for automated data transfer and management.

2. Methodology

The survey of sensors for underground mining is accomplished in University of Mining and Geology, a leading university on mining science in Bulgaria. Author performs the research after publishing the guide on safety technics in underground mining sponsored [4]. All parameters of the closed area, basic sensors, and measurement units are represented. In the next step the fast-developing sensors with their improved versions are depicted. The author uses specialized references. Assessment criteria for measurement devices are defined on the base of Computer System for Microclimate Management in Closed Areas of the Post-Mining Galleries and Greenhouses [1]. Sensors have analogue or digital output in this system. This leads to the important need of placement of an analog-to-digital converter (ADC) and building of digital interface (SPI, I2C, CAN, Ethernet, etc.) [12], [13]. Most of the built-in ADCs have a resolution of 10 or 12 bits and contain 8-16 input channels which is sufficient for these applications [14]. The core of the system is a microcontroller that samples data from the sensors, processes the information and sends it to a computer through interfaces. Since this system will sample data during long periods of time (minutes, hours, days or even months), a backup power supply is needed along with the main one [15]. The switching between these two power supplies must happen automatically; therefore, an electronic switch is needed too. The results from the measurements should be transferred to the computer, where the end-user can view them. This is done with the interface RS232 [16]. As more than one system is used, the interface could be Wi-Fi to speed up the process of data gathering [17]. The network of measurement elements consists of a base microcomputer station and several sensors, distributed in the space, and linked to the station by cables or wireless Ethernet. The measurement of all climatic elements is made within 1 sec in total with an interval between the separate measurements - 3 minutes. It is necessary to perform measurements with low speed, as most of the values are not changing rapidly and because of the inertia of the sensors. Figure 1 shows the existing Computer System for Microclimate Management in Closed Areas of the Post-Mining Galleries and Greenhouses [1]. Cortex-M_0 architecture is used. It has a 32-bit embedded microprocessor, especially designed for low-powered optimized microcontrollers.

Therefore, specialists could assess the new sensors by the following three criteria:

- Whether are they analogue?
- Whether are they digital?
- Whether are they connected to microcontroller for automated data transfer?

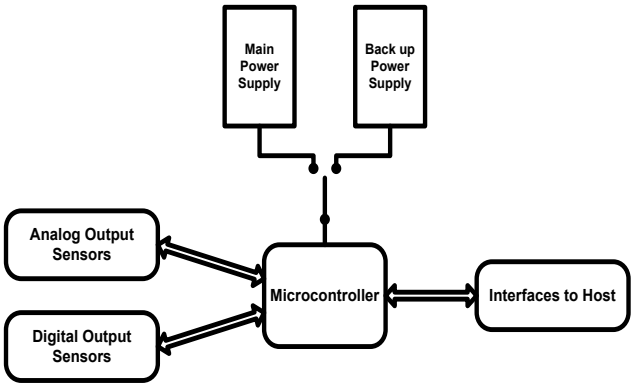


Figure 1: Computer system for microclimate management in closed areas of the post-mining galleries and greenhouses

Next steps are:

- Discussion of findings and their consequences and
- Proposal of future work about devices` investigation.

3. Results

The first result from the survey represents the systemized environmental parameters in underground mining, conventional devices for their measurement and the corresponding measurement units. It is shown in Table 1.

Table 1: Environmental parameters in underground mining, conventional devices for their measurement and the corresponding measurement units

Parameter	Device and measurement unit
1. Air Temperature	Quicksilver and Thermistor Thermometers, [ºC]
2.Temperature of surrounding building constructions, heated surfaces of technological machines and equipment.	Contact Thermometers, [ºC]
3.Heat flow, Heat irradiation	Actinometer, [W/m^2]
4.Relative Humidity	Psychrometer of Assmann, [%]
5.Velocity of air flow	Anemometer, Kata Thermometer, [m/s]
6.Noise	Noise meter (Sound meter) for exposition; integrating noise meter; stationary installations; installations, fitted to human (dosimeters), [dB]
7.Illumination	Luxmeter (Illuminance Meter), [lx] Luminance Meter [cd/m^2] Spectroradiometer, [nm]
8.Gas concentration	Gas analyzer, [%]
9.Dust level	Konimeter, [particles number/cm^3]
10.Ionizing radiations	Geiger-Muller counter, [Sv] Scintillation counter, [Sv] Pocket chambers and Pocket Dosimeters, [Sv] Film badge Dosimeter, [Sv] Thermoluminescent Dosimeters (TLDs), [Sv]

	Electronic Personal Dosimeter (EPD), [Sv]
11.Radon Concentration in Air	Hermetically closed chamber, [%]

The second result from the survey shows the fast-developing sensors according to specialized references. These are sensors for air temperature, heat flow, relative humidity, illumination, and gas concentration.

3.1. Sensors for Air Temperature

Classical measurement devices for air temperature are quicksilver and thermistor [18] thermometers, but they miss automated data transfer.

In [19], the authors represent smart temperature sensors and systems of temperature sensors. For temperature measurements in smart sensors and microelectromechanical systems the most frequently used elements are transistors, thermocouples, and thermopiles (several connected thermocouples), because these elements could be accomplished with integrated circuits technology. For the temperature range from -55ºC to 150ºC bipolar transistors are especially useful as temperature sensors because they could be produced by practically all standard technologies for integrated circuits together with other circuits. With similar circuits could be made a reference for a voltage gap. This allows performing of radiometric research, where temperature sensing voltages are compared with predefined voltage reference. By CMOS analogue circuits extremely high precision could be achieved by using error-reducing technics such as dynamic synchronization of elements and chopping (signal cutting). The smart temperature sensor is represented with output signal with modulated duty cycle. High precision is reached with remarkable low energy consumption. The output of sensor is digital, but without automated data transfer.

In [20], the authors portray a temperature sensor with bipolar transistor used as a diode. Thereby bigger temperature range, lower power consumption and smaller chip area are reached. The sensor works in time mode: The time is measured for a capacitor charging to activating voltage of the transistor. A comparator is used instead of conventional sigma-delta analog-to-digital converter. Therefore, output of the sensor is digitized, but without automated data transfer.

3.2. Sensors for Heat Flow

In underground mining a heat flow is measured, coming from surrounding building constructions, heated surfaces of technological machines and equipment. Classical measurement instruments for heat flow are with thermistors (thermopiles) [21]. The sensor consists of a set of thermistors and an array of converters produces voltage, proportional of the heat flow and allows direct measurement of heat flow. The availability of thermistors array allows to be made a comparison between heat flow assessment, obtained from converters, and calculated heat flow, using temperature differences and Fourier Low. The element is extremely sensitive with accuracy lower than 0.1 heat flow units (HFU) and could be used for long term readings. The output of element is digital, but without controller for data transfer.

The recent heat flux meters (HFM) suffer from low performance by weak and constant flows, non-uniformity of the heat flux in the measurement section and impossibility to control the heat flux at different temperatures.

In [22], the authors show a calibration system of a HFM in different operating conditions (low and middle heat fluxes) to design a new prototype with heat flux uniformity and to improve performance in a specific sub-region of the measuring section where the HFM is applied. Furthermore, a metrological characterization of the system allows a combined standard error to be better than 6% at low heat flows. The element is analogue, without controller for data transfer.

In [23] and [24], the authors report heat flow sensors integrated in textile. They determine the amount of heat exchanged between human body and environment. The aim is to improve comfort, efficiency, and sometimes safety of the wearing person. Passive heat flow sensors are used. Such sensors are still under developing and ideas for their appropriate calibration are proposed. Moreover, the existing heat flow sensors are impermeable and do not allow evaporation. This shows less precise results when phenomenon evaporation exists. A flexible heat flow sensor is developed, which is sensitive to water vapor. This sensor takes into consideration phenomenon evaporation and allows better energy measurement during heat exchange. The sensor is analogue without controller for data transfer.

3.3. Sensors for Relative Humidity

Classical devices for relative humidity measurements are psychrometers. Recent research [25] displays indoor air quality and the connected health effects. Complaints about sensory irritation in eyes and upper airways, also perception for dry air are among the most important symptoms in closed areas. In this line sensors for dynamics of particles, bacteria, and viruses in closed areas are necessary for improvement of labour quality and with a sharp focus on absolute humidity impact. The sensors could be analogue or digital and with microcontroller.

In [26], the authors consider a sensor for air humidity measurement using sensitive capacity-dependent crystal. Probe sensitivity is represented too. Moreover, new idea for excitation of entire humidity sensor with stochastic test signals is described, and the humidity measuring method is given. It includes the influence of test signals on weighting function uncertainty and on the A/D-D/A conversion. Error by humidity measurement is lower than 0.2% (t = 15ºC - 25ºC and humidity = 50% - 98%). There is a conversion ADC, DAC in sensor, but it misses a controller.

Research in [27] describes a program, which could simulate simultaneously heat and humidity transfer from walls and plates in closed area and its impact on indoor temperature and humidity. The aim is a sensor development. Recommendation is to be digital and connected to microcontroller.

3.4. Sensors for Illumination

Illumination on basic workplaces and galleries, where the miners move, is obligatory. This is not easy, because the atmosphere in coal mining is often accompanied with methane availability and is highly explosive. Humidity is high, but environment reflectance is extremely low. Therefore, special

explosion-proof illuminaters are used and their application is by hardest conditions – low reflection factor, methane availability, high humidity, low and narrow galleries etc.

What are the conditions in ore mining and coal extracting mining? Basically, there are not methane in the first and atmosphere is not explosive. In underground ore mining and these for extracting of non-metalliferous illumination is performed with illuminaters proof from dust and water. In such mining there is no danger from explosion of highly explosive compounds, but they are threatened from fire [5].

For illuminance tracking are applied illuminance meter (luxmeter) and luminance meter.

For measurement in mining galleries are suitable new generation luxmeters "ТКА-Lux" ("ТКА-Люкс") and "ТКА-РКМ-31" ("ТКА-ПКМ-31"). They have metrological characteristics on the level of the best world manufacturers. The illuminance measurement range is 10 – 200 000 lx, error – 6%.

They have the following performance:

- Viewing angle: 1.0 – 1.5º,
- Range of measurement: 10.0 – 2000 cd/m^2,
- Distance to measured object ~ 7m.

Luxmeters with dark scale are applied in coal mining, which is dangerous with gas and dust.

In underground mining exists blindness danger too. Therefore, it is necessary to measure luminance. It is measured with luminance meter. Basic measurement unit is candela [cd]. Recent luminance meter is TKA [5] and [28].

Illuminance assessment in underground mining must obligatory include:

- Illuminance assessment,
- Unevenness assessment,
- Blindness assessment,
- Discomfort assessment.

Special luxmeters are used for illuminance measurement on given surface (workplace). The most widespread luxmeters are photoelectric (objective) luxmeters, consisting of selenium element in series with galvanometer and micro amperemeter. The principle operation of photoelectric luxmeter is conducting a photocurrent between galvanometer`s electrodes by illumination of photoelement. The photocurrent is proportional of illuminance and this dependance could be assumed as linear, when the inner resistance is significant bigger than outer. This ratio is obtained by measurement of small illuminances and application of measurement devices with small inner resistance. The boundary electromotive force of selenium element (0.3 – 0.5 V) is obtained by illuminance 1000 lx and by bigger illuminance it does not change. Therefore, the photoelement illuminance does not have to exceed 1000 lx. For measurement of bigger illuminance is applied an absorbing filter, which decreases X-fold illuminance on photoelement. Selenium element is characterized by its spectral sensitivity too. It is graduated with the help of filament lamp. When an illuminance of other light sources (another light spectrum) is measured it is necessary to use correcting light filters or correcting coefficients, which by naturally illuminance is 0.8, and for luminescent lamps are from 0.88 to 1.2 according to their type [4].

In [29], the author publishes a "Luxmeter, showing light intensity, as it is received from a sensor".

The listed sensors are analogue and without controller for data transfer.

In [30], the authors show a spectroradiometer as a part of program-instrumental complex for management of illuminance installation in closed areas. Problems of spectroradiometer development are considered. Basic principles for building of automated illuminating system are formulated, based on criteria for quality and quantity assessment of illuminance effect. Scheme of automated illumination complex is considered. New development is represented – spectroradiometer, covering optical irradiation measurement and being a block of system for spectral and energetic management of LED illuminating installations. Attention is paid on the lack of metrological insurance of control of photometric characteristics of energy conservation technology, based on LEDs. It is foreseen that the sensor is digital and has a microcontroller.

3.5. Sensors for Gas Concentration

For measurement of gas concentration portable analyzers are widely used [4]: gas indicating tubes and gas analyzers. From gas analyzers most popular are: Gas Sense – 1000.L.EX; portable device for leakage discovery of explosive gases – Methane (CH_4), Propane (C_3H_8), Butane (C_4H_{10}), Liquified Petroleum Gas (LPG), Hydrogen (H_2), Ethanol (C_2H_5OH); Dräger Pac 5500 – measurement of Carbon Monoxide (CO), Hydrogen Sulfide (H_2S) or Oxygen (O_2); MSA EX-METER II – for measurement in explosive atmosphere; MSA SOLARIS 4 – Four channel gas analyzer, analogue, without microcontroller.

The trend is semiconductor sensors to be doped with metal oxides (e.g., tin) and their application for discovery and protection of environment from different gases like Carbon Monoxide (CO), Carbon Dioxide (CO_2), Hydrogen sulfide (H_2S), Ammonia (NH_3), swamp gas methane, LPG, and many others pollutant gases [31]. Different characteristics of gas sensor in composite form by different temperatures are also displayed. Evidence is produced that sensors in composite form are more resilient than these with a single material. A research about sensor`s output – analogue or digital – is not represented. About its ability for data transfer too.

The author in [32] considers manufacturer methods of zinc oxide nanorods and their application for gas discovery with low concentration due to their various conductance range, reaction of oxidative and reductive gases and extremely high sensitivity and selectivity. A research about sensor output - analogue or digital – is not given. For data transfer ability too.

The survey in [33] describes an electronic gas sensor from nano graphene, which could discover extremely low quantity gas (harmful chemical agents). Nano graphene does not have grainy structure and offers long term stability. Moreover, material has high quality crystal lattice, thus charge carriers move with high

velocity and generate low noise. All this supposes design of digital output of the sensor. Due to their specific properties carbon nanomaterials have potential for manufacturing of automated sensors.

The authors in [34] portray high selective gas sensors for smart monitoring and control of air quality, search and save of people etc. Substance choice in complex gas composite is the main challenge. The article represents a guide for material engineering for absorbing, size selecting and catalytical filters. Accent is laid on materials design with purposeful gas separation, portable elements integration and performance. All this shows that gas sensors could be designed with digital output and microcontroller for data transfer.

4. Discussion

The performed short review leads to the findings that the new measurement devices in underground mining are not suitable for automation. Small part of them has a digital output, and most of them are without microcontroller for data transfer. To accelerate and facilitate automation sensors must be *delivered with controller.* The availability of controller also presumes availability of management services, united in Program Library. The resent achievements are represented with articles with low level detail [35] or high-level concepts [36] and [37]. Other authors go even further away proposing sensors for inspecting of underground mining to be moveable and to be carried from a drone, but without giving detail about their management [38]-[41]. Therefore, Articles for sensors with microcontrollers and management programs must describe comprehensive:

- Integration of sensors and controller;
- Algorithms for management of integrated sensor-controller;
- Function for bearing of sensors for timely reading of fast changing parameters at different places in space aiming sensor security;
- Algorithms for management of sensors bearing, as some problems are considered: signal propagation; life, size, and safety of battery; security of bearing system etc.

5. Conclusion

Main conclusions from this research are as follows:

1. Criteria for assessment of measurement devices in underground mining are synthesized. Methods on which devices perform their functions are specified.
2. Devices and measurement units which cover all environmental parameters in underground mining are listed, namely: air temperature; temperature of surrounding building constructions, heated surfaces of technological machines and equipment; heat flow, heat irradiation; relative humidity; velocity of air flow; noise; illumination; gas concentration; dust level; ionizing radiations; radon concentration in air.
3. Representations of fast-developing devices are accomplished: for air temperature; for heat flow; for relative humidity; for illumination and for gas concentration. Quality method is used. Assessments organized according to synthesized criteria are defined.
4. On the base of discussion about availability of controller to sensor the future work is proposed in following aspects:

a. Integration of existing elements sensor and controller and defining of its management;

b. Moving of new elements and synthesis of algorithms for its management.

Formulated aim is complex, and this article concentrate on devices` assessment. Representation level and assessments have the necessary depth to illustrate the needed innovations. Basic conclusion from this article is that the developed assessments are eligible base for development of sensors and their management in closed areas, structured as underground mine workings. This will lead to more precise assessment of industrial risk and improvement of safety activities.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

This article is a part of project “Survey and Management of Sensors for Microclimate in Mining Galleries” in University of Mining and Geology “St. Ivan Rilski”, Bulgaria.

References

[1] M. Ilieva-Obretenova, "Computer System for Microclimate Management in Closed Areas of the Post-Mining Galleries and Greenhouses." in 7th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE), Ruse, 2020, 1-4, doi: 10.1109/EEAE49144.2020.9278785.

[2] Ministry of labour and social policy. Healthy Workplaces Campaign, Pictograms 2018-2019, Sofia. (In Bulgarian)

[3] https://russian.worldbuild365.com/news/nbs9ontsc/hvac/mikroklimat-po-gostu-chto-dolzhen-uchitivat-kazhdyy-proektirovschhik. Accessed June 2021.

[4] B. Vladkova, Техническа безопасност, University of Mining and Geology “St. Ivan Rilski” – Sofia, 2020. (In Bulgarian)

[5] G. Ganchev, Книга за осветлението. Минното осветление е тъмна работа, University of Mining and Geology - Sofia, 2018. (in Bulgarian)

[6] D. Pressyanov, et al., "Passive radon monitors with part-time sensitivity to radon". Radiation measurements, **118** (2018) 72-76, doi: 10.1016/j.radmeas.2018.08.014

[7] H. Mischo, Radioactivity Practical Course. Institute of Mining and Special Civil Engineering. Chair of Underground Mining Methods, Technical University Bergakademie Freiberg, 2018.

[8] B. Vladkova, Цианидите в добивната промишленост – технологични особености и безопасност при работа с тях. University of Mining and Geology, Sofia, 2016. (In Bulgarian)

[9] Z. Dinchev, Основи на проектирането на промишлена вентилация. University of Mining and Geology, Sofia, 2016. (in Bulgarian)

[10] "Wireless Round-the-Clock Observation of Coal Stacks, No Matter where Measurement Points May Move." 2016, https://www.yokogawa.com/es/library/resources/applicatoin-notes/wireless-round-the-clock-observation-of-coal-stacks-no-matter-where-measurement-points-may-move/.

[11] Y. Anastasova, "Internet of Things in the mining industry – security technologies in their application." Sustainable Extraction and Processing of Row Materials Journal, **1**, 2020, 7-10, doi: 10.5281/zenodo.4275895

[12] I. Nikolov, "Development of a warehouse microclimate managing system." Innovation and entrepreneurship, **6**(3), 2018, 166-174

[13] S. Velinova, "Possibilities for using of mining galleries for growing plants on artificial lighting." International Scientific Conference UNITECH 2017, 148-153

[14] N. Kolev, Evaluation of main elements of the soil energetic balance by electronic means. Doctor of Science thesis, Sofia, 1996.

[15] N. Kolev, et al., "Organisation of long-term microcomputer systems evaluation of field water balance elements." Proceedings of the International

Conference Energy efficiency and agricultural engineering, Russe, 2006, 347-353
[16] Z. Nenova, G. Georgiev, St. Ivanov. "Computer based system for monitoring of the air parametars in the closed premises." Russe University`s scientific publications, **54**(3), 186-190, 2015 (in Bulgarian)
[17] A. Levi, R. Ivanov, I. Bogdanov, N. Kolev. "Microcomputer systems network for plant growing management - approaches and structure." Electrotechnica & Electronica (E+E), 2013, **48**(7-8), 12-16, doi: 10.1007/s10973-019-08321-6.
[18] N. Nakayama, K. Misumi H., Koike T. TEMPERATURE SENSOR AND TEMPERATURE SENSOR SYSTEM, United States Patent, Patent No US 8,419,275 B2, Apr. 16, 2013
[19] C.M. Gerard, Smart temperature sensors and temperature sensor systems, Smart Sensors and MEMs (Second Edition), Woodhead Publishing, 2018, doi: 10.1016/B978-0-08-102055-5.00003-6.
[20] D. Zhu, C. Koh. TEMPERATURE SENSOR. United States Patent, Patent No US 10,378,969 B2, Aug. 13, 2019
[21] C.R. Carrigan, TRIAXIAL THERMOPILE ARRAY GEO-HEAT-FLOW SENSOR, PATENT-US-A7516399, 1992, DE92 004606
[22] G. Cortellessa, et al. "Experimental and numerical analysis in heat flow sensors calibration." J Therm Anal Calorim **138**, 2901 – 2912, 2019. doi: 10.1007/s10973-019-08321-6.
[23] T.C. Codau, et al., "Embedded textile heat flow sensor characterization and application." Sensors and Actuators A: Physical, **235**, 2015, 131 – 139, doi: 10.1016/j.sna.2015.10.004.
[24] E. Onofrei, et al., "Textile sensor for heat flow measurement." Textile Research Journal, **87**(2), 165-174, doi: 10.1177/0040517515627167.
[25] P. Wolkoff, "Indoor air humidity, air quality, and health – An overview." International Journal of Hygiene and Environmental Health, **22**1(3), 2018, 376 – 390, doi: 10.1016/j.ijheh.2018.01.015.
[26] V.D. Matko, "Sensor for high-air-humidity measurement." in IEEE Transactions on Instrumentation and Measurement, **45**(2), 561 – 563, April 1996, **doi:** 10.1109/19.492787
[27] M. Transfer, "Dynamik Modelling of Indoor Air Humidity. (PDF) Dynamic Modelling of Indoor Air Humidity (researchgate.net) Accessed April 2021
[28] G. Ganchev, et al., Оценка на заслепяване и намаляване на усещането за заслепяване. University of Mining and Geology - Sofia. 2017. (in Bulgarian)
[29] C Eng Faruk Bin Poyen. Luxmeter, showing light intensity, as it is received from a sensor. January 2017, (PDF) Luxmeter displaying light intensity as received by sensor (researchgate.net), Accessed April 2021
[30] S. Baev, "Спектрорадиометър в състав програмно-апаратного комплекса управления облучательными установками в закрытом грунте." SVETOTEHNIKA, SPECIAL ISSUE, 2019, 55-58, 2019.
[31] R. Nikam, et al., "Tin Doped Gas Sensors in Semiconductor Metal Oxide Form and Their Scientific Applications: A Review." J. Biol. Chem. Chron. 2019, **5**(3), 53-56, 2019
[32] Y. Wang, "ZnO Nanorods for Gas Sensors. Nanorods and Nanocomposites. Morteza Sasani Ghamsari and Soumen Dhara, IntechOpen, 2020 DOI: 10.5772/intechopen.85612.
[33] M. Miandehy, Nano graphen gas sensor (performance and applications). January 2021
[34] V. Broek, Jan; Weber, Ines C.; Güntner, Andreas T.; Pratsinis, Sotiris E. "Highly selective gas sensing enabled by filters." Materials Horizons, **8**(3), 661-684, 2019, doi: 10.1039/D0MH01453B.
[35] J. Miller, Proof-of-concept: The day after. Mentor®. A Siemens Business. 2018. www.mentor.com
[36] W. Ruh, "Drilling Deep into Digital Industrial Transformation will determine who survives and thrives." – Internet of Things Magazine, September 2018.
[37] C. Perera, et al., "Designing the Sensing as a Service Ecosystem for the Internet of Things." – Internet of Things Magazine, December 2018, 1, 2.
[38] R. Gonzalez, et al., "Navigation technics for mobil robots in greenhouses." Applied Engineering in Agriculture. 2009, Vol. 25, No. 2, pp. 153-165
[39] roboticsbiz.com. Drones in underground mines – Applications and benefits, Drones in underground mines - Applications and benefits (roboticsbiz.com) September 28 2020, Accessed April 2021
[40] J. Shahmoradi, et al., "A Comprehensive Review of Applications of Drone Technology in the Mining Industry." Drones. 2020. 4. DOI: 10.3390/drones4030034.
[41] University of Utah. A Safer Battery to Power Drones in Underground Mines. A Safer Battery to Power Drones in Underground Mines | Technology Org, April 9, 2021

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 401-405 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Quantum Secure Lightweight Cryptography with Quantum Permutation Pad

Randy Kuang*, Dafu Lou, Alex He, Alexandre Conlon

Quantropi Inc., Ottawa, K1Z 8P9, Canada

ARTICLE INFO

Article history:
Received: 18 June, 2021
Accepted: 17 August, 2021
Online: 28 August, 2021

Keywords:
AES
Quantum Permutation Gates
Quantum Permutation Pad
Permutation Matrix
Quantum Algorithm
QPP
Shannon entropy
Lightweight Cryptography
Streaming Cipher
Block Cipher

ABSTRACT

Quantum logic gates represent certain quantum operations to perform quantum computations. Of those quantum gates, there is a category of classical behavior gates called quantum permutation gates. As a quantum algorithm, quantum permutation pad or QPP consists of multiple quantum permutation gates to be implemented both in a quantum computing system as a quantum circuit operating on n-qubits' states for transformations and in a classical computing system represented by a pad of n-bit permutation matrices. Since first time proposed in 2020, QPP has been recently applied to create a quantum safe lightweight block cipher by replacing SubBytes and AddRoundKey with QPP in AES called AES-QPP. In AES-QPP, QPP consists of 16 selected 8-bit permutation matrices based on the shared classical key materials. For quantum safe, the key length can be any size from 256 bits to 4 KB. That means, this QPP holds up to 4 KB of Shannon information entropy. Its code size is less than 2 KB with 4 KB of RAM memory. In this paper, we propose to apply QPP for a streaming cipher and carry out its encryption performance and the randomness analysis of this streaming cipher. The proposed QPP streaming cipher demonstrates not only good randomness in its ciphertexts but also huge performance improvement: 13x faster than AES-256, with an overall runtime space (6.8 KB).

1. Introduction

Since the U.S. National Institute of Standards and Technology (NIST) announced the standardization of Advanced Encryption Standard or AES in 2001 [1], AES has been widely accepted as secure data encryption for data in transit or at rest. As a standard block cipher, AES accepts a fixed block size of 128 bits for three key lengths: 128, 192, and 256 bits with 10, 12, and 14 rounds respectively. Each AES round includes four steps: SubBytes, ShiftRows, MixColumns, and AddRoundKey. Over the past decade, the internet of things or IoT has captured the great attentions cross the world due to its potential to transform our daily lives through varieties of aspects such as smart home, smart city, autonomous vehicles, connected devices, etc. IoT devices are generally considered as resource constrained systems. They are often battery-powered, low computing power, and limited storages. These limitations put certain pressures on the standard AES to run in IoT devices, especially with high security requirements. The authors published their NIST report on lightweight cryptography [2], covering lightweight bloch ciphers, lightweight hash functions, lightweight message authentication codes, and lightweight streaming ciphers. In [3], the authors proposed their implementation of modified lightweight AES in FPGA, with a parallel manner for achieving better latency.

On the other hand, varieties of symmetric lightweight cryptographic algorithms have been proposed. In 2018 [4], the author have reviewed those algorithms to benchmark them on executing time, RAM memory and binary code sizes. those algorithms support the block sizes from 64 bits to 128 bits with key lengths from 80 bits to 128 bits.

AES generally faces three types of attacks: differential, linear, and integral [5]-[8]. The single static S-box representing substitutions or non-linear-transformations enables the differential analysis attacks due to some characteristic of XOR differences between input blocks and output blocks, especially impossible differences found at round 4, also called impossible differential attacks [6,7]. The differential analysis attack can be further improved with sets or multisets of input and output XOR results to create a new integral attack [8]. AddRoundKey at the end of each round contributes the linear analysis attack due to the linear transformation between rounds.

*Corresponding Author: Randy Kuang, randy.kuang@quantropi.com

https://dx.doi.org/10.25046/aj060445

In 1994, Shor proposed an algorithm to use quantum systems or qubits to perform computations called quantum computing [9]. Shor's algorithm enables a new natural parallel computing mechanism arising from the fundamental characteristic of their superpositions. With quantum computers, the classical exponential difficulty of prime factorizations becomes polynomial time, shaking the foundation of classical public key cryptography. The recent advancements in quantum computing development speeds up the urgency of quantum safe cryptography for both asymmetric and symmetric. In September 2019, Google announced their 54-qubit quantum computer called "Sycamore", marked their quantum supremacy [10]. On the other hand, in [11] the author made a milestone achievement in prime factorization with D-Wave's annealing quantum computer.

In 1996, Grover proposed his new search algorithm by using quantum computing mechanism called Grover's algorithm [12]. Grover's algorithm can achieve a square root complexity $O(\sqrt{n})$ in the unstructured search problem of size n, while any classical algorithm needs $O(n)$ queries. The squared searching power from the Grover's algorithm requires any symmetric cryptography not only to double the key length from 128 bits to 256 bits, but also to be true random. For resource constrained IoT devices, standard AES cryptography itself is already heavy so the doubled key length requires extra four rounds from 10 rounds to 14 rounds. It is necessary to explore different lightweight cryptographic algorithms in the post-quantum era. In [13], the authors proposed to use quantum cryptography with One-Time-Pad or OTP for secure encryption for power grid data. This can be considered as a hybrid quantum secure encryption combining QKD with OTP. In [14], the authors proposed a new lightweight symmetric cryptographic algorithm called Saturnin by introducing two representations of AES 256-bit internal states: the 2-dimensional and 3-dimensional notations. In its 2-dimensional representation, a 256-bit state can be expressed by sixteen 16-bit registers; but in its 3-dimensional representation, it is expressed by a 4x4x4 cube of nibbles. Through those major changes, Saturnin established a new quantum resistant lightweight cryptography for 128-bit and 256-bit block ciphers and a 256-bit hash function.

Quantum mechanics allow us to have two implementations of quantum gates: physically for quantum computing power and digitally for quantum security. In [15], the authors first attempted to present classical information quantum mechanically with a state denoted by a Dirac ket over a quantum computational basis. The well-known symmetric group S_n containing entire group actions over a set of n items has its matrix representations or permutation matrices over the corresponding quantum computational basis. The extremely large size of the quantum permutation group, $2^8!$ (factorial), for an 8-qubit quantum computational basis holds huge equivalent Shannon information entropy, desirable for information security. For an 8-bit system, the permutation group is S_{256}. Kuang and Bettenburg [15] extend the Shannon perfect secrecy of the classical one-time-pad (OTP) over GF(2^n) [16], to their proposed quantum permutation pad (QPP) over a quantum computational basis. In contrast to the one-time-use nature of OTP, QPP retains the Shannon perfect secrecy over multiple uses, thanks to the general non-commutativity properties of the symmetric group.

In [17], the authors have applied QPP for a lightweight block cipher called AES-QPP, with 16 permutation matrices selected by using the shared classical random key to replace both SubBytes and AddRoundKey. AES-QPP has a footprint below 2KB and RAM memory 4KB, with performance improvement about 3x. In this paper, we extend the work [17] further for a quantum safe lightweight streaming cipher.

This paper is organized as follows: Section 2 is for the summary of lightweight block cipher AES-QPP. Then Section 3 describes the proposed streaming cipher. The randomness and performance of the proposed streaming cipher is presented in Section 4. We will draw a conclusion at the end.

2. Quantum safe Lightweight Block Cipher AES-QPP

2.1. Quantum Permutation Pad

QPP is a pad of quantum permutation matrices randomly selected from the n-bit permutation group [15, 17] as shown in Fig. 1 for AES with 16 8-bit permutation matrices. They are all 2^8x2^8 square matrices and they are unitary and reversable so their reverse transformations are their transposes. This characteristic is great for lightweight cryptography, especially for resource constrained IoT devices. At the encryption side, we can use the selected QPP and then at receiving side we use their transposes or QPP^T. Therefore, transformations are exactly symmetric with the same computational performance for encryption and decryption. The QPP selection is a process to map classical key materials into a QPP pad over the 8-bit computational basis. There are several algorithms to be used such as RC4 key scheduling algorithm or Fisher-Yates random shuffling algorithm. We use the Fisher-Yates algorithm to map classical key into a QPP pad as shown in

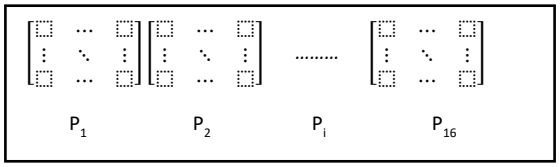


Figure 1: QPP is illustrated.

Algorithm 1.

Algorithm 1: Pseudo code of mapping a key to permutation matrix

```
Result: Permutation matrix P[256][256]
Input: 256 bytes of random key k[256]
Initialization: set S[256] and P[256][256] to 0
for i = 0 to 255
    S[i] = i;
end for
i = 255
while i-- > 1 do
    j = k[i];
    swap S[j] and S[i];
end
for i = 0 to 255
    P[i][S[i]] = 1;
end for
```

For each permutation matrix, we need 256 bytes of random numbers as shown in the pseudo-code. For a QPP pad with 16

permutation matrices, we would need to supply total 4 KB of random numbers. That means, the selected QPP holds total 4 KB of entropy. It is up to the desired security level to choose a right key length from 256 bits to 32,768 bits. To support this variable key length, a key scheduling is required to extend a variable key length to 4 KB.

2.2. AES-QPP

SubBytes in AES performs a substitution with a static S-box which is a 16x16 matrix. S-box can be converted to a 256x256 permutation matrix. The substitution can be considered as the permutation matrix multiplication with an 8-bit state vector over the 8-bit computational basis. AddRoundKey step performs byte-by-byte XOR operations between the output block from MixColumns step and a round key. XOR operation is a special case of permutation transformations. In our early block cipher, we use the QPP to replace both SubBytes and AddRoundKey in an AES round with the ShiftRows and MixColumns steps as follows:

1. 16 8-bit QPP;
2. ShiftRows;
3. MixColumns;
4. the same QPP as in the step 1.

We illustrate AES-QPP in Figure 2. In standard AES, each byte in a16-byte block is supplied to the single static S-box, but in AES-QPP, each byte in a block is supplied to its corresponding permutation matrix. The last step is performed in the same way as in the first step, unlike the standard AES in the AddRoundKey step with each byte from MixColumns to be XORed with the corresponding byte in a round key. This design of AES-QPP enables us to use variable key length without change the implementation of the cryptography.

The decryption is the same process as in the encryption with transposed QPP^T.

2.3. AES-QPP Rounds

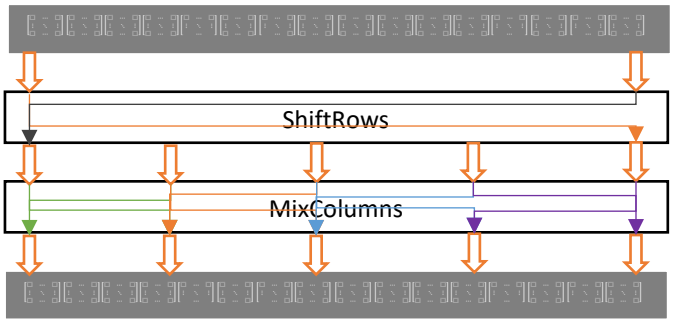


Figure 2: The proposed AES-QPP is illustrated

In comparison with AES rounds, QPP increases the diffusion capability at least 16x with 16 permutation matrices. This extra strengthened diffusion ability helps to reduce the number of rounds. AES-256 needs 14 rounds to achieve good randomness in ciphertexts. In our early report, the number of rounds in AES-QPP was reduced to 5 rounds. The ciphertext still demonstrates excellent randomness from NIST and ENT random testing suites.

2.4. Cipher Randomness with Shannon Entropy Distribution

Cipher randomness is a good measure for a cryptosystem to avoid statistical analysis attacks. We have demonstrated randomness analysis with NIST random test suite, especially with ENT testing tool to identify any byte level and bit level biases. AES-QPP shows excellent randomness in its ciphertexts. In NIST testing suites, AES-QPP ciphers with 5 rounds pass all 15 testing cases. In the sensitive testing suite ENT, the AES-QPP ciphers not only pass 6 testing cases but also demonstrate excellent Chi Square value, arithmetic means, Monte Carlo $\boldsymbol{\pi}$, as well as serial correlation. Here we want to add analysis for Shannon entropy distribution for AES-QPP in comparison with the standard AES. Shannon entropy distribution performs analysis of each 16-bit, entropy per 16-bit random data, as well as how close to the Gaussian distribution.

Figure 3 plots the Shannon entropy distribution for AES-QPP with 5 rounds, using a 16 bits Shannon Entropy calculator [18]. . The entropy is 15.999072 per 16 bits of ciphertexts, very close to the ideal entropy 16 bits. The left side of the graph displays the frequency of each 16-bit integer. The graph displays a nice symmetric behavior around an average count. The analysis shows that the median counts are 781, minimum count 656, and maximum count 896 for 100 MB files. The right-hand side of Fig. 3 shows a good Gaussian distribution with a nice symmetric shape around the median count, indicating a good randomness in AES-QPP-5 ciphertexts.

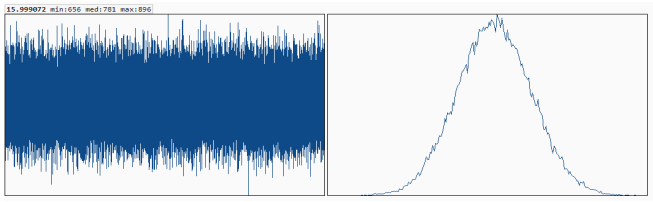
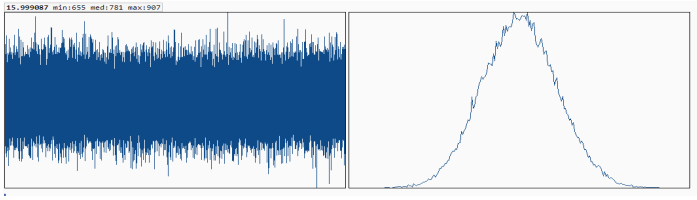


Figure 3: Shannon entropy distribution is plotted for AES-QPP-5.

Figure 4 plots the Shannon entropy distribution for AES-256. The same size of AES-256 ciphertext file as AES-QPP-5 is used. It demonstrates a very close relationship to Figure 5 with median counts 781, minimum counts 655 comparing with 656 in AES-QPP, and maximum counts 907, slightly better symmetry than AES-QPP. The Shannon entropy of AES-256 is 15.999087 per 16 bits of ciphertexts, extremely close to AES-QPP-5.



3. Quantum Safe Lightweight Streaming Cipher with QPP

In the last section, we discussed the block cipher implementation with QPP. We can also implement it in a streaming cipher with a pre-randomized dispatcher as shown in Figure 5.

Seed is an input classical key material, Init box is the scheduling to map the classical key material into a QPP pad as shown in Algorithm 1. PRNG box is a pseudo random number generator used to pre-randomize input plaintexts with a directly XOR operation before dispatching them to QPP. Dispatcher box also takes a PRNG byte and performing a 4-bit right shift as an index to the permutation matrix inside the QPP pad. A ket $|m\rangle$ represents a plaintext byte. The ciphertext denoted by a ket $|c\rangle$ can be created byte-by-byte in the same way as input plaintexts. At the receiving side, the same shared classical key material is used to establish the same QPP pad in a transposed mode because permutation matrix is unitary and reversable. The decryption is straightforward.

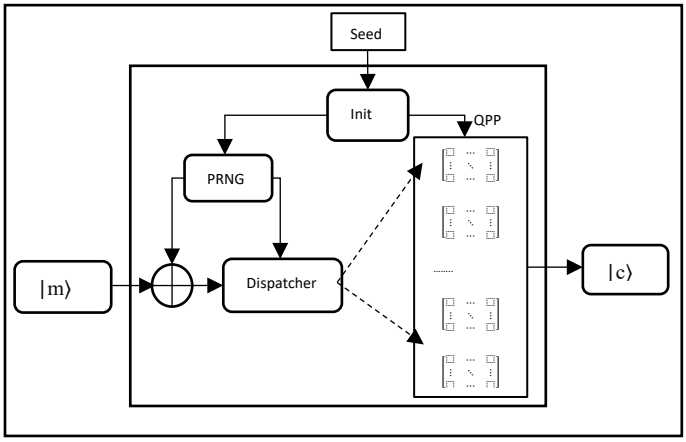


Figure 5: A streaming cipher is illustrated

QPP streaming cipher implementation eliminates the repeated rounds for better randomness ciphers, replacing with a pre-randomized process under a consideration of bijective quantum permutation transformations. We can roughly estimate the executing time for the encryption at the level of a single round AES encryption. This would dramatically improve its performance for latency and battery consumptions.

The footprint of QPP streaming cipher is about 2.5 KB, In comparison with 1.1 KB in AES-QPP because of the pre-randomization process in QPP streaming. RAM memory is the same as AES-QPP at 4KB because we still use a QPP pad with 16 permutation matrices.

4. Discussions of QPP Streaming Cipher

For the proposed QPP streaming cipher shown in Figure 5, we create a plaintext file of 120 MB by a paragraph of English sentences. Then QPP streaming encrypts the plaintext file and stores the ciphertexts into ciphertext file. The ciphertext file is passed all NIST 15 randomness test cases. We should be very interesting to see the ENT testing reports listed in Table 1, together with the results for the input plaintext file. The plaintext file comes with 4.49 bits of entropy per 8-bits, huge Chi Square value, totally wrong arithmetical mean, as well as a wrong Monte Carlo π value. All those results show that the input plaintext file is totally biased. The ciphertext file produced from our proposed QPP streaming cipher demonstrates excellent randomness: 8 bits of entropy per 8 bits of ciphertexts, Chi Square value 237.64 with a p-value 0.775, arithmetical mean 127.49 compared to 127.50, very nice Monte Carlo π value 3.141771788 compared to 3.14159265, and serial correlation value 8.2×10^{-5}. The overall ENT testing results from QPP streaming cipher is very similar to AES-QPP-5 [17].

Table 1: ENT randomness testing reports for QPP streaming cipher and plaintext files of size 120 MB

ENT	Plaintext	QPP Streaming
Entropy (bits)	4.491422	7.999999
Chi Square	1736817422.90	237.64
p-Value	0.0001	0.775
Arith. Mean	95.7060	127.49
Monte Carlo π	4.000000000	3.141771788
Serial Corr.	0.047905	0.000082

Figure 6 plots the Shannon Entropy distribution with 120 MB ciphertext file from a streaming encryption implementation of 16 permutation matrices. The Shannon entropy is 15.999244 bits per 16 bits of ciphertexts. The distribution demonstrates a nice Gaussian type with a med 957, minimum 826 and a maximum 1085, slightly better symmetry than AES-QPP-5 [17].

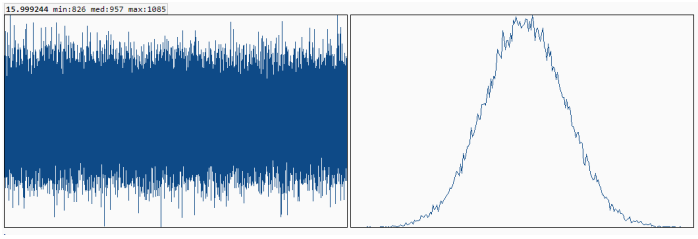


Table 2 illustrates performance comparisons among AES-256, AES-QPP-5, and QPP streaming cipher. For AES-256, we take the open-source implementation or Tiny-AES. AES-256 demonstrates a fastest key scheduling with 0.01 ms, then AES-QPP-5 with 0.120 ms, and QPP streaming with 0.235 ms, respectively. It is understandable that QPP initialization takes longer time because it processes 4KB key materials to select 16 permutation matrices, unlike in AES where key is only 32 bytes long. Also, in QPP streaming cipher, we take a special handling to increase confusion capability. This special handling would help to produce totally different QPP pad even with a single bit change in the supplied key materials.

Table 2: Performance comparisons are tabulated for key scheduling and encryptions for AES-256, AES-QPP-5, and QPP streaming cipher. Encryption speeds are tested with 16 bytes blocks in the same computer: MacBook Pro, 2.6 GHz 6-Core Intel Core i7.

	AES-256	AES-QPP-5	QPP Streaming
Key Schedule	0.01 ms	0.120 ms	0.235 ms
Encryption MB/s	51.3	115.1	672.3
Ratio	1.0	2.24	13.1
Code Footprint	11.5 KB	1.39 KB	2.5 KB
Run Time Space	12.0 KB	5.39 KB	6.5 KB

It can also be seen from Table 2 that AES-256 is the slowest among three ciphers with an encryption speed 51.3 MB/s for 16 bytes blocks, then AES-QPP is faster than AES-256 with 115.1 MB/s, finally QPP streaming cipher is the fastest cipher with 672.3 MB/s. AES-QPP is 2.24x faster than AES-256, slightly slower than what we expected 2.8x, that may be the fact that overall permutation transformations with 16 permutation matrices take longer than each step of SubBytes and AddRoundKey in AES. However, the discrepancy is very acceptable. QPP streaming cipher is 13x faster than AES-256, indicating that pre-randomization with randomly dispatching together is almost equivalent to a single round in AES. In comparison with AES-QPP-5, QPP streaming cipher is 5x faster than AES-QPP-5.

In comparison with code sizes, compiled footprints are 11.5 KB, 1.39 KB, and 2.5 KB for AES-256, AES-QPP, and QPP streaming respectively, and runtime memory spaces are 0.47 KB for AES-256, 4 KB for both AES-QPP and QPP streaming cipher. As for overall runtime space, AES-QPP takes the least runtime space at 5.39 KB, then QPP streaming cipher at 6.5 KB, after then AES-256 needs the most runtime space at 12 KB.

5. Conclusion

We have applied quantum permutation pad or QPP to establish both lightweight quantum safe block cipher and streaming cipher. In a block cipher implementation, QPP replaces both SubBytes and AddRoundKey in a standard AES or called AES-QPP. In addition to cipher randomness analysis in [17], we perform the Shannon entropy distribution for a more complete randomness analysis of this quantum safe block cipher.

In this paper We explored the QPP algorithm for a streaming cipher with a straightforward pre-randomization and random distribution process. The randomness analysis of the QPP streaming cipher demonstrates very good randomness, especially with ENT randomness testing tool. The very promising encryption speed plus overall memory space makes QPP streaming be a good candidate for quantum safe lightweight streaming cipher and AES-QPP for quantum safe lightweight block cipher.

In the future, we may extend the exploration both ciphers for 4-bit permutation matrix pad to further reduce their runtime memory spaces.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Information Technology Laboratory (National Institute of Standards and Technology), Announcing the ADVANCED ENCRYPTION STANDARD (AES), Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, MD 2001.

[2] K. McKay, L. Bassham, M. Sonmez, N. Mouha, Report on Lightweight Cryptography, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, [online], 2017, doi:10.6028/NIST.IR.8114 (Accessed August 23, 2021).

[3] M. James, D. S. Kumar, "An Implementation of Modified Lightweight Advanced Encryption Standard in FPGA，" Procedia Technology, **25**, 582-589，2016, doi:10.1016/ j.protcy.2016.08.148.

[4] D. Dinu, Y. L. Corre, D. Khovratovich, L. Perrin, J. Großschädl, A. Biryukov, "Triathlon of lightweight block ciphers for the Internet of things," Journal of Cryptographic Engineering, **9** (3), 283–302, 2018, doi:10.1007/s13389-018-0193-x. S2CID 1578215.

[5] J. Daemen, V. Rijmen, The Design of Rijndael, *AES - T*he Advanced. Encryption Standard, Springer-Verlag 2002.

[6] H. M. Heys, S. E. Tavares, "Substitution-permutation networks resistant to differential and linear cryptanalysis," Journal Cryptology **9,** 1–19, 1996. doi:10.1007/BF02254789.

[7] L. O'Connor, On the distribution of characteristics in bijective mappings, J. Cryptology **8,** 67–86, 1995, doi:10.1007/ BF00190756.

[8] J. Lu, O. Dunkelman, N. Keller, J. Kim, "New Impossible Differential Attacks on AES," In: Chowdhury D.R., Rijmen V., Das A. (eds) Progress in Cryptology - INDOCRYPT 2008. INDOCRYPT 2008. Lecture Notes in Computer Science, **5365**, Springer, Berlin, Heidelberg. doi：10.1007/978-3-540-89754-5_22.

[9] K. Autre, K. Arya, , *et al*， "Quantum supremacy using a programmable superconducting processor，" Nature **574** (7779)，505–510. 2019.

[10] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring，" in Proceedings 35th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press: 124–134, 1994.

[11] B. Wang, F. Hu, H. Yao, *et al.*, "Prime factorization algorithm based on parameter optimization of Ising model," Sci Rep **10,** 7106, 2020, doi:10.1038/s41598-020-62802-5

[12] L. Grover, "A fast quantum mechanical algorithm for database search," In: Proceedings of the 28th ACM STOC, Philadelphia, Pennsylvania, 212–219, ACM Press, 1996.

[13] Y. Li, P. Zhang and R. Huang, "Lightweight Quantum Encryption for Secure Transmission of Power Data in Smart Grid," in IEEE Access, **7**, 36285-36293, 2019, doi: 10.1109/ACCESS.2019.2893056.

[14] A. Canteaut, S. Duval, G. Leurent, M. Naya-Plasencia, L. Perrin, T. Pornin, A. Schrottenloher, "Saturnin: a suite of lightweight symmetric algorithms for post-quantum security," IACR Transactions on Symmetric Cryptology, **2020**(S1), 160-207, doi:10.13154/tosc.v2020.iS1.

[15] R. Kuang, N. Bettenburg, "Shannon Perfect Secrecy in a Discrete Hilbert Space," 2020 IEEE International Conference on Quantum Computing and Engineering (QCE), Denver, CO, USA, 249-255, 2020, doi: 10.1109/QCE49297.2020.00039.

[16] C. E. Shannon, "Communication Theory of Secrecy Systems?" Bell System Technical Journal， **28** (4): 656–715， October 1949.

[17] R. Kuang, D. Lou, A. He and A. Conlon, "Quantum Safe Lightweight Cryptography with Quantum Permutation Pad," 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), 790-795, 2021, doi: 10.1109/ICCCS52626.2021.9449247.

[18] Server Test, https://servertest.online/entropy.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 406-413 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Personalized Clinical Treatment Selection Using Genetic Algorithm and Analytic Hierarchy Process

Olena Nosovets, Vitalii Babenko*, Ilya Davydovych, Olena Petrunina, Olga Averianova, Le Dai Zyonh

Department of Biomedical Cybernetics, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, 03056, Ukraine

ARTICLE INFO

Article history:
Received: 13 July, 2021
Accepted: 24 August, 2021
Online: 28 August, 2021

Keywords:
Clinical Treatment
Genetic Algorithm
Analytic Hierarchy Process
Modeling

ABSTRACT

The development of Machine Learning methods and approaches offers enormous growth opportunities in the Healthcare field. One of the most exciting challenges in this field is the automation of clinical treatment selection for patient state optimization. Using necessary medical data and the application of Machine Learning methods (like the Genetic Algorithm and the Analytic Hierarchy Process) provides a solution to such a challenge. Research presented in this paper gives the general approach to solve the clinical treatment selection task, which can be used for any type of disease. The distinguishing feature of this approach is that clinical treatment is tailored to the patient's initial state, thus making treatment personalized. The article also presents a comparison of the different classification methods used to model patient indicators after treatment. Additionally, special attention was paid to the possibilities and potential of using the developed approach in real Healthcare challenges and tasks.

1. Introduction

This paper is an extension of work originally presented in the 15th International Conference on Computer Sciences and Information Technologies, held in Zbarazh (Ukraine) in September 2020 [1].

It is well known that human treatment is a delicate moment, as any wrong decision can radically affect the person's state [2], for which doctors will be primarily responsible [3]. The word 'treatment' can mean different things, and it should be understood that not all treatments are equal. For example, treating a person for influenza [4], wherein most cases it suffices to prescribe a few medications for a couple of weeks, is not equivalent to treating cancer [5], which can last for years [6-7]. Undoubtedly, this does not mean that influenza poses no risk to human health [8], but if influenza progresses to a critical stage, medical supervision and prescription of clinical treatment are mandatory [9-11].

Clinical treatment refers to the process during which the patient stays in a medical institution under the strict supervision of clinicians and specialists, and undergoes all stages of therapy (including experimental treatments) to eliminate the symptoms of disease or complications. One of the major objectives in Healthcare is to treat the disease itself (not the symptom), so treatment must be chosen carefully and without hesitation. It is therefore the regular practice to use clinical protocols (guidelines) [12-14] while prescribing treatment.

Briefly, clinical protocols are systematically developed statements, which assist clinicians and patients in making decisions about appropriate treatment for specific conditions based on the best scientific evidence at the time of development [15-17]. These protocols are the primary medico-technological documents that specialists must follow in any given clinical situation, choosing the most effective solution to cure a patient. A clinical protocol is a manual for a doctor, and it contains guidelines to treat a specific disease.

In Ukraine, Ministry of Healthcare allows international clinical protocols under Order no. 1422 of the by 29 December 2016, which entered into force on 28 April 2017 [18-19]. To get these protocols, the Ministry of Healthcare has signed an agreement with "Duodecim Medical Publications Ltd", a Finnish medical-scientific company specializing in comprehensive solutions for evidence-based medicine [18]. Since then, about a thousand international clinical protocols in English have been available online for registered users in Ukraine. Using new clinical protocols in medical practice has become one of the most important ways of implementing evidence-based medicine in Ukraine [19].

Many researchers are now focusing on clinical protocols to model the clinical treatment process [10] because these protocols do not guarantee a complete cure for the patient. The reason is that doctors do not have full knowledge of medicines, their types, and their correct use. This responsibility lies with clinical pharmacists (or clinical provisors) [20-21]. Specialist training in this field in

*Corresponding Author: Vitalii Babenko, vbabenko2191@gmail.com

Ukraine was established in 1999 by the National University of Pharmacy in Kharkiv [22]. Despite more than 20 years of training such specialists, the country simply does not have enough clinical pharmacists, which may be due to the limited financial resources of local medical institutions. This problem, and the rapid digitization of Healthcare in Ukraine [23], make it an urgent task to develop an automated decision support system for clinical treatment selection.

The making of decision support system was chosen because the selection of optimal clinical treatment should not be made solely by a machine. Medical staff is responsible for causing harm to humans [24], which is why Artificial Intelligence (AI) software must be used with utmost care for these challenges.

The basic requirements for such systems are:

- Usage of clinical protocol standards as a basis for decision support in treatment selection.
- As set out in right no. 12 of the European Charter of Patients' Rights [25], the individual characteristics of the patient must be considered.
- The availability of AI and Machine Learning (ML) techniques to provide high-quality assistance for medical staff.

This research aims to develop a general approach to selecting optimal clinical treatment based on the requirements listed above. The first thing to consider is how treatment outcomes are assessed, as that enables AI- and ML-based applications to improve efficiency in decision making.

2. Assessing the Quality of Treatment: An Overview

2.1. Scales of Patient Severity

The quality of clinical treatment should be assessed foremost by results, bearing in mind that the aim of providing medical care is to ease the patient's state. At all stages of treatment in the various fields of medicine, the determination of a patient severity objective assessment is essential for clinical decision-making. The close association of state severity with the prognosis of mortality risk further extends the application of such tools at different stages of care. For example, intensive care units (ICU) must provide prognoses for patients within the first 24 hours of admission.

The stratification of patients into risk groups according to the severity of their state is necessary to compare treatments and their quality, establish differences between different medical institutions, and evaluate the results of clinical trials in evidence-based medicine.

Establishing a prognosis comprises estimating the probability of death using indicators that are commonly used to diagnose and treat critically ill patients [26]. Severity scales are the classic tools used to establish such prognoses.

Rating systems quantify or qualify the severity of a state and classify the patient into specific risk groups, based on the analysis of anatomical, physiological and biochemical abnormalities [26-27]. Over two dozen severity scales have been developed, but only some of them can be considered universally accepted.

The most commonly used (in ICUs of USA and EU countries) scoring systems for assessing patient severity are: SAPS II, APACHE II and III, GCS, MPM II, SOFA, MODS, and LODS [28].

2.2. Models of Patient Severity

Most morbidity estimation models are based on Logistic Regression [29]. Authors of [30] analyzed papers that use SOFA models to predict mortality in ICUs. Only ten studies (56%) applied logistic regression models, and five of them had validated models with independent tests. The following models were also considered: combined with other assessment scales (APACHE, MOD) and additional measures of organ failure, time models (sequential SOFA scores), and automatically detected from the data SOFA templates. For example, the predictive ability of APACHE II was assessed using Logistic Regression analysis. For the Logistic Regression model based on the APACHE II score, the AUC of the ROC curve [31] was 0.863. Authors note that although there is heterogeneity across studies, it is impossible to say which SOFA-based model is optimal.

Decision Tree [32] methods have recently become more widespread in medical research. Clinical practitioners swallow them because they are illustrative and can turn into logical conditions (classification rules). Classification Trees have been used in critical situations, e.g., to calculate the probability of death from coronary pathologies [33], intracerebral hemorrhage [34] or craniocerebral injury [35], to predict persistent autonomic states [36], and to stratify patient groups by the likelihood of mortality in the general population of ICUs [37-38].

In [39], the author predicted the probability of hospital mortality using three Decision Tree classification algorithms: CART, CHAID, and C4.5. All models are based on estimating the severity of patients within the first 24 hours of admission only (2864 patients, 70:30). Authors of [39] point out that the chief advantages of Decision Trees are that the resulting decision rules can be easily interpreted and the composition of the patient group obtained at each final node of the tree is relatively homogeneous. All Decision Tree models achieved the AUC of 0.75-0.76, which was close to the AUC for APACHE II (0.77) but lower than the Logistic Regression AUC (0.81).

It is worth noting that such multidimensional models are not designed to handle streaming data. Modern morbidity estimation models are based on ML methods such as Support Vector Machine (SVM) [40], Bayesian models [41], Artificial Neural Network (ANN) [42], etc.

The predicted length of stay (LOS) [43], which is based on monitoring data, is seen as a target that helps plan resuscitation resources and make ICU care individualized.

2.3. Information Systems of State Assessment and Forecasting

INTCare [44] is an intelligent decision support system for intensive care medicine. It is a system based on both collecting data from monitors at the bedside and updating the model, reducing the need for human intervention. INTCare currently provides predictions about organ failure and the likelihood of in-hospital death. Reliable prognostic results contribute to improving the

quality of service. The system presents functional and structural aspects. It aims to automate the knowledge discovery process.

The most important feature of INTCare's intelligent decision support system is the ability to operate autonomously and in real-time. Two approaches were used to model and predict 2 targets (survival and LOS):

- Collected data and physiological features during the first 24 hours of inpatient treatment.
- Collected patient clinical data in real-time.

In the first approach, the achieved results were poor (73% accuracy). However, when predictions (dwell time) were made using data collected in real-time, the results became higher (model sensitivity 96.1%). Researchers in their work used these models: SVM, Decision Trees, and Naïve Bayes. For survival prediction, the Decision Trees method had the best result with a sensitivity of 87.32%.

A systematic review of the literature (2008-2018) [45], which aims to investigate the use of ML to improve patient health, analyzed 78 such studies. The conclusion is that AI techniques can analyze and learn useful standards from clinical datasets (which are stored in electronic medical records) to provide better evidence for supporting health professionals' decisions.

More recently, the work has appeared that uses Deep Learning (DL) techniques, namely, Recurrent Neural Networks (RNN) with the Long Short-Term Memory (LSTM) architecture [46]. Experiments have shown that it is possible to predict vital signs in advance with good accuracy (more than 80%) to the deterioration of the patient's state. Predicting a patient's vital signs and using them to calculate a prognostic index makes it possible to predict future severe diagnoses that would not be possible using only the patient's current vital signs (50%-60% of cases were not detected).

Also noteworthy is the Ukrainian work by Nastenko et al. [47], which used Group Method of Data Handling (GMDH) [48] models and Simplex Method [49] optimization algorithm to select the treatment strategy.

3. Mathematical Background

Let's devise the general Healthcare challenge of this research. It is to find the optimal clinical treatment for the patient. In fact, in the mathematical space of objects, the patient (object) can be represented as a multidimensional vector, where the parameters are his/her indicators. In the simplest case, 2 patient states are possible in a given space: an initial state (before treatment) and a final state (after treatment). On this basis, 2 vectors of data are given for the challenge:

- $X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}^T$ – describes the initial state of the patient (x_1, x_2, …, x_n are the patient's indicators before clinical treatment).
- $Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{pmatrix}^T$ – describes the final state of the patient (y_1, y_2, …, y_m are the patient's indicators after clinical treatment).

Finding the optimal treatment involves finding the optimum of vector Y. The clinical treatment, which is applied to a patient for getting the Y-vector, can be described by a vector $I = \begin{pmatrix} i_1 \\ i_2 \\ \dots \\ i_k \end{pmatrix}^T$, where i_1, i_2, …, i_k – are the different types of drugs (the influence parameters on the patient's state). Therefore, Y-vector directly depends on vectors X and I (Fig. 1).

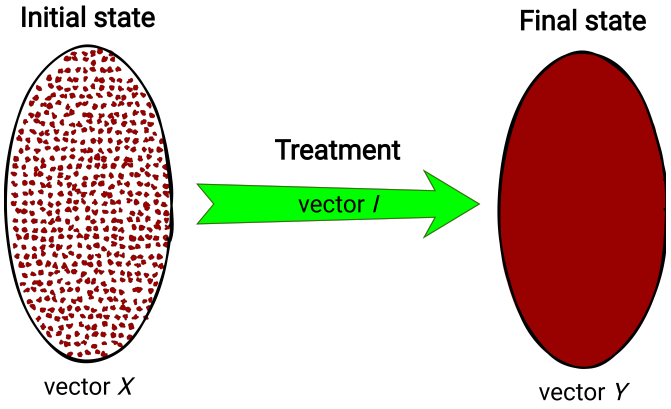


Figure 1: Visual representation of treatment process

This dependency can be described by the next general equation:

$$Y = X \times I \Leftrightarrow \begin{cases} y_1 = \sum_{j=1}^{n}\sum_{z=1}^{k} x_j i_z \\ y_2 = \sum_{j=1}^{n}\sum_{z=1}^{k} x_j i_z \\ \dots \\ y_m = \sum_{j=1}^{n}\sum_{z=1}^{k} x_j i_z \end{cases} \quad (1)$$

Thus, it can be said that $X = Y$ if no treatment is given (without regard to externalities).

Equations for $y_1, y_2, \dots, y_m$ (1) can be both linear or non-linear, parametric or non-parametric. With their usage, it seems possible to simulate (modeling) the clinical treatment process. Consequently, a multi-criteria optimization problem arises, where it is necessary to find such values of I-vector that will give the optimum of Y-vector. Values of X-vector are set by default, so the personalized solution search will be done by considering the patient's initial state indicators (this idea was proposed in [47] for a single-criteria problem). So, it is necessary to create an algorithm that will solve such a task.

4. Personalized Clinical Treatment Selection Algorithm

As mentioned earlier, finding the optimal clinical treatment is a multi-criteria optimization problem. However, creating the algorithm for solving this problem raises the following issues:

- Simultaneous optimization of the *Y*-vector parameters (patient's indicators after treatment).
- Searching the values of *I*-vector (influence parameters on the patient's state), which gives the global optimum (NP-complete problem).

The first issue can be solved using Multi-Criteria Decision Making (MCDM) methods [50]. Since in most cases the final state of a patient is described by two or more *Y*-vector parameters (in this problem – the criteria for optimization), it is worthwhile to assess the patient's state after treatment in the right way. MCDM methods allow getting a convolution of several criteria into one so-called "supercriterion". Apart from solving simultaneous optimization, this supercriterion can be used as an assessment metric to describe the final state of the patient.

One of the simplest and most easily interpreted methods of MCDM is the Analytic Hierarchy Process (AHP) [51-52], invented by Thomas L. Saaty in the 70s. This method allows getting a function of additive convolution by pairwise comparison of criteria priorities. The comparison mechanism by AHP in general form is shown in Table 1.

Table 1: The General Form of Criteria Pairwise Comparison

	y_1	y_2	…	y_m
y_1	$\frac{v_1}{v_1}$	$\frac{v_2}{v_1}$	…	$\frac{v_m}{v_1}$
y_2	$\frac{v_1}{v_2}$	$\frac{v_2}{v_2}$	…	$\frac{v_m}{v_2}$
…	…	…	…	…
y_m	$\frac{v_1}{v_m}$	$\frac{v_2}{v_m}$	…	$\frac{v_m}{v_m}$

where: v_i – the sequential number in the criteria list of *Y*-vector, ranked by importance.

The above table interprets the matrix of criteria pairwise comparison. To obtain a metric for the final patient state using AHP, the geometric mean for each matrix row is calculated. Then, the obtained values should be normalized; they will be the weights (w) of each criterion of the *Y*-vector, so the metric can be represented as follows:

$$F_{ac} = w_1 y_1 \pm w_2 y_2 \pm \ldots \pm w_m y_m \quad (2)$$

This metric is a function of additive convolution (F_{ac}) of the criteria. It has the advantage of flexibility because it depends on priorities set out in Table 1 by the decision-maker. The signs in (2) are placed depending on whether it is necessary to maximize (then the "+" sign) or minimize (then the "-" sign) y_i.

That solves the first issue of the given multi-criteria optimization problem, which allows converting it to the single-criteria optimization problem, where it is necessary to find the maximum of F_{ac}. Solving it can be done by many optimization approaches. One of the most famous methods is the Genetic Algorithm [53-55] – a stochastic method for finding the required solution. The ideas of natural selection and genetics provide a fast search for the global optimum, thus solving the second issue of the NP-complete problem. The algorithm is shown schematically in Fig. 2.

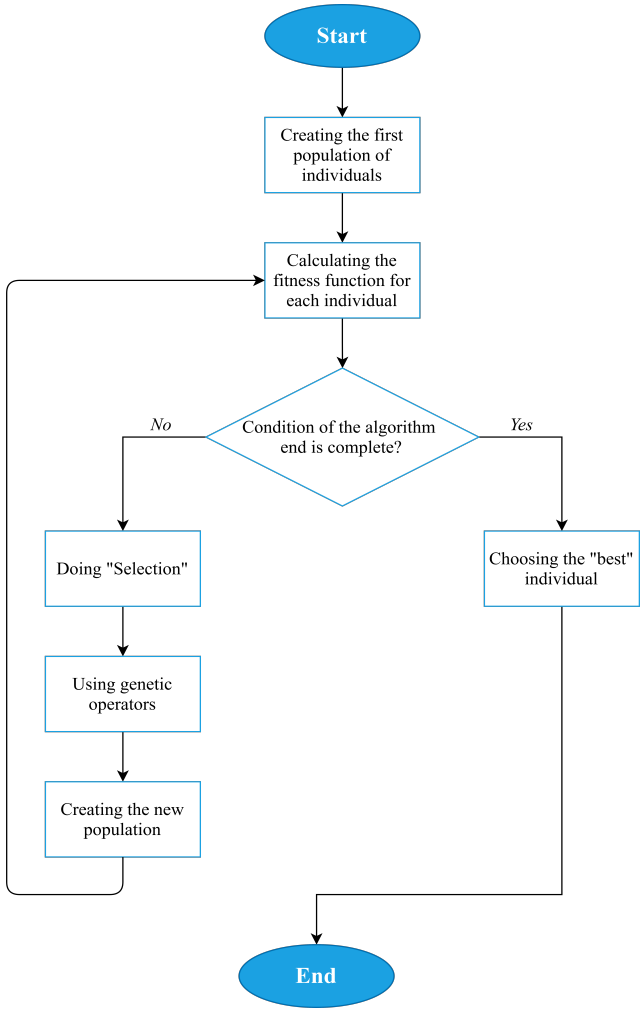


Figure 2: Genetic Algorithm scheme

To describe the Genetic Algorithm in more detail:

1. A random sample ("population") of *N* arrays ("individuals", or "chromosomes") that contain values of *I*-vector parameters ("genes") is created (Figure 3).

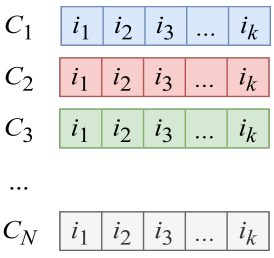


Figure 3: Population in general form

The number N, as well as the boundaries in which the values of genes will lie, are set directly by the researcher.

2. "Fitness function" for each individual is calculated. In the current research, the fitness function is F_{ac} (2).

3. The condition of the algorithm end is checked (it can be the presence of the preassigned value of F_{ac} or exceeding the time limit of the algorithm).

3.1. If the condition is complete, the Genetic Algorithm returns the "best" individual (optimal clinical treatment strategy).

3.2. If the condition is incomplete, the formation of a new population begins.

3.2.1. The "selection" [53-55] of individuals from the current population is carried out. This procedure aims to select individuals for the next generation creation, and the chance of selecting each individual directly depends on the value of his fitness function. The selected individuals form N pairs, which will then give "new" individuals.

3.2.2. The usage of "crossover" (one of the genetic operators) [53-55] for crossing the resulting pairs of individuals. A mixing of "genes" (parameters of I-vector) occurs between a pair of individuals, thus forming a new individual that stores the information about his "ancestors". In the general case, randomly the "crossover point" is determined, which allows mixing a pair of "parents": the genes of the first parent are before the crossover point, and after it – the genes of the second parent:

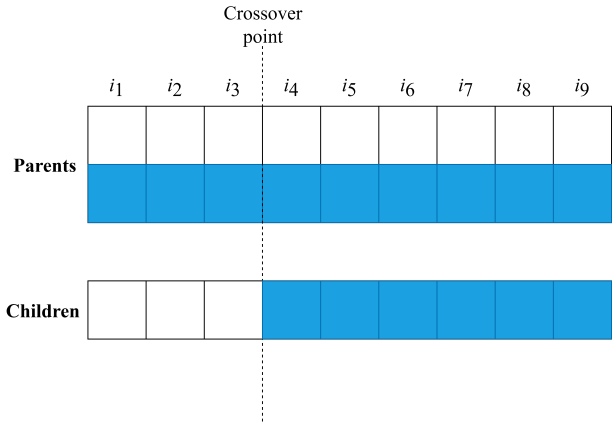


Figure 4: Example of crossover

That creates a new individual. To ensure the diversity of the population during the entire algorithm's operation, another genetic operator called "mutation" [53-55] is also used. This operator can be triggered with low probability instead of crossover, and its main purpose is to replace randomly selected genes of individuals with completely new ones.

3.2.3. Back to point number 2.

4. Using the best individual as a recommendation for personalized treatment. Multiple choices can be derived so that the doctor has a choice.

As a result, the Genetic Algorithm has been got, where F_{ac} is used as a fitness function, derived from the AHP ideology. The idea of using the convolution function (obtained by the MCDM method) as a fitness function of the Genetic Algorithm is not new. In [56], the authors used Weight Sum Approach and Tchebycheff Approach to get the convolution function. The authors of [57] were comparing Non-Dominated Sorting Genetic Algorithm II (NSGA-II), Multi-Objective Differential Evolution (MODE), and Multi-Objective Particle Swarm Optimization (MOPSO) algorithms. Such approaches are rather difficult to interpret, which makes it more complex to explain to the doctor the principle of the algorithm for finding the optimal clinical treatment. Therefore AHP was chosen to obtain a convolution function.

5. Statement of Findings

5.1. Description of Clinical Data

To test the performance of the algorithm, 2 clinical databases of patients with congenital heart defects [58] were used, which were provided by Amosov National Institute of Cardiovascular Surgery [59].

The first database ("DB1") has 128 patients from 3 to 28 years. They underwent a total cavopulmonary connection (TCPC) in an extracardiac conduit modification as the final stage of hemodynamic correction between January 2005 and September 2016. Patients were treated in two phases: surgical treatment (various types of surgery were performed, including TCPC) and conservative treatment (use of medication). Only conservative treatment is considered for the research. With that in mind, the database has the following variables:

- **7** patient indicators before conservative treatment – the vector X.
- **22** types of drugs (that were used to treat patients) – the vector I.
- **38** patient indicators after treatment – the vector Y.

It is worth mentioning that patient indicators before treatment were selected with the help of doctors specifically for the research.

The second database ("DB2") has 144 patients from 1 to 18 years. As in the first case, patients were treated in two stages (the only difference was the methodology). Variables of this database:

- **10** patient indicators before conservative treatment.
- **10** types of drugs (that were used to treat patients).
- **9** patient indicators after treatment.

It is also worth emphasizing that the variables' names are not given so that people cannot use this research for self-medication.

5.2. Modeling the Patient Final State

As mentioned earlier, to perform an optimal clinical treatment selection, it is necessary to obtain models of the patient final state parameters (1). Both patient indicators before and after conservative treatment are either quantitative or qualitative features. Regression methods can be used to model quantitative features. However, doctors are not so much interested in what a particular value will equal a feature whether it will be in the normal range. Therefore, the authors of this research proposed another unique approach, namely, the binarization of quantitative features.

In this way, it is necessary to get models of the binary features (0 – patient indicator after treatment is normal, 1 – patient indicator is abnormal). Patient indicators before treatment and drugs indicators will be used as models' predictors. That requires the use of classification methods. The following algorithms were chosen:

- Linear Discriminant Analysis (LDA) [60].
- Logistic Regression [29].
- Naïve Bayes [40].
- Linear SVM [41].
- SVM with Radial Basis Function (RBF) kernel [41].
- Gaussian Process Classifier (GPC) [61].
- Random Forest Classifier (RFC) [62].
- Adaptive Boosting (AdaBoost) [63].
- Multilayer Perceptron (MLP) [64].

All these algorithms were implemented using the *Python* programming language. Models were built for all final patient parameters in the "DB1" (38 indicators) and "DB2" (9 indicators) databases. To evaluate the models more adequately, the total data samples were split into training (80%) and test samples (20%). The models were evaluated according to their:

- accuracy (percentage of correctly classified objects):

$$\frac{TP+TN}{TP+FP+FN+TN} \quad (3)$$

- sensitivity (share of correctly classified objects of the first class):

$$\frac{TP}{TP+FP} \quad (4)$$

- specificity (share of correctly classified objects of the second class):

$$\frac{TN}{TN+FN} \quad (5)$$

- Matthews Correlation Coefficient (MCC):

$$\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

where: TP – true positives; FP – false positives (type I error); FN – false negatives (type II error); TN – true negatives.

The last metric (6) is a measure of binary classification quality. Its peculiarity is to consider positive and negative results, both true and false. MCC is a balanced measure that is used even for unbalanced classification. It is a correlation coefficient between real and predicted objects: it returns a value from -1 (complete mismatch) to 1 (perfect match). At 0, the classifier is considered to have made the prediction "by chance".

Tables 2 and 3 show results of the classification, namely, the average values of the model classification metrics for each algorithm. As seen from the tables below, the RBF SVM classification algorithm performed best, showing an average model accuracy of around 100% on the test sample.

The resulting models are mathematical equations for patient indicators after treatment (1), which can be substituted in formula (2) to get F_{ac}. This allows using the Genetic Algorithm to derive the best clinical treatment options for patients. These options will be personalized as each of the models is substituted for patient indicators before treatment.

Table 2: Final State Indicators Classification Results ("DB1")

Classification algorithm	Accuracy	Sensitivity	Specificity	MCC
Training sample (80%)				
LDA	89.6%	0.89	0.7	0.611
Logistic Regression	73.7%	0.742	0.779	0.321
Naïve Bayes	74.4%	0.807	0.758	0.429
Linear SVM	79.8%	0.801	0.838	0.474
RBF SVM	100%	1	1	1
GPC	98.6%	0.949	0.715	0.679
RFC	99.9%	1	0.996	0.998
AdaBoost	98.7%	0.985	0.98	0.97
MLP	87.1%	0.862	0.391	0.323
Test sample (20%)				
LDA	87.7	0.859	0.673	0.508
Logistic Regression	71.6	0.722	0.717	0.288
Naïve Bayes	75.3	0.811	0.74	0.412
Linear SVM	77.1	0.77	0.82	0.431
RBF SVM	99.2	0.987	0.99	0.982
GPC	97.6	0.936	0.728	0.661
RFC	99.1	0.995	0.972	0.975
AdaBoost	96.4	0.957	0.95	0.918
MLP	86.3	0.856	0.413	0.314

Table 3: Final State Indicators Classification Results ("DB2")

Classification algorithm	Accuracy	Sensitivity	Specificity	MCC
Training sample (80%)				
LDA	79.6	0.639	0.713	0.413
Logistic Regression	65.6	0.656	0.665	0.242
Naïve Bayes	54.5	0.321	0.901	0.215
Linear SVM	68.6	0.699	0.691	0.307
RBF SVM	100	1	1	1
GPC	100	1	1	1
RFC	99.9	0.999	0.996	0.997
AdaBoost	95.1	0.917	0.972	0.891
MLP	90.1	0.892	0.604	0.511
Test sample (20%)				
LDA	76.4	0.6	0.646	0.267
Logistic Regression	63	0.625	0.616	0.192
Naïve Bayes	55.4	0.364	0.843	0.172
Linear SVM	64.1	0.644	0.618	0.204
RBF SVM	99.6	1	0.989	0.992
GPC	98.5	1	0.972	0.967
RFC	99	0.991	0.971	0.973
AdaBoost	89.8	0.852	0.886	0.748
MLP	88.8	0.846	0.598	0.46

6. Conclusions and Future Work

The research described the development of an algorithm for personalized clinical treatment selection by using the principles of the Genetic Algorithm (for quick treatment variant searching) and the Analytic Hierarchy Process (for patient final state indicators simultaneous optimization). It was detailed from start to finish how the algorithm performs the selection of optimal treatment, including such steps as binarizing the quantitative features of the patient after treatment, and further modeling them with different classification algorithms (in the conference paper [1] only Group Method of Data Handling algorithm was used for modeling). A comparative analysis of classification algorithms showed that the best option for obtaining patient indicators after treatment models is the Support Vector Machine classifier with Radial Basis Function kernel.

The resulting classification models are substituted into the function of additive convolution formula (obtained by Analytic Hierarchy Process), which is used as an optimization function that estimates the final state of the patient. The values of this function range from 0 to 1, and the higher the function value, the better the patient's state. Such a feature could be used as a state-of-the-art metric for patient assessment.

This paper can be described as the beginning of creating a decision support system for personalized clinical treatment selection in Ukraine. It is necessary to carefully elaborate on all stages of the system to provide effective support for the doctor in deciding on a clinical treatment strategy. In this way, the system will be able to fill the absence of clinical pharmacists and optimize the work of medical institutions.

Also, despite the excellent results that were obtained, there were a few data to complete a full study. The problem is that Ukrainian medical institutions do not yet have much confidence in such Artificial Intelligence and Machine Learning methods, and with the medical liability legislation few institutions will provide data for comprehensive research and implementation. The decision support system requires a considerable number of resources and finances, which will be reviewed by authors for the future development of Medicine and Healthcare in Ukraine.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to express their gratitude to the experts from Amosov National Institute of Cardiovascular Surgery who provided clinical data and assist with the research.

References

[1] V. Babenko, O. Nosovets, "Calculating the Personalized Treatment Strategy by Genetic Algorithm Using Optimal Complexity Models," in 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), 1-4, 2020, doi: 10.1109/CSIT49958.2020.9321947.

[2] "Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013," **385**(9963), 117-171, 2015, doi: 10.1016/s0140-6736(14)61682-2.

[3] S. Holm, "Final responsibility for treatment choice: the proper role of medical doctors?," Health Expectations, **14**(2), 201-209, 2011, doi: 10.1111/j.1369-7625.2011.00673.x.

[4] H. Mousa, "Prevention and Treatment of Influenza, Influenza-Like Illness, and Common Cold by Herbal, Complementary, and Natural Therapies," Journal Of Evidence-Based Complementary & Alternative Medicine, **22**(1), 166-174, 2016, doi: 10.1177/2156587216641831.

[5] M. Arruebo, N. Vilaboa, B. Sáez-Gutierrez, J. Lambea, A. Tres, M. Valladares, Á. González-Fernández, "Assessment of the Evolution of Cancer Treatment Therapies," Cancers, **3**(3), 3279-333, 2011, doi: 10.3390/cancers3033279.

[6] C. Pucci, C. Martinelli, G. Ciofani, "Innovative approaches for cancer treatment: current perspectives and new challenges," Ecancermedicalscience, **13**, 2019, doi: 10.3332/ecancer.2019.961.

[7] E. Crimini, M. Repetto, P. Aftimos, A. Botticelli, P. Marchetti, G. Curigliano, "Precision medicine in breast cancer: From clinical trials to clinical practice," Cancer Treatment Reviews, **98**, 102223, 2021, doi: 10.1016/j.ctrv.2021.102223.

[8] E. Chow, J. Doyle, T. Uyeki, "Influenza virus-related critical illness: prevention, diagnosis, treatment," Critical Care, **23**(1), 2019, doi: 10.1186/s13054-019-2491-9.

[9] D. Hall, A. Prochazka, A. Fink, "Informed consent for clinical treatment," Canadian Medical Association Journal, **184**(5), 533-540, 2012, doi: 10.1503/cmaj.112120.

[10] X. Lu, Z. Huang, H. Duan, "Supporting adaptive clinical treatment processes through recommendations," Computer Methods And Programs In Biomedicine, **107**(3), 413-424, 2012, doi: 10.1016/j.cmpb.2010.12.005.

[11] Z. Huang, W. Dong, L. Ji, H. Duan, "Outcome Prediction in Clinical Treatment Processes," Journal Of Medical Systems, **40**(1), 2015, doi: 10.1007/s10916-015-0380-6.

[12] S. Woolf, R. Grol, A. Hutchinson, M. Eccles, J. Grimshaw, "Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines," BMJ, **318**(7182), 527-530, 1999, doi: 10.1136/bmj.318.7182.527.

[13] L. Hughes, M. McMurdo, B. Guthrie, "Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity," Age And Ageing, **42**(1), 62-69, 2012, doi: 10.1093/ageing/afs100.

[14] R. Rosenfeld, J. Shin, S. Schwartz, R. Coggins, L. Gagnon, J. Hackell et al., "Clinical Practice Guideline," Otolaryngology–Head And Neck Surgery, **154**(2), 201-214, 2016, doi: 10.1177/0194599815624407.

[15] L. Duff, E. Mclnnes, N. Cullum, A. Nelson, K. Luker, "Clinical guidelines," Primary Health Care, **9**(1), 28-30, 1999, doi: 10.7748/phc.9.1.28.s14.

[16] "Clinical Guidelines". Physiopedia. (2012). Retrieved 24 March 2021, from https://bit.ly/3hPPpFY.

[17] "National Clinical Effectiveness Committee Standards for Clinical Practical Guidance". Nursing and Midwifery Board of Ireland. (2015). Retrieved 24 March 2021, from https://bit.ly/3jY921d.

[18] "Clinical protocols" (in Ukrainian). Ministry of Healthcare of Ukraine. (2017). Retrieved 23 December 2020, from https://bit.ly/2T4LLj8.

[19] "Клінічні протоколи: що це і чи є вони в Україні" (in Ukrainian). Ministry of Healthcare of Ukraine. (2017). Retrieved 23 December 2020, from https://bit.ly/2T26BzF.

[20] J. Saseen, T. Ripley, D. Bondi, J. Burke, L. Cohen, S. McBane et al., "ACCP Clinical Pharmacist Competencies," Pharmacotherapy: The Journal Of Human Pharmacology And Drug Therapy, **37**(5), 630-636, 2017, doi: 10.1002/phar.1923.

[21] F. Khan, N. Waqas, A. Ihsan, P. Khongorzul, J. Wazir, W. Gang et al., "Analysis of the Qualities Matching New Classification of Clinical Pharmacist," Indian Journal Of Pharmaceutical Sciences, **81**(1), 2019, doi: 10.4172/pharmaceutical-sciences.1000473.

[22] "Clinical Pharmacy (Educational Programme)". National University of Pharmacy. (2015). Retrieved 15 October 2020, from https://bit.ly/3hZ0h46.

[23] "Ukraine to receive UAH 128 million for digitalization of services provided by Digital Transformation Ministry". UNDP. (2020). Retrieved 12 March 2021, from https://bit.ly/3e4czHm.

[24] "Відповідальність медичних працівників" (in Ukrainian). Legislation of Ukraine. (2011). Retrieved 14 March 2021, from https://bit.ly/3AOe9Hg.

[25] Ö. Emre, G. Sert, "European Charter of Patients' Rights," Turkish Journal Of Bioethics, **1**(4), 198-205, 2014, doi: 10.5505/tjob.2014.69775.

[26] "ASA Physical Status Classification System | American Society of Anesthesiologists (ASA)". American Society of Anesthesiologists. (2020). Retrieved 28 November 2020, from https://bit.ly/3xx4DGr.

[27] W. Owens, J. Felts, E. Spitznagel, "ASA Physical Status Classifications," Anesthesiology, **49**(4), 239-243, 1978, doi: 10.1097/00000542-197810000-00003.

[28] J. Havens, A. Columbus, A. Seshadri, C. Brown, G. Tominaga, N. Mowery, M. Crandall, "Risk stratification tools in emergency general surgery,"

Trauma Surgery & Acute Care Open, **3**(1), e000160, 2018, doi: 10.1136/tsaco-2017-000160.
[29] S. Sperandei, "Understanding logistic regression analysis," Biochemia Medica, 12-18, 2014, doi: 10.11613/bm.2014.003.
[30] L. Minne, A. Abu-Hanna, E. de Jonge, "Evaluation of SOFA-based models for predicting mortality in the ICU: A systematic review," Critical Care, **12**(6), R161, 2009, doi: 10.1186/cc7160.
[31] J. Muschelli, "ROC and AUC with a Binary Predictor: a Potentially Misleading Metric," Journal Of Classification, **37**(3), 696-708, 2019, doi: 10.1007/s00357-019-09345-1.
[32] C. Bulac, A. Bulac, "Decision Trees," Advanced Solutions In Power Systems: HVDC, FACTS, And Artificial Intelligence, 819-844, 2016, doi: 10.1002/9781119175391.ch18.
[33] P. Austin, "A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality," Statistics In Medicine, **26**(15), 2937-2957, 2007, doi: 10.1002/sim.2770.
[34] O. Takahashi, E. Cook, T. Nakamura, J. Saito, F. Ikawa, T. Fukui, "Risk stratification for in-hospital mortality in spontaneous intracerebral haemorrhage: A Classification and Regression Tree Analysis," QJM, **99**(11), 743-750, 2006, doi: 10.1093/qjmed/hcl107.
[35] A. Rovlias, S. Kotsou, "Classification and Regression Tree for Prediction of Outcome after Severe Head Injury Using Simple Clinical and Laboratory Variables," Journal Of Neurotrauma, **21**(7), 886-893, 2004, doi: 10.1089/0897715041526249.
[36] G. Dolce, M. Quintieri, S. Serra, V. Lagani, L. Pignolo, "Clinical signs and early prognosis in vegetative state: A decisional tree, data-mining study," Brain Injury, **22**(7-8), 617-623, 2008, doi: 10.1080/02699050802132503.
[37] A. Abu-Hanna, N. de Keizer, "Integrating classification trees with local logistic regression in Intensive Care prognosis," Artificial Intelligence In Medicine, **29**(1-2), 5-23, 2003, doi: 10.1016/s0933-3657(03)00047-2.
[38] L. Gortzis, F. Sakellaropoulos, I. Ilias, K. Stamoulis, I. Dimopoulou, "Predicting ICU survival: A meta-level approach," BMC Health Services Research, **8**(1), 2008, doi: 10.1186/1472-6963-8-157.
[39] J. Trujillano, M. Badia, L. Serviá, J. March, A. Rodriguez-Pozo, "Stratification of the severity of critically ill patients with classification trees," BMC Medical Research Methodology, **9**(1), 2009, doi: 10.1186/1471-2288-9-83.
[40] V. Chauhan, K. Dahiya, A. Sharma, "Problem formulations and solvers in linear SVM: a review," Artificial Intelligence Review, **52**(2), 803-855, 2018, doi: 10.1007/s10462-018-9614-6.
[41] T. Fragoso, W. Bertoli, F. Louzada, "Bayesian Model Averaging: A Systematic Review and Conceptual Classification," International Statistical Review, **86**(1), 1-28, 2017, doi: 10.1111/insr.12243.
[42] R. Houthooft, J. Ruyssinck, J. van der Herten, S. Stijven, I. Couckuyt, B. Gadeyne et al., "Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores," Artificial Intelligence In Medicine, **63**(3), 191-207, 2015, doi: 10.1016/j.artmed.2014.12.009.
[43] Moore, D., Keegan, T., Dunleavy, L., & Froggatt, K. (2019). Factors associated with length of stay in care homes: a systematic review of international literature. Systematic Reviews, **8**(1). doi: 10.1186/s13643-019-0973-0
[44] P. Gago, M. Santos, A. Silva, P. Cortez, J. Neves, L. Gomes, "INTCare: a Knowledge Discovery Based Intelligent Decision Support System for Intensive Care Medicine," Journal Of Decision Systems, **14**(3), 241-259, 2005, doi: 10.3166/jds.14.241-259.
[45] N. Kaieski, C. da Costa, R. da Rosa Righi, P. Lora, B. Eskofier, "Application of artificial intelligence methods in vital signs analysis of hospitalized patients: A systematic literature review," Applied Soft Computing, **96**, 106612, 2020, doi: 10.1016/j.asoc.2020.106612.
[46] D. da Silva, D. Schmidt, C. da Costa, R. da Rosa Righi, B. Eskofier, "DeepSigns: A predictive model based on Deep Learning for the early detection of patient health deterioration," Expert Systems With Applications, **165**, 113905, 2021, doi: 10.1016/j.eswa.2020.113905.
[47] I. Nastenko, V. Pavlov, O. Nosovets, K. Zelensky, O. Davidko, O. Pavlov, "Solving the Individual Control Strategy Tasks Using the Optimal Complexity Models Built on the Class of Similar Objects," Advances in Intelligent Systems and Computing IV, **1080**, 535-546, 2020, doi: 10.1007/978-3-030-33695-0_36.
[48] V. Vaishnav, J. Vajpai, "Assessment of impact of relaxation in lockdown and forecast of preparation for combating COVID-19 pandemic in India using Group Method of Data Handling," Chaos, Solitons & Fractals, **140**, 110191, 2020, doi: 10.1016/j.chaos.2020.110191.
[49] R. Vanderbei, "The Simplex Method," International Series In Operations Research & Management Science, 11-25, 2020, doi: 10.1007/978-3-030-39415-8_2.
[50] S. Panjwani, S. Naresh Kumar, L. Ahuja, "Multi-criteria decision making and its applications," International Journal of Innovative Technology Exploring Engeering **8**(9 Special Issue 4), 2019, doi: 10.35940/ijitee.I1122.0789S419.
[51] T. L. Saaty, "Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World," RWS Publications, 1990, doi: 10.1016/0377-2217(89)90066-0.
[52] T. L. Saaty, "THE ANALYTIC HIERARCHY PROCESS WITHOUT THE THEORY OF OSKAR PERRON," International Journal Of The Analytic Hierarchy Process, **5**(2), 2014, doi: 10.13033/ijahp.v5i2.191.
[53] D. Whitley, "A genetic algorithm tutorial," Statistics And Computing, **4**(2), 1994, doi: 10.1007/bf00175354.
[54] J. García, C. Acosta, M. Mesa, "Genetic algorithms for mathematical optimization," Journal of Physics: Conference Series, 5-5, 2020, doi: 10.1088/1742-6596/1448/1/012020.
[55] S. Katoch, S. Chauhan, V. Kumar, "A review on genetic algorithm: past, present, and future," Multimedia Tools And Applications, **80**(5), 8091-8126, 2020, doi: 10.1007/s11042-020-10139-6.
[56] Q. Zhang, H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," IEEE Transactions on Evolutionary Computation, **11**(6), 712–731, 2007, doi: 10.1109/TEVC.2007.892759.
[57] H. Monsef, M. Naghashzadegan, A. Jamali, R. Farmani, "Comparison of evolutionary multi objective optimization algorithms in optimum design of water distribution network," Ain Shams Engineering Journal, **10**(1),103–111, 2019, doi: 10.1016/j.asej.2018.04.003.
[58] A. Dydyk, O. Nosovets, V. Babenko, "Setting Up the Genetic Algorithm for the Individualized Treatment Strategy Searching," Herald of Advanced Information Technology, **3**(3), 125–135, 2020, doi: 10.15276/hait.03.2020.2.
[59] "Amosov National Institute of Cardiovascular Surgery". Retrieved 1 October 2018, from https://bit.ly/2T9vWrq.
[60] D. HU, X. LI, F. NIE, "Deep linear discriminant analysis hashing," SCIENTIA SINICA Informationis, **51**(2), 279-279, 2021, doi: 10.1360/ssi-2019-0175.
[61] D. Mackay, M. Gibbs, "Variational Gaussian process classifiers. IEEE Transactions On Neural Networks, **11**(6), 1458-1464, 2000, doi: 10.1109/72.883477.
[62] C. Mantas, J. Castellano, S. Moral-García, J. Abellán, "A comparison of random forest based algorithms: random credal random forest versus oblique random forest," Soft Computing, **23**(21), 10739-10754, 2018, doi: 10.1007/s00500-018-3628-5.
[63] D. Feng, Z. Liu, X. Wang, Y. Chen, J. Chang, D. Wei, Z. Jiang, "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Construction And Building Materials, **230**, 117000, 2020, doi: 10.1016/j.conbuildmat.2019.117000.
[64] J. Tang, C. Deng, G. Huang, "Extreme Learning Machine for Multilayer Perceptron," IEEE Transactions On Neural Networks And Learning Systems, **27**(4), 809-821, 2016, doi: 10.1109/tnnls.2015.2424995.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 409-417 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Factors Affecting the Decision of Selecting Banking to Save Money of Individual Customers - Experimental in Da Nang City

Le Anh Tuan[1,2,*], Mai Thi Quynh Nhu[1,2], Nguyen Le Nhan[1,2]

[1]*Faculty of Accounting, Duy Tan University, Da Nang, 550000, Vietnam*

[2]*Institute of Research and Development, Duy Tan University, Da Nang, 550000, Vietnam*

ARTICLE INFO

Article history:
Received: 08 March, 2021
Accepted: 09 May, 2021
Online: 27 June, 2021

Keywords:
Savings deposits
Individual customers
Da Nang

ABSTRACT

The purpose of this study is to evaluate the factors affecting the decisions of factors affecting the decision of individual customers to choose a savings bank at commercial banks in Da Nang city. . Through the use of appropriate research methods, the authors have found that there are 5 factors with 5 groups of factors that greatly affect the decision of individual customers to choose a savings bank. Service quality, safety, stakeholder influence, financial benefit, convenience. Through this result, it will help opinions, orientations and solutions to improve individual customers' decisions on choosing a savings bank at joint stock commercial banks in Da Nang city.

1. Introduction

In today's competitive open economy, commercial banks always find ways to use many attractive and novel savings products to attract customers to their banks. On the part of customers using the service, they are also interested in factors related to the benefits of saving money at banks. This is also something that bank managers are interested in understanding, so identifying the factors that affect individual customers' choice of banks when saving money is a necessary job to help. Banks promote their strengths, enhance their image in the eyes of customers to attract more idle money from the population, contributing to improving business efficiency.

In 2018-2019, the macroeconomic situation and Da Nang continue to face many challenges. However, with the efforts of the whole political system, Da Nang has achieved many successes in the realization of local socio-economic development goals; in which, the banking sector has made an important contribution to the socio-economic development of Da Nang. Through general assessment of the banking performance of the city. In Da Nang in 2018-19, it is remarkable that the operations of the banking system in the area continue to grow compared to the previous years. By the end of 2018, the total mobilized capital in the province was about VND 125,994 billion, up 9.92% compared to the end of 2017, the highest growth rate in recent years (up 25.84%). In which, 36/57 credit institutions have increased mobilized capital compared to the end of 2017, contributing greatly to the recovery and development of the business community in the city.

Each bank has its own capital mobilization policy, depending on the needs and purposes of its business. To have a large capital source requires commercial banks to have appropriate mobilization policies to attract the necessary amount of capital in the economy to serve their business and development. Although there is a high credit growth rate compared to the whole country (about 14%), the bad debt rate is low at 1.62%. At the same time, at the end of 2018, banks increased deposit and lending rates to stimulate demand, attract capital, and the impact of exchange rates and liquidity pressure. Techcombank raised interest rates for 12-36-month term from 6.4% a year to 6.8%. VIB applies 7.3% for the 24-month term. ACB also applies a new interest rate schedule for deposits with terms over 18 months, ranging from 7-7.2% a year, 0.1-0.3% higher than May.

2. Background Study

Up to now, there have been many practical studies on the adoption and use of e-banking services, including studies on the factors influencing the decision to use services. Typical electronic banking services. In [1], the author presented that intention is considered to include motivational factors that influence the behavior of each individual. These factors indicate the willingness or effort that each individual will put in the behavior. The above

*Corresponding Author: Le Anh Tuan, Email: latuan0507@gmail.com.

definition explains behavior that is influenced by a person's confidence in his or her ability to perform the behavior. Therefore, behavior is affected by controlling cognitive and intent behaviors as well as attitudes, subjective criteria and perceptions of behavior controlling indirectly through intentions. In [2] the author expand the application of two Mobile banking service by adding the element of trust ("Feeling of trust") and 2 source factors ("Self-grasping ability" and "financial cost perception") into the model, and focus on the position of these factors in the existing structure of the TAM model. . In [3], the author define that behavioral intent is the ability of the consumer to use an innovation. With higher behavioral intentions, consumers will be more likely to use a new technology. There are a number of premises that can affect an individual's behavioral intentions. Applying the UTAUT model to study the behavior of 3G mobile telecommunications users in [4]. The author used the UTAUT model to conduct research surveys. The research results show that the factors that influence the "behavioral intent" include: "expectation of the effect", "social influence" and "favorable conditions", while the factor "expectations" in terms of effort " has no effect. In addition, three unrecognized relationships were discovered during the SEM analysis, modifying the UTAUT model for 3G telecommunications services. The biggest difference between this study and UTAUT theory is the issue of time of study and the significance of the determinant by external variables. Studying the factors affecting the intention to use online shopping services using the UTAUT model in [5]. The study identified 6 factors affecting the intention to use e-shopping services, including: (1) Expectation about price, (2) Perceived convenience, (3) Perception ease of use (4) Feel the pleasure, (5) Social influence, (6) Feel the risk of use. In addition, the model will also consider the influence on the intended use of three demographic variables: gender, income, and age. Inheriting the point of view of the above researchers, in [6], the author describes customers' savings deposit behavior expressed in choosing a bank and choosing a deposit period. Research on factors affecting the acceptability of Mobile Banking of individual customers applying the UTAUT model in [4]. Using the Unified Theory of Technology Acceptance and Use (UTAUT) to investigate the effects on people through Mobile banking, this study concludes that individuals' intentions on mobile services Banking is significantly affected by the following factors: (1) social influence, (2) perceived financial costs, (3) performance expectation, and (4) perceived trust in order of influence. Behaviors are significantly influenced by personal intentions and favorable conditions. This study found that gender controls the effect of performance expectations and financial cost perceptions on behavioral intent, age of adjustment, effects of favorable conditions, and perceived self-grasping. to real acceptance behavior.

Data were collected from 165 questionnaires and used regression to analyze relationships in [7]. The results have shown that all factors, except the "perceived cost" factor, have a significant impact on behavioral intent in using Mobile Banking. "Feeling useful" is the most influential factor in explaining the user's intent to accept. For users, "feel the ease of use" is the most important factor, while "feel efficient" significantly affects the intention of accepting the service. The results from these analyzes help banking organizations to have a suitable strategy for Mobile banking to expand their acceptability of this service. Background theory related to rational action theory (TRA) developed in [8] and in [9] shows the inclusion and coordination of the components of attitudes in a structure. which are designed to better predict and explain consumer behavior in society based on two basic concepts: (i) consumer attitudes toward behavior and (ii) subjective standards of consumers. In addition, the intended behavioral theory (TPB) was developed in [9] by adding a cognitive element of behavior control to the TRA model. According to TPB, human behavior is guided by 3 factors: behavioral beliefs, normative beliefs and controlling beliefs. Thus, beliefs about behavior create an attitude of like or dislike about behavior; normative beliefs create subjective social or normative stress and controlling beliefs increase cognitive behavioral control.

Thus, most studies using TAM model are conducted based on the adoption and use of information / technology systems. In studies, the TAM model was found to be valuable in predicting the adoption of variables in the user system. However, in many areas, as mentioned earlier, the original TAM model does not fully explain behavioral intent towards technology adoption and adoption, and there is a need for research to add factors to improve predictability for this model. Up to now, there have been many studies on individual customer banking choices conducted in many different countries around the world, with many different survey subjects. This is a rich source of documents, the basis of orientation to conduct research on factors affecting the choice of banks of individual customers at joint stock commercial banks. Not to mention the research on the earliest bank selection criteria in the US, in [10], the author showed that the recommendation of friends is the most important factor, followed by reputation, reputation. of the bank and large loan capital. In the American university student article, in [11], the author presented that convenience is still the main reason most students choose their bank. In addition, family tradition and loyalty to the bank are also important factors for college students. In [12], the following factors are identified as the decisive attributes affecting the choice of banks: service fee collection policy, reputation, loan interest rate, disbursement time and friendly bank staff. In [13], it is reported that economic factors such as monthly costs and deposit interest play a decisive role in student's choice of bank. In [13], The author found fast and efficient service delivery, friendly, enthusiastic staff and a bank reputation are important factors in choosing select the student's bank. In [14], the author showed that the key factors influencing customer choice in Greece include: convenience, reputation of a bank, quality of products / services. , interest rates and fees, staff communication skills and qualifications, facilities, the environment of the points of sale, and satisfaction with after-sales service and services.

There have been many empirical studies at home and abroad on the factors affecting the decisions of individual customers to deposit money at banks in [15]-[19].The important factors affecting the deposit behavior of customers are financial interests, brand awareness, influence of relatives, fast and safe, marketing, service style of employees, delivery Convenient translation and safe feeling. In [20], the banking selection criteria of Business Administration students in Delhi, India concluded that convenience is an important determinant, Includes: parking, free delivery on demand, phonebanking, free home money transfer service because customers want to save time. In [21], the author conducted in Malaysia found three factors that most influence a

student's banking choice: sense of security, ATM service, and financial benefits. In [16], the author conducted a similar study to determine the factors influencing the choice of commercial banking of university students in South Africa. The study was based on a sample of 186 students from Fort Hare University (Alice campus). The researchers found that the important factors that determine a college student's choice of banking (ranked in diminishing importance) are: service, convenient location, attractiveness, gender referrals of others, marketing activities and prices. In [22], the author studied the factors affecting customer satisfaction in the banking industry in Pakistan to find the relationship between service quality, satisfaction and loyalty. customers. The research results show that 54% change in customer loyalty is due to service quality and customer satisfaction. The study concludes that, service quality affects customer satisfaction and customer satisfaction affects customer loyalty. In today's competitive marketplace, banks can gain a competitive edge by providing quality services that meet the needs of their customers.

For studies in this area in Vietnam, there has been a study on factors affecting the choice of banks of individual customers in [23] surveyed above. 350 individual customers in Da Lat city concluded that brand awareness has the strongest impact on the trend of choosing a bank, followed by factors: Convenience in terms of location, troubleshooting, photos enjoyment of a loved one, appearance and attitude toward marketing. Meanwhile, in [24], the author surveyed 200 customers using banking products and services showed that there is a positive relationship between customer satisfaction and reliability factors. feedback, accessibility, service capacity and quality of products and services. No correlation has been found between information and customer satisfaction. Based on the research results, a few recommendations have been proposed to improve the satisfaction of individual customers with Vietinbank East Hanoi branch in particular and other commercial banks in general. In [25], the author presented on the factors affecting the choice of savings banks of individual customers in Ho Chi Minh City, researched and surveyed over 500 customers. In [25], the author focused on testing the factors of service quality affecting satisfaction and loyalty. In which, testing a number of impacts that regulate the relationship of satisfaction - loyalty of customers for banking services in Hau Giang. Based on a survey sample from 333 customers, Cronbach's Alpha coefficients, exploratory factor analysis, confirmatory factor analysis and structural equation model were used for analysis. The results show that empathy and price are the factors that strongly impact on customer satisfaction. In [26], the author studied the factors affecting the satisfaction of customers with the savings deposit service at Agribank Binh Minh, Vinh Long. Research results show that, 4 main factors affecting customer satisfaction including responsiveness, facilities, service capacity and peace of mind. The relationship between the service quality and the satisfaction of customers participating in the transaction is very important to the operation of the retail bank.

The results of this study have identified 8 groups of factors that affect individual customers' banking choices, including: financial benefits, sense of security, referral, troubleshooting, service. Bank, brand, staff, convenience. At the same time also evaluate their importance. This is the guiding basis for commercial banks to take action measures that focus on key factors to influence customers' choice of banks and thereby attract more. customers using their bank's retail products and services. Based on the research results, the author gives some suggestions to help banks maintain existing customers and attract new customers. From there, the leaders of banks can refer to make appropriate policies to increase the competitiveness of banks in the retail market. In [27], the author showed on the factors affecting the decision of individual customers to deposit savings - Research at Joint Stock Commercial Bank for Industry and Trade of Vietnam, Lam Dong province, based on into data collected from 184 observations processed by SPSS tool, with Cronbachs Alpha scale reliability testing techniques, factor analysis (EFA), correlation analysis and multiple regression analysis . Research results have identified 5 factors that directly affect individual customers' decisions to save money, including: brand awareness, influence of relatives, financial benefits, fast, safe and marketing. Based on the research results, governance implications have been proposed to better meet customer needs and expand markets to attract new customers. In [23], the author studied the factors affecting individual customer satisfaction with deposit products at Lien Viet Post Joint Stock Commercial Bank - Soc Trang Branch. Research results have found out 5 factors affecting the satisfaction of customers who save money at this bank are: Reliability; Tangible Media; guarantee; sympathy; Transaction office network. In which, tangible means, assurance, sympathy, and transaction office network are the factors that strongly affect the satisfaction of savings customers. As can be seen, the factor that is considered important by the customer when choosing banks varies from country to country and cultures due to differences in economic, cultural and banking systems.

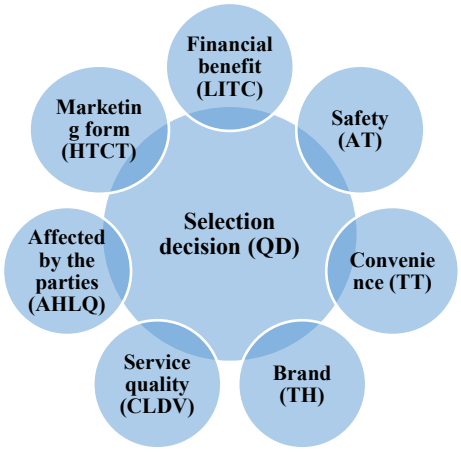


Figure 1: Proposed research model
(Source: Author compiled, 2020)

3. Research Methods

After consulting domestic and foreign research models, survey experts with the list and content as outlined in the discussion paper. The authors have formed a preliminary scale of factors affecting the decision of individual customers to choose a savings bank at commercial banks in Da Nang city to be used when surveying ideas. the experts. After synthesizing the answers of experts about the factors in the preliminary scale. The results of qualitative expert survey showed a high rate of agreement (over 80%) on 7 groups of factors and experts did not add any new groups of factors. After collecting information from the most knowledgeable and experienced staff about deposit services and customers who are conducting transactions at joint stock commercial banks who

have sent savings and intend to save money in the near future. With the model study on an overview of the research situation related to the factors that influence the decision to choose a bank to deposit a savings, we expect an analytical framework with 5 common factors affecting decide to choose a bank to save individual customers as follows:

Based on the research objectives, the survey and research model of the proposed factors influencing the decision to buy Korean facial care products by female university students in Da Nang city, based on the background theory presented above, the research hypotheses are determined as follows:

Hypothesis H1: Financial benefits (LITC) and decision to choose a savings bank of individual customers have positive relationship.

Financial benefits are assessed through two criteria: High saving interest rate and reasonable service fee. Savings deposit interest rate is an indispensable question of customers when looking for information about savings deposit service at a certain bank. Interest is the rate of interest a customer receives in addition to the principal, after a certain period of time at the bank. The bank uses a currency that customers commit to deposit with a fixed term to do business and seek profits, and pay interest on that deposit currency. Interest can be paid at the beginning of the term, paid periodically, or at the end of the term. In the banking sector in Vietnam, competition by savings interest rates is a traditional competitive measure to attract deposits. Interest rate is also the basic factor when customers choose a bank to save money.

Service fee is the cost that a customer has to pay to use banking products and services. When an individual customer comes to the bank not only to save money but also transact in many other areas such as proof of finance, transfer ... - maybe equal to the amount of money that has been saved. If products and services do not have big differences in characteristics and benefits, customers tend to choose products and services with lower costs. Therefore, service costs may also be a concern for customers when choosing a bank to save money.

Hypothesis H2: The safety (AT) and decision of individual customers to choose a savings bank have a positive relationship.

Transactions at banks are often sensitive because they are directly related to customers' financial resources, so a sense of security is a factor worth considering. In [15], the author pointed out that a sense of security reflects a desire of banking users to be stable and secure when making financial transactions. This means a sense of security that includes both security at banks and bank financial stability. Therefore, there are three criteria: Secure customer information, Strong banking financial foundation, Security conditions of transaction points are used to assess the sense of security of customers about the performing bank.

Hypothesis H3: Convenience (TT) and the individual customer's decision to choose a savings bank has a positive relationship.

Another characteristic of individual customers when conducting financial transactions is that they prioritize services that are provided quickly and can be done anywhere, at any time. This convenience is measured through 7 criteria: the large network of transaction points, the location of convenient transaction points, having a transaction point near home / workplace, suitable bank transaction time, there is ample parking, e-banking available and free home deposit service. The network and location of transaction points provide customers with convenient mobility. Customers can deposit money in one city and withdraw money, transfer money ... in another city. In other words, individuals can perform transactions with the bank at any point in the system thanks to the development of information technology in the banking sector without spending much time and effort on going. again. Often times, the transaction points near home and work are arguably the most attractive. In addition, the spacious parking not only brings a sense of safety but also helps customers more convenient when dealing directly with the bank. Another factor that is considered is the bank's working time in accordance with the individual transaction schedule. Clients with flexible hours of work may not appreciate the fact that banks are open for transactions outside of office hours. On the contrary, for civil servants and office workers - who are often forced by strict working hours, the bank's acceptance of overtime will help them a lot in transactions.

Nowadays, when science and technology are developing rapidly, electronic transactions are expected to be one of the factors attracting customers for convenience. Individuals can perform transactions with the bank whether at home or on the street, with only one device with network connection such as laptop, desk phone, mobile phone ... Instead of having to directly go to bank transactions. Customers, individual customers can transfer, pay electricity and water bills, pay credit card bills or even save money, pay off loans, manage accounts at home, work with just few simple steps. Some other utilities such as free home deposit service should also be deployed for maximum convenience for customers.

Hypothesis H4: Brand (TH) and individual customers' decision to choose a savings bank has a positive relationship.

Combining a number of studies and factors Brand awareness is evaluated through the following observed variables: Recognition of the bank's brand name, logo, image, signature music, Community programs, activities Big sponsorship that the bank participates in and Promotions. In [28], the author showed that Brand awareness has a positive correlation with Brand Desire. Brand awareness is the first component of an attitude, an emotion. Consumers have a certain emotion towards a brand that is first and foremost about being aware of that brand among competing brands. Brand desire consists of two components: preference and preference. A consumer's interest in a brand is expressed through an emotion of interest. And when it comes to choosing between competing, interchangeable brands, consumers tend to choose the brand that causes the most emotion. The authors in [23] have shown that Brand awareness is a factor affecting the trend of choosing a bank, but the awareness aspect is only two points of view “brand is a system. identity "(AMA, 1960) and" brand is symbolic "(in [29]& In [30]), the level of brand awareness is limited only to recognizing the name, logo, image, music sign. characteristics of the brand, advertising and promotion programs of the bank ...Community programs, big sponsorships that the bank participates in and special, unique and attractive promotions contribute to helping consumers to identify banks among competing banks.

Hypothesis H5: Service quality (CLDV) and decision to choose a savings bank of individual customers have positive relationships.

The influence of the factor Products and services of the bank are assessed by easily opening a deposit account or a savings account, providing the service quickly and effectively to help customers not lose much. waiting in line for transactions and products and services are diversified and plentiful, meeting the different needs of consumers. In the era of industrialization and modernization, saving time is one of the criteria that customers are aiming for. Individual customers who come to the savings bank do not want to spend a lot of time on paperwork and procedures with a series of cumbersome and complicated processes. Therefore, the element of ease of opening an account and providing services quickly and efficiently can affect the choice of a bank.

Besides, as mentioned above, the service needs of individual customers are often diverse. A customer can both have a need for both short-term savings and long-term savings. In addition, customers come to certain banks first to save money, but later can use many other services such as wire transfer, payment for electricity and water bills, receiving remittances, borrowing money ... and vice versa. Therefore, diversified and diversified products and services with clear and accurate information are a factor worth considering when using a savings service at a bank.

Hypothesis H6: Related person influence (AHLQ) and decision to choose a savings bank of an individual customer with a positive relationship.

In [31], when people start to want to save money, they collect information about the brand and savings deposit products. of many different banks, thereby making decisions on choosing a bank to save money. Information about brands and product characteristics can be based on their own experiences, or on the recommendations or judgments of others. Therefore, the referral of a third party can be one of the factors influencing the decision-making process of choosing a bank to deposit a customer's savings. The recommendation of family members, friends and the bank staff who are transacting are 3 observed variables used to evaluate the Referral factor. Friends and relatives who have been using the bank's savings deposit service are a useful basis for new customers to choose which bank to save, based on practical experience, understanding and feeling.

In addition, the bank's employees who are transacting are also a valuable source of reference for customers. Usually, individual customers come to the bank not only to use the savings deposit service. They may have made transfers, withdraw money remittances, pay for goods ... before and have a rough feeling about the bank is transacting. At that time, the bank's staff was the bridge leading customers to their bank's savings service.

Hypothesis H7: The form of marketing promotion (HTCT) and the decision to choose a savings bank of individual customers have a positive relationship.

The role of marketing activities in the banking business is not much different from other business areas in conveying information from banks to customers and vice versa. In [24], the author has shown that attitude towards marketing has an influence on the brand desires of consumers. If they have a good attitude and are interested in the branding of a product or brand; distinguish it from competing products and brands; and when there is a demand, the ability to choose for that product or brand is very high.

Thus, the above factors have been tested in many previous studies. However, since most studies have been done in foreign countries, in many different territories, at different times, with separate observers, the results need to be checked for conformity. such as adjustments and supplements through pre-use research to determine the factors that actually affect the choice of banks to deposit money of individual customers.

Table 1: Summary of factors in the proposed research model

Factor	Encode	Hypothesis	Related studies
Financial benefits	LITC	+	[21]; [16]
The safety	AT	+	[31]; [14]
Convenience	TT	+	[20];[16];[14];[32]
Brand	TH	+	[23]; [14]
Service quality	CLDV	+	[14]
Related person influence	AHLQ	+	[16]; [32]
The form of marketing promotion	HTCT	+	[14]
Selection decision	QD		[16]; [14]; [32]

(Source: Author compiled, 2020)

3.1. Research model

The multivariate linear regression equation of this study has the form:

Overall regression function:

$QD = \beta 0 + \beta 1LI + \beta 2AT + \beta 3TT + \beta 4TH + \beta 5CLDV + \beta 6AHLQ + \beta 7HTCT + ei$

Inside:

- QD: Decided to choose individual customer's savings bank
- Financial benefits (LI)
- Safety (AT)
- Convenience (TT)
- Brand (TH)
- Quality of service (QoS)
- Affected person effects (AHLQ)
- Marketing form (HTCT)

The regression model will find out the independent factors that affect the dependent factor. At the same time, the model also describes the impact level, thereby helping us predict the value of the dependent factor.

4. Result

On the basis of influencing factors that were discovered in the research phase, the survey questionnaire was deployed to the identified survey subjects in the form of direct to individual customers. After giving out 400 surveys to individual customers in the form of live broadcast. Results of 305 votes were collected, processed again, of which 285 votes were met. Data used in the research of the topic are data obtained directly from the survey tables that have been cleaned to eliminate incomplete or inadequate survey responses to the research requirements.

4.1. Age

The results show that out of a total of 285 survey samples meeting the requirements, there are some couples 18-25 years old accounting for 5.3%, 25-40 years old accounting for 39.3%, the age from 40-60 accounts for 37.2% and Finally, the age over 60 accounts for 18.2%. The analysis results show that people aged 25-60 are often more interested in choosing a bank to save money, this is also a group of people with the ability to generate income and want to preserve their numbers. your money is safe and look for income from your savings.

4.2. Education level

Surveying 285 customers on factors affecting the decision to choose savings banks at commercial banks in Da Nang city, 1.1% of customers have high school education, 61.4% and 36.5% of clients have college / university degrees, 1.1% of customers have graduate degrees.

4.3. Income

Surveying 285 customers about factors affecting the choice of savings banks of individual customers at commercial banks in Da Nang city, customer groups with income from 5-10, customer groups goods with income from 10-50 million accounts and group of over 50 million accounts for the main proportion, this is also potential customers of the bank in mobilizing savings deposits.

4.4. Type of agency

Thus, there are 53% of customers working in state agencies, 18.9% of customers working in joint stock companies, 28.1% of customers working in joint ventures and the rest working in joint ventures. private companies and self-employed. Thus, the group of customers working at state agencies or joint ventures is the group most likely to deposit savings at banks.

After removing the unreliable observed variables, including: TH4, CLDV4, AHLQ2, CT1, CT2, CT3 according to the evaluation results above, re-test the Reliability of the scale with variables. The remaining results show as follows:

Table 2: Results of reliability analysis and factor analysis

No	Group variables		Number of observation variables	Cronbach's Alpha
1	Financial benefits	LITC	3	0.823
2	The safety	AT	4	0.835
3	Convenience	TT	4	0.782
4	Brand	TH	3	0.844
5	Service quality	CLDV	3	0.843
6	Selection decision	QD	3	0.939

(Source: Analysis results from SPSS 16.0 software)

Table 3: KMO coefficients and Bartlett's tests of independent factors

KMO coefficient		.800
Bartlett's test	**Approx. Chi-Square**	4250.315
	Df	171
	Sig.	.000

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
CLDV2	.798				
CLDV1	.789				
CLDV3	.787				
TH2	.646				
TT3	.623				
TH1	.578				
TH3					
AT2		.832			
AT3		.791			
AT4		.679			
AT1		.649			
TT1		.593			
AH3			.827		
AH1			.788		
LITC1				.772	
LITC3				.771	
LITC2				.734	
TT2					.841
TT4					.638

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 7 iterations.

(Source: Analysis results from SPSS 16.0 software)

The multivariate linear regression equation of this study has the form:

Overall regression function:

QD = β0 + β1F1 + β2F2 + β3F3 + β4F4 + β5F5 + β6F6 + Ui

Sample regression function:

QD = β^ 0 + β^ 1F1 + β^ 2F2 + β^ 3F3 + β^ 4F4 + β^ 5F5 + β^ 6F6 + ei

Inside:

QD: - Decided to choose the individual customer's savings bank

- F1 Service quality, includes variables CLDV2, CLDV1, CLDV3, TH2, TT3, TH1.

- F2 Safety, includes variables AT2, AT3, AT4, AT1, TT1.

- F3 Related effects, includes variables AH3, AH1.

- F4 Financial benefit, includes variables LITC1, LITC3, LITC2.

- F5 Convenience, includes variables TT2, TT4.

The regression model will find the independent factors that affect the dependent factor. At the same time, the model also describes the impact level, thereby helping us predict the value of the dependent factor.

The regression results in SPSS are as follows:

Table 4: Regression coefficients

Factor	Unstandardized Coefficients		Standardized Coefficients	T value	Sig.
	B	Standard error	Beta		
Constant	-1.905	.040		.000	1.000
F1 Service quality	.632	.040	.632	15.942	.000
F2 Safety	.314	.040	.314	7.915	.000
F3 Related effects	.218	.040	.218	5.509	.000
F4 Financial benefit	.092	.040	.092	2.324	.021
F5 Convenience	.085	.040	.085	2.146	.033

(Source: Analysis results from SPSS 16.0 software)

The regression results show that 5 independent factors from F1, F2, F3, F4, F5 are statistically significant, sig <0.05 is satisfactory, so they will be retained in the research model.

Table 5: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.749[a]	.561	.554	.66814717

(Source: Analysis results from SPSS 16.0 software)

Based on the results in the table above, ANOVA has value Sig = 0.000 <0.05, so the hypothesis H0 is rejected, assuming the hypothesis H1. That means the model exists.

In other words, with 5% significance, it can be concluded that the decision to choose a savings bank of an individual customer is influenced by at least 1 of the remaining 5 factors:

The regression results showed that the Durbin-Watson Statistics was 2,182. Thus, the model exists and with d = 2,219 □ 2,0, we can conclude that the model does not exist negative or positive autocorrelation.

So, the regression estimation model would be:

QD = -1.905 + 0.632F1 + 0.314F2 + 0.218F3 + 0.092F4 + 0.085F5

The standardized regression would be:

QD * = 0.632F1 * + 0.314F2 * + 0.218F3 * + 0.092 F4 * + 0.08 F5 *

And the model of quantitative research results takes the form:

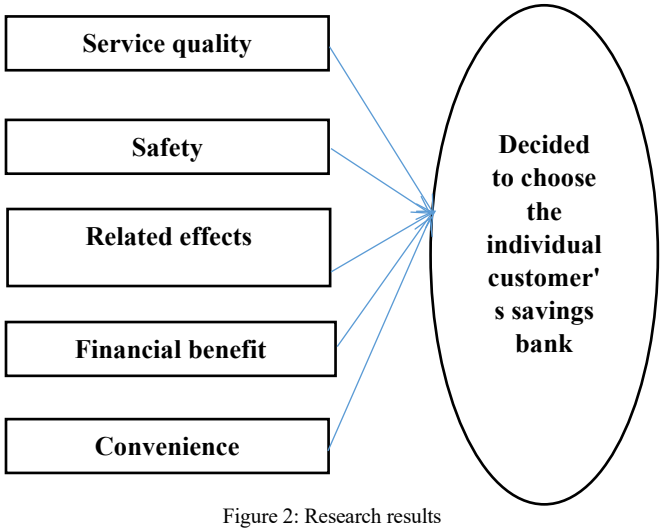


Figure 2: Research results

Linear regression model using Enter method is made with a number of assumptions and the model only really makes sense when these assumptions are guaranteed. Therefore, to ensure the reliability of the model, the topic must also evaluate the violation of necessary assumptions in linear regression. From the result in Figure 4.5, we have 1 <d = 2.219 <3 so we can conclude that the remainder are independent of each other and that the independence of the remainder is guaranteed. [33]

Finally, we will consider the multi-collinearity violation of the model. In the above correlation coefficient analysis, we have seen that the dependent variable has a fairly clear correlation with the independent variables, but we can also see that there is correlation between the independent variables. This will create the model's multicollinearity capability. Therefore, we have to detect multi-collinearity phenomenon by calculating the acceptability of the variable (Tolerance) and the variance inflation factor (VIF).

Thus, the linear regression model built according to the above equation does not violate the necessary assumptions in linear regression.

5. Conclusion

With the survey results including the number of samples of 285 for a regression model including 7 groups of factors affecting the decision of individual customers to choose a savings bank at commercial banks in Da Nang city. . The quantitative research results show that there are 5 groups of factors that greatly affect the decisions of individual customers to choose a savings bank: Service quality, Safety, Related effects, Benefits financial benefits, convenience

5.1. Improving the quality of banking services

Besides expanding the banking network, increasing investment in facilities and techniques, developing many convenient products to attract customers; Banks are constantly making efforts to perfect products, improve the quality of care and serve customers. Products and services are factors affecting customers' choice of banks to save savings. Because customers come to the bank, apart from the purpose of saving, have many other needs such as transferring money, paying bills ... Providing enough services according to customers' needs is also one of the attractive measures. Personal savings deposit. Therefore, the improvement of aspects related to the retail products and services that the bank offers to customers, including product variety, simplification of procedures, and time savings. , and there are different preferential policies depending on different subjects.

5.2. Increase safety for customers when sending money

With specific regulations on insured deposits, rights and obligations of the sender, payment limit, time of interest payment ... has shown the Bank's commitment and the State to ensure security. safety of depositors' deposits aims to best protect the legitimate rights and interests of depositors, ensure the safety and health of banking operations and handle financial crisis. In addition, the Bank needs to build support channels to help customers easily use smart services, tools and utilities to check their savings accounts at the Bank anytime, anywhere like channel. at the counter (account balance confirmation service), Internet Banking channel (for customers registered to use iPay service) or automatic answering switchboard channel (TPIN) ... As recommended by the Bank, customers Customers must also protect themselves when depositing money through the implementation of bank recommendations. Customers register SMS Banking with the Bank, enter the correct customer number or ID card to automatically look up the following information: Account balance, account number, last 5 transactions of the account account accounts, savings book balance, card status, exchange rate, savings interest rate, information on products, services, promotions….

5.3. Improve financial benefits for depositors

The research results show that among the factors individual customers will consider when choosing a bank to save money, financial benefits are the most important factor. In which, the observed variable "financial benefits" received a high consensus. It can be said that savings depositors appreciate the bank's financial capacity. Therefore, improving financial benefits is one of the conditions for banks to attract customers.

Financial benefits are shown through three aspects, including interest rates on savings deposits, fees applied to products and services and regular promotions. In general, this is one of the top concerns when customers use the bank's products and services. However, in Vietnam, issues related to interest rates are strictly managed by the State Bank of Vietnam through the ceiling deposit rate applicable to commercial banks. As more and more customers need convenience, the online savings channel becomes preeminent, especially for the group of 'digital customers'. Therefore, banks should encourage customers to deposit online savings with bonus interest rates, which is seen as a positive move proactively following customer needs and is a competitive factor in the market when The convenience of digital banking is becoming more and more popular in Vietnam. Both the Bank and the customers save translation time, and have established themselves in the retail industry with recognition from international organizations.

5.4. Increased convenience for customers to send money

Bank savings is always the first choice for unused idle money. However, many customers, when saving money, do not deposit all their idle money because they are afraid that they cannot withdraw their money when they need money, or they must withdraw their interest before their due date. Therefore, the bank needs to offer a solution for customers that should split the savings book instead of just depositing all the money in one book. Along with that, the bank needs to expand its network through increasing the number of transaction points, the banks also invest heavily in renovating the transaction space of the existing branches, providing a spacious transaction space. page, spacious, creating maximum comfort for customers. At the same time, banks also promote digital banking and gradually affirmed themselves with recognition from international organizations. On the digital banking platform, customers easily perform hundreds of types of services, from paying bills for electricity, water, telecommunications, monthly insurance to fast money transfers within 5 minutes for bank accounts. other goods

References

[1] I. Ajzen, "The theory of planned behavior." Organizational behavior and human decision processes, **50**(2), 179-211, 1991

[2] P. Luarnand and H.H. Lin, "Toward an understanding of the behavioral intention to use mobile banking." Computers in human behavior, **21**(6), 873-891, 2005.

[3] V. Venkatesh, S. M. Walton, and J. Y. L. Thong, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology." MIS quarterly, 157-178, 2012

[4] C.-S. Yu, "Factors affecting individuals to adopt mobile banking: Empirical evidence from the UTAUT model." Journal of electronic commerce research **13**(2), 104, 2012

[5] H. Q. Cuong, Research on factors affecting the intention to use online shopping services (Nghiên cứu các yếu tố ảnh hưởng đến ý định sử dụng dịch vụ mua hàng điện tử qua mạng), National University of Ho Chi Minh City, 2020.

[6] L. T. T Hang, Research on bank savings behavior of individual customers in 2012 (Nghiên cứu hành vi gửi tiền tiết kiệm ngân hàng của khách hàng cá nhân năm 2012). Doctoral thesis, Academy of Social Sciences, 2012

[7] B. K. Jeong, M. Khouja, and K. Zhao, "The impacts of piracy and supply chain contracts on digital music channel performance." Decision Support Systems **52**(3), 590-603, 2012.

[8] M. Fishbein and I. Ajzen, "Belief, attitude, intention, and behavior: An introduction to theory and research.", 1977.

[9] I. Ajzen, "From intentions to actions: A theory of planned behavior." Action control. Springer, Berlin, Heidelberg, 11-39, 1985

[10] W. T. Anderson, E. P. Cox and D. G. Fulcher, "Bank Selection Decisions and Market Segmentation: Determinant attribute analysis reveals convenience-and sevice-oriented bank customers." Journal of Marketing, **40**(1): 40-4, 1976. DOI: 10.1177/002224297604000107

[11] D. D Schram, C. C. Darling Milloy, and W. E. Rowe, "Juvenile sex offenders: A follow-up study of reoffense behavior". Community Protection Research Project, Washington State Institute for Public Policy, 1991.

[12] K. Khazeh, and W. H. Decker, "How customers choose banks." Journal of Retail banking **14**(4), 41-45, 1992.

[13] D. H. Tootelian and R. M. Gaedeke, "Targeting the college market for banking services." Journal of Professional Services Marketing **14**(2), 161-172, 1996.

[14] E. Mylonakis, E. L. Hohmann, and S. B. Calderwood, "Central nervous system infection with Listeria monocytogenes. 33 years' experience at a general hospital and review of 776 episodes from the literature." Medicine **77**(5), 313-336, 1998.

[15] S. Mokhlis, N. Hazimah, N. Mat, and H. S. Salleh, "Commercial bank selection: the case of undergraduate students in Malaysia." International Review of Business Research Papers **4**(5), 258-270, 2008.

[16] C. Chigamba, and O. Fatoki, "Factors influencing the choice of commercial banks by university students in South Africa." International Journal of Business and Management, **6**(6), 66, 2011

[17] H H. Siddique, "Bank selection decision criteria employed by indian expatriates in sultanate of oman: an empirical analysis." International Journal of Business and Management Studies **4**(2), 55-64, 2012.

[18] S. S. Almejyesh and K. S. Rajha, "Behavioral determinants and their impact on customer savings deposits in Islamic banks in Saudi Arabia." Journal of Islamic Banking and Finance **2**(1), 163-186, 2014

[19] C. I. Enyinda, "Evaluation of relationship marketing in Islamic banks in the UAE: Empirical evidence based on sensitivity analysis algorithm." Journal of Economic & Financial Studies **2**(3), 01-12, 2014.

[20] A. Sajeevan Rao, and R. K. Sharma, "Bank selection criteria employed by MBA students in Delhi: An empirical analysis." Journal of business studies Quarterly **1**(2), 56-69, 2010.

[21] S. Mokhlis, H. S. Salleh, and N. H. N. Mat, "What do young intellectuals look for in a bank? An empirical analysis of attribute importance in retail bank selection." Journal of Management Research **3**(2), 1-15, 2011.

[22] R. I. Sabir, Ghafoor, O., Akhtar, N., Hafeez, I., and Rehman, A. U, "Factors affecting customer satisfaction in banking sector of Pakistan." International review of management and business research **3**(2), 1014, 2014

[23] P. T. Tam, and P. N. Thuy, Factors affecting individual customers' tendency to choose the Bank (Yếu tố ảnh hưởng đến xu hướng lựa chọn Ngân hàng của khách hàng cá nhân). Journal of Banking Science and Training, 103, April 2010.

[24] N. T. Gam, Research on factors affecting individual customer satisfaction for Joint Stock Commercial Bank for Industry and Trade East Hanoi Branch, (Nghiên cứu các yếu tố ảnh hưởng đến sự thỏa mãn của khách hàng cá nhân đối với ngân hàng thương mại cổ phần Công thương Chi nhánh Đông Hà Nội). Science & Technology Magazine, **81**, 33-40, 2011.

[25] L. M. Linh, Research on customer satisfaction and loyalty to banking services in Hau Giang province (Nghiên cứu sự hài lòng và trung thành của khách hàng đối với dịch vụ ngân hàng trên địa bàn tỉnh Hậu Giang), Journal of Social Sciences and Humanities, **12**, 34 – 41, 2014.

[26] P. D. Khoi, Factors affecting customer satisfaction with savings deposit services: The case of Agribank Binh Minh, Vinh Long (Các nhân tố ảnh hưởng đến mức độ hài lòng của khách hàng đối với dịch vụ tiền gửi tiết kiệm: Trường hợp Agribank Bình Minh,Vĩnh Long). Scientific Journal of Can Tho University, **40**, 50-57, 2015. Doi: 10.1263/03697854100265

[27] N. N. D. Phuong, Factors influencing individual customers' decisions to save money - Research at Joint Stock Commercial Bank for Industry and Trade of Vietnam (Các nhân tố ảnh hưởng đến quyết định gửi tiền tiết kiệm của khách hàng cá nhân - Nghiên cứu tại Ngân hàng Thương mại cổ phần Công Thương Việt Nam). Industry and Trade Magazine, No. 01, June 2014

[28] N. D. Tho, and N. T. M. Trang, Research the components of brand equity and measure them in the consumer goods market in Vietnam (Nghiên cứu các thành phần của giá trị thương hiệu và đo lường chúng trong thị trường hàng tiêu dùng tại Việt Nam). Ministry-MSB-level scientific research project, 22-33, 2002

[29] J. N. Kapferer, Die Marke, Kapital des Unternehmens. No. hal-00788645. 1992.

[30] D. A. Aaker, "Measuring brand equity across products and markets." California management review 38.3 1996.

[31] P. Gerrard and J. B. Cunningham, "Singapore's undergraduates: how they choose which bank to patronise." International Journal of Bank Marketing, 2001. Doi: 10.1108/02652320110388531

[32] J. Tank and K. Tyler, "UK student banking revisited: influences and the decision-making process." Journal of financial services marketing **10**(2), 152-164, 2005. Doi: 10.1177/002224297604068954

[33] H. Trong and C. N. M Ngoc, Research Data Analysis with SPSS (Phân tích dữ liệu nghiên cứu với SPSS (Tập 1)), Statistical Publishing House, Ho Chi Minh City, 2008

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 418-423 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Identification of Genetic Variants for Prioritized miRNA-targeted Genes Associated with Complex Traits

Isabella He[1], Zhaohui Qin[2], Yongsheng Bai[*,3]

[1]*Pittsford Mendon High School, 472 Mendon Road, Pittsford, NY 14534, USA*

[2]*Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA*

[3]*Next-Gen Intelligent Science Training, Ann Arbor, MI 48105, USA*

ARTICLE INFO

Article history:
Received: 30 April, 2021
Accepted: 05 June, 2021
Online: 27 June, 2021

Keywords:
GWAS
BMI
miRNA
variant
3'UTR

ABSTRACT

Genome-wide association studies, or GWAS, have reported associations between SNPs and specific diseases/traits. GWAS results contain variants located in different genomic regions, including variants in the 3'UTR. MicroRNAs, or miRNAs, are small noncoding RNAs that bind to the 3'UTRs of genes to regulate gene expression. However, variant(s) that are located in the 3'UTR could impact miRNA binding, thus affecting expression of its targeted gene(s). To specifically elucidate miRNA targeting pairs and binding site variants associated with a specific trait, well-designed downstream analysis along with careful experimental design are necessary. Currently, there is no available state-of-the-art methodology for identifying miRNA targeting pairs and associated variants that could contribute to phenotypes using GWAS. Moreover, it is unrealistic to conduct experiments for elucidating all possible miRNA targeting pairs and binding site variants across the entire genome. In this project, we developed a bioinformatic pipeline to computationally identify genes and their targeting miRNA pairs that are enriched over the miRNA-gene tissue expression network for the studied genetic traits and examined the binding site variants' impact on Body Mass Index (BMI).

1. Introduction

GWAS is a powerful approach to identify common variants associated with common diseases/traits in studied populations [1, 2]. Traits may be activated due to genetics or changes in environment, or both.

MicroRNAs (miRNAs), noncoding RNAs with a length between 17-22 nucleotides, exert their impact on gene regulation through targeting mRNAs. To achieve their function, miRNAs bind to the 3' untranslated regions (3'UTRs) of target mRNAs which can often result in potential suppression of mRNA translation and/or gene expression.

Single-nucleotide polymorphisms (or SNPs) are the most common mutations found in the human genome [3, 4]. SNPs located in the 3'UTRs can potentially disrupt miRNA regulation [5]. A DNA motif is a conserved DNA sequence segment with biological consequences. The identification of a motif located in the 3'UTR could suggest how miRNAs potentially target genes in that region. Analysis of GWAS data by identifying genes or miRNAs that harbor genetic variants could elucidate trait-associated genetic variants [6].

In our experiment, we analyzed multiple traits compiled from GWAS and found BMI to be the most significant trait in the context of the miRNA-gene tissue expression target network. We further investigated several genes located within the region(s) associated with BMI, *FADS1* and *FADS2* due to their significant relevance to obesity and related conditions [7]. The overarching goal of this study is to identify how the interaction between genes and their targeting miRNAs change due to genetic variations and reveal any underlying biological mechanisms that contribute to the development of complex traits for humans.

2. Experimental Methods

To perform our experiment, we proposed the following approach:

*Corresponding Author: Yongsheng Bai, 2373 Foxway Drive, Ann Arbor, MI 48105, USA Contact No: (817)559-3448
Email: bioinformaticsresearchtomorrow@gmail.com

https://dx.doi.org/10.25046/aj060346

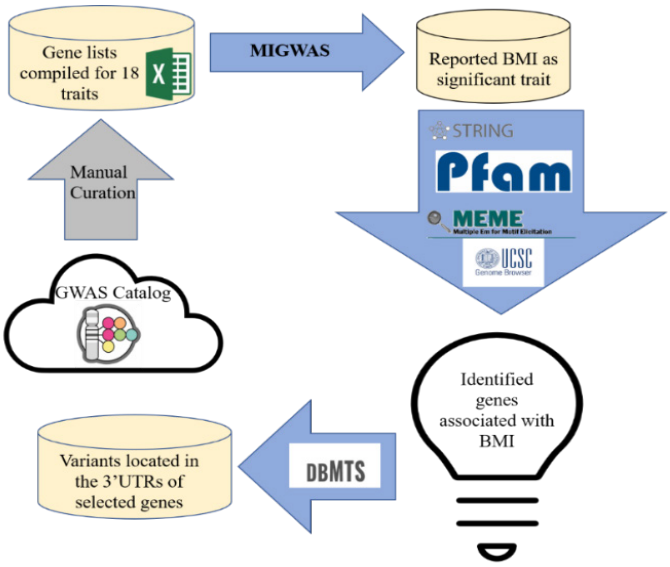


Figure 1: Workflow of variant identification for complex trait BMI

2.1. GWAS Traits Selection

We downloaded the annotation data of GWAS results from 18 traits, all of which had at least 100 associated genes as compiled by PheGenI [6]. All traits were previously described [8]. These traits ranged from autoimmune disorders such as Rheumatoid Arthritis and Asthma to behavior traits such as smoking.

2.2. Trait Relevant miRNA-gene Pairs Identification

We employed MIGWAS software [9, 10] using default running parameters (2018 Github version: https://github.com/saorisakaue/MIGWAS) to strengthen our analysis regarding all 18 traits and identify significant traits and miRNA-gene pair candidates over the gene tissue expression network. In order to run MIGWAS, we converted the variant genomic coordinate annotations (hg38) provided by PheGenI to the hg19 version. Figure 2 shows the Python pipeline of running MIGWAS.

Step 1:

```
python3 minimgnt.py hg19_BMI.txt --out hg19_BMI_sumstats_trans --cpus 10 --not-remove-HLA --remove-NA --no-rsid
```

Step 2:

```
python3 migwas.py --phenotype hg19_BMI_sumstats_trans --out miRA_hg19_BMI_sumstats_trans --cpus 10 --iterations 500 --output-candidate
```

Figure 2: Python code of running MIGWAS

2.3. Protein-Protein Interaction Examination and Domain Elucidation

We ran STRING database to look for evidence of interaction between the ten candidate genes associated with BMI at the protein level [11]. We also used Pfam to identify if there were any possible domains associated with BMI [12].

2.4. Motif Discovery for miRNA-targeted Genes Relevant to BMI Trait

We applied MEME software (version 5.3) to identify enriched motifs for the 3'UTR sequences retrieved from Ensembl Biomart Database [13, 14]. To run MEME, we stipulated the software parameter for the searched motif width to be between six and twenty base pairs. Also, the number of searched motifs was restricted to report the top six motifs. All other running parameters were kept as default.

2.5. Variant Identification for Selected Genes Associated with BMI

We checked the dbMTS database for reported candidate genes associated with BMI through MIGWAS and pinpointed SNPs located inside each gene's respective 3'UTRs [5].

3. Results

3.1. Selection of GWAS Results

We acquired the GWAS results of 18 traits with more than 100 associated genes that were reported in our recent study [8]. All selected traits are displayed in Table 1 below.

Table 1: Selected 18 traits as reported by PheGeni. All traits listed were associated with at least 100 genes.

Trait	# SNPs	# Genes
Alcoholism	802	657
Alzheimer's	873	1037
Arthritis, Rheumatoid	374	276
Asthma	712	818
Autistic Disorder	982	505
BMI	3601	1971
Coronary Artery Disease	493	471
Coronary Disease	440	345
Crohn's	633	403
T1D	309	235
T2D	542	416
Hypertension	314	270
NSCLC	129	131
Obesity	426	434
Osteoporosis	77	102
Psoriasis	425	292
Schizophrenia	1530	1453
Smoking	781	938

3.2. MIGWAS Analysis

We identified in particular that one out of the 18 traits, BMI, has a significant association with all tissues (p-value of .0313). There are only three traits which have a p-value reported for trait

and tissue association, including BMI, Crohn's, and Type 1 Diabetes (T1D) (Figure 3).

We have identified ten genes that are targeted by five miRNAs as candidate biomarkers in the BMI trait. For Crohn's, there was only one pair while T1D had two pairs (Table 2).

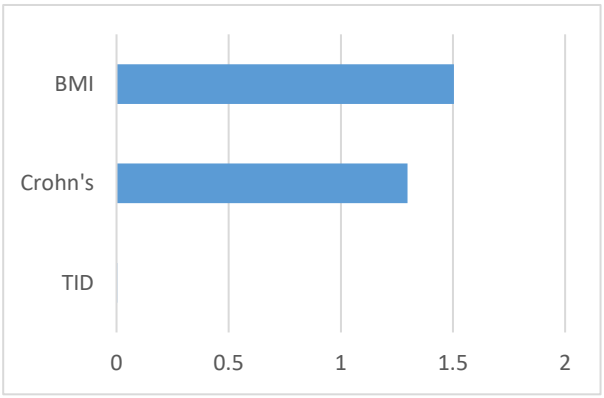


Figure 3: MIGWAS reported p-values (-log base 10 scale) for BMI, Crohn's disease, and Type 1 Diabetes.

Table 2: MIGWAS results for candidate genes/miRNAs identified for three traits

Trait name	Gene	miRNA
BMI	*ADAMTS9*	*hsa-mir-330*
BMI	*AMH*	*hsa-mir-615*
BMI	*CELSR2*	*hsa-mir-615*
BMI	*DNM3*	*hsa-mir-330*
BMI	*FADS1*	*hsa-mir-615*
BMI	*FADS2*	*hsa-mir-615*
BMI	*HOXC5*	*hsa-mir-615*
BMI	*MED17*	*hsa-mir-4531*
BMI	*NEGR1*	*hsa-mir-4721*
BMI	*OTUD7B*	*hsa-mir-642a*
Crohn's	*SYNGR1*	*hsa-mir-4728*
T1D	*HIPK1*	*hsa-mir-1236*
T1D	*TMEM116*	*hsa-mir-4646*

3.3. String Database Results

Out of the ten genes researched, there are three genes—*FADS2, FADS1*, and *AMH*—that interact with each other as reported by the STRING database, as shown in Figure 4. These three genes are also targeted by the same miRNA: *hsa-mir-615*. The STRING database also reported BMI's significant associations (FDR < .01) with fatty acid metabolism (hsa01212) and biosynthesis of unsaturated fatty acids (hsa01040) under KEGG pathways. Likewise, the BMI-related GO terms—linoleic acid metabolic process (GO:0043651), alpha-linolenic acid metabolic process (GO:0036109), and unsaturated fatty acid biosynthetic process (GO:0006636)—are also reported by the STRING database.

3.4. Pfam Database Results

The Pfam database reported two domains associated with BMI: Fatty Acid Desaturase (PF00487) and Cytochrome b5-like heme/steroid binding domain (PF00173), as shown in Figure 5.

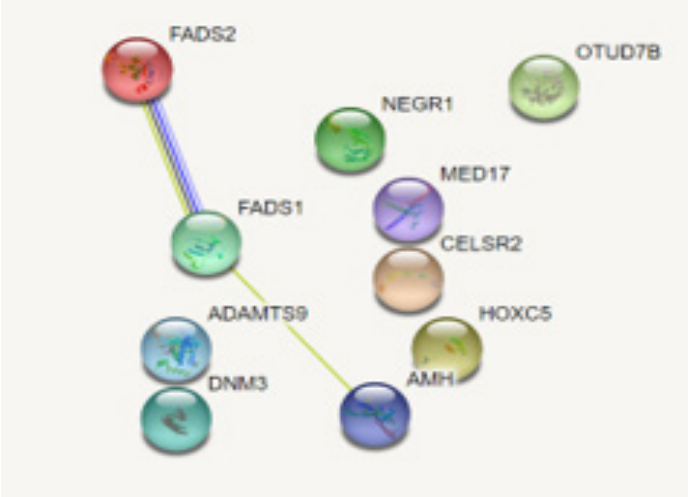


Figure 4: Protein-protein interaction results for all ten genes associated with BMI. Only three genes were found to have interactions with each other (*FADS1, FADS2, AMH*).

These results were consistent with the ones reported by the INTERPRO database.

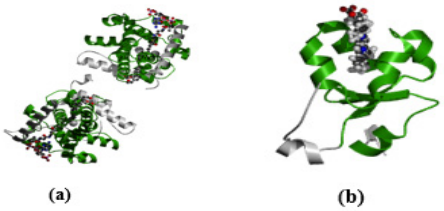


Figure 5: Protein structures of two BMI-associated domains. (a) displays the Fatty Acid Desaturase Domain. (b) displays the Cytochrome b5-like heme/steroid binding domain.

3.5. Motif Identification

Using MEME, we ran all ten of the genes associated with BMI. A motif was discovered containing nine out of these ten genes (*ADAMTS9, CELSR2, DNM3, FADS1, FADS2, HOXC5, MED17, NEGR1, OTUD7B*), with an e-value of .047 (Figure 6).

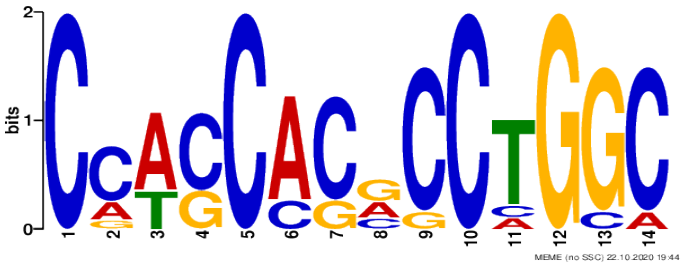


Figure 6: Motif conservation across nine (out of ten) genes associated with BMI

3.6. Variant Identification Results

To report known common variants contained in miRNA-targeted genes associated with BMI, we checked the UCSC Genome Browser [15]. This region contains multiple known variants (dbSNP 153 Track) that are located inside both *FADS1* and *FADS2* (see Figure 7 below). Interestingly, *FADS1*, *FADS2*, and *FEN1* (a neighboring gene) are highly expressed in the adrenal gland. As indicated by GTEx data, *CELSR2*, *FADS1*, and *FADS2* all are expressed with low amounts in both types of adipose tissue. Since all three genes are targeted by the same miRNA, *hsa-miR-615*, low expression of these genes in adipose tissues could be due to the repression effect of this miRNA.

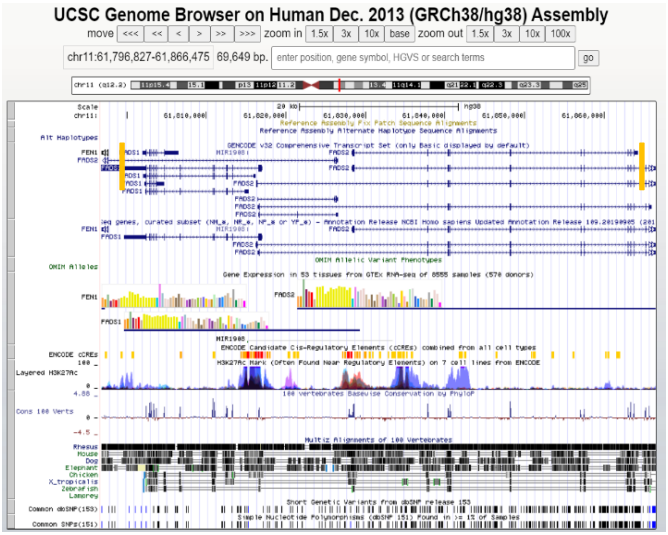


Figure 7: UCSC Genome Browser view for variants located in the genomic regions of *FADS1/FADS2.* The orange vertical bars placed within the view pinpoint the general motif coordinates defined by MEME. The left bar represents the identified motif region for *FADS1* while the right bar represents *FADS2's* identified motif region.

3.7. dbMTS Results

To identify miRNA binding-site variants, we checked the dbMTS database [5]. Table 3 shows that multiple variants, which could alter the binding relationship of miRNAs and their targets, are often present in the 3'UTRs of miRNA-targeted genes associated with BMI.

4. Discussion and Interpretation

In this study, we applied cutting-edge bioinformatics tools and databases to select candidate genes/miRNAs associated with complex genetic trait(s), such as BMI.

Using MIGWAS, we analyzed three traits: BMI, Crohn's, and T1D. While both BMI and Crohn's had significant p-values in the gene expression network, Crohn's yielded only one gene/miRNA pair. On the contrary, T1D had a non-significant p-value but yielded 2 gene/miRNA pairs (see Figure 3 and Table 2). We concluded that a significant p-value does not necessarily indicate the existence of more miRNA-gene targeting pairs for specific traits. As a result of this MIGWAS analysis, we focused on investigating BMI as it had many gene/miRNA pairs in comparison to the other two traits mentioned. When performing

Table 3: Variant results reported in dbMTS

rs_dbSNP150	chr	pos	ref	alt	VEP_ensembl_Gene_Name
rs7528419	1	109274570	A	G	*CELSR2*
rs11102967	1	109274623	C	G	*CELSR2*
rs1277917	1	109274857	G	C	*CELSR2*
rs12740374	1	109274968	G	T	*CELSR2*
rs660240	1	109275216	T	C	*CELSR2*
rs57677983	1	109275536	C	A	*CELSR2*
rs658435	1	109275648	C	T	*CELSR2*
rs629301	1	109275684	G	T	*CELSR2*
rs10410	1	109279782	G	A	*CELSR2*
rs14000	1	109279887	T	C	*CELSR2*
rs1061955	1	149942458	G	A	*OTUD7B*
rs2757500	1	172309881	G	A	*DNM3*
rs2586422	1	172309937	A	G	*DNM3*
rs2586423	1	172310272	G	A	*DNM3*
rs2421994	1	172310769	G	A	*DNM3*
rs10489282	1	172407906	G	A	*DNM3*
rs4916255	1	172411735	C	A	*DNM3*
rs7528296	1	172412243	T	C	*DNM3*
rs11714364	3	64650182	C	A	*ADAMTS9*
rs9851598	3	64650452	G	A	*ADAMTS9*
rs4246215	11	61796827	G	T	*FADS1, FADS2*
rs174544	11	61800281	C	A	*FADS1, FADS2*
rs174545	11	61801834	C	A	*FADS1, FADS2*
rs174545	11	61801834	C	G	*FADS1, FADS2*
rs174546	11	61802358	C	T	*FADS1*
rs174568	11	61826344	C	T	*FADS1, FADS2*
rs482548	11	61865710	C	A	*FADS2*
rs17764288	11	61866264	G	A	*FADS2*
rs11539526	11	61866298	C	G	*FADS2*
rs77167250	11	61866475	C	T	*FADS2*
rs676586	11	93788239	C	T	*MED17*
rs606087	11	93789476	A	G	*MED17*
rs17652258	11	93814518	G	C	*MED17*
rs677019	11	93818291	T	C	*MED17*
rs654896	11	93818330	A	T	*MED17*
rs17573753	11	93818630	C	T	*MED17*
rs679241	11	93818748	A	G	*MED17*
rs585768	11	93819689	C	T	*MED17*
rs2071450	12	54034748	C	T	*HOXC5*
rs104894666	19	2250667	C	T	*AMH*

this study, we noticed that different tools/databases can analyze data with complementary results from different aspects.

For example, MIGWAS identified *CELSR2* containing 3'UTR variants also reported by the PheGenI database. MIGWAS identified two genes (out of nine)—*DNM3* and *FADS1*—that don't contain 3'UTR variants in the data provided by PheGenI. The other seven genes were not reported by PheGenI [6, 9].

After comparison, dbMTS reported results on nine genes total. Six genes (*OTUD7B, DNM3, ADAMTS9, MED17, FADS1,* and *FADS2*) contain 3'UTR variants also reported by MIGWAS. The remaining three genes (*CELSR2, HOXC5,* and *AMH*) do not contain 3'UTR variants as dbMTS states.

Using DIANA Tools, we identified two genes (*CELSR2* and *ADAMTS9*) that are targeted by *hsa-miR-615* [16]. Although MIGWAS stated that *ADAMTS9* was targeted by *hsa-miR-330*, this is likely due to this gene being targeted by multiple miRNAs. However, wet-lab experiments are needed to validate this conclusion. In addition, the targeting outcome of multiple genes associated with BMI by *hsa-miR-615* is significant. Studies show that increased levels of *hsa-miR-615* may help inhibit palmitate-induced hepatocyte apoptosis [17]. If saturation levels of palmitate, a type of fatty acid, are increased, this increases the likelihood of developing diseases such as non-alcoholic fatty liver disease, which is correlated with obesity (or a high BMI) [18]. Indeed, *CELSR2* was recently discovered to have significant high expression in cancerous tissue than normal tissue. Specifically, hepatocellular carcinoma, a primary liver cancer whose major risk factors include non-alcoholic fatty liver disease [19].

Current literature evidence supports that *FADS2* has been discovered to be a drug target gene as well as an miRNA-target gene for rheumatoid arthritis (RA) [10]. However, such evidence reports *hsa-miR-4728* as *FADS2'*s target miRNA instead of *hsa-miR-615*. This suggests the involvement of both miRNAs in the development of BMI-related illnesses. In addition, multi-omics analysis proves *FADS2* is a potential biomarker for BMI. This is consistent with the conclusion that a high BMI leads to increased risk of developing RA, even though the biological mechanism responsible has yet to be identified [20, 21]. In addition, a recent study confirms that the *FADS1/2* cluster are significantly associated with fatty acid levels (see Figure 4). This study also reported *FADS1* alongside rs174546, matching our corresponding row in Table 3. Unlike our experiment, however, the mentioned study did not employ dbMTS to generate this match [22]. The conservation of both *FADS1* and *FADS2* in the identified motif (see Figures 6 and 7) further emphasizes the potential importance of these two genes in the association with BMI/development of obesity.

In a recent study, Kuryłowicz and colleagues report that two miRNAs—*hsa-miR-615* and *hsa-miR-330*—and their target genes (of which none are listed in Table 2) are involved in the oncogenesis pathway, otherwise known as the process where healthy cells become cancerous [23]. Shared target miRNAs between a genetic trait (BMI) and cancer suggests that the development of illness as a result of either share similar causal/biological factors. In addition, *hsa-miR-615* was found to be highly expressed in the subcutaneous adipose tissue (SAT) of obese patients after surgery-induced weight loss than the SAT of patients with a normal weight, suggesting that there is a difference in molecular pathways/miRNA expression between losing weight and consistently maintaining a healthy weight [23]. This is consistent with the finding that *hsa-miR-615* facilitates palmitate production which is correlated with a higher likelihood of developing obesity.

In our previous study, we found a statistically significant association between BMI and the PCDHA10 pathway [8]. PCDHA10 is a gene which codes for different protocadherins (Pcdhs), which play an important role in neural generation [24]. As neuronal development and brain structure have been linked to BMI through existing literature, this could be a reason why the PCDHA10 pathway is associated so strongly with BMI [25]. Although obesity is not identified by MIGWAS as a significant trait in our study, it has been reported to have been strongly associated with BMI and has some significant associations with body temperature [4, 26]. For future research, we would like to examine the popularity of variants located in the binding sites of 3'UTR for the miRNA-gene pairs of additional traits to further investigate the underlying biological mechanisms and causal factors of other genetic diseases, traits, and even cancers.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to express our deep gratitude to Tianli Jiang and Aryan Thakur for their contributions in curating the comprehensive list of traits from PheGenI.

References

[1] K. Norrgard, "Genetic variation and disease: GWAS," Nature Education **1**(1):87, 2008

[2] https://www.genome.gov/about-genomics/fact-sheets/Genome-Wide-Association-Studies-Fact-Sheet

[3] A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korbel, J.L. Marchini, S. McCarthy, G.A. McVean, G.R. Abecasis, "A global reference for human genetic variation," The 1000 Genomes Project Consortium, *Nature* 526**,** 68–74, 2015, https://doi.org/10.1038/nature15393

[4] D. Altshuler, P. Donnelly, "A haplotype map of the human genome," The International HapMap Consortium, *Nature* 437**,** 1299–1320, 2005, https://doi.org/10.1038/nature04226

[5] C. Li, M.D. Swartz, B. Yu, Y. Bai, X. Liu, "dbMTS: a comprehensive database of putative human microRNA target site SNVs and their functional predictions," bioRxiv. 2019, doi: https://doi.org/10.1101/554485

[6] E.M. Ramos, D. Hoffman, H.A. Junkins, D. Maglott, K. Ohan, S.T. Sherry, M. Feolo, L.A. Hindorff. "Phenotype-Genotype Integrator (PheGenI): synthesizing genome-wide association study (GWAS) data with existing genomic resources," European Journal of Human Genetics, **22**(1):144-147, 2014, doi:10.1038/ejhg.2013.96

[7] A. Maguolo, C. Zusi, A. Giontella, E.M.D.G. Miraglia, A. Tagetti, C. Fava, A. Morandi, C. Maffeis, "Influence of genetic variants in FADS2 and ELOVL2 genes on BMI and PUFAs homeostasis in children and adolescents with obesity," International Journal Observations (Lond). Aug 25, 2020, doi: 10.1038/s41366-020-00662-9. PMID: 32843713.

[8] I. He, T. Jiang, A. Thakur, Y. Bai, X. Qin, "Systematic and Comprehensive Survey of Genomic Loci Associated with Complex Diseases and Traits," Proc. of the 2020 14th IEEE International Conference on Bioinformatics and Biomedicine (BIBM), **1**, 1447-1452, 2020 doi: Bookmark: 10.1109/BIBM49941.2020.9313115

[9] S. Sakaue, J. Hirata, Y. Maeda, E. Kawakami, T. Nii, T. Kishikawa, K. Ishigaki, C. Terao, K. Suzuki, M. Akiyama, et al, "Integration of genetics and miRNA-target gene network identified disease biology implicated in tissue specificity," Nucleic Acids Reserves, **46**(22):11898-11909, 2018, doi: 10.1093/nar/gky1066

[10] Y. Okada, T. Muramatsu, N. Suita, M. Kanai, E. Kawakami, V. Iotchkova, N. Soranzo, J. Inazawa, T. Tanaka, "Significant impact of miRNA-target gene networks on genetics of human complex traits," Sci Rep, 6:22223, 2006.

[11] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, et al, "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," Nucleic Acids Res, Jan; **47**, 607-613, 2019, doi: 10.1093/nar/gky1131

[12] R.D. Finn, A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, et at, "Pfam: the protein families database," Nucleic Acids Research, **42**(Database issue):D222-D230, 2014, doi:10.1093/nar/gkt1223

[13] T.L. Bailey, M. Bodén, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, "MEME SUITE: tools for motif discovery and

searching," Nucleic Acids Research, **37**, :W202-W208, 2009, doi: 10.1093/nar/gkp335

[14] A.D. Yates, P. Achuthan, W. Akanni, J. Allen, J. Allen, J. Alvarez-Jarreta, M.R. Amode, I.M. Armean, A.G. Azov, R. Bennett, et al. "Ensembl 2020," Nucleic Acids Research, **48**(D1), 08 January 2020, Pages D682-D688, https://doi.org/10.1093/nar/gkz966

[15] W.J. Kent, C.W. Sugnet, T.S. Furey, K.M. Roskin, T.H. Pringle, A.M. Zahler, D. Haussler, "The human genome browser at UCSC," Genome Res. **12**(6):996-1006, 2002, https://genome.ucsc.edu/index.html

[16] D. Karagkouni, M.D. Paraskevopoulou, S. Chatzopoulos, I.S. Vlachos, S. Tastsoglou, I. Kanellos, D. Papadimitriou, I. Kavakiotis, S. Maniou, G. Skoufos, et al, "DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions," Nucleic Acids Res. 2018;**46**(D1):D239-D245, 2018, doi:10.1093/nar/gkx1141

[17] Y. Miyamoto, A.S. Mauer, S. Kumar, J.L. Mott, H. Malhi, "Mmu-miR-615-3p Regulates Lipoapoptosis by Inhibiting C/EBP Homologous Protein," PLoS ONE **9**(10), 2014: e109637. https://doi.org/10.1371/journal.pone.0109637

[18] E. Fabbrini, S. Sullivan, S. Klein, "Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications," Hepatology. **51**(2):679-689, 2010 doi:10.1002/hep.23280

[19] M. Xu, S. Zhu, R. Xu, N. Lin, "Identification of CELSR2 as a novel prognostic biomarker for hepatocellular carcinoma," BMC Cancer **20**, 313, 2020, https://doi.org/10.1186/s12885-020-06813-5

[20] X. Feng, X. Xu, Y. Shi, X. Liu, H. Liu, H. Hou, L. Ji, Y. Li, W. Wang, Y. Wang, et al, "Body Mass Index and the Risk of Rheumatoid Arthritis: An Updated Dose-Response Meta-Analysis," Biomed Res Int, 2019:3579081, doi:10.1155/2019/3579081

[21] M.D. George, J.F. Baker, "The Obesity Epidemic and Consequences for Rheumatoid Arthritis Care," Curr Rheumatol Rep, **18**(1):6, 2016, doi:10.1007/s11926-015-0550-z

[22] Z. He, R. Zhang, F. Jiang, H. Zhang, A. Zhao, B. Xu, L. Jin, T. Wang, W. Jia, W. Jia, et al, "FADS1-FADS2 genetic polymorphisms are associated with fatty acid metabolism through changes in DNA methylation and gene expression," Clin Epigenet **10**, 113, 2018, https://doi.org/10.1186/s13148-018-0545-5

[23] A. Kuryłowicz, Z. Wicik, M. Owczarz, M.I. Jonas, M. Kotlarek, M. Swierniak, W. Lisik, M. Jonas, B. Noszczyk, M. Puzianowska-Kuznicka, "NGS Reveals Molecular Pathways Affected by Obesity and Weight Loss-Related Changes in miRNA Levels in Adipose Tissue," Int J Mol Sci, **19**(1):66. Published 2017 Dec 27, 2017, doi: 10.3390/ijms19010066

[24] K.M. Mah, J.A Weiner, "Regulation of WNT Signaling by Protocadherins," Seminars in Cell and Developmental Biology, August 2017, doi: 10.1016/j.semcdb.2017.07.043

[25] P. Vakli, R. Deak-Meszlenyi, T. Auer, Z. Vidnyanskzy, "Predicting Body Mass Index From Structural MRI Brain Images Using a Deep Convolutional Neural Network," Front. Neuriform, March 2020, https://doi.org/10.3389/fninf.2020.00010

[26] F. Bastardot, P. Marques-Vidal, P. Vollenweider, "Association of body temperature with obesity. The CoLaus study," International Journal of Obesity, **43**, 1026-1033, 2019

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 431-438 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Design of Anthropomorphic Hand based on Human Finger Anatomy

Zixun He[1], Yousun Kang[2], Duk Shin[3,*]

[1]*Department of Human Centered Science and Biomedical Engineering, Tokyo Institute of Technology, Yokohama, 226-8503, Japan*

[2]*Department of Applied Computer Science, Tokyo Polytechnic University, Kanagawa, 243-0297, Japan*

[3]*Department of Electronics and Mechatronics, Tokyo Polytechnic University, Kanagawa, 243-0297, Japan*

ARTICLE INFO

Article history:
Received: 21 April, 2021
Accepted: 09 August, 2021
Online: 28 August, 2021

Keywords:
Human hand biomechanics
Prosthetic hand
EMG signals
Quality of life

ABSTRACT

In the past decade, multiple anthropomorphic prosthetic hands have been developed to replace the role of human hands. Prostheses should not only replace the functions of human hands in functionality but also replicate human hands in appearance and sense of body-belonging intuitively. Human fingers have very delicate and complex structures, and it is these complex structures that make our hands dexterity. This study proposes a design based on the anatomical characteristics of the human hand. The proposed design replicates human fingers from bones, ligaments, extensor hoods, and extensor mechanism of tendon, intended to develop a prosthesis that has the same flexibility and appearance as human hand. To evaluate the performance of the proposed prosthetic in daily life, we conducted grasping experiments on common objects. It is successfully proved that the proposed design helps to improve the grasping performance of the artificial hand and has a natural appearance. In this paper, our design succeeds to improve the grasping performance of the artificial hand and gain natural appearance.

1. Introduction

This paper is an extended paper of our work initially presented in International Symposium on Community-centric Systems (CcS 2020) [1]. The human hand is an important tool for us to interact with and perceive the physical environment. Upper limbs loss is one of the most disabling diseases that a person may experience, which can severely affect amputees Activities of daily living (ADLs) and working abilities. In the United States alone, there are approximately 50,000 people who have lost upper limbs, and it is predicted that the numbers of upper limb loss will increase over time [2]. According to the investigation [3], different amputees may have different requirements for prostheses. Besides performance of prosthetics devices, a high level of personification and simplified operating system are also required. Prosthetic developers need to consider these different requirements, which bring challenges to the development of prosthetic hands.

In order to improve the acceptability of the prosthesis [4,5], various artificial hand designs and models have been proposed. Commercially available advanced prosthetic hands, such as Michelangelo hand (by Otto Bock), are easy to use and can help amputee grip the objects. However, the price of commercially available prosthetics is very expensive. It is difficult for people who have lost upper limbs to afford this financial burden. In recent years, with the development of new manufacturing technologies such as fused deposition modeling (FDM) 3D printers, low-cost and easy-to-make open-source prostheses have also been designed [6,7].

Most research groups aim to replicate the movement mechanism of human hands in a mechanized design, improving the flexibility and performance of the prosthetics. In the past two decades, the research community has developed some novel artificial hands [8-10]. Lee et al. proposed the design of a 9- DOF bio-mimetic robot hand. This hand has four under-actuated fingers, each of finger is equipped with a tactile sensor and driven by two linear actuators coupled (together) [11]. In order to further reduce the weight of the prosthetic hand, Mohammadi et al. used soft materials and developed a lightweight soft robotic prosthetic hand that has synergy-based motion and cable-driven actuation system. In addition, successfully performed three different grips with high dynamic power (21.5N) and fast finger bending speed. Although these artificial hands have excellent grip and performance, they

*Corresponding Author: Duk Shin, 1583, Iiyama, Atsugi, Kanagawa, 243-0297, Japan, Email: d.shin@eng.t-kougei.ac.jp

have not yet reached the level of human dexterity. The reason for that is these designs adopt the structure of the human body in the mechanized way, the incomplete human hand-like motion mechanism limited to the motion performance of the manipulator [12].

In [13], the author developed the anatomically corrected tested (ACT) hand by replicating the bone structures of the fingers, which is the first robotic hand developed based on human hand interpretation. Although its mechanism is still based on hinges and universal joints, it still exhibits a more flexible grip than previous manipulators. In [14], the author designed a robot hand through high bionics, and a flexor mechanism for an elastic pulley mechanism to realize the joint movement of human fingers. These studies of highly bionic anthropomorphic robotic hands have showed the possibility of replicating the structure of the human hand on the prosthesis to make the movements more dexterity.

Electromyography signal (EMG) is a signal that can reflect the user's movement intention. In recent years, it has been widely used to control prostheses or robots. Although this kind of EMG-based control is widely used for pattern recognition and rehabilitation, the nature of EMG signal is random or non-stationary and will be affected by nearby muscles that generate additional noise [15]. Therefore, in order to control the prosthetic hand accurately, it is necessary to extract the features related to the selected action patterns, and then use support vector machine (SVM), linear discriminant analysis (LDA), Artificial neural network (ANN) and other machine learning to classify the action. The correct selection of features is critical to classification performance. However, because of different electrode placement schemes, number of channels, and EMG signal preprocessing methods in different studies, it is difficult to compare clearly which features can significantly improve the accuracy of classification [16]. Deep learning methods have recently been used in the classification of EMG signals and have shown strong performance [17]. Côté-Allard et al. use Convolutional Neural Networks (CNN) to classify the actions of 7 forelimbs with an accuracy of 97.9% [18]. It can be said that in a laboratory environment, such a classifier with high computational cost and powerful performance is satisfactory, but it is an important practical consideration to ensure the effectiveness while reducing the computational burden when building the system. In addition, reducing the number of electrodes to improve comfort also needs to be considered for the prosthetic hand system. Tavakoli et al. use only 2 EMG channels to classify four types of gestures, and the classification accuracy rate exceeds to 90% [19]. In daily life, we use hand movements to grasp objects or interact with the environment [20,21]. Therefore, by defining the minimum number of daily grasping patterns, the number of electrodes can be reduced, and the simplicity of the system can be maintained.

In this article, we designed an anthropomorphic prosthetic hand based on anatomical structure. The four-finger and thumb model of the prosthetic hand is modeled based on laser scanned hand bone data on cadaveric. This can reduce the rejection of the amputee to the prosthesis after the artificial skin is worn on the prosthetic hand, and the movement angle of each finger will not exceed the movable range of the original finger. Incorporate soft tissue structures such as ligaments, tendons and tendon sheaths that have a significant impact on the flexibility of the fingers at each joint to replicate the movement characteristics of the human hand. Then he conducted a grasping experiment on the prosthetic hand to test the performance of our prosthesis in daily life. In order to control the prosthetic hand using Electromyogram (EMG) signals, an EMG control system is designed for the artificial hand by an artificial neural network (ANN). The EMG control system is used to control the four main hand positions used in daily life: Power grasps, Precision grasps, Lateral grasps and relax state [21].

2. Development of prosthetics

2.1. Hand structures and the Bones Models

We designed highly anthropomorphic prosthetic hand based on the anatomy. First, the skeletal structure of the human finger movement mechanism is mainly determined by the metacarpal and phalangeal bones. The human finger comprises the distal, middle, proximal phalanges and metacarpals. These four bones make up the distal interphalangeal (DIP), proximal interphalangeal (PIP) and metacarpophalangeal (MCP) joints of the hand. The thumb finger has only two phalanges, so it has only one interphalangeal joint. Every finger joint has soft tissues called ligaments [22].

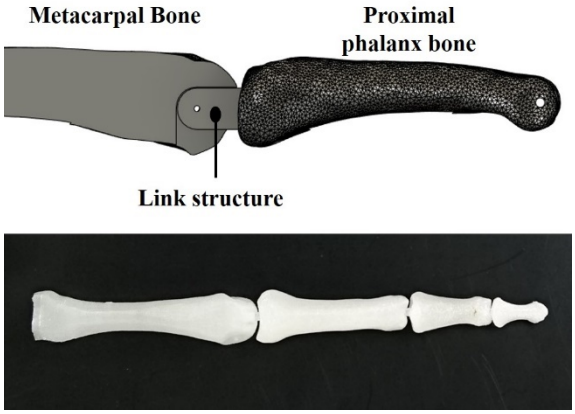


Figure 1: Designed Finger model

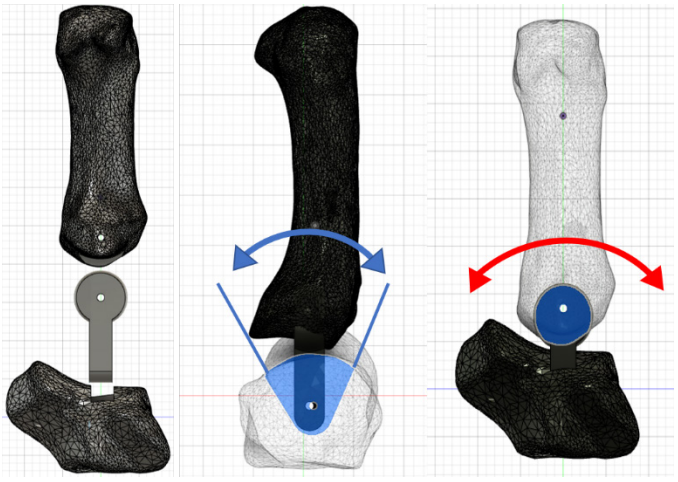


Figure 2: Proposed thumb design

We used laser-scanned bone data as the basic model of the prosthesis [23]. Since the size of the prosthetic hand skin in the market vendor is larger than the bone model of this hand, we adjusted the size of the model. In order to improve the efficiency of power transmission, a link structure is designed between each joint as shown as Figure 1.

Table 1: Finger Flexion degree [deg]

DIP Joint	PIP Joint	MCP Joint
0 - 77	0 - 105	0 - 90

The range of motion of the finger joints will vary between different fingers and different people, depending on the bone shape, tendons and muscle structure characteristics of the hand. By referring to the previous research on joint structure [24], in this research, their range of motion is uniformly designed, as shown in Table 1.

Although the thumb comprises only two bones, it has great freedom of movement. The reason is attributed to the carpometacarpal (CMC) joint composed of the thumb metacarpal bone and trapezium bone. The trapezium bone is shaped like a saddle used for horse riding, allowing the thumb to move in a larger range. The movement of the thumb is caused by sliding the metacarpal bone of the thumb along the trapezium bone at the CMC joint back and forth (extension / flexion), left and right (abduction / adduction) or both at the same time [25]. The following method is used to reproduce the thumb structure to ensure correct thumb movement. In order to reproduce the adduction and abduction functions of the thumb, the connecting rod structure on the metacarpal model additionally uses a rotary joint, as shown in Figure 2. And a groove is designed on the trapezium bone to achieve flexion and extension movement.

2.2. Design of Joint Ligament and Tendon

Ligaments are a type of fibrous connective tissue that connects bone to bone. These tissues can be found in all joints. The range of motion of each finger joint is limited by the length of the ligament. There are collateral ligaments on both sides of the hand bones, which are attached to the volar plates. Cartilage structures called volar plates are inserted on both sides of the bone joints. Volar plates and ligaments form a joint pack to prevent joint overextend and enhance joint stability. There are two sets of tendons in the human hand to stretch and bend the fingers. They are the extensor and flexor tendons. The flexor tendons extend from the forearm and, finally, branches become the flexor digitorum superficialis (FDS) tendon and the flexor digitorum profundus (FDP) tendon. The FDS tendon is fixed to the intermediate phalange bones of the PIP joint. The FDP tendon is inserted into the bottom of the palm of the distal phalanges. When the finger is bent, the PIP joint attached to the FDS tendon will bend first. As the movement progresses, the FDP tendon will play a role in bending the DIP joint. These two flexor tendons are firmly attached to the phalanx through the tendon sheath. The extensor tendons are combined with extensor hood that regulate joint extension and flexion. Extensor hood is a complex mesh structure that wraps the phalanx directly from the back of the hand. Its structure is shown in Figure 5 (a) [26]. The first layer of the dorsal aponeurosis of the extensor hood is inserted into the base of the DIP joint and divided into two small tendons at the PIP joint. The second layer of the dorsal aponeurosis of the extensor hood is located at the bottom of the PIP joint.

According to the ligament, the replicated model of the structure of the volar plate is shown in Figure 3. The volar plate is made of High Elasticity Rubber Plate. Anchor the rubber volar plate to each joint with screws. The ligament is made of rubber thread, and the natural length of the ligament is adjusted when inserted into the volar plate and both sides of the hand bone, so that the fingers of the artificial hand can be naturally bent like a human finger when they are naturally relaxed. Thus, the prosthetic hand will not be kept in a straight and rigid motion, and the prosthetic hand will be more personified.



Figure 3: Structure in anatomy and reproduced model

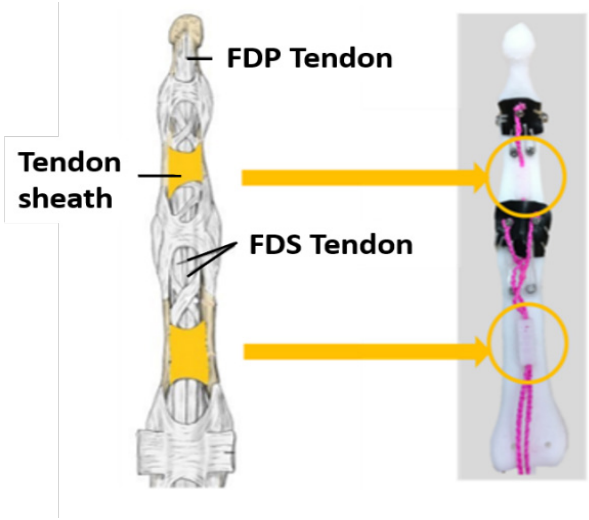
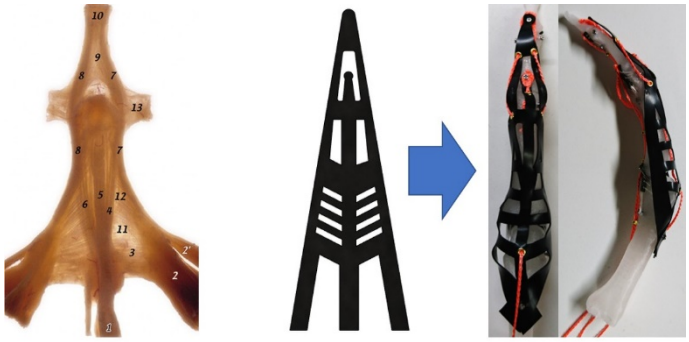


Figure 4: Tendon and Tendon sheaths

The flexor and extensor tendons composed of FDS tendons and FDP tendons are made of 0.33 mm polyethylene wire (Figure 4), with chief strength (250 N breaking strength) and good flexibility. The tendons are woven to prevent the volar plate and tendon sheath from being worn out because of the long-term movement of the tendons.



(a) (b)

Figure 5: Hood parts; (a) Extensor hood Structure [26] (b) Reproduced extensor hood sheath

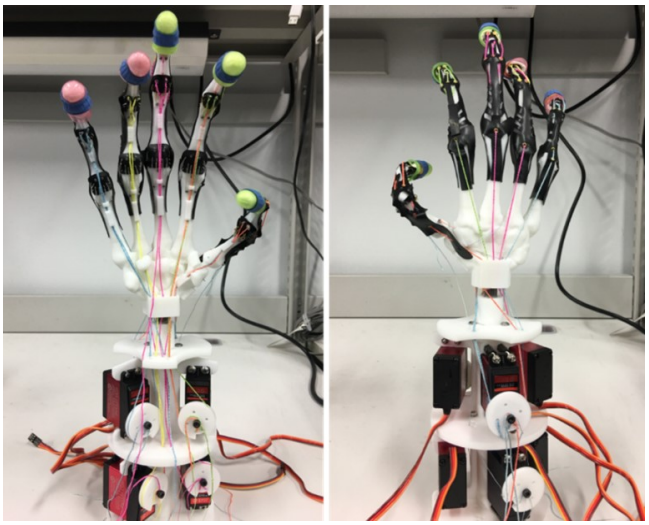


Figure 6: Configuration of the prosthesis

Channel	Muscle
1	Flexor digitorum superficialis muscle
2	Flexor digitorum profundus muscle
3	Extensor pollicis longus muscle
4	Extensor digitorum muscle

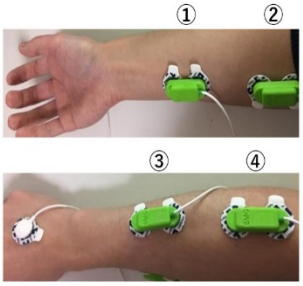


Figure 7: Location of the EMG electrodes

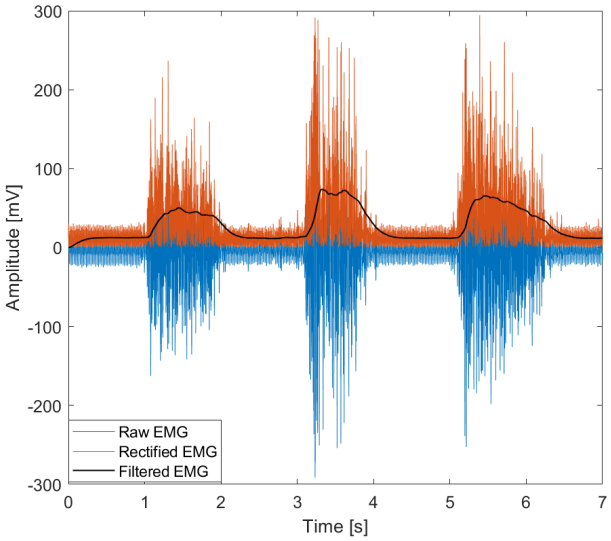


Figure 8: Pre-Processing of EMG signal

Table 2: The specifications of DS3218 Servo

Dimensions	40 x 20 x 40.5 mm
Weight	60 g
Stall Torque	21.5 kg・cm (6.0V)

The extensor tendon has a central fiber bundle that reaches the base of the intermediate phalanges while two outer fiber bundles are anchored at the base of the distal phalanx. We simplified the structure of the extensor hood and made the extensor hood using a highly elastic rubber sheet to reproduce the function of the extensor hood to adjust joint extending or bending. Figure 5 shows the finger after all components have been assembled.

2.3. Configuration of the prosthesis and Control Method

After all fingers are assembled, combined them with the wrist part. The complete prosthesis is shown in Figure 6. The tendon of the finger is fixed on the pulley, and the pulley is connected to the servo. By pulling the tendon with the servo motor, the finger joint bent or extended. Six Servo motors (Ds servo DS3218) are used to control the prosthetic hand. The servo specifications are shown in Table 2. Considering Thumb has 2 sets of movement, we use 2 servos to control Adduction or Abduction, and Flexion or Extension of the thumb, respectively. Other fingers are each controlled by one servo. In order to use the intuitive operation method to control the prosthetic hand in the grasping experiment, an operating system was developed using the right-hand glove part of Perception Neuron (Noitom, Beijing, China). Perception Neuron's joint obtain the movement of the finger detected in three-dimensional coordinates. The servo is controlled by the DIP joint angle detected by the sensor neuron. Calculate the bending angle of the finger according to formula (1).

$$\theta = 180 - arccos\frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad (1)$$

Here, x, y, and z are three-dimensional coordinates detected from the sensory neuron of the DIP joint. The calculated angle data will be sent to Arduino UNO to control the servo.

2.4. EMG signals Recording and Processing

The author showed that increasing the number of electrodes does not always improve the accuracy of gesture classification [27]. In addition, problems such as using many electrodes or the long time for classification can also make amputee annoying.

Based on these basic demands, we decided on the number of electrodes and the classification time. The connection points of the EMG sensor and the surface muscles selected to measure the EMG signal are shown in Figure 7. These muscles control the four movements (power grasp, precision grasp, lateral grasp and relax) used for classification. The reason for choosing these 4 kinds of motions is because power grasps, precision grasps, Lateral grasps are most used in activities of daily living and user will also need a relax mode for an inactive position.

We gained EMG signals using the BITalino EMG sensor. By using MATLAB's Bitalino'Toolbox, the EMG signal is recorded on MATLAB at a sampling rate of 1000 Hz [28]. Before calculating the features of the EMG signal, we rectified and filtered the four EMG signal channels. All preprocessing are performed using MATLAB. A second-order Butterworth low-pass filter with a cutoff frequency of 5 Hz is used to filter the EMG signal to reduce the noise of the EMG signal [29]. Figure 8 shows the raw EMG signal and filtered EMG signal. During training and experiment, all features were analyzed using a sliding time

window of 200 (ms) with an overlap of 100 (ms). The calculation processing delay is about 300 (ms). The control is quasi real-time.

2.5. Feature Extraction

Three types of signal features: time-domain features, frequency-domain features, and time-frequency features are usually used for EMG classification. In order to improve the accuracy of classification, two or more features are usually combined. We tested 8 EMG signal features proposed by Phinyomark et al. in the previous work to select features [30]. These features are: Integrated EMG (IEMG), Mean Absolute Value (MAV), Simple Square Integral (SSI), Variance of EMG (VAR), Root Mean Square (RMS), Waveform Length (WL), Mean Frequency (MNF) and Median Frequency (MDF). Among them, we found that the combination of the three features of RMS, WL and MNF(o) showed better results. These results are consistent with previous research results, root mean square (RMS) and wavelength (WL) can get good classification results and require less computational cost [31]. Compared with time-domain features such as RMS and WL, frequency-domain features require more computational costs. In [32], the author reported that the frequency-domain features are better than the time-domain features when predicting the angle and flexion of the thumb. Therefore, this study uses a combination of two time domain features RMS, WL and a frequency-domain features Mean Frequency (MNF) for classification.

The WL, RMS, MNF are defined as follows:

$$WL = \sum_{n=1}^{n-1} |x_{n+1} - x_n| \tag{2}$$

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{3}$$

$$MNF = \sum^{M} F\,P \Big/ \sum^{M} P \tag{4}$$

here, x_n denotes the n^{th} of EMG signal, N is the length of signals. F is the frequency, P is the Power spectrum.

Table 3: The list of objects used in the Experiment

Number	Name	Weight (g)	Size(cm)
1	TV Remote	75	length 13 width 5 thickness 3
2	Pill	1	diameter 1
3	Telephone	137	length 17 width 5 thickness 2
4	Bottle	40	length 13 width 5 thickness 3
5	Glasses	38	length 15 width 3 thickness 3
6	Spoon	5	length 15 width 3 thickness 2
7	Phone	190	length 16 width 7 thickness 1
8	Tooth Paste	238	length 13 width 5 thickness 3
9	Envelope	14	length 23 width 12
10	Drink (500ml)	534	length 21 width 7 thickness 7
11	Bowl	57	length 11 width 11 thickness 5
12	Key	9	length 7 width 3 thickness 0.5
13	Pen	15	length 14 width 1 thickness 1
14	Paper box	73	length 18 width 11 thickness 8
15	Bill	1	length 15 width 7
16	Straw	2	length 10 width 0.5
17	Coin	5	diameter 2.5
18	Bag	897	length 40 width 25 thickness 15
19	DVD	16	diameter 11
20	Small Ball	12	diameter 3

3. Evaluation Experiment

3.1. Grasping experiment

In order to evaluate the applicability of the proposed prosthesis to daily life, we used the items in the list of items that amyotrophic lateral sclerosis (ALS) patients are difficult to pick up in daily life for experiments [33]. We tested the feasibility of the prosthetic hand grasped and pick up easily some items as shown in Table 3. The gripping and dexterity must be needed the equivalent strength of strings driven like an actual hand. We have selected 20 items in the item list above. Among them, 19 items can be easily found in the laboratory. In addition, a school bag of about 1 kg was selected as the object to observe the performance of the artificial hand against heavy objects in Appendices. In grasping an item, if the prosthesis can grip for the object over 5 seconds, it will be recorded as a success, otherwise, it will be recorded as a failure.

3.2. Classification experiment

In order to classify four gestures, we used MATLAB's neural network toolbox to build a three-layer artificial neural network (ANN) as a classifier. This ANN consists of twelve (four channels x three features) nodes of input layers, one hidden layer composed of twelve nodes, and four nodes output layers. The activation functions of the middle layer and the output layer use hyperbolic tangent functions. The training method uses a learning algorithm called Levenberg-Marquardt backpropagation (LMBP). LMBP is a general non-linear least-squares optimization method, its major advantage lies in the speed of convergence. Before starting the Classification Experiment, we will train their own ANN for each subject in advance. The dataset used for training ANN contains 10 sets of data for each motion, and each set of data contains 5 seconds of EMG signal. In order to prevent over fitting, the training of ANN will stop if Mean Square Error (MSE) $\leqq$ 0.001. Before performing the Classification experiment, use 10-fold cross-validation on the collected dataset to test the accuracy of the classifier as shown as Figure 8. For each subject's dataset, 80 % of the data is used to train the classifier, and the remaining 20 % are used to validate the performance of the classifier. Figure 9 shows the average accuracy of the four motions of 4 subjects after using 10-fold cross-validation. The highest accuracy rate is 94.7% (Subject4), the lowest accuracy rate is 91.4% (Subject2).

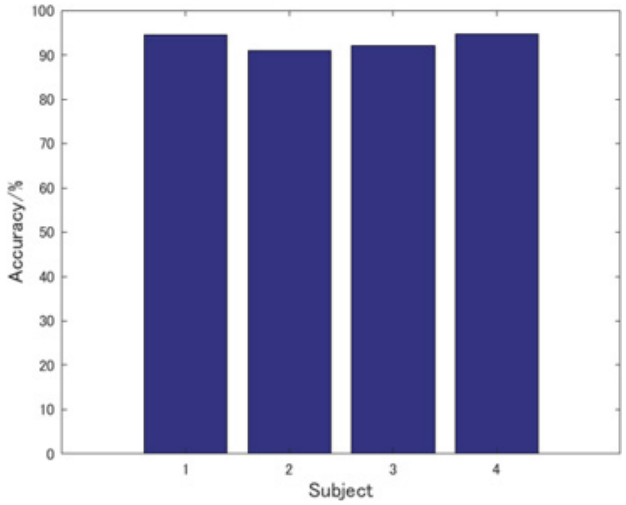


Figure 9: 10-fold cross validation results

4. Results

4.1. Result of Grasping experiment

In the experiment, our prosthetic hand successfully grasped selected 20 objects (Figure 10). Different from the traditional ON/OFF clip-shaped prosthetic hand or the prosthetic hand that grasps in a fixed posture, the proposed anthropomorphic design allows our prosthetic hand to grasp objects in the suitable grasping posture.



Figure 10: Result of grasping experiment

4.2. Result of Classification experiment

It showed the results of the movement Classification experiment in Table 4. The highest average accuracy rate was 93.7%. The lowest average accuracy rate was 87.5%. All subjects achieved 100% accuracy when Classifying "Power grasp" Motion. In addition, the classification of "Precision grasp" and "Lateral grasp" is slightly lower than the accuracy rate of other motions, with a minimum of 80%. During the experiment, the researchers observed that the subject's 'Relax' movements were classified as 'Lateral grasp' movements multiple times. This might because the

Table 4: Classification accuracy

Subject / Motion	1	2	3	4
Power grasp	20/20 100%	20/20 100%	20/20 100%	20/20 100%
Precision grasp	18/20 90%	18/20 90%	17/20 85%	16/20 80%
Lateral grasp	20/20 100%	16/20 80%	18/20 90%	17/20 85%
Relax	17/20 85%	19/20 95%	18/20 90%	17/20 85%
AVG.	93.7%	91.2%	91.2%	87.5%

EMG signal generated by thumb adduction is difficult to collect from superficial muscles.

5. Discussion

In the Grasping experiment, we tested the gripping performance of the proposed prosthesis against objects commonly found in daily environments. When grasping objects such as cylindrical objects or large objects, we usually use the Power grasp posture. The 500ml PET bottle and the 18x11x8cm paper bag are very large for the hand, so it is necessary to grasp them with your fingertips. In addition, the artificial hand can lift a bag of about 1 kg and hold the posture for over 5 seconds. From the above results, it can be found that the anatomically reproduced finger structure sufficiently contributes to the transmission of force and transmits the torque of the motor to the fingers and fingertips. Although a strong torque moment applies to the fingertips, the shape of the finger joints is not deformed, and the proposed artificial hand has sufficient force and rigidity to grasp heavy objects. By using precision grasp and Lateral grasp, the prosthetic hand can pick up small objects such as balls and coins with fingertips during the experiment.

These results shown that our proposed design shows highly coordinated motions of the thumb and index finger, suggested that the proposed thumb structure has the same kinetic features as a human thumb. It is also one of our design goals to use bones as a model to give the artificial hand a natural appearance. As shown in Figure 11, our prosthetic hand also holds Natural-Appearance when wearing a silicon skin glove. It should be noted that the silicon skin glove worn is only a general commercial product.

For healthy subjects, the ANN classification system can achieve high-precision classification of 4 motions. But before the system is used in amputees, it is still necessary to continue to improve performance. From the experimental results, the accuracy of the two motions of "precise grasp" and "side grasp" is lower than other motions. The reason may be that "precise grip" and "side grip" include thumb movement, so the feature sets of these two motions are similar. In our experiment, the EMG signal is measured by two EMG sensors in the flexor group and two EMG sensors in the extensor group. However, it may be difficult to classify detailed finger movements from the extracted EMG signals. Because there are many muscles in the forearm (Flexor carpi radialis, flexor carpi ulnar is, Flexor digitorum superficialis muscle, Deep digital flexor muscle, flexor digitorum longus) located deep in the measurement position. This will affect the accuracy and precision of classification.

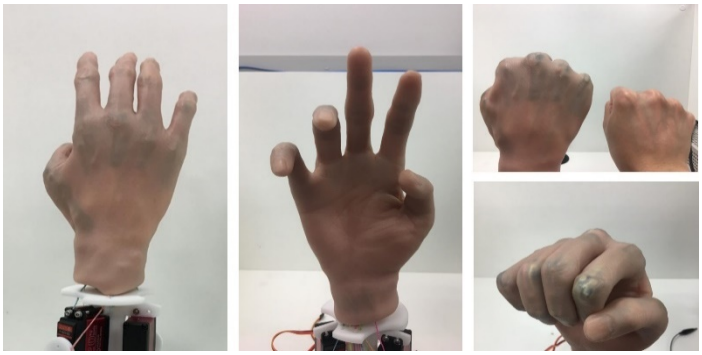


Figure 11: Appearance of prosthetic hand wearing silicon skin glove

6. Conclusion

In this study, we developed a biomimetic prosthetic hand based on human anatomy. By reproducing critical soft tissues such as tendons and ligaments in the human hand, the developed artificial hand has the same dexterity as a human finger. For this reason, our prosthetic hand can stably grasp all the object types proposed in the experiment. In order to control the prosthetic hand by EMG signal, an ANN has been designed for classifying four motions. Four subjects tested the classification performance of ANN. The classification accuracy average rate was 91%.

The limitations of this study are as follows. The overall system of the proposed bionic prosthesis is about 560 grams. A survey of prosthetic users pointed out that if the weight of the prosthetic hand exceeds 400 grams, and it is considered being heavy [34]. Currently, our biomimetic prosthesis does not include the wrist part. In order to use artificial hands in daily life, the ability of the wrist part also plays a vital role. Currently, the proposed EMG control method only tests the performance of healthy subjects, and needs to be tested on amputees to investigate whether the performance will change. In the next phase of this research, we will make improvements to address these issues.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 19K11428.

References

[1] Z. He, R.R. Yurievich, S. Shimizu, M. Fukuda, Y. Kang, D. Shin, "A Design of Anthropomorphic Hand based on Human Finger Anatomy," in 2020 International Symposium on Community-centric Systems (CcS), IEEE: 1–5, 2020, doi:10.1109/CcS49175.2020.9231423.

[2] K. Ziegler-Graham, E.J. MacKenzie, P.L. Ephraim, T.G. Travison, R. Brookmeyer, "Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050," Archives of Physical Medicine and Rehabilitation, **89**(3), 422–429, 2008, doi:10.1016/J.APMR.2007.11.005.

[3] F. Cordella, A.L. Ciancio, R. Sacchetti, A. Davalli, A.G. Cutti, E. Guglielmelli, L. Zollo, "Literature Review on Needs of Upper Limb Prosthesis Users," Frontiers in Neuroscience, **0**(MAY), 209, 2016, doi:10.3389/FNINS.2016.00209.

[4] E.A. Biddiss, T.T. Chau, "Upper limb prosthesis use and abandonment: A survey of the last 25 years," doi:10.1080/03093640600994581.

[5] I. Vujaklija, D. Farina, O.C. Aszmann, "New developments in prosthetic arm systems," Orthopedic Research and Reviews, **8**, 31, 2016, doi:10.2147/ORR.S71468.

[6] G. Langevin. InMoov – open-source 3D printed life-size robot, http://inmoov.fr/, Aug. 2021.

[7] Exiii Inc. HACKberry |3D-printable open-source bionic arm, http://exiii-hackberry.com/, Aug. 2021.

[8] J. Zuniga, D. Katsavelis, J. Peck, J. Stollberg, M. Petrykowski, A. Carson, C. Fernandez, "Cyborg beast: a low-cost 3d-printed prosthetic hand for children with upper-limb differences," BMC Research Notes 2015 8:1, **8**(1), 1–9, 2015, doi:10.1186/S13104-015-0971-9.

[9] P. Weiner, J. Starke, F. Hundhausen, J. Beil, T. Asfour, "The KIT Prosthetic Hand: Design and Control," IEEE International Conference on Intelligent Robots and Systems, 3328–3334, 2018, doi:10.1109/IROS.2018.8593851.

[10] P. Weiner, C. Neef, Y. Shibata, Y. Nakamura, T. Asfour, "An Embedded, Multi-Modal Sensor System for Scalable Robotic and Prosthetic Hand Fingers," Sensors 2020, Vol. 20, Page 101, **20**(1), 101, 2019, doi:10.3390/S20010101.

[11] S. Lee, S. Noh, Y.K. Lee, J.H. Park, "Development of bio-mimetic robot hand using parallel mechanisms," 2009 IEEE International Conference on Robotics and Biomimetics, ROBIO 2009, 550–555, 2009, doi:10.1109/ROBIO.2009.5420706.

[12] A. Mohammadi, J. Lavranos, H. Zhou, R. Mutlu, G. Alici, Y. Tan, P. Choong, D. Oetomo, "A practical 3D-printed soft robotic prosthetic hand with multi-articulating capabilities," PLOS ONE, **15**(5), e0232766, 2020, doi:10.1371/JOURNAL.PONE.0232766.

[13] A.D. Deshpande, Z. Xu, M.J.V. Weghe, B.H. Brown, J. Ko, L.Y. Chang, D.D. Wilkinson, S.M. Bidic, Y. Matsuoka, "Mechanisms of the anatomically correct testbed hand," IEEE/ASME Transactions on Mechatronics, **18**(1), 238–250, 2013, doi:10.1109/TMECH.2011.2166801.

[14] Z. Xu, E. Todorov, "Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration," Proceedings - IEEE International Conference on Robotics and Automation, **2016-June**, 3485–3492, 2016, doi:10.1109/ICRA.2016.7487528.

[15] M.B.I. Reaz, M.S. Hussain, F. Mohd-Yasin, M.B.I. Raez, "Techniques of EMG signal analysis: detection, processing, classification and applications," Biol. Proced. Online, **8**(1), 11–35, 2006, doi:10.1251/bpo115.

[16] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," Computer Methods and Programs in Biomedicine, **161**, 1–13, 2018, doi:10.1016/J.CMPB.2018.04.005.

[17] A. Kalantari, A. Kamsin, S. Shamshirband, A. Gani, H. Alinejad-Rokny, A.T. Chronopoulos, "Computational intelligence approaches for classification of medical data: State-of-the-art, future challenges and research directions," Neurocomputing, **276**, 2–22, 2018, doi:10.1016/J.NEUCOM.2017.01.126.

[18] U. Côtéallard, F. Nougarou, C.L. Fall, P. Gigu'ere, C. Gosselin, F. Laviolette, B. Gosselin, "A Convolutional Neural Network for robotic arm guidance using sEMG based frequency-features," IEEE International Conference on Intelligent Robots and Systems, **2016-November**, 2464–2470, 2016, doi:10.1109/IROS.2016.7759384.

[19] M. Tavakoli, C. Benussi, P. Alhais Lopes, L.B. Osorio, A.T. de Almeida, "Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier," Biomedical Signal Processing and Control, **46**, 121–130, 2018, doi:10.1016/J.BSPC.2018.07.010.

[20] I.M. Bullock, J.Z. Zheng, S. De La Rosa, C. Guertler, A.M. Dollar, "Grasp frequency and usage in daily household and machine shop tasks," IEEE Transactions on Haptics, **6**(3), 296–308, 2013, doi:10.1109/TOH.2013.6.

[21] C.C. Scuola, S. Sant'anna, M.C. Scuola, M. Chiara, C. Scuola, C. Cipriani, M. Controzzi, M. Chiara Carrozza, "Objectives, criteria and methods for the design of the SmartHand transradial prosthesis Sensory feedback and closed loop control of artificial limbs View project Artificial arms components View project Objectives, criteria and methods for the design of the SmartHand transradial prosthesis," Robotica, **28**, 919–927, 2010, doi:10.1017/S0263574709990750.

[22] C.L. Taylor, R.J. Schwarz, "The Anatomy and Mechanics of the Human Hand," Artificial Limbs VOL. 2 MAY 1955 NO. 2 CONTENTS, 2, 1955.

[23] Human Hand Bones - Thumb by siderits - Thingiverse, https://www.thingiverse.com/thing:15342, Aug. 2021.

[24] J.C. Becker, N. V. Thakor, "A Study of the Range of Motion of Human Fingers with Application to Anthropomorphic Designs," IEEE Transactions on Biomedical Engineering, **35**(2), 110–117, 1988, doi:10.1109/10.1348.

[25] I. Komatsu, J.D. Lubahn, "Anatomy and Biomechanics of the Thumb Carpometacarpal Joint," Operative Techniques in Orthopaedics, **28**(1), 1–5, 2018, doi:10.1053/J.OTO.2017.12.002.

[26] EDUCATION EXHIBIT Extensor Mechanism of the Fingers: MR Imaging-Anatomic Correlation 1, Aug. 2021, doi:10.1148/rg.233025079.

[27] A. Gailey, P. Artemiadis, M. Santello, "Proof of Concept of an Online EMG-Based Decoding of Hand Postures and Individual Digit Forces for Prosthetic Hand Control," Frontiers in Neurology, **0**(FEB), 7, 2017, doi:10.3389/FNEUR.2017.00007.

[28] MathWorks Instrument Control Toolbox Team (2021). BITalino Toolbox, https://www.mathworks.com/matlabcentral/fileexchange/53983-bitalino-toolbox, Aug. 2021.

[29] Y. Koike, M. Kawato, "Estimation of dynamic joint torques and trajectory formation from surface electromyography signals using a neural network model," Biological Cybernetics 1995 73:4, **73**(4), 291–300, 1995, doi:10.1007/BF00199465.

[30] A. Phinyomark, C. Limsakul, P. Phukpattaranont, "A Novel Feature Extraction for Robust EMG Pattern Recognition," Journal of Medical Engineering and Technology, **40**(4), 149–154, 2009.

[31] W. Chen, M. Liu, S. Yue, W. Chen, C.-H. Xiong, W.-R. Chen, B.-Y. Sun, M.-J. Liu, S.-G. Yue, W.-B. Chen, "Design and Implementation of an Anthropomorphic Hand for Replicating Human Grasping Functions," IEEE TRANSACTIONS ON ROBOTICS, **32**(3), 2016, doi:10.1109/TRO.2016.2558193.

[32] A.R. Siddiqi, S.N. Sidek, "Estimation of continuous thumb angle and force using electromyogram classification," Aug. 2021, doi:10.1177/1729881416658179.

[33] S.C. Young, T. Deyle, T. Chen, J.D. Glass, C.C. Kemp, "A list of household objects for robotic retrieval prioritized by people with ALS," 2009 IEEE International Conference on Rehabilitation Robotics, ICORR 2009, 510–517, 2009, doi:10.1109/ICORR.2009.5209484.

[34] J.T. Belter, A.M. Dollar, "Performance characteristics of anthropomorphic prosthetic hands," IEEE International Conference on Rehabilitation Robotics, 2011, doi:10.1109/ICORR.2011.5975476.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 559-565 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Application of Piecewise Linear Approximation of the UAV Trajectory for Adaptive Routing in FANET

Kuzichkin Oleg R.[1,*], Vasilyev Gleb S.[1], Surzhik Dmitry I.[2], Kurilov Igor A.[2]

[1]*Belgorod State Research University, Belgorod, 308015, Russia*

[2]*Vladimir State University named after Alexander Grigoryevich and Nikolai Grigorievich Stoletovs, Vladimir, 600000, Russia*

ARTICLE INFO

Article history:
Received: 03 February, 2021
Accepted: 04 March, 2021
Online: 20 March, 2021

Keywords:
Unmanned aerial vehicle
UAVs
Flying Ad Hoc Network
FANET
Routing
Trajectory approximation
Piecewise linear function

ABSTRACT

A significant problem of routing protocols in the Flying Ad Hoc Networks (FANET) is a significant overhead cost due to the high mobility of networking nodes. The problem is caused by a need to send information messages about locations of unmanned aerial vehicles (UAVs). In order to reduce the amount of service information, the following trajectory approximation algorithms have been investigated: an algorithm for conjugating courses and an algorithm based on continuous piecewise-linear functions (CPLF). Four modifications of the CPLF-based algorithm are considered, which differ in the type of piecewise linear function used: basic CPLF, generalized CPLF, generalized CPLF with a compact notation form, and adaptive CPLF. The disadvantages of each algorithm are analyzed. The CPLF approximation of a fragment of an aircraft trajectory consisting of two straight sections and a curved section with variable steepness between them is performed. It is established that adaptive CPLF with variable step reduces the error of trajectory approximation due to the location of most points on the curved sections of the aircraft maneuvering. The modified version of ADV routing protocol has shown a lower overhead value (the gain for small pause time values reaches 23 %). Thus, the effectiveness of the proposed approximation-based routing in FANET is shown.

1. Introduction

This paper is an extension of work originally presented in the 7th International Conference on Control, Decision and Information Technologies (CoDIT' 20) [1]".

Networks built on the basis of unmanned aerial vehicles (UAVs) (Flying Ad Hoc Network, FANET) [2] are characterized by an unstable network topology due to the movement of network nodes. This aspect should be taken into account when developing and applying routing algorithms. The well-known network layer protocols do not fully meet the requirements of the FANET, so the development of new routing protocols for flying networks remains relevant.

Proactive protocols calculate routes between all network nodes beforehand and maintain a relevant routing table at each node [3]. An integral part of such routing protocols is a connection management mechanism that establishes logical connections between network nodes that are within earshot of each other, and also closes logical connections. All stations distribute information about the established logical connections over the network. Thus, each of the stations decentralizes the logical topology of the network and fills in its own routing table. In addition, proactive routing protocols provide minimal delay in the delivery of data packets, and when the network topology changes, broadcast messages about these changes are initiated [4].

The author proposed a reactive protocol Dynamic Source Routing (DSR) [5]. In reactive routing protocols, routes exist only when they are needed, that is, when data is being transmitted over them. When the route is initially determined, packets are sent in all possible directions, and information about the passed node is added to the header. As a result, when the goal is reached, the packet header contains a fully formed route between the specified nodes. In the event of loops, i.e. re-reception of the first packet, the node destroys this packet. The high mobility of FANET nodes when using proactive protocols requires frequent updates of the routing table, which is their disadvantage.

Hybrid routing protocols combine the advantages of reactive and proactive routing protocols, but require a relatively complex implementation. Their use is associated with the need to divide the

*Corresponding Author: Kuzichkin O.R, Contact No: 910-171-3945, Email: oldolkuz@yandex.ru

network structure into segments, which reduces the efficiency of routing (for example, HWMP-Hybrid Wireless Mesh Protocol).).

Geographic routing protocols use the spatial coordinates of network nodes and are more efficient for FANET, especially when the network is large. An example is the Greedy Perimeter Stateless Routing protocol (GPSR) [6]. Directional antennas-based protocols reduce the number of transit nodes and, as a result, lower the delay time (DOLSR-Directional Optimized Link State Routing Protocol [7]).

A significant problem of routing protocols in the Flying Ad Hoc Networks (FANET) is a significant overhead cost due to the high mobility of networking nodes. The problem is caused by a need to send information messages between nodes about locations of separate unmanned aerial vehicles (UAVs) [8]. Also the use of directional antennas on board the UAV [9-11] necessitates additional transmition of the angular coordinates of each node, since when the airplane turns, the receiving antenna may deviate from the maximum of the radiation pattern, which will reduce the useful signal.

In [1] the application of UAV trajectory approximation for adaptive routing in FANET networks is justified. However, the expressions are given in a fairly brief form. In addition, the approximation example of a trajectory fragment in the form of a sinusoidal function does not allow us to fully evaluate the advantages of the proposed approach, for which it is necessary to model a trajectory consisting of a set of straight sections and curved sections with variable steepness.

2. The algorithm for conjugated courses

Approximation of the UAV trajectory will reduce the overhead caused by the transmission of additional information about the position of network nodes. The use of trajectory approximation algorithms will allow the vehicle coordinates to be transmitted only at the beginning or end of each maneuver. With this approach, it is not necessary to transmit data about intermediate points of the trajectory between maneuvers, which reduces overhead costs.

The creation of the trajectory of an flying vehicle consists of several stages [12]. The first is to set the motion kinematics, which determines the spatial position of the object at an arbitrary time t. For this purpose, the object is assigned a trajectory that can be set by a set of reference vertices with a velocity vector defined at each vertex. Next, the trajectory of the air object is plotted along the reference verteces. The main disadvantage of the method of conjugating courses is a rather large amount of calculations and, accordingly, a decrease in the speed of the algorithm.

The course matching algorithm is based on dividing the trajectory into segments, and each segment is described as a mathematical expression [13]. The input data of the algorithm are the coordinates of the node points of the trajectory. The rectilinear motion of the aircraft is described by parametric equations:

$$x = x_0 + \vartheta_{0x} t + \frac{a_x t}{2}, y = y_0 + \vartheta_{0y} t + \frac{a_y t}{2}, \quad (1)$$

where x and y are current spacial coordinates, x_0 and y_0 are initial values of spacial coordinates, $\vartheta_{0x,y}$ are projections of velocity on x and y axis, a_x and a_y are projections of acceleration on x and y axis, accordingly.

When calculating straight segments of the trajectory, the minimum trajectory curvature radius with a nominal (acceptable) level of overload, it is necessary to calculate the coefficient $K(\varphi_n)$

$$K(\varphi_n) = \begin{cases} \dfrac{1{,}017456}{\sqrt{tg\dfrac{\varphi_n}{2}}} \cdot \left(1 + \dfrac{1}{3 \cdot tg^2 \dfrac{\varphi_n}{2}}\right) & at \quad \varphi_n \geq \dfrac{\pi}{2} \\ \left(0{,}272166 + \dfrac{1{,}014303}{tg\dfrac{\varphi_n}{2}}\right) & at \quad \varphi_n < \dfrac{\pi}{2} \end{cases} \quad (2)$$

where φ_n is course at the segment n of a trajectory.

For circular segments of the trajectory, when calculating speed, one might need to calculate the sine, cosine:

$$\vartheta(t) = \frac{(\varepsilon \cdot (t - t_0) + \omega_0)^2}{\sqrt{x^2 + y^2 + z^2}} \cdot \begin{bmatrix} \cos(\psi(t)) \cdot (x \cdot A_x + y \cdot A_y + z \cdot A_z) - \\ -\sin(\psi(t)) \cdot (x \cdot B_x + y \cdot B_y + z \cdot B_z) \end{bmatrix} \quad (3)$$

where $\vartheta(t)$ is current radial velocity of an aircraft, ε is circular acceleration, t and t_0 is current and initial time, ω_0 is initial circular velocity, x, y, and z are spatial coordinates, ψ is current current manuevering angle, $A_{x,y,z}$ and $B_{x,y,z}$ are coefficients of the movement.

For parabolic segments when calculating the maximum move in a parabola, there is a need also calculate the arc sine:

$$\psi_n = 2\arcsin\left[3tg\frac{\varphi_n}{2}\sqrt{\frac{tg^2\dfrac{\varphi_n}{2} + \dfrac{1}{9}}{\left(tg^2\dfrac{\varphi_n}{2} - \dfrac{7}{72}\right)^2 + 36tg^2\dfrac{\varphi_n}{2} \cdot \left(tg^2\dfrac{\varphi_n}{2} + \dfrac{31}{144}\right)^2}}\right], \quad (4)$$

where ψ_n and φ_n are current manuevering angle and course at the segment n of a trajectory.

Expressions (1-4) describe a smooth trajectory with high accuracy, which is necessary in automatic control systems [14] and simulation systems of full-scale modeling of aircraft [12, 15]. However, routing algorithms do not require such high accuracy in describing the path. The main requirement for the routing algorithm is the choice of optimal transit nodes for a specific network topology, which can be successfully performed with small errors in the representation of the trajectory. Thus, for routing algorithms the very detailed trajectory model is redundant, computationary complex and not required.

3. Algorithms based on continuous piecewise linear functions

Studies have shown that continuous piecewise linear functions (CPLF) based on the modulus of a linear function are effective for approximating nonlinear characteristics, trajectories [16-18]. Calculating the approximation coefficients and the function itself

requires less computational resources compared to the course conjugation algorithm.

A smaller value of the error is achieved at the expense of equality of error in every area of approximation. The use of this approach when approximating straight lines leads to a decrease in the error, since the approximating function "adjusts" (adapts) to the change in the nature of the nonlinearity of the original characteristic.

The approximation of the trajectory using the adaptive CPLF is schematically shown in Figure 1. The function section is divided into segments. The approximation nodes are calculated as follows: a fixed error value is set, then the number of segments (approximation nodes) is determined by the algorithm. The length of segment xi varies depending on the value of approximation error ε.

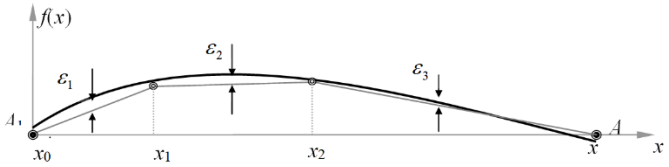


Figure 1: The principle of trajectory approximation using CPLF

The CPLF approximation assumes the replacement of the object's trajectory along an arc with the trajectory of movement along a straight line. The time and velocity modules are selected according to the reference algorithm [12]. As a result, in both cases (trajectories $f(x)$ and F_i) start and end speed is the same and time spent is also the same. This means that each point of the reference segment is projected onto the corresponding point of the CPLF segment. For example, consider a section of the trajectory of Figure 1 in the interval $[x_0, x_1]$. It is shown in Figure 2. It can be seen that the error in determining the position of the object on the segment is equal to the maximum length of the projection line. In the adaptive CPLF, the errors ε_i are aligned with each other. The geometric meaning is shown in Figure 3.

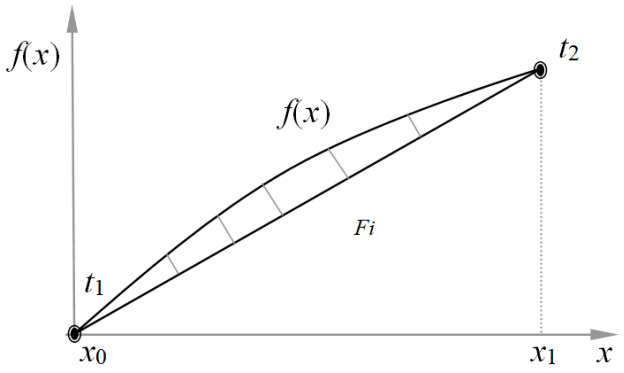


Figure 2: Fragment of the trajectory

The actual error is the distance between the position of the object on the reference curve and its position on the approximating curve. Figure 3 shows the discrepancy between the actual error and the approximation error. Knowing that $\varepsilon_1 = \varepsilon_2 = ... = \varepsilon_n = ... = \varepsilon_{N-1}$ we estimate the range of changes in the error value δ_i. From the right triangle ABC in Figure 4, we obtain $\sin(\pi/2 - \alpha_i) = \delta_i / \varepsilon_i$, where α_i is the angle between the CPLF and the reference trajectory. Thus

$$\delta_i = \varepsilon_i \sin\left(\frac{\pi}{2} - \alpha_i\right). \quad (5)$$

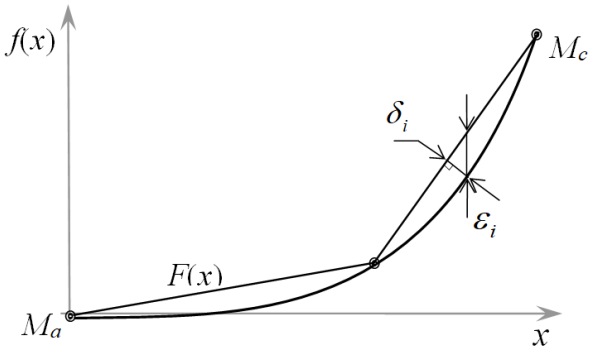


Figure 3: Determination of the approximation error

To estimate the error δ_i we need to know the error α_i. The minimum value $\alpha_0 = 0^o$ (at the origin in Figure 4), $\delta_{i\min} = \varepsilon_i \sin\left(\frac{\pi}{2}\right)$. The maximum value $\alpha_i = \frac{\pi}{2} - \frac{\varphi}{2}$ and $\delta_i = \varepsilon_i \sin(\varphi/2)$, therefore the actual error δ_i is determined by the inequalities $\delta_{i\min} < \delta_i < \delta_{i\max}$ and $\sin\frac{\pi}{2} < \delta_i < \varepsilon_i \sin\left(\frac{\varphi}{2}\right)$. It follows from the inequality that the greater the value of φ, the smaller the error δ_i, and δ_i (the side of the triangle ABC) does not exceed ε_i. This means that ε_i is the maximum error of the approximation of the reference trajectory using the CPLF.

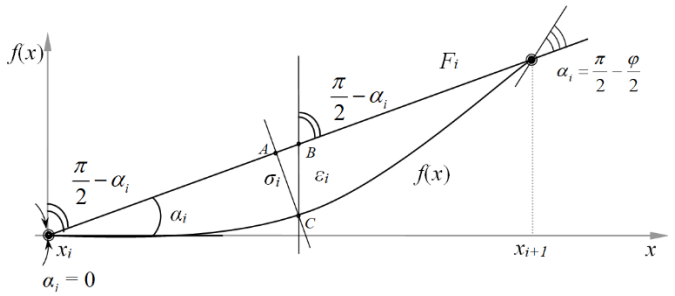


Figure 4: Determination of the linear and angular approximation error

3.1. Basic CPLF

It has a compact form and contains only one approximation coefficient. The basic CPLF is convenient for representing functions and nonlinear characteristics approximated with a fixed argument step, with a small number of approximation nodes. The base CPLF has the form:

$$f_M(t) = b_M \left| t - t_0 - \Delta t \cdot M \right|, \quad (6)$$

where t_0 is an initial shift of the CPLF relative to the origin;

Δt is a shift of the modules relatively to each other;

M=0,1,2,3 is a count of the number of modules;

b_M is approximation coefficient.

Summing up the four functions shifted relative to each other by a fixed value, we get the basic CPLF (Figure 5):

$$\begin{aligned} &b_0 \cdot |t - t_0| + b_1 \cdot |t - t_0 - \Delta t| + b_2 \cdot |t - t_0 - 2 \cdot \Delta t| + \\ &+ b_3 \cdot |t - t_0 - 3 \cdot \Delta t| = f(t) \end{aligned} . \quad (7)$$

We determine the coefficients b_M. To do this, using the four points $(t_0;0)$; $(t_0 + \Delta t; y_1)$; $(t_0 + 2 \cdot \Delta t; y_2)$; $(t_0 + 3 \cdot \Delta t;0)$ shown in the figure, we make a system of four equations:

$$\begin{cases} b_0 \cdot |0| + b_1 \cdot |-\Delta t| + b_2 \cdot |-2 \cdot \Delta t| + b_3 \cdot |-3 \cdot \Delta t| = 0 \\ b_0 \cdot |\Delta t| + b_1 \cdot |0| + b_2 \cdot |-\Delta t| + b_3 \cdot |-2 \cdot \Delta t| = y_1 \\ b_0 \cdot |2\Delta t| + b_1 \cdot |\Delta t| + b_2 \cdot |0| + b_3 \cdot |-\Delta t| = y_2 \\ b_0 \cdot |3\Delta t| + b_1 \cdot |2\Delta t| + b_2 \cdot |\Delta t| + b_3 \cdot |0| = 0. \end{cases} \quad (8)$$

After the transformations, the system takes the form:

$$\begin{cases} 0 + \Delta t \cdot b_1 + 2 \cdot \Delta t \cdot b_2 + 3 \cdot \Delta t \cdot b_3 = 0 \\ \Delta t \cdot b_0 + 0 + \Delta t \cdot b_2 + 2 \cdot \Delta t \cdot b_3 = y_1 \\ 2 \cdot \Delta t \cdot b_0 + \Delta t \cdot b_1 + 0 + \Delta t \cdot b_3 = y_2 \\ 3 \cdot \Delta t \cdot b_1 + 2 \cdot \Delta t \cdot b_1 + \Delta t \cdot b_2 + 0 = 0. \end{cases} \quad (9)$$

Solving system (9) by the Gauss method [20], we obtain:

$$b_0 = \frac{y_1}{2 \cdot \Delta t};\ b_1 = \frac{y_2 - 2 \cdot y_1}{2 \cdot \Delta t};\ b_2 = \frac{y_1 - 2 \cdot y_2}{2 \cdot \Delta t};\ b_3 = \frac{y_2}{2 \cdot \Delta t}, \quad (10)$$

where y_1, y_2 are the function values in the approximation nodes.

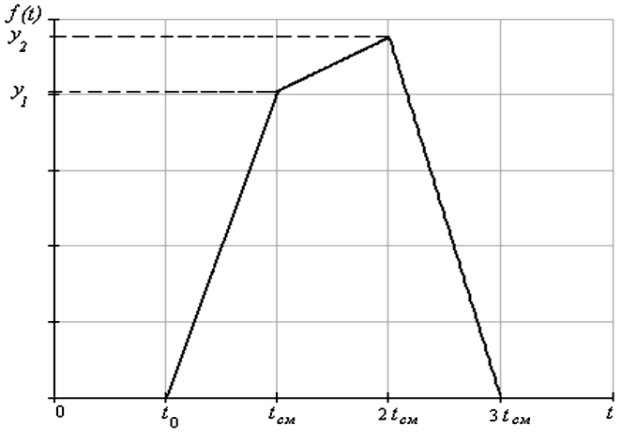


Figure 5: Basic CPLF

In order not to use separate expressions for each coefficient b_0, b_1, b_2, b_3, we present them as a single coefficient B_M:

$$\begin{aligned} B_M = &\frac{(M-1)(M-2)(M-3)}{-3} \cdot \frac{y_1}{2 \cdot \Delta t} + \\ &+ \frac{M(M-2)(M-3)}{1} \cdot \left(\frac{y_2 - 2 \cdot y_1}{2 \cdot \Delta t} \right) + \\ &+ \frac{M(M-1)(M-3)}{-1} \cdot \left(\frac{y_1 - 2 \cdot y_2}{2 \cdot \Delta t} \right) + \\ &+ \frac{M(M-1)(M-2)}{3} \cdot \frac{y_2}{2 \cdot \Delta t}. \end{aligned} \quad (11)$$

We transform equation (11):

$$B_M = y_2 \frac{5M^3 - 21M^2 + 19M}{6 \cdot \Delta t} - y_1 \frac{5M^3 - 24M^2 + 28M - 3}{6 \cdot \Delta t}. \quad (12)$$

Substituting the coefficient B_M, we finally get:

$$f(t) = \sum_{M=0}^{3} B_M |t - t_0 - \Delta t \cdot M|. \quad (13)$$

The disadvantages of the basic CPLF include a cumbersome expression for determining the approximation coefficients B_M. In addition, to approximate functions over a large number of points (more than three), it is necessary to add up individual base CPLFS shifted between each other.

3.2. Generalized CPLF

The generalized CPLF reduces the computational cost of approximation using the basic CPLF. This function is the sum of a few basic CPLF. Using the generalized CPLF, we can approximate functions of any kind based on a single expression for the approximating function, rather than a large number of expressions:

$$F(t) = \sum_{n=1}^{N} \sum_{M=0}^{3} B_{M,n} |t - [t_0 + 2 \cdot \Delta t(n-1)] - \Delta t \cdot M|, \quad (14)$$

where

$$B_{M,n} = y_{n2} \frac{5M^3 - 21M^2 + 19M}{6 \cdot \Delta t} - y_{n1} \frac{5M^3 - 24M^2 + 28M - 3}{6 \cdot \Delta t}$$

is the approximation coefficient, n is the number of basic CPLF for the approximation; $t_0 + 2 \cdot \Delta t$ is the initial offset of the second approximating function; $y_{n1} = y(t_0 + \Delta t \cdot (2n-1))$ are values of the initial characteristic in odd approximation nodes;

$y_{n2} = y(t_0 + \Delta t \cdot 2n)$ are values of the initial characteristic in even approximation nodes.

The disadvantages of the generalized CPLF include additional terms under the module and two sum counters, which complicates the computational process. A generalized CPLF can have a compact notation form and include only one sign of the sum.

3.3. Generalized CPLF with compact notation

We design a function that has a compact expression and allows us to perform an approximation with an arbitrary step. It is the sum of the modules of linear functions with different coefficients in front of them. Expressions under modules represent shifts of linear functions in the time domain. These shifts determine the approximation step. Thus, shifting the individual linear functions by a different value t_n we get the CPLF:

$$F(t) = \sum_{n=1}^{N} K_n \cdot |t - t_n|, \quad (15)$$

where n is the sum counter; t_n is the approximation step; K_n is the approximation coefficient.

We define an expression for finding the coefficient K_n. To do this, we approximate the function $y = f(t)$ $n = N$ points. Let the first approximation node correspond to the value $t = t_1$,

and the value of the approximated function $y_1 = f(t_1)$. The expression for the CPLF at this point:

$$F(t_1) = \sum_{n=1}^{N} K_n \cdot |t_1 - t_n| \quad (16)$$

or

$$\sum_{n=1}^{N} K_n \cdot |t_1 - t_n| = y_1. \quad (17)$$

The law of motion along each straight segment can be given by a function of the parameter t:

$$x(t) = v_{0x}t + \frac{a_x t^2}{2}, \quad y(t) = v_{0y}t + \frac{a_y t^2}{2}, \quad z(t) = v_{0z}t + \frac{a_z t^2}{2}. \quad (18)$$

By decomposing the expression for the CPLF into separate modules we get:

$$K_1 \cdot |t_1 - t_1| + K_2 \cdot |t_1 - t_2| + ... + K_{N-1} \cdot |t_1 - t_{N-1}| + K_N \cdot |t_1 - t_N| = y_1. \quad (19)$$

We perform a similar operation for the second node $t = t_2$, which corresponds to the value of the approximated function $y_2 = f(t_2)$ and the expression for the adaptive CPLF at this node:

$$F(t_2) = \sum_{n=1}^{N} K_n \cdot |t_2 - t_n| \quad (20)$$

or

$$\sum_{n=1}^{N} K_n \cdot |t_2 - t_n| = y_2. \quad (21)$$

As a result, for N approximation nodes we obtain a system of N equations:

$$\begin{cases} K_1|t_1 - t_1| + K_2|t_1 - t_2| + ... + K_{N-1}|t_1 - t_{N-1}| + K_N|t_1 - t_N| = y_1 \\ K_1|t_2 - t_1| + K_2|t_2 - t_2| + ... + K_{N-1}|t_2 - t_{N-1}| + K_N|t_2 - t_N| = y_2 \\ \vdots \\ K_1|t_{N-1} - t_1| + K_2|t_{N-1} - t_2| + ... + K_{N-1}|t_{N-1} - t_{N-1}| + K_N|t_{N-1} - t_N| = y_{N-1} \\ K_1 \cdot |t_N - t_1| + K_2|t_N - t_2| + ... + K_{N-1}|t_N - t_{N-1}| + K_N|t_N - t_N| = y_N. \end{cases} \quad (22)$$

We find the values of the coefficients $K_1 - K_N$ by the inverse matrix method. In matrix form, the system will take the form:

$$[K] \cdot [T] = [y]. \quad (23)$$

Hence, the desired coefficients $K_1 - K_N$ have the form:

$$[K] = [y] \cdot [T]^{-1}. \quad (24)$$

3.4. Adaptive CPLF

Approximating trajectories with a fixed step is impractical if the trajectories contain long straight sections. To represent each linear fragment, two approximation points are sufficient, and it is advisable to place most of the approximation points on curved sections. Thus, it is promising to apply an adaptive CPLF with a variable approximation step.

As a criterion for the quality of the approximation, we use the minimum mean square error (MSE) criterion. The expression for the approximation MSE in the range (t_0, t_N) has the form

$$\sigma^2 = \int_{t_0}^{t_N} [y(t) - F(t)]^2 dt. \quad (25)$$

The approximation is carried out based on finding the vector $\mathbf{T} = (t_1 \ \dots \ t_{N-1})^T$ which fits to the condition of the mimumum MSE value:

$$\sigma^2(\mathbf{T}) \rightarrow min. \quad (26)$$

Adaptive CPLF allows to approximate nonlinear characteristics and functions with a lower error than CPLF with a fixed approximation step. This is achieved by obtaining an equal error in each approximation section. The use of this approach when approximating with lines leads to a decrease in the error, since the approximating function "adjusts" (adapts) to the change in the nature of the nonlinearity of the original characteristic.

Equal approximation error in each section will reduce the total error. To obtain an equal approximation error in each section, the following condition must be met:

$$\delta_1 = \delta_2 = ... = \delta_n. \quad (27)$$

Based on the MSE expressed by Equation (25), we write:

$$\int_{t_0}^{t_1} [y(t) - F(t)]^2 dt = \int_{t_1}^{t_2} [y(t) - F(t)]^2 dt = ... = \int_{t_{N-1}}^{t_N} [y(t) - F(t)]^2 dt. \quad (28)$$

Let's represent the integrand expressed in Equation (28) as follows:

$$g(t) = [y(t) - F(t)]^2. \quad (29)$$

$$\delta = \int_{t_l}^{t_k} g(t)dt = G\Big|_{t_l}^{t_k} = G(t_k) - G(t_l), \quad (30)$$

where G is antiderivative of the function $g(t)$.

We find the points for the approximation of an arbitrary function based on the adaptive CPLF with a lower error. To find these points, we define a system of equations:

$$\begin{cases} G(t_1) - G(t_0) = G(t_2) - G(t_1) \\ G(t_1) - G(t_0) = G(t_3) - G(t_2) \\ \vdots \\ G(t_1) - G(t_0) = G(t_N) - G(t_{N-1}). \end{cases} \quad (31)$$

From the system of equations (31), the values of t_n are determined, which approximate the studied characteristic with an equal error in each section. The values of the function $y(t_n)$ at the points t_n are determined from the known expression for the original function y (t).

4. Approximation of the trajectory fragment of the aircraft based on the CPLF

The solution of the minimization problem (26) was performed on the basis of the nonlinear conjugate gradient method [19,20],

the value of convergence tolerance parameter of 10^{-6} was accepted. Evaluation of the approximation MSE based on Equation (25) is shown in Fig. 3.

The approximation result of the aircraft trajectory fragment $\widetilde{F}_{1,2}(t)$ by N=16 nodes and the absolute approximation error $\Delta_{F1,2}(t)$ are shown in Figures 6 and 7. Here, the lower index 1 corresponds to the uniform law for the approximation nodes, and the index 2 corresponds to the optimal law. The dots in the figures show the approximation nodes. The values of the source function and the approximating function at the nodes are equal, so the error at these points is zero. The MSE values are σ^2_{F1}=1,187·10^{-3}; σ^2_{F2}=1,982·10^{-5}. With the optimal nodes, most of them are located on a curved section, where the trajectory changes most quickly, which gives a significant decrease of MSE (by 59.9 times).

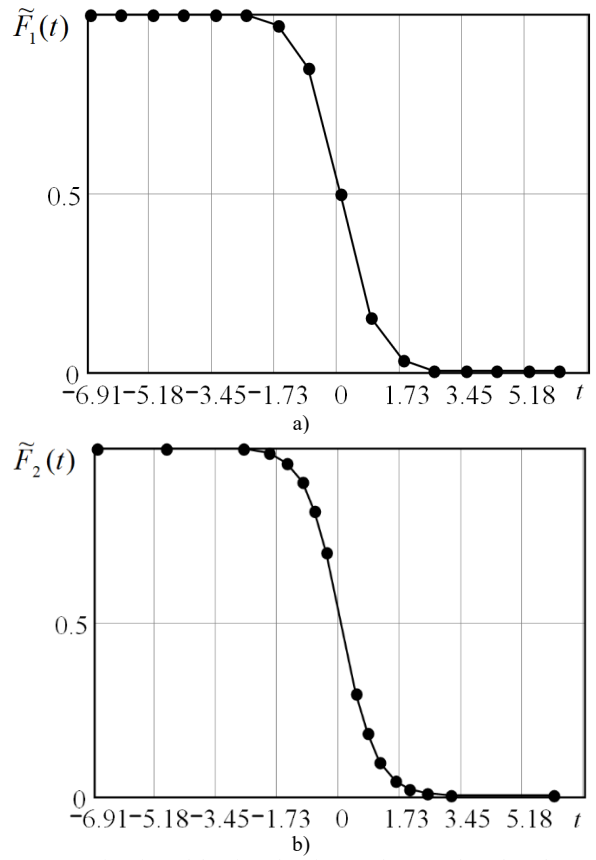
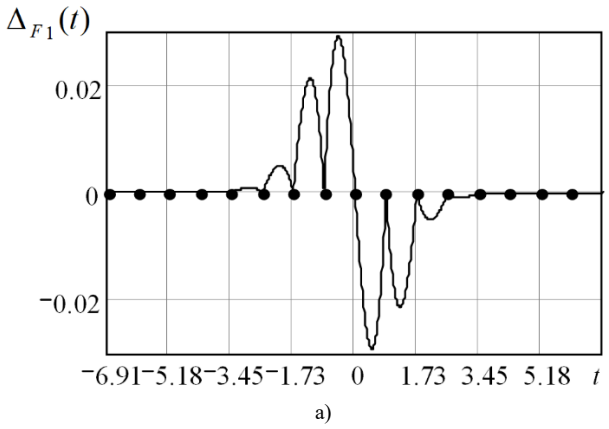


Figure 6: Approximation of the aircraft trajectory fragment based on the CPLF: a) with a uniform step; b) with an adaptive step



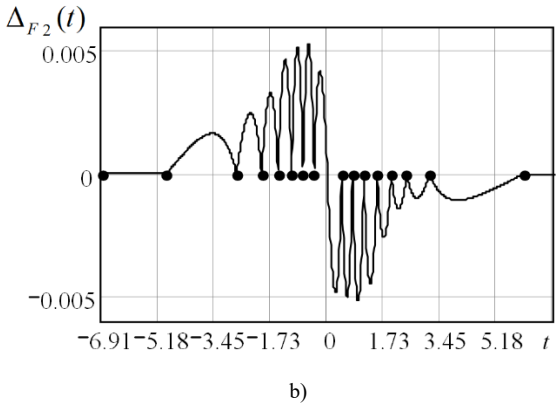


Figure 7: Approximation error of the aircraft trajectory fragment based on the CPLF: a) with a uniform step; b) with an adaptive step

5. Modeling of approximation-based routing in FANET

A comparative simulation of the UAV network with two routing protocols was performed: the standard ad-hoc on-demand vector (AODV) and the modified one. In the modified version of the protocol, Route Request packets were not transmitted at a fixed time interval, as in the standard version, but only when the course changed by more than 200. Otherwise, Route Request packets were not transmitted, and the coordinates of the nodes obtained from the past coordinate values by the piecewise linear approximation were used to update the route table. Changes in the course of individual nodes were taken into account in the node mobility model (the OPNET modeler 14.5 simulator was used). Thus, the modified routing protocol is a hybrid of the standard AODV protocol and geographic-based routing. When modeling the network, the following parameters were obtained:

- the number of network nodes - 20;
- channel bandwidth – 2 Mbit/s;
- the power of the UAV receiving and transmitting device is 10 mW;
- threshold receiving power-minus 90 dB;
- the maximum data packet size is 1024 bytes;
- MAC layer protocol - CSMA/CA (carrier sense multiple access with collision avoidance)
- network size 10x10 km;
- average speed of mobile nodes is 10 m/s.

Mobility model – Random-Waypoint [21]

During the simulation, the Pause time parameter of a mobility model (time between changes in direction and/or speed) was varied and the routing overhead in percent was determined for the standard and modified AODV. The simulation results are shown in Figure 8. The overhead value of the standard AODV is in good agreement with the results in [22]. The modified version of AODV routing protocol shows a lower overhead value (the gain for small pause time values reaches 23 %). Thus, the effectiveness of the proposed approximation-based routing in FANET is shown.

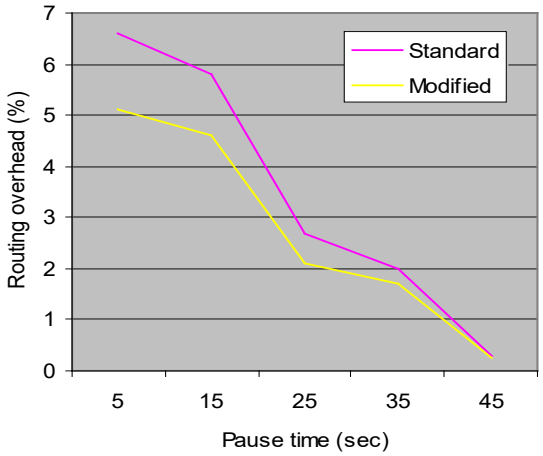


Figure 8: Routing overhead vs Pause time for a standard and modified AODV

6. Conclusion

A serious problem of routing protocols in FANET is the significant overhead caused by the high mobility of network nodes. The problem is caused by a need to send information messages about locations of unmanned aerial vehicles (UAVs). The perspective of algorithms for constructing the UAV trajectory to reduce the amount of service information about the UAV position transmitted via the FANET routing protocols is substantiated. When using such algorithms, information about the coordinates of the nodes is directly transmitted only in the areas of maneuvering of the aircraft, and in the areas of rectilinear motion is calculated based on the algorithms for approximating the trajectory. In order to reduce the amount of service information, the following trajectory approximation algorithms have been investigated: an algorithm for conjugating courses and an algorithm based on continuous piecewise-linear functions (CPLF). Four modifications of the CPLF algorithm are considered which differ in the type of function used: basic CPLF, generalized CPLF, generalized CPLF with a compact notation form, and adaptive CPLF. The advantages and disadvantages of each algorithm are analyzed. Adaptive CPLF with variable approximation step allows lower computational resources compared to the rate-matching algorithm, as well as more accurate compared to the CPLF with a fixed step.

An approximation of a trajectory fragment consisting of two straight sections and a curved section with a variable steepness between them (approximately described by the arctangensoidal function) was performed. The computational experiment has shown the effectiveness of the direct approach: in the considered example with long straight sections, the approximation MSE by the adaptive CPLF is 59.9 times less compared to the fixed step CPLF. The modified version of ADV routing protocol shows a lower overhead value (the gain for small pause time values reaches 23 %). Thus, the effectiveness of the proposed approximation-based routing in FANET is shown.

Acknowledgment

The work was supported by the RFBR grant 19-29-06030-mk "Research and development of a wireless ad-hoc network technology between UAVs and smart city control centers based on the adaptation of transmission mode parameters at various levels of network interaction". The theory was prepared within the framework of the state task of the Russian Federation FZWG-2020-0029 "Development of theoretical foundations for building information and analytical support for telecommunications systems for geoecological monitoring of natural resources in agriculture".

References

[1] D.I. Surzhik, G.S. Vasilyev, O.R. Kuzichkin, "Development of UAV trajectory approximation techniques for adaptive routing in FANET networks," 7th International Conference on Control, Decision and Information Technologies (CoDIT'2020), June 29-July 2, 2020, 1226-1230, doi: 10.1109/CoDIT49905.2020.9263944.

[2] K. Palan, P. Sharma, "FANET Communication Protocols: A Survey," International Journal of Computer Science & Communications (IJCSC), **7**(1), 219-223, 2016, doi: 10.090592/IJCSC.2016.034.

[3] A. Kiryanov, "Analysis of mechanisms for constructing logical topology in MANET networks," Information Processes, **15(2)**, 183-197, 2015.

[4] D.E. Prozorov, "Protocol of hierarchical routing of a ad-hoc mobile network," Radio Engineering and telecommunications systems, **15**(3), 74-80, 2014.

[5] T. Brown, S. Doshi, S. Jadhav, D. Henkel, R. Thekkekunnel, "A full scale wireless ad hoc network test bed," Proc. of International Symposium on Advanced Radio Technologies, 50–60, 2005.

[6] K. Brad, K. Hsiang, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," Proceedings of the Annual International Conference on Mobile Computing and Networking (MOBICOM), 2020, doi:10.1145/345910.345953.

[7] A.I. Alshabtat, L. Dong, J. Li, F. Yang, "Low latency routing algorithm for unmanned aerial vehicles ad-hoc networks," International Journal of Electrical and Computer Engineering, **6**(1), 48–54, 2010.

[8] R.V. Kirichek, "Development and research of a complex of models and methods for flying sensor networks", Doct. thesis of technical sciences, St. Petersburg, 2017.

[9] A.I. Alshabtat, Cross-Layer Design For Mobile Ad-Hoc Unmanned Aerial Vehicle Communication Networks, PhD Thesis, Western Michigan University, 2011.

[10] N. Ahmed, S. Kanhere, S. Jha, "Link characterization for aerial wireless sensor networks," GLOBECOM Wi-UAV Workshop, 1274–1279, 2011, doi: 10.1109/JSAC.2013.130825.

[11] T. Samil, B. Ilker, "LODMAC: Location Oriented Directional MAC protocol for FANETs," Computer Networks, **83**(4), 76–84, 2015, doi: 10.1016/j.comnet.2015.03.001.

[12] M.S. Bobrov, A.M. Averyanov, G.G. Piskunov, V.V. Chekushkin, "Implementation of air object movement routes in simulator-modeling systems," Questions of radio electronics, **4**, 157-166, 2009.

[13] V.N. Burkov, 'Adaptive predictive flight control systems", Moscow, Nauka. Phys.-mat. lit., 1987.

[14] A. Tewar, "Advanced Control of Aircraft, Spacecraft and Rockets", John Wiley & Sons, 2011.

[15] I.A. Kurilov, D.V. Pavelev, D.N. Romanov, "Review of methods for constructing the trajectory of an air object in a simulator-modeling complex," Radio Engineering and Telecommunications Systems, **1**, 42-45, 2011.

[16] G.S. Vasilyev, O.R. Kuzichkin, D.I. Surzhik, I.A. Kurilov, "Algorithms for analysis of stability and dynamic characteristics of signal generators at the physical level in FANET networks," MATEC Web of Conferences 309, 03019 (2020), doi: https://doi.org/10.1051/matecconf/202030 903019.

[17] I.A. Kurilov, G.S. Vasiliev, S.M. Kharchuk, "Analysis of dynamic characteristics of signal converters based on continuous piecewise linear functions," Scientific and technical Bulletin of the Volga region, **1**, 100-104, 2010.

[18] I.A. Kurilov, G.S. Vasilyev, S.M. Kharchuk, D.I. Surzhik, "Research of static characteristics of converters of signals with a nonlinear control device," 2011 International Siberian Conference on Control and Communications (SIBCON), 93 – 96, 2011.

[19] Yu.A. Maksimov, E.A. Fillipovskaya, 'Algorithms for solving problems of nonlinear programming", Moscow: MEPhI, 1982.

[20] Y.-H. Dai, Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property, " SIAM J. Optim, **1**, 177–182, 1999.

[21] J. Yoon, M. Liu, B. Noble, "Random Waypoint Considered Harmful," 0-7803-7753-2/03, IEEE INFOCOM, 2003, doi: 10.1109/INFCOM.2003.1208967.

[22] M. Sheetal, "Performance Comparison of Ad-hoc Routing Protocols," International Journal of Advance research, Ideas and Innovations in Technology, **2**(5), 1-8, 2016, doi: 10.3923/itj.2005.278.283.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 719-725 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A Recommendation Approach in Social Learning Based on K-Means Clustering

Sonia Souabi*, Asmaâ Retbi, Mohammed Khalidi Idrissi, Samir Bennani

RIME TEAM-Networking, Modeling and e-Learning Team- MASI Laboratory- Engineering 3S Research center. Mohammadia School of Engineers (EMI), Mohammed V University, Rabat, 10000, Morocco

ARTICLE INFO

Article history:
Received: 01 December, 2020
Accepted: 14 January, 2021
Online: 05 February, 2021

Keywords:
Social Networks
Hybrid Recommendation System
K-means
Correlation
Co-occurrence

ABSTRACT

E-learning, among the most prominent modes of learning, offers learners the opportunity to attend online courses. To improve the quality of online learning, social learning through social networks promotes interaction and collaboration among learners. As part of the learning process management in these environments, the implementation of recommendation systems facilitates the provision of content adapted to the needs and requirements of learners and generates recommendations likely to arouse their interest. Many researchers have been involved in several recommendation techniques such as the development of Machine Learning algorithms and the incorporation of social interactions between learners. However, the behavior within a learning environment can diverge from one learner to another. This must therefore be taken into consideration when generating recommendations, i.e., it is initially important to form groups of homogeneous learners prior to proposing recommendations. In this respect, the recommendations generated will be more appropriate to the learners' profiles and level of interaction. On this basis, we raise an important issue which is the importance of grouping learners into homogeneous groups in a recommendation system. In the recommendation system we advocate, we group learners based on the degree of interaction within the learning environment before generating the recommendation list based on a hybrid approach for each cluster. The overall system is, therefore, based on the identification of communities based on the k-means algorithm and the generation of recommendations list for each community separately. Finally, we compare the results of the system integrating the classification of learners as a preliminary step to the system excluding the k-means algorithm. The results reveal that the integration of the clustering algorithm leads to improvements in terms of performance and accuracy.

1. Introduction

This paper is an extension of the work originally presented at the Fourth International Conference on Intelligent Systems and Computer Vision [1]. Currently, e-learning is emerging on a large scale around the world. When face-to-face learning is not available, e-learning allows students to bring order to their ideas and to continue to lead students in their learning process [2], [3]. Social networks, a modernized version of learning, fully promotes interaction and collaboration among learners [4], [5]. When a traditional learning platform is sometimes unable to provide effective collaboration between learners, social networks offer a multitude of options for learners encouraging responsiveness between them. However, the availability of online courses is not the only factor that contributes to the success of the learning process. Learners require a more highly organized learning environment responsive to their needs and profile. In this respect, recommendation systems are the perfect medium to perform this task and to manage learning resources within e-learning [6], [7]. The role of recommendation systems is to filter the information that aims to present different learning objects. Through recommendation systems, a certain organization is implemented within the learning environment and learners are more able to explore what is most interesting to them. Many recommendation systems have been proposed in the literature, including: the content-based approach, collaborative filtering and hybrid approaches [8]-[10]. Within the e-learning context, each proposed recommendation system addresses specific questions and handles an underlying issue identified at the recommendation system level. By way of example, several researchers address collaborative filtering to generate recommendations to learners [11]. Others

*Corresponding Author: Sonia Souabi, soniasouabi@research.emi.ac.ma

focus their attention on hybrid approaches to improve the performance of their recommendation system. Machine Learning algorithms have also been proposed as part of learning recommendation systems, including supervised and unsupervised algorithms [12]. Most recommendation systems use explicit feedback from learners, i.e., assessment by learners to generate recommendations. However, explicit feedback is sometimes an unreliable indicator for recommendations. On the other hand, an e-learning environment can include hundreds or thousands of learners with varying degrees of interactivity, constituting divergent groups, interactive and non-interactive learners. Most studies conducted on recommendation systems do not consider that point with respect to online learning. In this regard, we propose to group learners into homogeneous groups before generating recommendations. Our work aims at implementing a preliminary step in the calculation of recommendations, namely the detection of communities with the same level of interactivity. In our previous work, we proposed a hybrid recommendation system based on the activities performed by the learners. Within the same framework, the current work presents a continuity of the previous work while improving the old system with a new recommendation strategy integrating the k-means algorithm as a preliminary step. In this sense, we propose to classify learners according to well-defined criteria of interactivity within a social learning network, and then generate recommendations for each community separately. In this way, each community will receive its own recommendations adapted to its specific needs. The approach we propose brings considerable added value as it categorizes learners according to their level of interactivity before calculating recommendations.

The document is divided as follows: Section 2 deals with the research work related to the recommendation systems proposed in e-learning, especially those that consider Machine Learning algorithms. Section 3 explains the global recommendation approach based on correlation, co-occurrence and the k-means algorithm, section 4 presents the results of the tests performed on the approach without k-means and the approach with k-means with interpretation, and section 5 summarizes the work done in the paper.

2. Literature review

2.1. Recommender systems in E-learning

There are several types of recommender systems: content-based recommender systems, collaborative filtering recommender systems and hybrid recommender systems. Hybrid recommender systems generally combines between content-based techniques and collaborative filtering approaches. In E-learning, all these types of recommendations were addressed, but differently from one proposition to another. Some researchers discuss the analysis of activities in collaborative filtering within the educational field [13]. This work aims to use ontology and the semantic web to provide efficient recommendations. Others propose a recommendation approach to guide learners in developed countries to select more appropriate resources [14]. Calculations are based on developed knowledge and rating predictions. A recommendation system was suggested based mainly on two primary steps: pre-processing and prediction [15]. Multiple algorithms was used: SVM, KNN, Random Forest, and Naïve Bayes. A personalized recommendation system was developed based mainly on ontology using the java programming language [16]. Learning style and level of knowledge were addressed in a proposed recommender system [17]. The approach contains four modules relating to courses and learners.

2.2. Related works

In Machine Learning, unsupervised learning includes algorithms that should perform based on unannotated examples. K-means, one of the important unsupervised algorithms, is used to identify clusters with the similar characteristics and properties by working with gravity centres. The optimal number of clusters is usually obtained through several methods, including: elbow method, average silhouette method, gap statistic method. In online learning, k-means is handled for many purposes, for instance classifying learners according to their attitudes, their performance, interaction rates (...). In the literature, there is a multitude of works addressing k-means as an algorithm for classifying not only learners, but also items, users, in different contexts and areas, such as recommender systems.

A learning content recommendation system was implemented within a learning platform to generate recommendations in an intelligent way based on the learners' interest [18]. The proposed system aims to combine ontology and clustering algorithms using the collected ontology. A new recommendation system that combines the content-based approach and the k-means algorithm simultaneously was proposed [19]. The approach consists of performing a transformation on user data. Then, learners are grouped into clusters based on the content approach and the clustering algorithm, and thus recommendations will be generated from the detected clusters. Some researchers focus on methods to optimize k in the clustering algorithms in order to maintain the variance of each cluster [20]. The clustering was based on movie genres and tags. The objective is to correctly evaluate the recommendation algorithm based on user classification and use sophisticated measures such as mean square error, proximity centrality. Abnormal profiles in the context of recommendation systems was detected. A hybrid recommendation system that integrates both the k-means algorithm and ant colony optimization was suggested [21]. The evaluation was based on several measures including precision, recall and accuracy. Several approaches are based on clustering techniques in recommendation systems that are mainly based on group preferences. Several algorithms were compared to identify which algorithm gives better accuracy and higher results [22]. Some use k-means, others use k-NN, and the results were compared in terms of performance and number of clusters obtained. Instead of working with a traditional clustering algorithm, a mutli-clustering approach working on a set of clusters was proposed [23]. The advantage of this algorithm is its ability to visualize the neighbourhood in a clearer and more refined way, and the time efficiency is high. Contextual information was used to generate more relevant recommendations [24]. The proposed technique seeks to use the k-means algorithm to cluster contexts and generate new user matrices.

Based on researches in terms on recommender systems based on clustering algorithms , it comes out that k-means has been approached in many areas, such as e-learning, e-commerce, movies in order to generate more appropriate recommendations. Thus, we can notice that clustering algorithms has been adopted in recommender systems for different purposes in all the cited works.

On the other hand, clustering algorithms have not been adequately addressed in e-learning. In our context, we propose to use the k-means algorithm within a pure learning context and to combine it with a hybrid recommendation approach in order to generate more relevant recommendations. Our goal is to classify learners according to their degree of interactivity before calculating recommendations to generate a more relevant list of recommendations.

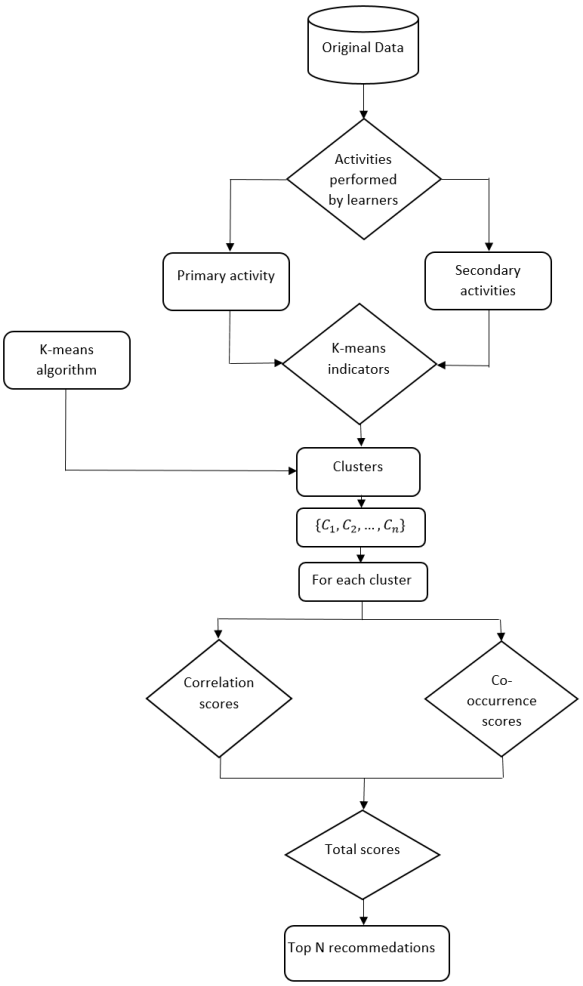


Figure 1: The recommendation system model proposed

3. Our proposed approach

3.1. K-means algorithm

Before proceeding to the description of our recommendation system, we will first define the k-means algorithm. The k-means algorithm is considered among the most widespread clustering algorithms. It consists of the analysis of a dataset containing several descriptors. The final goal is to group similar data in the same clusters. The idea is straightforward, all you need to do is to calculate the distance between the descriptors of two data to decide whether they are part of the same cluster. In a set of data, and to implement the k-means algorithm, there are the so-called centroids. Centroids are the key elements of the algorithm. The k-means relies on the centroids to detect the data closest to each centroid, and this loop continues until there is no further improvement in the distances between similar data, i.e. until the algorithm converges. The algorithm below summarizes the different parts of the k-means algorithm (algorithm 1).

Algorithm 1: Concept of K-means Algorithm
Input
K optimum number of clusters
Output
Clusters or groups
Method
Selecting randomly K points. These are the centroids.
REPEAT
(1): Assign each point to the closest group according to centroids.
(2): Modifying centroids according to new clusters.
UNTIL CONVERGENCE

3.2. The proposed recommender system

The following flowchart (Figure 1) summarizes our recommendation approach.

The flowchart consists of several parts core to the proposed approach:

- The learners' actions in the database (sharing, communicating, downloading ...).
- The reconsideration of the initial data.
- The k-means algorithm.
- The different learning resources.
- The recommendation part.

In this part, we will describe each phase of the proposed approach separately:

a- First phase: The first phase consists in detecting the learners' actions in the database that have been carried out within the learning environment, for instance: sharing learning resources, sending private messages to other learners, downloading resources, exploiting resources. These actions are then grouped into two main categories: primary and secondary actions. The primary action is the action that can directly reflect learners' real preferences, for example, liking content, searching for content. It is through this indicator that we can ascertain learners' preferences. Secondary actions are also indicators of learner preferences, but are less consistent than the primary action.

b- Reconsideration of the initial data: This part consists in converging the data table into data that translate the interaction between learners and each learning object with respect to the selected actions, i.e. it is mandatory to restructure the database so that it meets our requirements for calculating recommendations.

c- The k-means algorithm: After restructuring the database, the next step consists in grouping the learners in clusters by selecting a number of descriptors. Each learner has a number of descriptors which are used to build the most similar clusters. Therefore, the choice of descriptors is an obvious prerequisite for applying the k-means algorithm.

d- Recommendations part: the last part consists of calculating the recommendations after detecting the clusters of the most similar learners. We will therefore generate the recommendations for each cluster separately. The idea is to consider the notions of correlation and co-occurrence in the calculation of recommendations, and thus to generate the final recommendations.

The hybrid recommendation algorithm combined with k-means algorithm is described below (algorithm 2).

Algorithm 2: Recommendation algorithm with k-means
Input
Learners data
Actions performed by learners
Output
Recommendations generated for each group
Method
1: Describing each learner by descriptors.
2: implementing the K-means algorithm based on previous list of learners and their descriptors.
3: Assign each learner to a specific cluster.
For each cluster
1: Calculate correlation and co-occurrence matrices.
2: Calculate recommendation scores based on previous step.
3: Find top N recommendations for each learner.
End for each

3.3. First stage: the k-means algorithm

➢ Define the descriptors

The database contains the learners' history, including the actions that are performed by the learners. We restructure the database to group learners according to two main descriptors:

- The first descriptor describes the level of interaction based mainly on the number of times an action has been performed.
- The second descriptor reflects the discrepancy between the actions performed by a learner.

Based on these two descriptors, it is possible to group learners into several clusters using the k-means algorithm within the learning platform. The indicators or descriptors are calculated as follows:

For a particular learner, the history of actions $\{a_1, a_2\}$ such as a_1 presents the primary action et a_2 is the secondary action, and the following resources $\{i_1, i_2\}$ (table 1 and table 2).

Table 1: Activities vs learning items structure

	a_1	a_2
i_1	x_1	x_2
i_2	x_3	x_4

Table 2: K-means descriptors

Learner 1	
First descriptor	$x_1 + x_3$
Second descriptor	$(x_1 + x_3) - (x_2 + x_4)/2$

For each learner, we will perform this operation, which allows us to classify learners according to these two indicators.

➢ Optimum number of K:

In order to properly apply the k-means algorithm, it is of paramount importance to define the optimal number k of clusters to be identified. One of the most successful methods for calculating the optimal number of k is the elbow method. It actually allows to specify the most adequate number of k to be considered since the parameter k can change and can take several values. The elbow method consists in describing the variation according to the number of clusters. But in this case, what would be the optimal number? The optimal number k is simply the value from which the curve starts to take a constant value. So we will ensure that the value of k we selected represents the most accurate value for our context.

3.4. Second stage: The hybrid recommendation part

Now that the k-means algorithm is applied and clusters are detected, we now have the different groups of learners with similar behaviour in terms of interaction level. The rest of the work consists in generating recommendations through the calculation of recommendation scores for each learner pertaining to a given class. The proposed recommendation system is based on the calculation of correlation and co-occurrence scores through the analysis of the actions performed by the learners. In the following section, we will describe the mathematical model of our recommendation approach in detail:

$$The\ history\ matrix\ of\ learners =$$

	i_1			...			i_n		
l_1	a_1	...	a_m	a_1	...	a_m	a_1	...	a_m
⋮	⋮			⋮			⋮		
l_p	a_1	...	a_m	a_1	...	a_m	a_1	...	a_m

$\{i_1, \dots, i_n\}$: $Learner\ objetcs.$

$\{a_1, \dots, a_m\}$: $History\ actions\ performed\ by\ learners.$

$\{l_1, \dots, l_p\}$: $Learners.$

Referring to the basic matrix, which presents the learners' interaction with contents according to the selected actions, we generate recommendations for each cluster detected separately through the calculation of recommendation scores.

➢ Correlation

Correlation allows to identify the existing connection between two variables. Our context consists in calculating the correlation between the primary action and the secondary actions. The purpose is to situate the importance of the secondary actions in relation to the primary action to be used in the calculation of recommendations. We opted for the spearman correlation since it is the type of correlation that best suits our case, since it provides a way to evaluate the correlation in case of a non-normal distribution. We explain the notion of correlation in what follows (table 3 and eq. 1).

Table 3: General structure of correlation matrix

	a_1	...	a_m
a_1	1	...	$cor(a_1, a_m)$

⋮	⋮	⋮	
a_m	$cor(a_1,a_m)$	...	1

$$Correlation\ score = Total\ history\ matrix \times \begin{bmatrix} 1 \\ \vdots \\ cor(a_1, a_m) \end{bmatrix} = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \cdots & \cdots \\ s_{l1} & \cdots & s_{ln} \end{bmatrix} \quad (1)$$

$\{(s_{11}, \dots, s_{1n}), \dots, (s_{l1}, \dots, s_{ln})\}$ are scores of correlation for each particular learner towards each learning resource.

➢ Co-occurrence

Co-occurrence is defined as a measure to determine how many times two actions occur simultaneously in the same database. Learners generally do not react to the same actions in the same manner. They can be active in one action without engaging in another action. In this case, the co-occurrence between these two actions might be very weak. The role of co-occurrence is to define where two actions are linked at this level (table 4 and eq. 2).

Table 4: General structure of co-occurrence matrix

	a_1	...	a_m
a_1	$occurence\ (a_1)$	...	$co - occ(a_1,a_m)$
⋮	⋮	⋮	
a_m	$co - occ(a_1,a_m)$	...	$occurrence\ (a_m)$

$$Co - occurrence\ score = Total\ binary\ history\ matrix \times \begin{bmatrix} occurrence\ (a_1) \\ \vdots \\ co - occ(a_1, a_m) \end{bmatrix} = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \cdots & \cdots \\ p_{l1} & \cdots & p_{ln} \end{bmatrix} \quad (2)$$

To generate the total recommendation scores, we proceed as follows (eq. 3).

$$Total\ score = Correlation\ score + co - occurrence\ score = \begin{bmatrix} s_{11} + p_{11} & \cdots & s_{1n} + p_{1n} \\ \vdots & \cdots & \cdots \\ s_{l1} + p_{l1} & \cdots & s_{ln} + p_{ln} \end{bmatrix} \quad (3)$$

From the scores obtained, we extract the top N recommendations according to scores ranking.

4. Tests and results

In this work, we propose a recommendation system that aims to integrate the k-means algorithm within the system to generate more pertinent and reliable recommendations. Our intention is to prove that the k-means was efficient in generating more relevant recommendations and results. In order to test our recommendation approach for both cases, without k-means and with k-means, we selected two groups from a social network containing data on learners' interaction with the learning environment:

- The very first database consists of a set of 100 learners with many learning objects. We limited the number of learning objects to 10 and two actions, the primary action being the number of likes and the secondary action being the number of shares.
- The second database contains the same 100 learners from the first database, but interacting in a different social learning group. We selected some learning objects for the aforementioned actions: the number of likes and the number of shares.

The purpose of choosing these two databases is to highlight the performance of our recommendation system which incorporates the k-means algorithm, and which emphasizes the importance of grouping learners with similar attitudes in the same cluster. Homogeneity will be a major advantage in this case in terms of generated recommendations.

➢ Recommendation system evaluation:

The evaluation of the recommendation system is a prerequisite to measure its performance. In this respect, we divide the database into two major parts:

- The first part which consists in creating the recommendation model by referring to 80% of the data.
- The second part consists in measuring the accuracy and performance of the system by referring to 20% of the data.

We highlight the following measures (eq. 4 and eq. 5) according to specific parameters (Table 5).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Table 5: Accuracy and precision parameters

	Recommended	Not recommended
Preferred	True positive (TP)	False positive (FP)
Not preferred	False positive (FP)	True positive (TP)

Before moving on to the application of k-means, it is important to identify the optimal number of clusters. In this sense, we use the elbow method which allows to plot the variation as a function of the K number. The optimal value we reach for both databases through the elbow method is 2.

In what follows, we will compare the results obtained through the application of the recommendation system not considering the k-means and the recommendation system identifying the number of clusters in the database before generating the recommendations (Table 6).

Table 6: Accuracy and precision for both recommender systems in the two databases

Database		Accuracy	Precision
First database	Without clustering	0,375	0,5
	Cluster 1	0,4	0,52
	Cluster 2	0,416	0,526
	Average between	0,408	0,523

	cluster 1 and cluster 2		
Second database	Without clustering	0,3	0,44
	Cluster 1	0,40625	0,55
	Cluster 2	0,406	0,525
	Average betwen cluster 1 and cluster 2	0,406	0,5375



Figure 2: Accuracy and precision according to the two types of recommender systems (with k-means and without k-means)

Based on the results obtained (Figure 2), it appears that the recommendation system based on the classification of learners provides a higher performance compared to the recommendation system where community detection is not considered. A difference of 0.408-0.375=0.033 in terms of accuracy and 0.526-0.5= 0.026 in terms of precision. Apparently, the difference seems insignificant since the database is not of considerable size. Therefore, if the test is performed on a larger database the difference will be more noticeable. The same holds true for the second database where the difference in accuracy is 0.406-0.3=0.106 and the difference in precision is 0.5375-0.44=0.0975. These results highlight the importance of grouping learners according to their degree of interactivity before proceeding with the assignment of recommendations. Grouping learners with similar patterns of behaviour within the same community allows the generation of more adapted recommendations to their needs. Generating recommendations to learners with no prior classification reduces the performance of the recommendation system and provides recommendations that do not necessarily fit the needs of each community separately from the other. We recognize that the k-means algorithm brings considerable added value to the recommendation system due to the accuracy and performance improvements achieved. Grouping learners in homogeneous clusters facilitates the generation of recommendations.

5. Conclusion

The following work proposes a hybrid recommendation approach merging learner classification as a preliminary step before proceeding with recommendation calculations. In addition to the implicit feedback from learners that was integrated into the initial recommendation system, our approach proposes to group learners into communities with the corresponding degree of interactivity to improve the quality of the recommendations generated. Our approach consists in classifying learners according to specific interactivity criteria, then calculating recommendations for each community based on all the activities performed by the learners within the learning social network. On the one hand, the two notions of correlation and co-occurrence are combined in a unique hybrid recommendation system valuing the implicit feedback from the learners. On the other hand, the approach fosters the detection of homogeneous communities and clusters in order to generate recommendations more adapted to each community separately. The test was carried out on two databases of two different social learning groups, each consisting of 100 learners. The objective was to collect implicit feedback from learners within the learning environment. The results indicated that the recommendation approach based on the k-means algorithm leads to higher performance than the system without the k-means algorithm. This highlights the importance of classifying learners into homogeneous groups before generating recommendations. Our future work will consist of:

- Testing our recommendation approach on a larger database.
- Applying other algorithms other than k-means.

References

[1] S. Souabi, A. Retbi, M.K. Idrissi, S. Bennani, "A Recommendation Approach in Social Learning Based on K-Means Clustering," in 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), IEEE, Fez, Morocco: 1–5, 2020, doi:10.1109/ISCV49265.2020.9204203.

[2] D.K. Abdul-Rahman Al-Malah, S. Ibrahim Hamed, H.Th.S. Alrikabi, "The Interactive Role Using the Mozabook Digital Education Application and its Effect on Enhancing the Performance of eLearning," International Journal of Emerging Technologies in Learning (IJET), **15**(20), 21, 2020, doi:10.3991/ijet.v15i20.17101.

[3] E. Edelhauser, L. Lupu-Dima, "Is Romania Prepared for eLearning during the COVID-19 Pandemic?," Sustainability, **12**(13), 5438, 2020, doi:10.3390/su12135438.

[4] A. Muradi, H. Hasanzada, "The Role of Social Networks in Learning English Language," **7**(1), 10, 2020.

[5] C.G. Brinton, M. Chiang, "Social learning networks: A brief survey," in 2014 48th Annual Conference on Information Sciences and Systems (CISS), IEEE, Princeton, NJ, USA: 1–6, 2014, doi:10.1109/CISS.2014.6814139.

[6] A. Sekhavatian, M. Mahdavi, "APPLICATION OF RECOMMENDER SYSTEMS ON E-LEARNING ENVIRONMENT," 10, 2011.

[7] M.H. Ansari, M. Moradi, O. NikRah, K.M. Kambakhsh, "CodERS: A hybrid recommender system for an E-learning system," in 2016 2nd International Conference of Signal Processing and Intelligent Systems (ICSPIS), IEEE, Tehran, Iran: 1–5, 2016, doi:10.1109/ICSPIS.2016.7869884.

[8] D.M. Kandakatla, K. Bandi, "A Content Based Filtering and Negative Rating Recommender System for E-learning Management System," 7, 2018.

[9] S.K. Gorripati, V.K. Vatsavayi, "Community-Based Collaborative Filtering to Alleviate the Cold-Start and Sparsity Problems," **12**(15), 9, 2017.

[10] A.N. Ngaffo, W.E. Ayeb, Z. Choukair, "A Bayesian Inference Based Hybrid Recommender System," IEEE Access, **8**, 101682–101701, 2020, doi:10.1109/ACCESS.2020.2998824.

[11] S. Sharma, A. Mahajan, "A Collaborative Filtering Recommender System for Github," **5**(8), 7, 2017.

[12] S.S. Khanal, P.W.C. Prasad, A. Alsadoon, A. Maag, "A systematic review: machine learning based recommendation systems for e-learning," Education and Information Technologies, 2019, doi:10.1007/s10639-019-10063-9.
[13] M. Brik, M. Touahria, "Contextual Information Retrieval within Recommender System: Case Study 'E-learning System,'" TEM Journal, 1150–1162, 2020, doi:10.18421/TEM93-41.
[14] J.-P. Niyigena, Q. Jiang, A Hybrid Model for E-Learning Resources Recommendations in the Developing Countries, Association for Computing Machinery, New York, NY, USA: 21–25, 2020.
[15] K. Chaudhary, N. Gupta, "E-Learning Recommender System for Learners: A Machine Learning based Approach," International Journal of Mathematical, Engineering and Management Sciences, **4**(4), 957–967, 2019, doi:10.33889/IJMEMS.2019.4.4-076.
[16] O.C. Agbonifo, M. Akinsete, "Development of an Ontology-Based Personalised E- Learning Recommender System," **38**(1), 12, 2020.
[17] Sunil, M.N. Doja, "An Improved Recommender System for E-Learning Environments to Enhance Learning Capabilities of Learners," in: Singh, P. K., Panigrahi, B. K., Suryadevara, N. K., Sharma, S. K., and Singh, A. P., eds., in Proceedings of ICETIT 2019, Springer International Publishing, Cham: 604–612, 2020.
[18] Saintgits College of Engineering, India, A. S Nath, "A Personalized Course Recommender System for E-Learning," International Journal of Networks and Systems, 11–14, 2019, doi:10.30534/ijns/2019/03832019.
[19] P. Cao, D. Chang, "A Novel Course Recommendation Model Fusing Content-Based Recommendation and K-Means Clustering for Wisdom Education," in: Zhang, J., Dresner, M., Zhang, R., Hua, G., and Shang, X., eds., in LISS2019, Springer Singapore, Singapore: 789–809, 2020.
[20] D. Cintia Ganesha Putri, J.-S. Leu, P. Seda, "Design of an Unsupervised Machine Learning-Based Movie Recommender System," Symmetry, **12**(2), 185, 2020, doi:10.3390/sym12020185.
[21] M.S. Kumar, J. Prabhu, "A hybrid model collaborative movie recommendation system using K-means clustering with ant colony optimisation," 18, 2020.
[22] Y. Yang, D. Hooshyar, J. Jo, H. Lim, "A group preference-based item similarity model: comparison of clustering techniques in ambient and context-aware recommender systems," Journal of Ambient Intelligence and Humanized Computing, **11**(4), 1441–1449, 2020, doi:10.1007/s12652-018-1039-1.
[23] U. Kużelewska, "Dynamic Neighbourhood Identification Based on Multi-clustering in Collaborative Filtering Recommender Systems," in: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., and Kacprzyk, J., eds., in Theory and Applications of Dependable Computer Systems, Springer International Publishing, Cham: 410–419, 2020.
[24] E. Kannout, "Context Clustering-based Recommender Systems," in 2020 15th Conference on Computer Science and Information Systems (FedCSIS), 85–91, 2020, doi:10.15439/2020F54.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 907-914 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Real-time Target Human Tracking using Camshift and LucasKanade Optical Flow Algorithm

Van-Truong Nguyen*, Anh-Tu Nguyen, Viet-Thang Nguyen, Huy-Anh Bui, Xuan-Thuan Nguyen

Faculty of Mechanical Engineering, Hanoi University of Industry, Hanoi, SO298, Vietnam

ARTICLE INFO

Article history:
Received: 31 January, 2021
Accepted: 24 March, 2021
Online: 10 April, 2021

Keywords:
Human tracking
Camshift algorithm
Lucas-Kanade Optical Flow
Kalman Filter

ABSTRACT

In this paper, a novel is proposed for real-time tracking human targets in cases of high influence from complexity environment with a normal camera. Firstly, based on Oriented FAST and Rotated BRIEF features, the Lucas-Kanade Optical Flow algorithm is used to track reliable keypoints. This method represents a valuable performance to decline the effect of the illumination or displacement of human targets. Secondly, the area of the human target in the frame is determined more precise by using the Camshift algorithm. Compared to the existing approaches, the proposed method has some merits to some extents including rapid calculations in implementation, high accuracy in case of similar objects detection, the ability to deploy easily on mobile devices. Finally, the effectiveness of the proposed tracking algorithm is demonstrated via experimental results.

1. Introduction

These days, with the continuous development of high precision, cost-effective computer imaging hardware equipment and corresponding auxiliary detection algorithm, real-time target human tracking technology [1]-[8] has become a research hotspot in various fields such as interactive video games, military, robotic, etc. Human tracking is the process of locating moving targets by performing tracking algorithms and analyzing that trajectory in many applications over sequential video frames [9]-[12] The models of the target could be a single human or multiple humans based on different applications. However, human tracking is a difficult task with many challenges such as the system needs to be fast enough to carry out real-time applications or operate in different environments with high accuracy. The impact of external factors affects the accuracy of the algorithm such as moving obstacles, illuminations, or similar targets, etc. Hence, researches of higher accuracy and better computer vision detection algorithm are still developed.

With the development of technology, various methodologies have been proposed to solve real-time human tracking. In the literature review, the Kalman Filter algorithm [13] could only be used in a linear system in case of mimic real scenarios. In [14] - [16], Laser Range Finder to track human is proposed to determine the human' leg and analyze the human tracking motion. However, the measurement of this method is limited in case of long distance between human' leg and sensor. The multi-camera systems [17] - [21] are also designed to track human, but those approaches only process in a small area which installs cameras. The hand-held cameras system is equipped for the sports player in [22] to synthesize a stroboscopic image of a moving target. Nevertheless, its computational speed and accuracy could not be optimal in case of the elaborate environments. In [23]-[25], Particle filter-based vehicle tracking via HOG features is very robust when fast-moving human target but this algorithm is hard to integrate with a mobile system. In [26] and [27], the Camshift algorithm is well-known as popular tracking human method. The advantages of this methodology are low-cost computation, easy to manipulate, etc. The Camshift algorithm follows the track of the targets based on histogram back projection. The stability of this method is impacted by color, illumination, and noise. Hence, this method suffers from series of errors in case of similar color background.

In light of the remarkable importance and advantages previously mentioned, a real-time tracking human system is proposed for normal cameras under the circumstances of high influence from complicated environments. Lucas-Kanade Optical Flow Algorithm is a well know method for tracking humans for a decade [28]-[30] The method is designed based on the combination between Camshift algorithm and Lucas-Kanade Optical Flow Algorithm (LK-OFA) [31], [32] with Oriented FAST and Rotated BRIEF features [33] Compared to the existing works, the proposed approach has several contributions as follows:

(1) The reliable keypoints which are less affected by illumination or displacement of the human targets are tracked based on the LK-OFA features. It is noted that optical follow algorithm is very good at tracking points in the next frame from

*Corresponding Author: Van-Truong Nguyen, Hanoi, Vietnam, nguyenvantruong@haui.edu.vn

the locations of them in the previous frame. The Camshift algorithm is carried out to determine the area of the human target with similar appearances in the frames more precisely.

(2) Since the proposed method reduces number of loops that need to converge the center of the search window and centroid of human target, it is easy to implement on mobile devices.

(3) The designed system generates a small amount of data to be processed; therefore, this could be simply done in real-time, and a powerful computer is not required.

The rest of the paper is organized as follows. The next section analyses the detailed structure of the proposed approach. Section 3 summarizes the experiment platform and its result, followed by the brief conclusions and the outlook in section 4.

2. Proposed Methodology

In this section, a short overview of the human tracking algorithm is illustrated in the flowchart as Fig.1. To begin with, for choosing the tracking target person, a region of interest (ROI) is created manually. After that, HSV histogram of ROI is computed to prepare for CAMSHIFT later. At the following stage, the system finds ORB features in both ROI and initial frame. In the subsequent step, 50 best keypoints are chosen from best matched features based on Bruce-Force matcher. Those keypoints are traced by using Optical flow with Lucas-Kanade algorithm. At the next stage, the failed tracked points are totally removed by utilizing sum square of difference (SSD) and Backward tracking method. If the number of the keypoints is smaller than three, the process comes back to the finding ORB features step. In contrast, the process continues with good keypoints left, the center of them is determined using first and zeros moment. This center is used to create the initial search window of Camshift. By applying Camshift algorithm, the area of the human target can be found. Finally, the accuracy of tracking human's center is improved by Kalman filter.

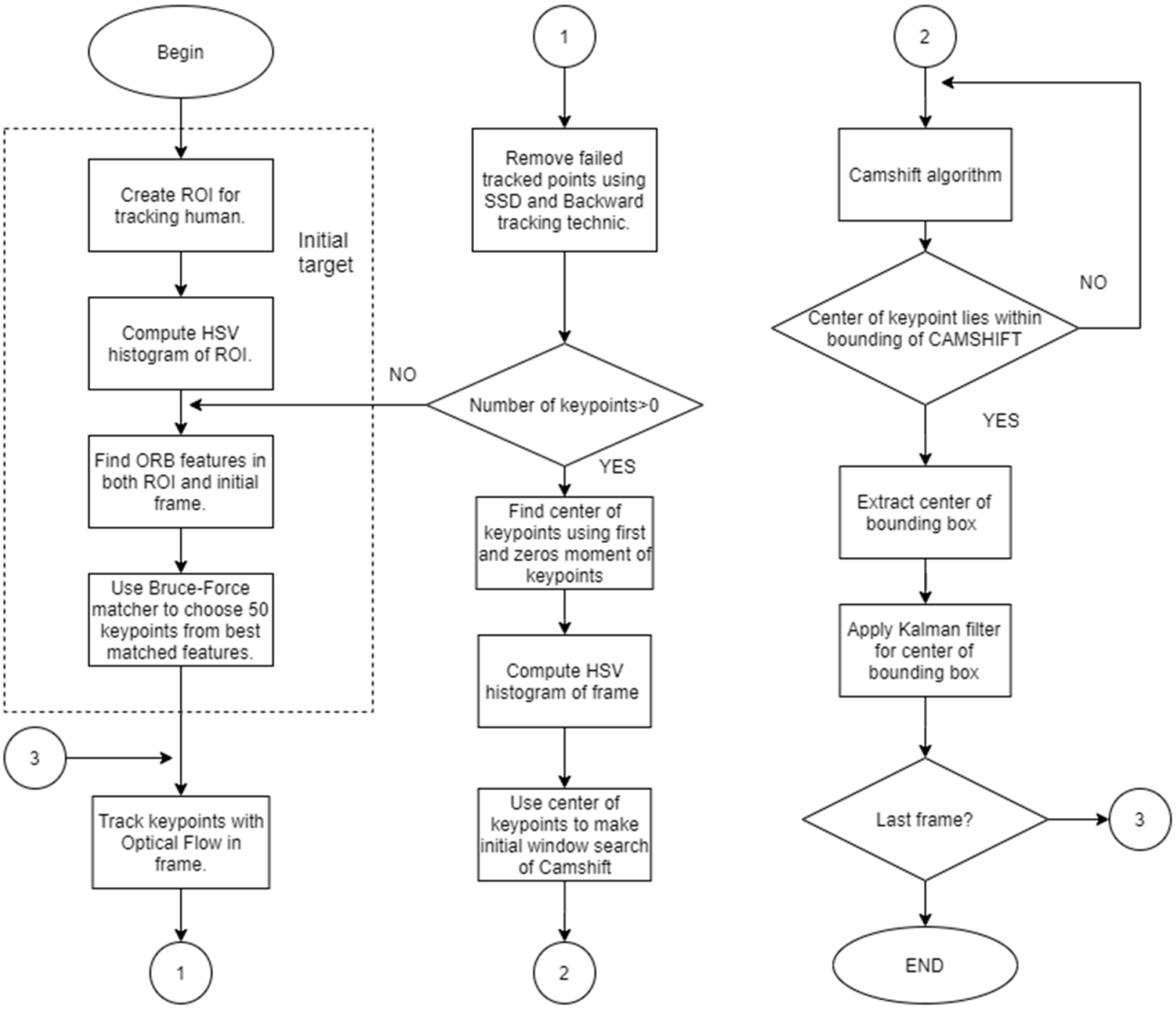


Figure 1: Flow diagram of system

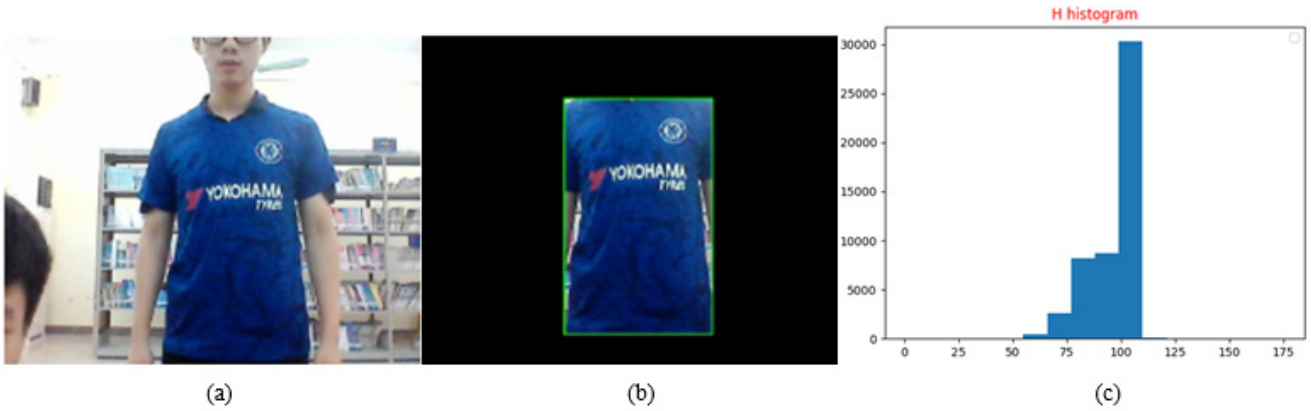


Figure 2: Create ROI and calculate H histogram of ROI

2.1. Create ROI and compute H histogram

Before tracking human, an area (ROI) on the target person is established (Figure 2 (b)). This ROI is used as a template of the target. Then, the HSV histogram of ROI is computed, but only the H channel is used with 16 bins to get the best result (Figure 2 (c)).

2.2. ORB feature and matching

For the purpose of tracking and creating the initial window for Camshift, this paper selected ORB feature detector (Oriented FAST and Rotated BRIEF) to find the best points. ORB algorithm modifies both FAST [34] keypoint detector and BRIEF [35] descriptor to reach the best result. ORB algorithm detects FAST points in the image, then applies Harris corner measure to discover top N points among them. After that, the algorithm searches for the intensity weighted centroid of the patch with the located corner at the center in order to calculate the orientation of points. The orientation value is computed by using direction of the vector from this corner point to centroid. After finishing the implementation of FAST, ORB algorithm modifies BRIEF descriptors or rBRIEF to create a descriptor for each keypoint.



Figure 3: Matching keypoints between ROI and initial Frame

In our system, ORB algorithm is implemented by extracting keypoints from both ROI and frame. The keypoints are searched in Gray image extracting from the original image and the archetype of ROI. Bruce-Force matcher is used to calculate the distance value between each pair of keypoints in both Gray images. By sorted pair of keypoints based on distance value, some keypoints best matched will be kept to use later as shown in Figure 3. To be more specific, the algorithm could operate with higher accuracy in case of more keypoints selection. However, if the number of keypoints incline too big, the calculation process would be taken long time to finish, especially, in case of mobile robots. In this paper, based on trial and errors method, having 50 keypoints enabled us to ensure the balance between the calculation time and the accuracy of our algorithm effectively.

2.3. Track keypoints and remove failure points

The positions of the previous keypoints are tracked by using Optical Flow with Lucas-Kanade Method. Here, a 3x3 patch around the points with an assumption that all the neighboring pixels have similar motion is applied to estimate the new locations of keypoints in the frame based on their previous positions by using equation (1).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i fx_i^2 & \sum_i fy_i \sum_i fx_i \\ \sum_i fx_i \sum_i fy_i & \sum_i fy_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_i fx_i . ft_i \\ \sum_i fy_i . ft_i \end{bmatrix} \quad (1)$$

where $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$ and f_x, f_y are gradient of keypoints and surrounding points along x, y axis.

After updating, some of those keypoints in the frame might be failed. In order to remove these failure points, this article combines two methods at the same time: SSD and the Backward tracking technic. Firstly, the SSD in RGB color-space for consecutive frames is expressed as:

$$SSD_R = \sum_{m,n} \begin{pmatrix} R_{t-1}\left(k_x^{t-1} + m, k_y^{t-1} + n\right) \\ -R_t\left(k_x^t + m, k_y^t + n\right) \end{pmatrix}^2 \quad (2)$$

$$SSD_G = \sum_{m,n} \begin{pmatrix} G_{t-1}\left(k_x^{t-1} + m, k_y^{t-1} + n\right) \\ -G_t\left(k_x^t + m, k_y^t + n\right) \end{pmatrix}^2 \quad (3)$$

$$SSD_B = \sum_{m,n} \begin{pmatrix} B_{t-1}\left(k_x^{t-1} + m, k_y^{t-1} + n\right) \\ -B_t\left(k_x^t + m, k_y^t + n\right) \end{pmatrix}^2 \quad (4)$$

where k^t is the position of keypoint, m and n are width and height of the compared window. The idea of SSD is to compute the total error between a patch around keypoint in the current frame

with the location of them in the previous frame. Figure 4 and Figure 5 illustrate errors between the location of keypoints in consecutive frames corresponding to two cases: small error and big error. After reckoning the SSD values in three channels, the total error is expressed as follows:

$$Error = \alpha SSD_R + \beta SSD_G + \gamma SSD_B \quad (5)$$

where α, β, γ are three positive coefficients and the sum of α, β, γ is equal to one:

$$\alpha, \beta, \gamma > 0, \alpha + \beta + \gamma = 1 \quad (6)$$

In this paper, the coefficients are chosen as follows: $\alpha = 0.3, \beta = 0.4, \gamma = 0.3$. The error value of all keypoints are illustrated in Figure 6.

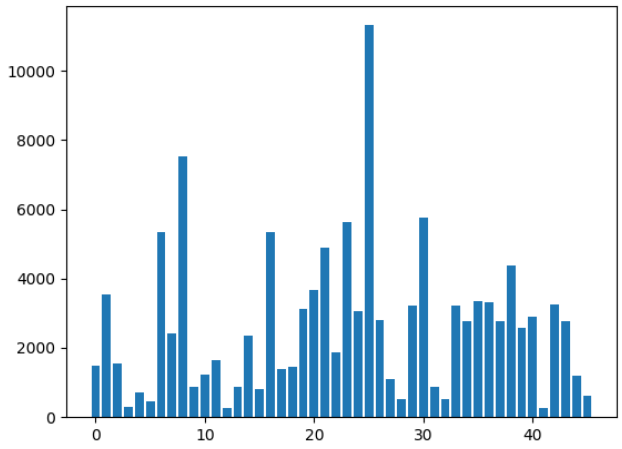


Figure 4: Small error between the location of key points in consecutive frames

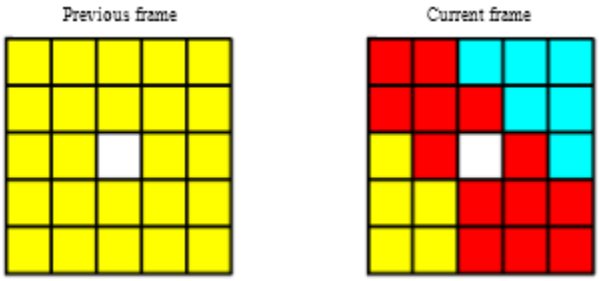


Figure 5: Big error between the location of key points in consecutive frames

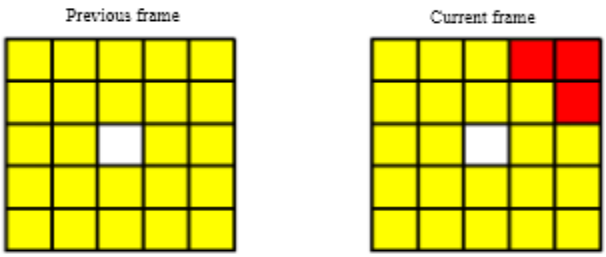


Figure 6: Error value of all key points

After applying SSD for the first filter, the Backward tracking technic examines all keypoints again and removes the failed points. As shown in Figure 7, the keypoints in the current frame are tracked back to the previous frame by applying Optical Flow. Let denote r^{t-1} is the estimated position of k^t in the previous frame. The distance d between r^{t-1} and k^{t-1} is calculated by using the Euclid distance method. All keypoints whose distance value is smaller than a thresh hold are kept.

$$d = \sqrt{\left(k_x^{t-1} - r_x^{t-1}\right)^2 + \left(k_y^{t-1} - r_y^{t-1}\right)^2} \quad (7)$$

By utilizing two mentioned method above, all failed keypoints are removed. If all keypoints are failed, our method has a viable improvement on current methods. To be more detailed, in this case, a new set of keypoints would be created inside Camshift bounding box of the target.

2.4. Initial search Window for Camshift algorithm

By applying first and zeros moment for all remain keypoints, the mean position of them is indicated as:

$$M_{00} = \sum_x \sum_y I(x, y) \quad (8)$$

$$M_{10} = \sum_x \sum_y xI(x, y) \quad (9)$$

$$M_{01} = \sum_x \sum_y yI(x, y) \quad (10)$$

$$X_c = \frac{M_{10}}{M_{00}}, Y_c = \frac{M_{01}}{M_{00}} \quad (11)$$

with I is the intensity of keypoints. Since we project all keypoints into a bitmask, so $I = 1$. Location of initial search window Camshift is created from the center points.

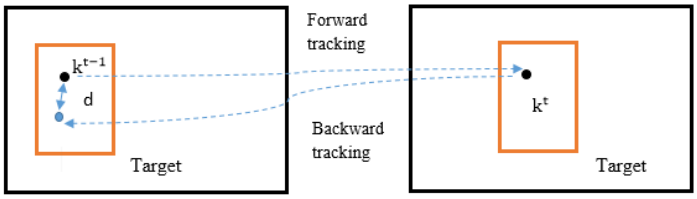


Figure 7: Backward tracking technic

2.5. Human tracking with Camshift

Camshift algorithm is an effective method to find color feature of target in frame. The principle of Camshift algorithm bases on Meanshift algorithm. However, it was modified to update the size of the tracking window in the next frame and find the orientation of the target. For more specific, Camshift computes H histogram of both ROI and Frame, and then exchanges pixel' value in the frame with the probability value of its color appearance. Then, it creates the color probability distribution image that performs the probability of appearance of each pixel within the range. Camshift algorithm uses a search window and loops it until the center of the search window converges with the centroid of points having high probability. By using an initial search window created with tracking keypoints, it reduces the number of iterations to find the area of the target. The size of tracking window is updated by zeros moment as follows:

$$S = 2 \times \sqrt{\frac{M_{00}}{256}} \quad (12)$$

Then, the orientation of the target is obtained by using the second central moments:

$$\theta = \frac{1}{2}\arctan\left(\frac{2\left(\frac{M_{11}}{M_{00}}\right) - X_c Y_c}{\left(\frac{M_{20}}{M_{00}} - X_c^2\right) - \left(\frac{M_{02}}{M_{00}} - Y_c^2\right)}\right) \quad (13)$$

The aspect ratio is expressed as:

$$Ratio = \frac{M_{20}}{X_c^2}\frac{Y_c^2}{M_{02}} \quad (14)$$

The width and the height of the window in the frame are estimated as follows:

$$w = 2.M_{00}.Ratio \quad (15)$$

$$h = 2.M_{00}.Ratio \quad (16)$$

2.6. Kalmal filter for the position of the human target

The center of human target finding by Camshift algorithm has some vibration and needs to be removed to improve the smoothness of the human tracking process. Kalmal filter is used in this paper to solve the vibration of center position. Kalman filter method estimates the location of the center based on value prediction and the new location of the center. It contains two main stages: prediction and correction.

Let denote $X_t = \left[x_t, y_t, v_{xt}, v_{yt}\right]$ as a state variable and $Z_t = \left[x, y\right]^t$ as measurement variable. Here, x_t, y_t are the positions of the center along x - axis and y - axis in the image. v_{xt} and v_{yt} are displaced in the direction of the horizontal position and vertical position. The equations of the motions on the x - axis and y - axis without the acceleration are calculated as follows:

$$\hat{x}_t = x_{t-1} + v_{xt}.\Delta t \quad (17)$$

$$\hat{y}_t = y_{t-1} + v_{yt}.\Delta t \quad (18)$$

The following equation is inferred.

$$\hat{X}_t = F.X_{t-1} \quad (19)$$

With $\hat{X}_t$ is the value prediction of X_t, F is translation matrix. The predictor covariance equation is expressed as follows:

$$\hat{P}_t = F.P_{t-1}.F^T + Q \quad (20)$$

In the equation, $\hat{P}_t$ is the value prediction of covariance $\hat{P}_t$, Q is the interference factor.

In correction step, there are three equations named as: the Kalman gain equation, the State update equation, and the Covariance update equation:

$$K_t = \hat{P}_t.H^T.\left(H.\hat{P}_t.H^T + R\right)^{-1} \quad (21)$$

$$X_t = \hat{X}_t + K_t\left(Z_t - H\hat{X}_t\right) \quad (22)$$

$$P_t = \hat{P}_t + K_t.H.\hat{P}_t \quad (23)$$

In these above equations, R is measurement noise covariance matrix, H is measurement matrix. The value of R, H, F and Q is implemented as follows:

$$F = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, R = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

2.7. Summary methodology

Our algorithm is carried out with the following step:

Step 1: Create ROI for human target and analysis the HSV Histogram of ROI.

Step 2: Compute ORB features in ROI.

Step 3: Compute ORB features in the first frame and match them with ORB features in ROI using BruceForce matching. Take 50 best matching points to track later.

Step 4: Tracking those keypoints in the previous step by using the Lucas-Kanade Optical Flow algorithm.

Step 5: Remove failed keypoints by using SSD and Backward tracking technic.

Step 6: Find the center k of remain keypoints in the frame with zeros and the first moment.

Step 7: Use the center k to create the initial window for Camshift algorithm.

Step 8: Find the bounding box and update centroid of the tracking human by applying Camshift algorithm.

Step 9: Use the Kalman filter to improve the smoothness of the human tracking process.

3. Experiment results

During the experimental process, the configuration of the computer is as follows: 4 GB ram, Intel(R) Celeron (R) CPU N3350 @ 1.1 GHz, webcam with 0.3 Mpx (480x640) and 2 Mpx Webcam. The algorithm is carried out in Python with the support of OpenCV library and Numpy Library.

For the purpose of showing the development in performance, this paper offers comparisons between the proposed method and the original Camshift [36] The comparisons are checked in different scenes as shown in Figure 8 and Figure 9. To be more specific, the blue bounding box is created by using Camshift and the green one is created by using the proposed method.

Table 1: Average Frame per second value

Criteria	Laptop webcam (0.3Mpx)	Webcam (2Mpx)
Set FPS	30	30
Real FPS	20-25	14-17
Processing time (s)	0.04-0.05	0.06-0.07

Table 2: Experiment results

Environmental conditions	Cases	Number of tests frames	Results of Original Camshift		Results of the proposed method	
			Number of false frames	Acuracy (%)	Number of false frames	Acuracy (%)
Good lighting room condition (500 lux)	One person	323	8	97.5	5	98.45
	Group of people wearing the same shirts	412	298	27.66	15	96.35
	Group of people wearing the different shirts	452	15	96.68	12	97.34
Poor lighting room condition (250 lux)	One person	255	24	90.58	22	91.37
	Group of people wearing the same shirts	542	420	22.5	49	90.95
	Group of people wearing the different shirts	526	47	91.06	38	92.77
Outdoor Condition (700 lux)	One person	552	53	90.39	45	91.84
	Group of people wearing the same shirts	587	401	31.68	55	90.06
	Group of people wearing the different shirts	479	48	89.97	44	90.81

Firstly, in the case of different color shirts as shown in Figure 8, it could be seen that there are almost none of the differences between two approaches. In the second case, we evaluate the performance of the proposed algorithm with undesired humans having the same shirt in random backgrounds. From Figure 9, it is noticeable that the original Camshift recognizes and tracks the wrong target. In contrast, the analyzing and human tracking of the proposed method is very clear and precise.

Table 1 provides the data about the Average Frame per second value of the proposed algorithm on both 0.3 Mpx webcam and 2 Mpx Webcam. On average, it could be seen that the system is fast enough to be utilized in real-time with FPS value of up to 25. Meanwhile, the effect of Kalman filter is evaluated over 400 frames and the results are shown in Figure 10. Furthermore, the proposed method indicates a significant decline in the chattering of the bounding box centre especially in x-axis when human target moves.

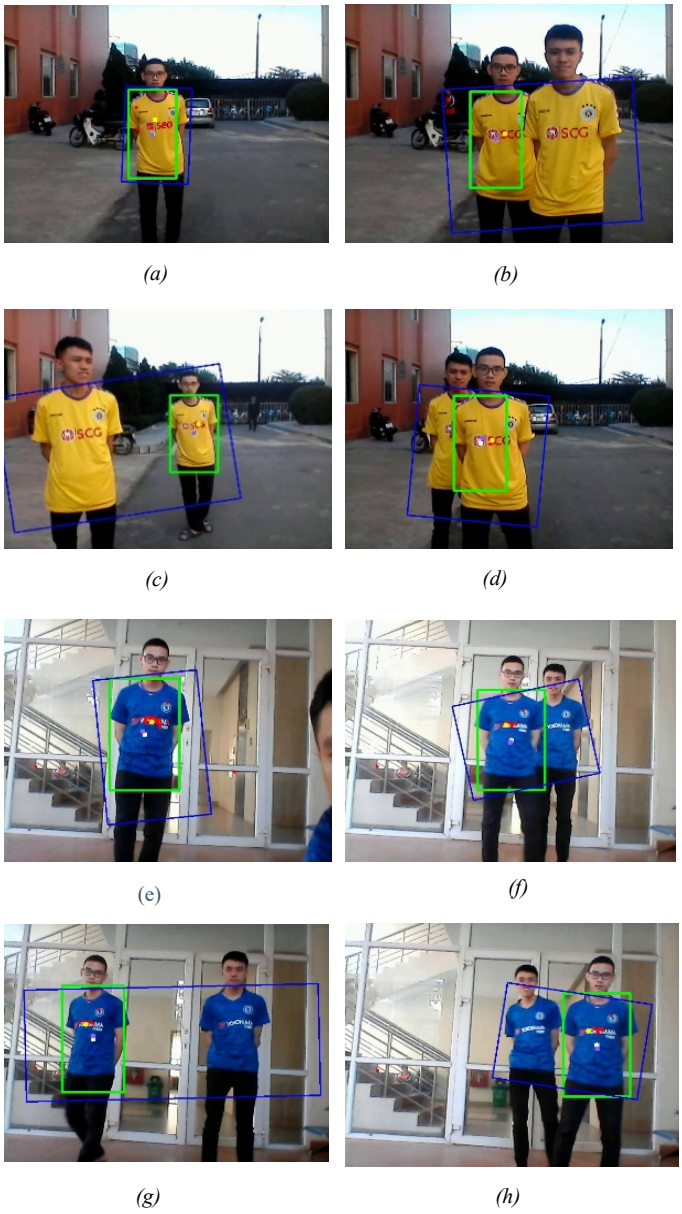


Figure 8: Tracking human in case of different color shirts

Table 2 compares the experimental results between the Original Camshift and the proposed approach among different types of the environmental conditions. In each scene, we test 60 times for 3 cases: One person, Group of people wearing the same shirts, Group of people wearing the different shirts. Out of three environmental conditions, it is clear that the false frames in poor lighting room condition of the Original Camshift is the highest, at about 420. By constract, the figure for the proposed method is only 49. Over different cases, the precision rate of the proposed method is quite higher compared to the Original Camshift. For instance, the accuracy of Original Camshift in good lighting room condition is just over 27.66%, while the proportion of the proposed method is nearly 96.35%. In addition, for tracking human wearing the same shirts, the tracking rates of the Original Camshift are noticeably low, at about 27% in average. On the other hand, the percentages of the developed method always maintain over 90%. Furthermore, the results illustrates that the proposed system could work effectively even in the indoor environments or the outdoor environments with the accuracy up to 98%.

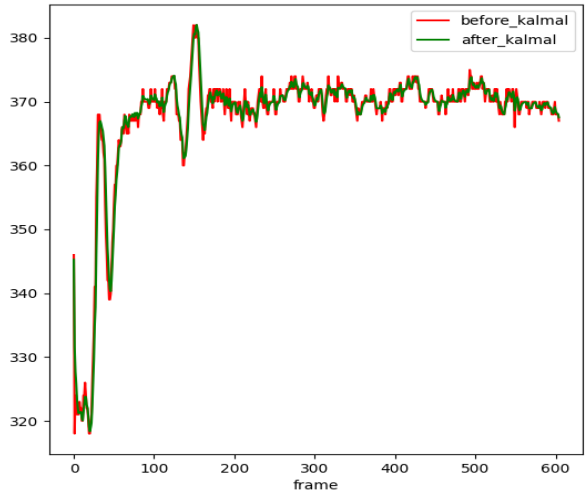
4. Conclusions

In this article, the new method of human tracking is analyzed in depth. The proposed method is developed based on Camshift algorithm and the Lucas-Kanade Optical Flow Algorithm with Oriented FAST and Rotated BRIEF features. The experiment results indicate that the proposed system could be well-adjusted in real-time applications. By comparing with the existing human tracking algorithms, it could be noted that the proposed algorithm reaches to the higher accuracy. Furthermore, its computing time is relatively faster. Hence, the adaptability of the proposed method is better in practical applications. In conclusion, the present results provide practical reference values about human tracking algorithm for the development of equipment with high anti-interference performance, the design of test plans, and the establishment of international standards in the future.

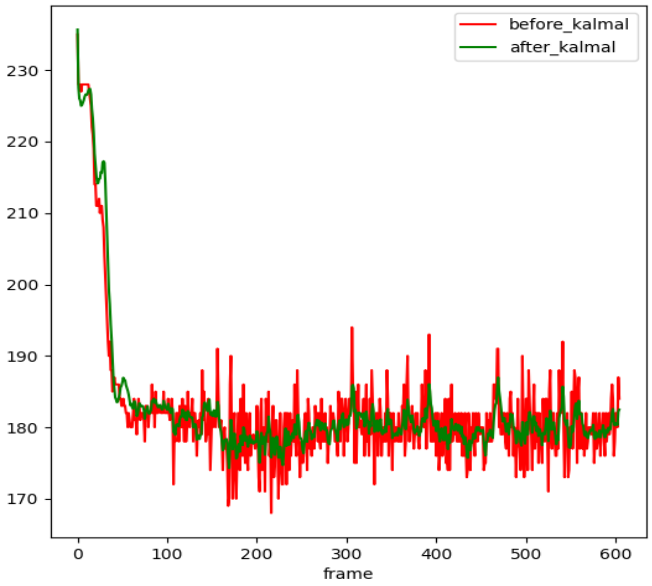


(a) (b) (c) (d) (e) (f) (g) (h)

Figure 9: The test results in case of similar color shirt



(a) Y-axis



(b) X-axis

Figure 10: The position center of bounding box before and after using Kalman filter: (a) Y-axis, (b) X-axis

Acknowledgement

This research was funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant No. “107.01-2019.311”, and in Hanoi University of Industry, Hanoi, Vietnam, Project no. 04-2020-RD/HĐ-ĐHCN.

References

[1] R. Henschel, T. Marcard, B. Rosenhahn, "Accurate Long-Term Multiple People Tracking Using Video and Body-Worn IMUs," IET Image Processing, **29**(11), 8476-8489, 2020, doi: 10.1109/TIP.2020.3013801.

[2] S. Damotharasamy, "Approach to model human appearance based on sparse representation for human tracking in surveillance," IET Image Processing, **14**(11), 2383-2394, 2020, doi: 10.1049/iet-ipr.2018.5961.

[3] F. Angelini, Z. Fu, Y. Long, L. Shao, S.M. Naqvi, "2D Pose-based Real-time Human Action Recognition with Occlusion-handling," IEEE Transactions on Multimedia, **22**(6), 1433- 1446, 2019, doi: 10.1109/TMM.2019.2944745.

[4] X. Yang, et al, "Human Motion Serialization Recognition With Through-the-Wall Radar," IEEE Access, **8**, 186879-186889, 2020, doi: 10.1109/ACCESS.2020.3029247.

[5] P. Vaishnav, A. Santra, "Continuous Human Activity Classification With Unscented Kalman Filter Tracking Using FMCW Radar," IEEE Sensors Letters, **4**(5), 1-4, 2020, doi: 10.1109/LSENS.2020.2991367.

[6] K.M. Abughalieh, S.G. Alawneh, "Predicting Pedestrian Intention to Cross the Road," IEEE Accesss, **8**, 72558 - 72569, 2020, doi: 10.1109/ACCESS.2020.2987777.

[7] G. Ren, X. Lu ,Y. Li, "A Cross-Camera Multi-Face Tracking System Based on Double Triplet Networks," IEEE Access, **9**, 43759-43774, 2021, doi: 10.1109/ACCESS.2021.3061572.

[9] Q.Q. Chen, Y.J. Zhang, "Sequential segment networks for action recognition." IEEE Signal Processing Letters, **24**(5), 712-716, 2017, doi: 10.1109/LSP.2017.2689921.

[11] B. Chung, C. Yim, "Bi-sequential video error concealment method using adaptive homographybased registration," IEEE Transactions on Circuits and Systems for Video Technology, **30**(6), 1535 - 1549, 2019, doi: 10.1109/TCSVT.2019.2909564.

[12] Z. Yu, L. Han, Q. An, H. Chen, H. Yin and Z. Yu, "Co-Tracking: Target Tracking via Collaborative Sensing of Stationary Cameras and Mobile Phones," IEEE Access, **8**, 92591-92602, 2020, doi: 10.1109/ACCESS.2020.2979933.

[13] B. Deori and D. M. Thounaojam, "A survey on moving object tracking in video," International Journal on Information Theory (IJIT), **3**(3), 31–46, 2014, doi: 10.5121/ijit.2014.3304.

[14] E.J. Jung, et al, " Development of a laser-rangefinder-based human tracking and control algorithm for a marathoner service robot," IEEE/ASME transactions on mechatronics, **19**(6), 1963-1976, 2014, doi: 10.1109/TMECH.2013.2294180.

[15] K. Dai, Y. Wang, Q. Ji, H. Du, C. Yin, "Multiple Vehicle Tracking Based on Labeled Multiple Bernoulli Filter Using Pre-Clustered Laser Range Finder Data," IEEE Transactions on Vehicular Technology, **68**(11), 10382-10393, 2019, doi: 10.1109/TVT.2019.2938253.

[16] D. Li, et al, "A multitype features method for leg detection in 2-D laser range data, " IEEE Sensors Journal, **18**(4), 1675-1684, 2017, doi: 10.1109/JSEN.2017.2784900.

[17] M. J. Domınguez-Morales, et al, "A. Bio-Inspired Stereo Vision Calibration for Dynamic Vision Sensors," IEEE Access, **7**, 138415-138425, 2019, doi: 10.1109/ACCESS.2019.2943160.

[18] B. Bozorgtabar and R. Goecke, "MSMCT: Multi-State Multi-Camera Tracker," IEEE Transactions on Circuits and Systems for Video Technology, **28**(12), 3361-3376, 2018, doi: 10.1109/TCSVT.2017.2755038.

[19] Y. Han, R. Tse and M. Campbell, "Pedestrian Motion Model Using Non-Parametric Trajectory Clustering and Discrete Transition Points," IEEE Robotics and Automation Letters, **4**(3), 2614-2621, 2019, doi: 10.1109/LRA.2019.2898464.

[20] V. Carletti, A. Greco, A. Saggese, M. Vento, "Multi-object tracking by flying cameras based on a forward-backward interaction," IEEE Access, **6**, 43905-43919, 2018, doi: 10.1109/ACCESS.2018.2864672.

[21] F.L. Zhang, et al, "Coherent video generation for multiple hand-held cameras with dynamic foreground," Computational Visual Media, **6**(3), 291-306, 2020, doi: https://doi.org/10.1007/s41095-020-0187-3.

[22] K. Hasegawa, H. Saito, "Synthesis of a stroboscopic image from a hand-held camera sequence for a sports analysis," Computational Visual Media, **2**(3), 277-289, 2016, doi: https://doi.org/10.1007/s41095-016-0053-5.

[23] X. Kong, et al, "Particle filter-based vehicle tracking via HOG features after image stabilisation in intelligent drive system," IET Intelligent Transport Systems, **13**(6), 942-949, 2018, doi: https://doi.org/10.1049/iet-its.2018.5334.

[24] S. Papaioannou, A. Markham, N. Trigoni, "Tracking people in highly dynamic industrial environments," IEEE Transactions on mobile computing, **16**(8), 2351-2365, 2016, doi: 10.1109/TMC.2016.2613523.

[25] Y. Zhai, et al, "Occlusion-aware correlation particle filter target tracking based on RGBD data," IEEE Access, **6**, 50752-50764, 2018, doi: 10.1109/ACCESS.2018.2869766.

[26] Y. Zhang, "Detection and Tracking of Human Motion Targets in Video Images Based on Camshift Algorithms," IEEE Sensors Journal, **20**(20), 11887-11893, 2020, doi: 10.1109/JSEN.2019.2956051.

[27] F. Li, R. Zhang, F. You, "Fast pedestrian detection and dynamic tracking for intelligent vehicles within V2V cooperative environment," IET Image Processing, **11**(10), 833-840, 2017, doi: 10.1049/iet-ipr.2016.0931.

[28] Q. Tran, S. Su and V.T. Nguyen, "Pyramidal Lucas—Kanade-Based Noncontact Breath Motion Detection," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **50**(7), 2659-2670, 2020, doi: 10.1109/TSMC.2018.2825458.

[29] B. Du, S. Cai and C. Wu, "Object Tracking in Satellite Videos Based on a Multiframe Optical Flow Tracker," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **12**(8), 3043-3055, 2019, doi: 10.1109/JSTARS.2019.2917703.

[30] Q. Tran, S. Su, C. Chuang, V.T. Nguyen and N. Nguyen, "Real-time non-contact breath detection from video using adaboost and Lucas-Kanade algorithm," in 2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS), Otsu, Japan, 1-4, 2017.

[31] T. Song, B. Chen, F. M. Zhao, Z. Huang, M. J. Huang, "Research on image feature matching algorithm based on feature optical flow and corner feature," The Journal of Engineering, **2020**(13), 529-534, 2020, doi: https://doi.org/10.1049/joe.2019.1156.

[32] J. de Jesus Rubio, Z. Zamudio, J. A. M. Campana, M. A. M. Armendariz, "Experimental vision regulation of a quadrotor," The Journal of Engineering, **13**(8), 2514-2523, 2015, doi: 10.1109/TLA.2015.7331906.

[33] J. Li, Z. Li, Y. Feng, Y. Liu, G. Shi, "Development of a human-robot hybrid intelligent system based on brain teleoperation and deep learning SLAM," IEEE Transactions on Automation Science and Engineering, **16**(4), 1664-1674, 2019, doi: 10.1109/TASE.2019.2911667.

[34] C. Xin, Z. Jingmei, Z. Xiangmo, W. Hongfei, C. Hui " Vehicle ego-localization based on the fusion of optical flow and feature points matching," IEEE Access, **7**, 167310-167319, 2019, doi: 10.1109/ACCESS.2019.2954341.

[35] S. K. Lam, G. Jiang, M. Wu, B. Cao, "Areatime efficient streaming architecture for FAST and BRIEF detector," IEEE Transactions on Circuits and Systems II: Express Briefs, **66**(2), 282-286, 2019, doi: 10.1109/TCSII.2018.2846683.

[36] N. Q. Nguyen, S. F. Su, Q. V. Tran, V. T. Nguyen, J. T. Jeng, "Real time human tracking using improved CAM-shift," in IEEE International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS), 1-5, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 915-920 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Efficient 2D Detection and Positioning of Complex Objects for Robotic Manipulation Using Fully Convolutional Neural Network

Dominik Štursa*, Daniel Honc, Petr Doležel

University of Pardubice, Department of Process Control, Faculty of Electrical Engineering and Informatics, Pardubice, 530 02, Czech Republic

ARTICLE INFO

Article history:
Received: 18 January, 2021
Accepted: 29 March, 2021
Online: 10 April, 2021

Keywords:
Machine Vision
Fully Convolutional Neural Net
U-Net
Machine Learning
Pick and Place

ABSTRACT

Programming industrial robots in a real-life environment is a significant task necessary to be dealt with in modern facilities. The "pick up and place" task is undeniably one of the regular robot programming problems which needs to be solved. At the beginning of the "pick and place" task, the position determination and exact detection of the objects for picking must be performed. In this paper, an advanced approach to the detection and positioning of various objects is introduced. The approach is based on two consecutive steps. Firstly, the captured scene, containing attentive objects, is transformed using a segmentation neural network. The output of the segmentation process is a schematic image in which the types and positions of objects are represented by gradient circles of various colors. Secondly, these particular circle positions are determined by finding the local maxima in the schematic image. The proposed approach is tested on a complex detection and positioning problem by evaluation of total accuracy.

1. Introduction

Automated systems have been developing rapidly for decades and they have increasingly helped to improve the reliability and productivity in all domains of industry. Machine vision, which is the family of methods used to provide imaging-based and image-processing-based inspections and analyses, is an indispensable element of automation. The implementations of machine vision approaches can be found in process control [1], automatic quality control [2], and especially in industrial robot programming and guidance [3].

Considering robot programming and guidance in an industrial environment, more and more intelligent machines are being utilized to deal with various applications. These days, a static and unchanging production environment is often being replaced by dynamically adapting production plans and conditions. Therefore, the assembly lines are managed on a daily basis and consequently, the automated systems and industrial robots need to be capable of dealing with more generalized tasks (general-purpose robotics). Although most robotic applications are still developed analytically or based on expert knowledge of the application approach [4], some industrial robotics producers have begun to implement deep learning methods in their applications like Keyence and their IV2 Vision Sensor. It is beginning to be generally recognized, that deep learning methods can play a significant role especially in the mentioned general-purpose robotics [5].

Deep learning consists of a family of modified machine learning methods aiming to solve the tasks that come naturally to human beings. Deep learning methods are performed directly on the available task-specific data in order to get a heuristic relation between the input data and the expected output. Various deep learning approaches have already been applied successfully to deal with various classification and detection tasks [6, 7] and are also utilized in other domains.

In general-purpose robotics, the "pick and place" task is the key problem to deal with. Generally, a "pick and place" task consists of a robotic manipulator (or group of manipulators) able to pick a particular object of attention and place it in a specific location with defined orientation.

In this contribution, the initial part of the "pick and place" issue is examined. To be more specific, the grasp point or grasping pose, which defines how a robotic arm end-effector should be set in order to efficiently pick up the object, is dealt with. Clearly, there is a broad group of grasp point detection techniques which can be listed either according to the type of perception sensor, or by a

* Corresponding Author: Dominik Štursa, nám. Čs. Legií 565, 530 02 Pardubice, Czech Republic, +420466037124, dominik.stursa@upce.cz

procedure used for object and grasp point detection and positioning. The most current methods are clearly described in survey [8]. Deep learning approaches for the detection of robotic grasping poses are summarized in review [4]. In this contribution, we propose a rapid, efficient, and accurate system for finding multiple object centers for flat and moving surfaces.

The key contribution of this article is:

- Proposal of an efficient grasp point detection method for a robotic arm capable of handling more types of objects. This method uses a classical industrial camera as the input data source.
- As an important part of the method, proposal of a deep learning-based approach to transform an RGB image of the scene of interest into a schematic grayscale frame. In this frame, the types of objects and the feasible positions of the grasp points are coded into gradient shapes of various colors. To the authors' knowledge, this is the first application of this approach to the grasp point detection.
- The proposed grasp point detection method is insensitive to changing light conditions and highly variable surroundings. In addition, it is efficient enough to be used in real time with specific edge computing tools, such as NVIDIA Jetson Nano, Google Coral or Intel Movidius.

The structure of the article is as follows. The aim of the work is formulated, and the goals are defined in the next section. Then, the solution based on deep learning approach is proposed. After that, the case study, which should demonstrate the main features of the proposed solution, is presented. Finally, the results are summarized, and the article is finished with some conclusions. This paper is an extension of work originally presented in the 24th International Conference on System Theory, Control and Computing (ICSTCC) [9].

2. Problem Formulation

As stated already in the preceding section, we deal with the essential task of industrial robotics called the “pick and place” problem. To be more specific, we are interested in the first challenge of this problem, i.e. detection and positioning of the objects. As a very necessary and attractive problem, detection and positioning of objects of interest has been examined and researched from many points of view [10, 11]. These days, due to a greater use of laser scanner technology (e.g. Photoneo Phoxi 3D Laser Scanner, Faro Focus 3D Laser Scanner), clouds of points are often used for object detection and positioning [12, 13]. Apparently, laser scanners in combination with robust 3D object registration algorithms, provide a strong tool to be applied in “pick and place” problems. Nevertheless, solutions based on clouds of points are generally costly and for some materials (shiny metal, glass, etc.), their accuracy decreases. In addition, a large number of laser scanners provide framerates too low to be used with moving objects of interest. Therefore, we propose an alternative solution based on a classical monocular RGB industrial camera as an image acquisition sensor. Besides, we suggest a novel approach based on a fully convolutional neural network in combination with a classical image processing routine, in order to analyze the signal from the industrial camera.

The proposed solution is supposed to meet the requirements of the industrial sector, i.e. stable performance, insensivity to light conditions, quick response and reasonable cost. In order to demonstrate these requirements, we perform a case study, which is summarized at the end of this article. This case study is supposed to fulfill the following conditions:

- Objects – up to four different complex objects should be detected and located.
- Conditions – Detection and positioning accuracy should not be affected by background surface change.
- Framerate – In order to be able to register moving objects, the framerate should exceed 10 frames per second.
- Hardware – The detection and positioning system should be based on hardware suitable for industrial applications. However, the costs should not exceed $500 to be economically viable.

The graphic plan of our task is depicted in Figure 1.

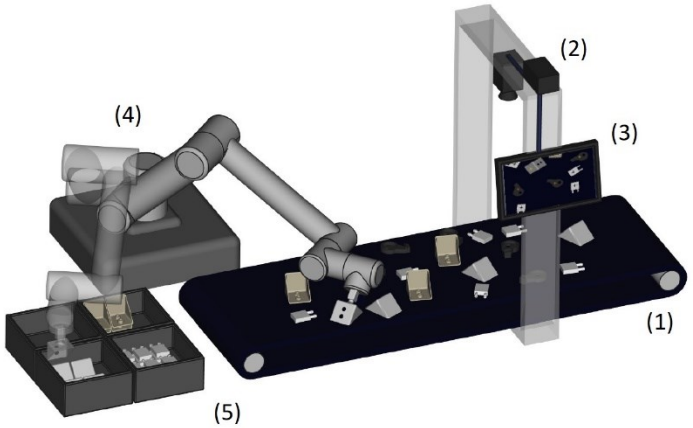


Figure 1: The arrangement of the task - the conveyor belt (1) brings the objects, the industrial camera (2) takes the image of the area, the detection and positioning system (3) determines the grasp points for manipulation and the robotic arm (4) puts the objects to the desired positions (5)

3. Proposed Approach

We propose that the approach for object detection and positioning is composed of two parts. The first part is designed to perform the scanning of the area and it provides the visual data to be processed by the following part. The second part then processes the data and provides particular detected objects and their positions. Using this information, a parent control system should be able to manipulate the objects in order to achieve the desired positions.

3.1. Camera sensor

In order to achieve a sufficient framerate, we propose to implement an ordinary industrial monocular RGB sensor equipped with a corresponding lens as the source of input image. Ostensibly, the camera and lens should be chosen according to the situation in the specific task (the scanned scene size, light, the distance of the camera from the objects, etc.). The tutorials of a camera sensor and lens combination selection are available at the vision technology producer information sources e.g. the Basler Lens Selector provided by Basler AG.

3.2. Detection and positioning system

Various studies published in recent years prove that the family of convolutional neural network topologies (CNN) outperforms classical image processing methods in tasks of object detection and classification benchmarked on various datasets [14, 15]. Following this fact, we propose a detection and positioning system, that uses CNN to transform the original RGB image of the monitored area into a specific schematic image. The main purpose of this particular operation is to create a graphic representation, where the positions of the detected objects are highlighted as gradient circles in defined colors, while the rest of the image is black. Specifically, each pixel in an RGB image representing the scene is labeled with a float number in the range <0; 1>, where 1 means the optimal and 0 means the most unsuited grasp point position. These labels are situated in the R layer, G layer, B layer or all layers of the output of the CNN, according to the type of object. Hence, the positions of the gradient circles represent not only the positions of the detected objects, but the exact points on the object body, which are optimal for manipulation by a robotic arm (grasp points). The proposed approach is described in Figure 2.

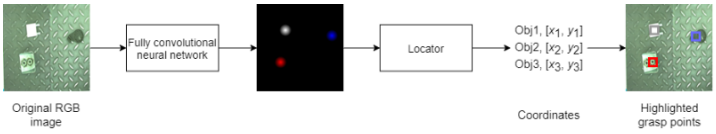


Figure 2: The proposed approach - the original RGB image of the scanned area is transformed with a convolutional neural network into the schematic image, where the optimal grasp point positions for manipulation are highlighted by gradient circles of various colors, each particular color then represents the type of the object

The first step of the proposed procedure, i.e. the transformation of the RGB image into a graphical representation of the object positions, is a key element of our approach. We propose to implement a fully convolutional neural network connected as an encoder-decoder processor. Such a processor provides encoding the original input into a small shape and restoring it using the decoder's capabilities. If the process is successful, the approach provides a correctly transformed image as the output from the decoder. We believe, that the correctly designed fully convolutional neural network is able to code gradient circles of defined colors on the exact positions of grasp points on the object bodies.

Therefore, in the next subsection, we introduce a specific fully convolutional neural network, that transforms an original RGB image into a graphical representation, where object grasp points are represented as radial gradients of defined colors.

3.3. Fully convolutional neural network for image transformation

Apparently, many different fully convolutional neural networks, such as ResNet [16], SegNet [17] or PSPNet [18], have been proposed to deal with various image processing tasks. From a wide family of neural network topologies, we select U-net as an initial point of development.

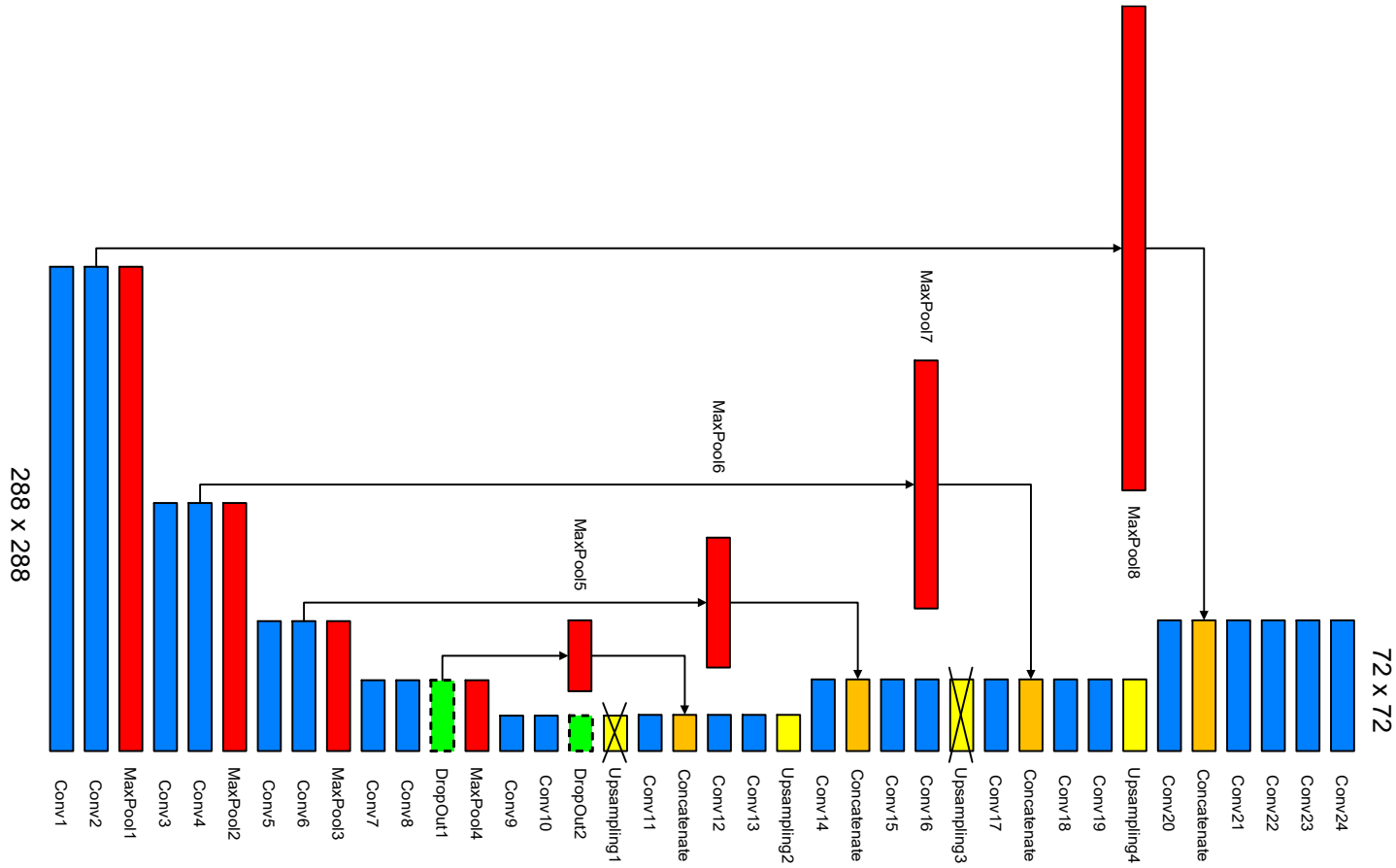


Figure 3: The development of the proposed topology of a fully convolutional neural network - layers removed from the original U-Net are crossed

U-Net is a fully convolutional neural network proposed initially for image segmentation tasks in biological and medical fields. However, it was then adapted to many other applications across different fields. This topology follows a typical encoder-decoder scheme with a bottleneck. In addition, it also contains a direct link between the parts of the encoder and the decoder, which is allowing the network to propagate contextual information to higher resolution layers. U-net topology is adopted from [19].

We streamline U-Net to fit our task in a more efficient way. To be more specific, U-Net originally provides the output data of the same dimensions as the input data. Such a detailed output is not required for a "pick and place" problem in industrial applications. Hence, we reduce the decoder part of U-Net topology. In our case, the output data is 16 times lower, which still provides an accuracy sufficient enough for object detection and positioning, and the topology itself is less computationally demanding. See Figure 3, where the changes applied to U-Net topology are shown in detail.

3.4. Locator for positioning of grasp points

Positioning of the gradient circles in a schematic image (last step in Figure 2) is a generic process of finding local maxima of each implemented color. These positions of the maxima directly represent the grasp points for manipulation using a robotic arm.

Generally, the process of finding local maxima in an array can be performed in several ways. The most obvious solution is to find the indices of the values, which are greater than all their neighbors. However, this approach is very sensitive to noise or small errors in the input data. Hence, it is more appropriate to implement a maximum filter operation, which dilates the original array and merges neighboring local maxima closer than the size of the dilation. Coordinates, where the original array is equal to the dilated array, are returned as local maxima. Clearly, the size of the dilation must be set. In the proposed locator, it is suitable to set it equal to the radius of the gradient circles.

4. Case Study

The aim of this section is to demonstrate the features of our detection and positioning system through the solution of a particular task. The task is properly defined in the next subsection. Subsequently, we propose a hardware implementation of the system and, as the final step of the procedure, a fully convolutional neural network is trained to be able to transform the original RGB images into a graphical representation of object types and positions.

4.1. Object detection and positioning task

For this case study, we need to develop a system for different object detection and positioning. The four object combinations placed on five different types of surfaces were used. The objects of interest are shown in Figure 4 and the surfaces are shown in Figure 5. The objects are of a similar size. Three of them are metallic and one is of black plastic.

4.2. Hardware implementation

The system is composed from a camera sensor and a processing unit, which should process data acquired by the RGB sensor, in order to determine the types and positions of the objects of interest. In this case study, we implement a Basler acA2500-14uc industrial RGB camera as a data acquisition tool. This sensor is able to provide up to 14 5-MPx RGB frames per second. The camera is equipped with a Computar M3514-MP lens in order to monitor the 300 x 420 mm scan area from above at a distance of 500 mm.

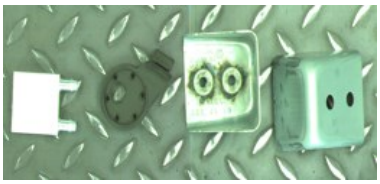


Figure 4: Objects of interest - objects are labelled as Obj1 to Obj4

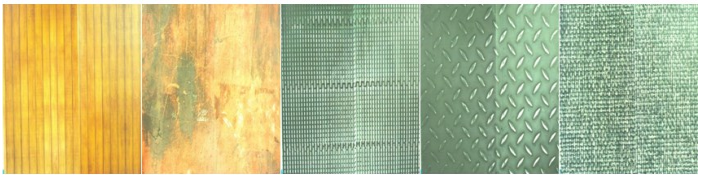


Figure 5: Various types of surfaces used in case study

The processing unit is supposed to be capable of processing images in real-time, as mentioned above. For this case study, the single-board NVIDIA Jetson NANO computer is used. It offers the NVIDIA Tegra X2 (2.0 GHz, 6 cores) CPU together with 8 GiB RAM. Furthermore, it provides wide communication possibilities (USB 2.0, 3.0, SATA, WiFi). The total price of components used for hardware implementation costs around $500.

4.3. Datasets

In this case study, we prepare 1021 original RGB images using various combinations of objects (Figure 4) and surfaces (Figure 5). In order to follow the topology of the CNN (Figure 3), we transform the images to 288 × 288 px. Then, we randomly divide the images into the training set (815 samples) and the testing set (206 samples).

After that, the trickiest part of the development follows. The target images for the training set (the graphical representations of the object types and positions) should be manually prepared. Hence, for any RGB image, we construct a target artificial image, where the optimal grasp point of each individual object in the original image is highlighted by a colored gradient circle. Four colors (red, green, blue and white) are implemented, since we consider four types of objects in this case study. We prepared a custom labeling application to prepare target images. In this application, each input image is displayed, and a human user labels all feasible grasp points using a computer mouse. The application then generates the target images. Several examples of input-expected output pairs are demonstrated in Figure 6.

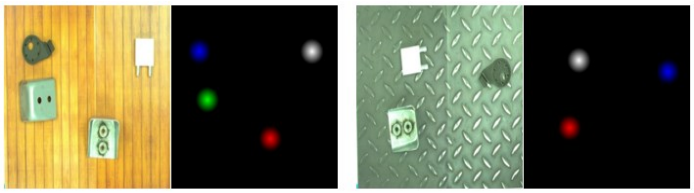


Figure 6: Two input-target pairs in training set - input image resolution is 288x288 px, target image resolution is 72x72 px

4.4. Fully convolutional neural network training

As the last step of the development, we train the network topology depicted in Figure 3.

We select the ADAM algorithm for the neural network weights and biases optimization as it is generally considered as an acceptable performing technique in most of the cases [20]. The random initial weights set with gaussian distribution was used. The experiments are run fifty times in order to reduce the stochastic character of training. The best instance is then evaluated. The training process and its parameters are depicted in Figure 7.

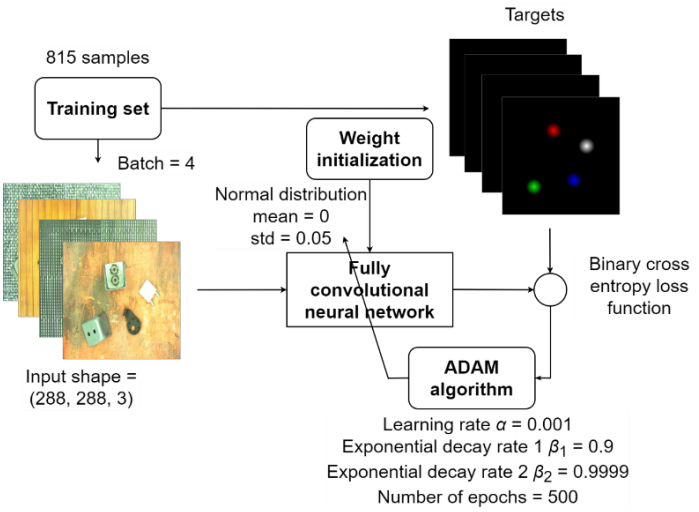


Figure 7: Training process

4.5. Results

We select a confusion matrix as an evaluation metric for detection and positioning system performance. The prediction of the type of the object and position of the grasp point is labeled as correctly predicted, if the local maximum position of the gradient circle of the defined color directly corresponds to the original position of the particular object using the 72 x 72 px map, i.e. the map defined by the target image.

The confusion matrix for the best network trained according to the previous paragraph for the testing set, is summarized in Table 1. Note that the number of correctly predicted free spaces in the image is not present, because it is essentially a black surface that cannot be explicitly evaluated. However, as seen in the table, the detection and positioning system provides 100 % accuracy over the testing set. In addition, implementing Jetson NANO described above, the detection and positioning system is capable of processing 13 frames per second, which is more than required at the beginning of the paper.

Table 1: Confusion matrix (206 images, 611 objects in total)

-	Obj1 pred.	Obj2 pred.	Obj3 pred.	Obj4 pred.	Free space pred.
Obj1 actual	149	0	0	0	0
Obj2 actual	0	153	0	0	0
Obj3 actual	0	0	158	0	0
Obj4 actual	0	0	0	151	0
Free space actual	0	0	0	0	Irrelevant

5. Conclusion

In this contribution, we introduced a novel engineering approach to object grasp point detection and positioning for "pick and place" applications. The approach is based on two consecutive steps. At first, a fully convolutional neural network is implemented in order to transform the original input image of the monitored scene. Output of this process is a graphical representation of the types and positions of the objects present in the monitored scene. Secondly, the locator is used to analyze the graphical representation to get the explicit information of object type and position.

We also performed a case study to demonstrate the proposed approach. In this study, the proposed system provided accurate grasp point positions of four considered objects for manipulation.

In future work, we will try to enhance the system in several ways. Critically, the approach should provide not only the positions of the grasp points, but also the required pose of the robotic arm. This feature will be advantageous especially for clamp grippers. Apart from that, we will try to optimize the processing unit, both from the hardware and software point of view, in order to get the close-to-optimal topology of the fully convolutional neural network and the hardware suitable for it.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The work was supported from ERDF/ESF " Cooperation in Applied Research between the University of Pardubice and companies, in the Field of Positioning, Detection and Simulation Technology for Transport Systems (PosiTrans)" (No. CZ.02.1.01/0.0/0.0/17_049/0008394).

References

[1] Y. Cheng, M.A. Jafari, "Vision-Based Online Process Control in Manufacturing Applications," IEEE Transactions on Automation Science and Engineering, **5**(1), 140-153, 2008, doi:10.1109/TASE.2007.912058.

[2] M. Bahaghighat, L. Akbari, Q. Xin, "A Machine Learning-Based Approach for Counting Blister Cards Within Drug Packages," IEEE Access, **7**, 83785-83796, 2019, doi:10.1109/ACCESS.2019.2924445.

[3] H. Sheng, S. Wei, X. Yu, L. Tang, "Research on Binocular Visual System of Robotic Arm Based on Improved SURF Algorithm," IEEE Sensors Journal, **20**(20), 11849-11855, 2020, doi:10.1109/JSEN.2019.2951601.

[4] S. Caldera, A. Rassau, D. Chai, "Review of Deep Learning Methods in Robotic Grasp Detection," Multimodal Technologies and Interaction, **2**(3), 2018, doi:10.3390/mti2030057.

[5] R. Miyajima, "Deep Learning Triggers a New Era in Industrial Robotics," IEEE MultiMedia, **24**(4), 91-96, 2017, doi:10.1109/MMUL.2017.4031311.

[6] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, J.A. Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," IEEE Transactions on Geoscience and Remote Sensing, **57**(9), 6690-6709, 2019, doi:10.1109/TGRS.2019.2907932.

[7] F. Xing, Y. Xie, H. Su, F. Liu, L. Yang, "Deep Learning in Microscopy Image Analysis: A Survey," IEEE Transactions on Neural Networks and Learning Systems, **29**(10), 4550-4568, 2018,

doi:10.1109/TNNLS.2017.2766168.
[8] A, Björnsson, M. Jonsson, K. Johansen, "Automated material handling in composite manufacturing using pick-and-place systems – a review," Robotics and Computer-Integrated Manufacturing, **51**, 222-229, 2018, doi:10.1016/j.rcim.2017.12.003.
[9] P. Dolezel, Petr, D. Stursa, D. Honc, "Rapid 2D Positioning of Multiple Complex Objects for Pick and Place Application Using Convolutional Neural Network," in 2020 24th International Conference on System Theory, Control and Computing (ICSTCC), 213-217, 2020, doi:10.1109/ICSTCC50638.2020.9259696.
[10] C. Papaioannidis, V. Mygdalis, I. Pitas, "Domain-Translated 3D Object Pose Estimation," IEEE Transactions on Image Processing, **29**, 9279-9291, 2020, doi:10.1109/TIP.2020.3025447.
[11] J. Pyo, J. Cho, S. Kang, K. Kim, "Precise pose estimation using landmark feature extraction and blob analysis for bin picking," in 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), 494-496, 2017, doi:10.1109/URAI.2017.7992786.
[12] J. Kim, H. Kim, J.-I. Park, "An Analysis of Factors Affecting Point Cloud Registration for Bin Picking," in 2020 International Conference on Electronics, Information, and Communication (ICEIC), 1-4, 2020, doi:10.1109/ICEIC49074.2020.9051361.
[13] P. Dolezel, J. Pidanic, T. Zalabsky, M. Dvorak, "Bin Picking Success Rate Depending on Sensor Sensitivity," in 2019 20th International Carpathian Control Conference (ICCC), 1-6, 2019, doi:10.1109/CarpathianCC.2019.8766009.
[14] S. Krebs, B. Duraisamy, F. Flohr, "A survey on leveraging deep neural networks for object tracking," in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 411-418, 2017, doi:10.1109/ITSC.2017.8317904.
[15] Y. Xu, X. Zhou, S. Chen, F. Li, "Deep learning for multiple object tracking: a survey," IET Computer Vision, **13**(4), 355-368, 2019, doi:10.1049/iet-cvi.2018.5598.
[16] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, 2016, doi:10.1109/CVPR.2016.90.
[17] V. Badrinarayanan, A. Kendall, R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, **39**(12), 2481-2495, 2017, doi:10.1109/TPAMI.2016.2644615.
[18] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, "Pyramid Scene Parsing Network," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 6230-6239, 2017, doi:10.1109/CVPR.2017.660.
[19] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241, 2015, doi:10.1007/978-3-319-24574-4_28.
[20] E.M. Dogo, O.J. Afolabi, N.I. Nwulu, B. Twala, C.O. Aigbavboa, "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," in 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 92-99, 2018, doi: 10.1109/CTEMS.2018.8769211.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 922-927 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

The Mediating Role of Entrepreneurial Orientation on the Knowledge Creation-Firm Performance Nexus: Evidence from Indonesian IT Companies

Desman Hidayat[1,*], Edi Abdurachman[2], Elidjen[2], Yanthi Hutagaol[3]

[1]*BINUS Entrepreneurship Center, Management Department, Bina Nusantara University, Jakarta, Indonesia 11480*

[2]*Management Department, BINUS Business School Doctor of Research in Management, Bina Nusantara University, Jakarta, Indonesia 11480*

[3]*International Accounting & Finance Program, Accounting Department, Faculty of Economics & Communication, Bina Nusantara University, Jakarta, Indonesia 11480*

ARTICLE INFO

Article history:
Received: 16 September, 2020
Accepted: 21 January, 2021
Online: 12 February, 2021

Keywords:
Knowledge Creation
Entrepreneurial Orientation
Firm Performance

ABSTRACT

Disruptive innovation has created fast changes in the business environment and competition among companies, especially on information technology companies. Knowledge creation and entrepreneurial orientation are two variables that can improve firm performance. There is still limited study on how knowledge creation and entrepreneurial orientation both affects firm performance. This study aims to discuss how to effectively apply knowledge creation and entrepreneurial orientation to develop firm performance. A questionnaire has been conducted to 55 medium-large IT companies in Jakarta, Indonesia, and analyzed using structural equation modeling (SEM). The result showed that knowledge creation did not directly affect firm performance but indirectly affected entrepreneurial orientation. Knowledge creation also had a positive and significant effect on entrepreneurial orientation, and so does entrepreneurial orientation towards firm performance. Therefore, IT companies should consider both variables to improve their performance. Future studies may consider using qualitative or mixed-method approaches, conducting research for small IT companies and in other countries.

1. Introduction

The world is changing fast with the innovation that happens in the world. Disruptive innovation, where a new market disrupts and replaces the old market, shows how quickly the world is changing [1]. This situation raises the competition among companies as well [2].

Disruptive innovation heavily impacts the information technology (IT) sector. As the largest economy in South East Asia, Indonesia has many growing industries, especially in the digital sector [3]. Nevertheless, IT companies need to prepare themselves for disruption and the competitive environment to survive.

According to dynamic capabilities theory, companies need to explore their knowledge assets to face rapid technological change [4]. Knowledge is an essential part of IT companies that focus more on intangible assets to compete. Knowledge creation is a way to create value for IT companies.

Technology advancement that changes rapidly can help companies find new opportunities to improve their performance [5]. One of the sectors that are affected by these changes is the IT sector. So, IT companies need to find a way to seek opportunities within this condition. Entrepreneurial orientation can be one way where they will be able to survive the competition and to increase their performance [6].

Although there are a lot of previous studies talking about how knowledge creation is related with entrepreneurial orientation [7, 8] and how each of those variables are related with firm performance [9], [10], but there is still limited study on how knowledge creation and entrepreneurial orientation both affects firm performance. It is also interesting to see how the variables will be related in IT sector that relies on intangible assets, such as knowledge. These explanations show the novelty of this study.

* Corresponding Author: Desman Hidayat, Email: d4906@binus.ac.id

This study aims to enrich the literature by investigating the interaction between knowledge creation, entrepreneurial orientation, and firm performance. Overall, this study makes two contributions: (1) It expands the understanding of the knowledge creation-firm performance relationship in IT companies, and (2) It explores how entrepreneurial orientation can be related to knowledge creation and firm performance.

2. Literature Review

2.1. Firm Performance

Firm performance is a variable that is often used to measure how good companies run. Researches often use firm performance as a dependent variable [11]. This study used firm performance to measure how good IT companies manage their businesses.

Firm performance is derived from the organizational effectiveness theory [12]. The performance of a company can show the effectiveness of that organization [13]. It takes more than just the financial factor to measure performance [12].

In this study, the firm performance was measured using financial and non-financial performance. It used five indicators: revenue, ROI, employees, products, and development [14, 15]. Financial performance was measured by using revenue and ROI, while non-financial performance was measured using employees, products, and development.

2.2. Knowledge Creation

Knowledge creation is an activity or process of developing new knowledge by sharing and combining tacit and explicit knowledge [16]. It enables firms to improve efficiency and create value [17]. Knowledge creation needs participation from individual members of an organization to be effective [18].

Knowledge-based view of the firm is the most common foundation used to define knowledge creation theory [19]. This view argued that firms' significant resources are mainly intangible and dynamic, such as knowledge [20]. This view is derived from dynamic capabilities.

Socialization, Externalization, Combination, and Internalization (SECI) are indicators to measure knowledge creation [18]. These indicators are representing the interaction between tacit and explicit knowledge [19]. This study used SECI as indicators to measure knowledge creation.

2.3. Entrepreneurial Orientation

Entrepreneurial orientation is one of the most researched topics in entrepreneurship literature [21]. It is the processes, practices, philosophy, and decision-making activities that help companies innovate [22]. Companies with the right entrepreneurial orientation continuously try to find new opportunities and strengthen their competitive positions [8].

The foundation of entrepreneurial orientation was based on entrepreneurship theory itself, where the main point of entrepreneurship is to understand how companies can seek and exploit opportunities [23]. Opportunities do not have to be related to something new, but they can focus on optimizing the existing framework. To discover the opportunities, companies must possess prior information related to the opportunities and cognitive properties to value them.

Five indicators are commonly used to measure entrepreneurial orientation. Those indicators are autonomy, innovativeness, risk-taking, proactiveness, and competitive aggressiveness [6]. The earlier concept of entrepreneurial orientation used three aspects: innovativeness, proactiveness, and risk-taking [24], but now it has been improved by adding the other two indicators. This study used the five indicators mention above to measure entrepreneurial orientation.

2.4. Hypothesis Development

Knowledge is one of the critical intangible resources that can help companies develop their performance [9, 25]. A lot of previous studies discuss how vital knowledge creation in relationship with performance [19]. High-tech companies need to have adequate knowledge resources in order to remain competitive [26]. Hence the hypothesis:

H1: There is a positive effect of knowledge creation towards firm performance.

Entrepreneurial orientation has been known to have a positive association with firm performance [8,9]. It has been tested in different contexts and countries [27]. This statement is also argued to be true for technological companies [28]. Therefore, hypothesis two is predicted as below.

H2: There is a positive effect of entrepreneurial orientation towards firm performance.

Knowledge creation and entrepreneurial orientation are often researched together. Previous studies showed that both variables are related [7, 8]. IT companies need to pay attention to both variables to maximize their performance. Therefore, hypothesis three states:

H3: There is a positive effect of knowledge creation towards entrepreneurial orientation.

Even though most of the previous research discussed how knowledge creation and entrepreneurial orientation separately affect firm performance, some studies talked about how they can simultaneously affect firm performance by having entrepreneurial orientation as the mediating variable [29]. This study argued that innovativeness and competitive aggressiveness as part of entrepreneurial orientation play the mediator between knowledge and performance. Hence the hypothesis:

H4: There is a positive indirect effect of knowledge creation towards firm performance mediated by entrepreneurial orientation.

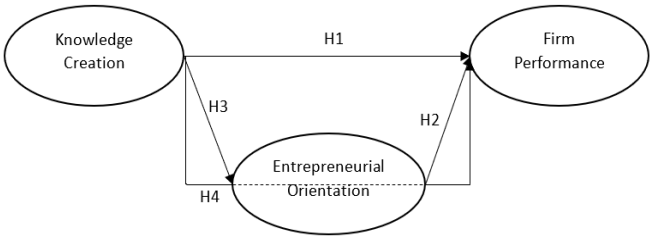


Figure 1: Research Model

3. Research Methodology

This study was conducted using an online survey with quantitative analysis. The unit analysis of this study is medium-

large IT companies in Jakarta. Indonesian Micro, Small, and Medium Enterprise Regulation No.20 (2008) defined medium enterprises as companies with 2.5-50 billion IDR revenues per year or 178-3.5 million USD and large enterprises as companies that have >50 billion IDR revenues per year or >3.5 million USD. According to 2016 Economic Census by Statistics Indonesia, the population of medium-large IT companies in Jakarta, the capital of Indonesia, is 303 companies [30].

This study used probability sampling design with simple random sampling. The samples for this study were 55 companies. The minimum sample size was 33 data, calculated using two arrows pointing at a construct, 5% significance level, and minimum R Square 0.25 [31]. Therefore, the samples are sufficient for this study. The survey was conducted for three months, from August to November 2020. The unit of observation in this study was a managerial level employee. Only one employee per company participated in this study.

Based on the companies' established period, most of the respondents (42%) were companies that have already been established for at least twenty years, and the smallest frequency came from new companies that have only been established for less than five years (7%). This data shows that many medium and large IT firms in Jakarta took a long time to develop their business into a medium-large company. Table 1 shows the details of the company establishment period.

Table 1: Companies' Established Period

Established Period	Frequency	Percentage
<5 years	4	7%
5-9 years	11	20%
10-14 years	9	16%
15-19 years	8	15%
>=20 years	23	42%
TOTAL	55	100%

Most of the company respondents have 11-50 employees (40%). Only one respondent has less than ten employees, while seven companies have more than five hundred employees. This data shows that medium or large companies do not mean that they must have many employees. Table 2 below shows more details on the numbers of employees.

Table 2: Companies' Number of Employees

Number of Employees	Frequency	Percentage
<=10 employees	1	2%
11-50 employees	22	40%
51-200 employees	20	36%
201-500 employees	5	9%
>500 employees	7	13%
TOTAL	55	100%

This study was done by using exploratory empirical research. The data analysis was done by partial least squares structural equation modeling (PLS-SEM). PLS-SEM is mainly used for the development of theories and exploratory research [31]. The analysis for this study was done using the smartPLS program. This study used a five-point Likert scale on the questionnaires. A five-point scale increases the response rate and quality of the responses while reduces the stress of the respondent [32].

4. Results and Discussion

4.1. Results

Validity was measured by removing items with less than 0.7 outer loadings. From the result, the risk-taking indicator measuring entrepreneurial orientation is not valid. Other indicators are all proven to be valid. Cronbach's Alpha, Composite Reliability (CR), and Average Variance Extracted (AVE) measure reliability. Variables should have Cronbach's Alpha >0.5, CR >0.6, and AVE >0.5 to be considered reliable [31]. The result showed that all the variables are reliable. More details on validity and reliability test result on entrepreneurial orientation (EO), firm performance (FP), and knowledge creation (KC) can be seen in table 3.

Table 3: Validity & Reliability Test Result

Variables	Indicators	Items	Instrument	Outer Loading	Cronbach's Alpha	CR	AVE
EO	Autonomy	EO11	Our company gives freedom to employees or team to express their business concept and vision, and oversees them until finish	0.783	0.892	0.913	0.569
		EO12	Our company has the self-directed ability and willingness to seek for opportunities	0.809			
	Innovativeness	EO21	Our company supports employees' creativity	0.771			
		EO22	Our company has a lot of new marketable products/services within the last five years	0.723			
	Proactiveness	EO41	Generally, our company's top managers have the tendency to lead the competition with new idea or product	0.721			
		EO44	Our company reacts quickly on the market demand	0.743			
	Competitive Aggressiveness	EO51	Our company is very aggressive	0.756			
		EO52	Our company is very competitive	0.723			
FP	Employee	FP12	Our company has good planning towards the future of employees	0.706	0.865	0.899	0.597

	Product	FP22	Our company's automation is higher than our competitors in the same industry	0.830			
	Development	FP31	Our company is very keen on investing in new market development	0.733			
		FP32	Our company is very keen on investing in new technology development	0.778			
	Revenue	FP41	Our company's income is higher than our competitors in the same industry	0.779			
	ROI	FP51	Our company's Return on Investment is higher than our competitors in the same industry	0.803			
KC	Socialization	KC11	Our company emphasizes on the creation of working environment that helps employees learn skills	0.700	0.930	0.939	0.586
		KC12	Our company arranges employee meetings to share and trade knowledge and experiences	0.742			
	Externalization	KC21	Our company uses collaborative learning tools	0.741			
		KC22	Our company usually develops working group discussion by using several techniques through internet	0.775			
		KC23	Our company usually share information, experiences, best practice, and learning to solve problems	0.732			
	Combination	KC31	Our company emphasizes on manual and document creation on products and services	0.835			
		KC32	Our company emphasizes content creation from management data gathering	0.791			
		KC33	Our company emphasizes content creation from technical information gathering	0.872			
	Internalization	KC41	Our company emphasizes on value seeking and sharing	0.764			
		KC42	Our company emphasizes on new thoughts seeking and sharing	0.709			
		KC43	Our company usually use on the job training to enrich knowledge	0.744			

Based on the result of Fornell-Lacker criterion, discriminant validity, all three variables are valid. Correlation between items and the square root of AVE shown no problem with the top numbers being the biggest one. Table 4 shows the details of discriminant validity.

Table 4: Fornell-Lacker Criterion

	EO	FP	KC
EO	**0.754**		
FP	0.752	**0.773**	
KC	0.689	0.570	**0.766**

The last part of discriminant validity measure beside loading factor and Fornell-Lacker criterion is Heterotrait-Monotrait Ratio (HTMT). The score for HTMT ratio should be <1.00 [31]. Based on the result, all variables are valid. Table 5 shows the HTMT result.

Table 5: HTMT Ratio

	EO	FP	KC
EO			
FP	0.848		
KC	0.703	0.616	

R Square shows the proportion of variation of dependent variables towards independent variables. Results showed that 57,1% of firm performance could be described through entrepreneurial orientation and knowledge creation, while 47.5% of entrepreneurial orientation can be described through knowledge creation. R square result can be seen in table 6.

Table 6: R Square

	R Square	R Square Adjusted
EO	0.475	0.465
FP	0.571	0.554

Goodness of fit can be seen from the value of Standardized Root Mean Square Residual (SRMR) and Normal Fit Index (NFI). On this study, the result of SRMR estimated model is 0.108 while the NFI score is 0.555. Although there are several PLS-SEM based model fit measures, but those measures are still in development [31].

After testing the validity and reliability, the next step was to test the path between variables. The test was done using SmartPLS. How the research model looked with the path coefficient and t-value can be seen in Figure 2.

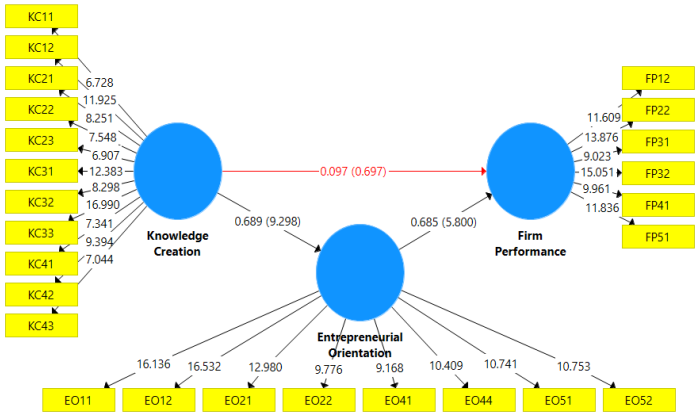


Figure 2: Research Model with Path Coefficient and T-Value

T-value should be >2.021 and p-value <0.050 for a significant path on a 5% error rate [33]. From the result, the direct effect of entrepreneurial orientation is positive and significant towards firm performance, and the direct effect of knowledge creation is positive and significant towards entrepreneurial orientation. However, knowledge creation has an insignificant direct effect on firm performance. By measuring the indirect effect, knowledge creation has a positive and significant effect on firm performance mediated by entrepreneurial orientation. Details on the structural equation model (SEM) test result can be seen in table 7.

Table 7: SEM Test Result

	Path Coef.	T-Stat	P-Value	Sig.
KC-> FP	0.097	0.697*	0.486*	Insignificant
EO -> FP	0.685	5.800	0.000	Positive & Significant
KC->EO	0.689	9.298	0.000	Positive & Significant
KC->EO->FP	0.472	4.873	0.000	Positive & Significant

4.2. Discussion

Based on this study, knowledge creation does not have a significant effect on firm performance. This argument supported previous research [34]. IT companies should not only be focusing on knowledge creation if they want to raise their performance. They should also think about other factors to make the knowledge creation effective. Therefore, hypothesis 1 is rejected.

Entrepreneurial orientation in this study is proved to have a positive and significant effect on firm performance. This statement strengthened the arguments done by previous researches [21,35,36]. By focusing on entrepreneurship, IT companies can find and exploit opportunities to improve their performance. Therefore, hypothesis 2 is accepted.

Knowledge creation has a positive and significant effect on entrepreneurial orientation, according to this study. This result strengthened previous research that argued the correlation between knowledge creation and entrepreneurial orientation [37] and supported the argument that showed knowledge creation affected entrepreneurial orientation [7]. IT companies must pay attention to knowledge creation to increase their entrepreneurial orientation. Therefore, hypothesis 3 is accepted.

This study showed that knowledge creation has a positive and significant effect on firm performance when it is mediated by entrepreneurial orientation. This result supported the previous research on the same topic [29]. Both knowledge creation and entrepreneurial orientation are essential factors to improve IT company performance. Therefore, hypothesis 4 is accepted.

4.3. Implication

IT companies generally rely on intangible resources such as knowledge as their important asset. Nevertheless, the result of this study showed that knowledge alone is not enough to make the company perform better. Therefore, medium-large IT companies in Jakarta should think about other factors while focusing on their knowledge creation.

Technology changes rapidly. IT companies should think about a way to face this challenge. Entrepreneurial orientation can help the companies to seek opportunities and adapt to changes. The result of this study showed that it is important to develop entrepreneurial orientation in the companies to increase their firm performance.

Another important result from this study is the relation between knowledge creation and entrepreneurial orientation. This study showed that knowledge creation is important to develop entrepreneurial orientation, which in the end raise the firm performance. Therefore, IT companies should develop both their knowledge creation and entrepreneurial orientation.

5. Conclusion

5.1. Conclusion

The results of this study indicate the factors that can increase IT firm performance. This study's main conclusion is that knowledge creation and entrepreneurial orientation are essential antecedents of firm performance, especially for companies in the IT sector. However, knowledge creation alone is not enough to raise firm performance. It must also be mediated by entrepreneurial orientation.

The interesting finding in this study is that entrepreneurial orientation mediates the relationship between knowledge creation and firm performance. It indicates that IT companies should create an environment that supports knowledge creation for the employees. Companies should also have entrepreneurial orientation to adapt with the fast-changing environment.

5.2. Limitation and Future Work

This study has several limitations. The analysis in this study was done by quantitative approach. Adding qualitative approaches by having interview or focus group discussion on future research will enrich this study. Mixed-method research is also a good approach to understand more about the industry.

The companies observed is also limited to medium-large IT companies. Future research can be done to small IT companies. The knowledge creation might be different in small companies and their way to face the changes can also be different. Therefore,

doing study in small IT companies or conducting comparative study between small, medium, and large IT companies will enrich this study.

Other limitation is the scope of this research. This study used IT companies in Jakarta, Indonesia. Therefore, it will be hard to generalize this result to other countries. Future research ought to be held on other countries as well.

References

[1] C.M. Christensen, The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail, Harvard Business School Press, Boston, Massachussets, USA, 1997.

[2] J.R.L. Kaivo-oja, I.T. Lauraeus, "The VUCA approach as a solution concept to corporate foresight challenges and global technological disruption," Foresight, **20**(1), 27–49, 2018, doi:10.1108/FS-06-2017-0022.

[3] K. Das, M. Gryseels, P. Sudhir, K.T. Tan, Unlocking Indonesia's Digital Opportunity, 2016.

[4] D.J. Teece, G. Pisano, A. Shuen, "Dynamic Capabilities and Strategic Management," Strategic Management Journal, **18**(7), 509–533, 1997.

[5] S. Venkataraman, "The Distinctive Domain of Entrepreneurship Research," Advances in Entrepreneurship, Firm Emergence and Growth, **3**, 119–138, 1997.

[6] G. Lumpkin, G.G. Dess, "Linking two dimensions of entrepreneurial orientation to firm performance," Journal of Business Venturing, **16**(5), 429–451, 2001, doi:10.1016/S0883-9026(00)00048-3.

[7] C. Weerakoon, A.J. McMurray, N. Rametse, P. Arenius, "Knowledge creation theory of entrepreneurial orientation in social enterprises," Journal of Small Business Management, **58**(4), 834–870, 2020, doi:10.1080/00472778.2019.1672709.

[8] N.A. Omar, K. Professional, M.A. Nazri, "The Effect of Entrepreneurial Orientation , Innovation Capability and Knowledge Creation on Firm Performance : A Perspective on Small Scale Entrepreneurs," Jurnal Pengurusan, **48**, 187–200, 2016.

[9] M.S. Aliyu, H.B. Rogo, R. Mahmood, "Knowledge Management , Entrepreneurial Orientation and Firm Performance : The Role of Organizational Culture Knowledge Management , Entrepreneurial Orientation and Firm Performance : The Role of Organizational Culture," Asian Social Science, **11**(23), 2015, doi:10.5539/ass.v11n23p140.

[10] B. Liu, J. Wang, "Demon or angel: an exploration of gamification in management," Nankai Business Review International, **11**(3), 317–343, 2020, doi:10.1108/NBRI-02-2018-0013.

[11] J.B. Santos, L.A.L. Brito, "Toward a subjective measurement model for firm performance," BAR - Brazilian Administration Review, **9**(SPL. ISS), 95–117, 2012, doi:10.1590/S1807-76922012000500007.

[12] K. Cameron, "A STUDY OF ORGANIZATIONAL EFFECTIVENESS AND ITS PREDICTORS," Management Science, **32**(I), 87–112, 1986.

[13] T. Connolly, E.J. Conlon, S.J. Deutsch, "Organizational Effectiveness: A Multiple-Constituency Approach.," Academy of Management Review, **5**(2), 211–218, 1980, doi:10.5465/amr.1980.4288727.

[14] R.L. Daft, Management, 10th ed., Cengage Learning, Ohio, 2012.

[15] S.M. Tseng, P.S. Lee, "The effect of knowledge management capability and dynamic capability on organizational performance," Journal of Enterprise Information Management, **27**(2), 158–179, 2014, doi:10.1108/JEIM-05-2012-0025.

[16] P.C. Izunwanne, "Developing an Understanding of Organisational Knowledge Creation: A Review Framework," Journal of Information & Knowledge Management, **16**(2), 2017, doi:10.1142/S0219649217500204.

[17] M. Tsai, Y. Li, "Knowledge creation process in new venture strategy and performance," **60**(1), 371–381, 2007, doi:10.1016/j.jbusres.2006.10.003.

[18] I. Nonaka, "A Dynamic Theory of Organizational Knowledge Creation," Organization Science, **5**(1), 14–37, 1994, doi:10.1287/orsc.5.1.14.

[19] D. Hidayat, E. Abdurachman, Elidjen, Y. Hutagaol, "Empirical Studies on Knowledge Creation and Performance : a Literature Review," in 2020 International Conference on Information Management and Technology (ICIMTech) 533, Bandung, Indonesia: 533–537, 2020.

[20] C. Curado, "The Knowledge Based-View of the Firm: From Theoretical Origins To Future Implications," Igarss 2014, (1), 1–5, 2006, doi:10.1007/s13398-014-0173-7.2.

[21] H. Montiel-Campos, "Entrepreneurial orientation and market orientation: Systematic literature review and future research," Journal of Research in Marketing and Entrepreneurship, JRME-09-2017-0040, 2018, doi:10.1108/JRME-09-2017-0040.

[22] M. Madhoushi, A. Sadati, H. Delavari, M. Mehdivand, R. Mihandost, "Entrepreneurial Orientation and Innovation Performance: The Mediating Role of Knowledge Management," Asian Journal of Business Management, **3**(4), 310–316, 2011.

[23] S. Shane, S. Venkataraman, "The Promise of Entrepreneurship as a Field of Research," The Academy of Management Review, **25**(1), 217, 2000, doi:10.2307/259271.

[24] D. Miller, "The Correlates of Entrepreneurship in Three Types of Firms," Management Science, **29**(7), 770–791, 1983, doi:10.1287/mnsc.29.7.770.

[25] P. Heisig, O.A. Suraj, A. Kianto, C. Kemboi, G. Perez Arrau, N. Fathi Easa, "Knowledge management and business performance: global experts' views on future research needs," Journal of Knowledge Management, **20**(6), 1169–1198, 2016, doi:10.1108/JKM-12-2015-0521.

[26] F. Zouaghi, M. Sánchez, M.G. Martínez, "Did the global financial crisis impact firms' innovation performance? The role of internal and external knowledge capabilities in high and low tech industries," Technological Forecasting and Social Change, **132**(April 2016), 92–104, 2018, doi:10.1016/j.techfore.2018.01.011.

[27] R. Basco, F. Hernández-Perlines, M. Rodríguez-García, "The effect of entrepreneurial orientation on firm performance: A multigroup analysis comparing China, Mexico, and Spain," Journal of Business Research, **113**(May), 409–421, 2020, doi:10.1016/j.jbusres.2019.09.020.

[28] Y. Liu, M. Wang, "Entrepreneurial orientation, new product development and firm performance: the moderating role of legitimacy in Chinese high-tech SMEs," European Journal of Innovation Management, 2020, doi:10.1108/EJIM-05-2020-0204.

[29] R. Zacca, M. Dayan, T. Ahrens, "Impact of network capability on small business performance," Management Decision, **53**(1), 2–23, 2015, doi:10.1108/MD-11-2013-0587.

[30] Badan Pusat Statistik, Sensus Ekonomi 2016 Direktori Usaha/Perusahaan Menengah Besar Kategori Informasi dan Komunikasi, 2017.

[31] J. Hair Jr, G.T. Hult, C. Ringle, M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2017.

[32] S. Sachdev, H. Verma, "Relative importance of service quality dimensions: a multisectoral study," Journal of Services Research, **4**(1), 93, 2004.

[33] U. Sekaran, R. Bougie, Research Methods for Business, 7th Editio, Wiley, West Sussex, UK, 2016, doi:10.1007/978-94-007-0753-5_102084.

[34] C.E. Lucier, J.D. Torsilieri, "Why knowledge programs fail: A CEO's guide to managing learning," Strategy & Business, **9**(4), 14–28, 1997.

[35] M.I. Hanif, F. Malik, A.B. Abdul Hamid, A.B.A. Hamid, "The effect of knowledge management and entrepreneurial orientation on organization performance," Journal of Entrepreneurship Education, **21**(4), 2651, 2018.

[36] R. Masa'deh, J. Al-Henzab, A. Tarhini, B.Y. Obeidat, "The associations among market orientation, technology orientation, entrepreneurial orientation and organizational performance," Benchmarking, **25**(8), 3117–3142, 2018, doi:10.1108/BIJ-02-2017-0024.

[37] F. Vidic, "Entrepreneurial Orientation (EO) and Knowledge Creation (KC)," International Journal of Economic Sciences and Applied Research, **6**(2), 103–124, 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 949-955 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Numerical Analysis for Feature Extraction and Evaluation of 3D Sickness

Kohki Nakane, Rentaro Ono, Hiroki Takada*

Department of Human & Artificial Intelligence Systems, Graduate School of Engineering, University of Fukui, 910-8507, Japan

ARTICLE INFO

Article history:
Received: 08 December, 2020
Accepted: 21 March, 2021
Online: 15 April, 2021

Keywords:
Generative adversarial network (GAN)
Statokinesigram (SKG)
Stochastic process
Double-Wayland algorithm

ABSTRACT

Artificial intelligence (AI) systems have been applied not only to numerical simulations of the economical sequences but also to the bio-signal, for instance, the statokinesigrams (SKGs). According to the nonlinear analysis of the bio-signal, we have considered that the motion process of the body sway is more random than that of the other bio-signal. In this study, we proposed a method for the numerical analysis of biological data using AI. The AI numerical solutions can indicate graphs that are very similar to the SKGs in degree of the determinism. In addition, we succeeded in extracting partial figure patterns that the AI regarded as a feature of 3D sickness. Compering with the properties resulting from the mathematical analysis, interpretations can be given for the black box processing in the AI.

1. Introduction

In recent years, artificial intelligence (AI) has been incorporated into the automatic algorithms for buying and selling stocks and bonds. NASDAQ-100, actual stock prices, Standard & Poor's 500 Stock Index, and historical exchange rates have been analyzed to develop an automated trading algorithm. However, when using past stock price fluctuations, there is a drawback in that the test pattern is limited. In addition, when the time series generated by the stochastic process is used, there is a drawback wherein the actual fluctuation of the stock price is not reflected. Therefore, neither case is sufficient to trust the test results. In this study, we attempted to create a mathematical model that can generate innumerable fluctuation patterns after reflecting the fluctuation characteristics of the time series. In previous studies, the generative adversarial network (GAN) was applied to the time-series generation of exchange rates. GAN is attracting attention because it has succeeded in generating high-precision images. Stationarity [1], fractalness [2], and determinism [3] were measured to compare the actual exchange rates with the pseudo exchange rates generated by GAN. Using these indices, we also measured the similarity between the actual exchange rates and the time series generated by the Winner process [4]. From the stationary perspective, the pseudo-exchange rates generated by GAN showed higher similarity than those of the Winner process, and from the deterministic point of view, both showed higher similarity.

On February 2, 2018, the Dow Jones Industrial Average had the largest decline to date [5]. The actual reason for this decline has not been clarified, but one of the possible causes cited by market participants is continuous selling, such as a chain of automatic loss cuts by the AI or automatic trading algorithms [6]. As a similar example, from March 9th to 18th, 2020, the decline in Standard & Poor's 500 Stock Index triggered four circuit breakers, temporarily suspending trading on the NYSE and NASDAQ [7–10]. This may indicate that the introduction of AI or automated trading algorithms is accelerating in financial markets. Some financial institutions, for instance, offer mechanical investment services to consumers. These services have solved the wealth-management problem of the general public.

However, mechanical investment services using AI cannot always protect assets. This is because reliability and robustness of the system is insufficient, owing to the limited number of tests. As mentioned above, the automatic trading algorithm uses past fluctuations in stock prices and exchange rates. In addition, even if the results look good at first glance, they may be overfitting to past fluctuations. To solve this problem, we propose an evaluation using a new simulated time series that does not depend only on historical data of past stock prices and exchange rates. It is conceivable to use the stochastic process as an idea to generate a simulated time series, but so far, many researchers have considered the stock price as the stochastic process [11, 12]. Another approach is the generation of a simulated time series by GAN. In previous

*Corresponding Author: Hiroki Takada, Department of Human & Artificial Intelligence Systems, Graduate School of Engineering, University of Fukui, 910-8507, Japan Email: takada@u-fukui.ac.jp

studies, the pseudo-exchange rate generated by GAN showed higher similarity to the actual exchange rate in terms of stationarity than the time series generated by the Winner process. In terms of determinism, both showed high similarity.

The body sway is of particular interest regarding the assessment of changes in the postural control, as it is highly related to the risk of falling [13]. In the stabilometry, displacement in the center of pressure (CoP) is recorded in the medial-lateral (x) and the anterio-posterior (y) directions as a statokinesigram (SKG). In the SKGs, variables x and y are regarded to be independent [14]. The linear stochastic differential equation (Brownian motion process) has been proposed as a mathematical model to describe the body sway [15–17].

Therefore, the purpose of this study was to establish a new numerical simulation method for biological signals using the same method as the numerical simulation in the exchange rate. We examined whether it is possible to generate a highly accurate simulation for the SKGs by applying the GAN to the SKGs with an unknown mechanism by which 3D images affect the CoP. In addition, if we were able to generate SKGs, we would also verify whether the characteristics of SKGs could be visualized.

2. Mathematical Model

Stochastic differential equations (SDEs)

$$\dot{x} = -gradU_x(x) + \mu_x w_x(t), \quad (1)$$

$$\dot{y} = -gradU_y(y) + \mu_y w_y(t), \quad (2)$$

have been proposed as mathematical models describing the body sway [18-20], where μ and ω represent the noise amplitude and the Gaussian white noise, respectively. In Eqs.(1.1)-(1.2), the time-average potential functions in the *x*- and *y*-directions are expressed as $U_x(x)$ and $U_y(y)$, respectively. These functions were estimated from

$$U_x(x) = -\frac{\mu_x^2}{2} lnG_x(x) + const., \quad (3)$$

$$U_y(y) = -\frac{\mu_y^2}{2} lnG_y(y) + const., \quad (4)$$

where distributions in the *x*- and *y*-directions are expressed as $G_x(x)$ and $G_y(y)$, respectively. Each distribution is determined as a histogram for each component in the experiment. In some studies, the SDEs have been applied to conduct the numerical analysis of the body sway, however, there is no research using the AI. Using the AI, we perform a numerical simulation in this research.

3. Model Design

3.1. GAN

In recent years, neural network models, for instance, convolutional neural networks (CNNs) have been used as image-classification models [21].

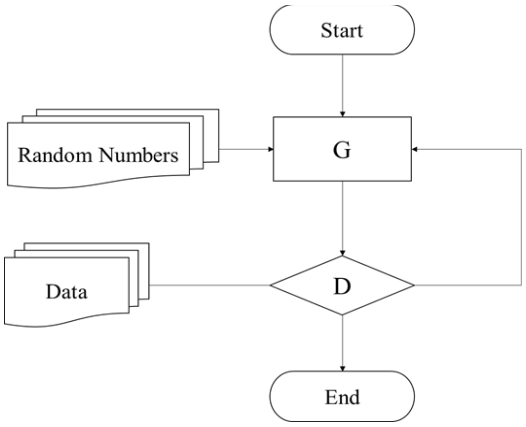


Figure 1: GAN model.

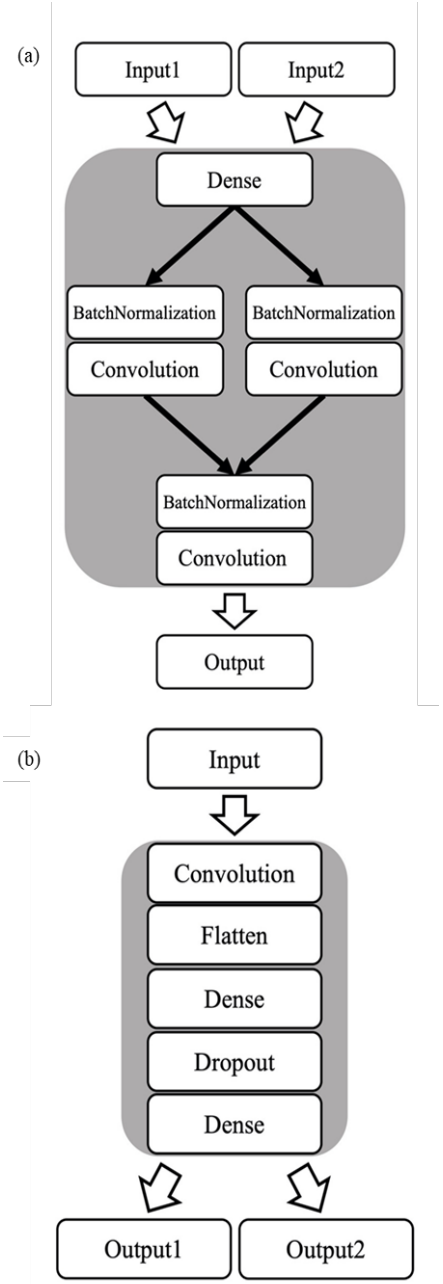


Figure 2: Constructed GAN model: generator (a), discriminator (b).

In addition, deep-neural-network, which are deeply stacked CNNs, has been used not only for image classification but also for a wide range of applications such as object detection [22], semantic segmentation [23], and the image generation [24]. Two advantages of the GANs use are that it is not necessary to define properties in the images to be generated beforehand, and that a number of images can be generated by using noise as the input to the network. In this study, the GAN was applied to the generator of time series. We can generate time series for the numerical simulation.

In this study, simulative time series (i.e., fake data) can be obtained from the generator G, which is defined as a network input noise. As the side note, the discriminator D distinguishes whether generated data are real (i.e., true data). "G" learns so that "D" determines the fake data as genuine. Inversely, "D" learns so that it can correctly determine whether the input data is true data or false data generated by "G". By repeatedly learning these networks (Figure 1), it can be generated for a large number of simulated time series if the output of "G" can generate data that is very close to true data.

3.1. Learning Statokinesigrams in GAN Model

It has been reported that it is difficult to stabilize in the GAN learning. Several models were, in fact, designed and trained to generate pseudo-exchange rates, however, few models could be found to exhibit stable learning. The way has been already suggested how to stabilize the GAN learning [25]. Based on these proposals, the hyperbolic tangent function (tanh) was herein used as an activation function in the output layer of the generator. LeakyReLU was used as an activation function, except for the output layer/fully connected layer in "D" [26]. The hyperbolic tangent function is defined as

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}},$$

where the derivative is given as

$$\tanh'(x) = 1 - \tanh(x)^2.$$

The advantage of using Eq. (4) is firstly that it can output any input value within the range of -1 to 1, thus reducing numerical divergence. Secondly, the hyperbolic tangent function does not have a large maximum value of the differential coefficient, which makes it stop vanishing gradients.

3.2. Parameter Optimization

To design a neural network model that achieves the desired performance, it is necessary to adjust the layer configuration to the optimum configuration. In general models for classification and prediction, the accuracy rate is often used as the objective variable [27]. However, because the aim of this study is to simulate the characteristics of SKGs, it cannot be evaluated using the accuracy rate. Therefore, it is necessary to define a new objective variable. In this study, we defined the following optimization function using the learning error for "G" and "D", which was developed in a previous study [28].

$$Opt.Function(G_{LOSS}, D_{LOSS}) = \ln \frac{G_{LOSS} + D_{LOSS}}{D_{LOSS} / G_{LOSS}}, \qquad (5)$$

where the training error for "G" and the training error for "D" are expressed as G_{LOSS} and D_{LOSS}, respectively.

In this study, parameters in the GAN model (see Appendix) were set to minimize the value of the optimization function [29]. The value was evaluated after the parameters were optimized. It should be noted that the parameters in "D" was fixed because of the high computational expense (Table 1). Figures 2 show the a generator model and a discriminator model developed in this study. The "G" uses the combination input of noise and the label. In the subsequent generation process, a branch structure was added so that the variables x and y could be generated independently. This is based on the idea that the variables x and y are considered to be independent in the body sway. The "D", in contrast to the "G", did not have a branched structure. This is because both series are referred to when determining the state of body sway.

Table 1: Network configuration in the discriminator.

Layers	Layer Name	Units	Kernel Size	Filters	Output Shape	Activation
0	Input	-	-	-	-	-
1	Convolution	-	1×3	32	2×600	LeakyReLU
2	Convolution	-	1×3	64	2×300	LeakyReLU
3	Convolution	-	1×3	128	2×150	LeakyReLU
4	Convolution	-	1×3	256	2×75	LeakyReLU
5	Convolution	-	1×3	512	2×38	LeakyReLU
6	Convolution	-	1×3	32	1×38	LeakyReLU
7	Flatten	-	-	-	1216	-
8-1	Dense	32	-	-	32	LeakyReLU
9-1	Dense	32	-	-	32	LeakyReLU
10-1	Dense	1	-	-	1	Sigmoid
8-2	Dense	32	-	-	32	LeakyReLU
9-2	Dense	32	-	-	32	LeakyReLU
10-2	Dense	1	-	-	1	Sigmoid

4. Experiment

In this study, we measured the sway of the CoP of the elderly for the GAN learning. Next, pseudo SKGs were generated by the numerical simulation. Finally, SKGs was colored using a GAN discriminator to visualize the characteristics of the sickness.

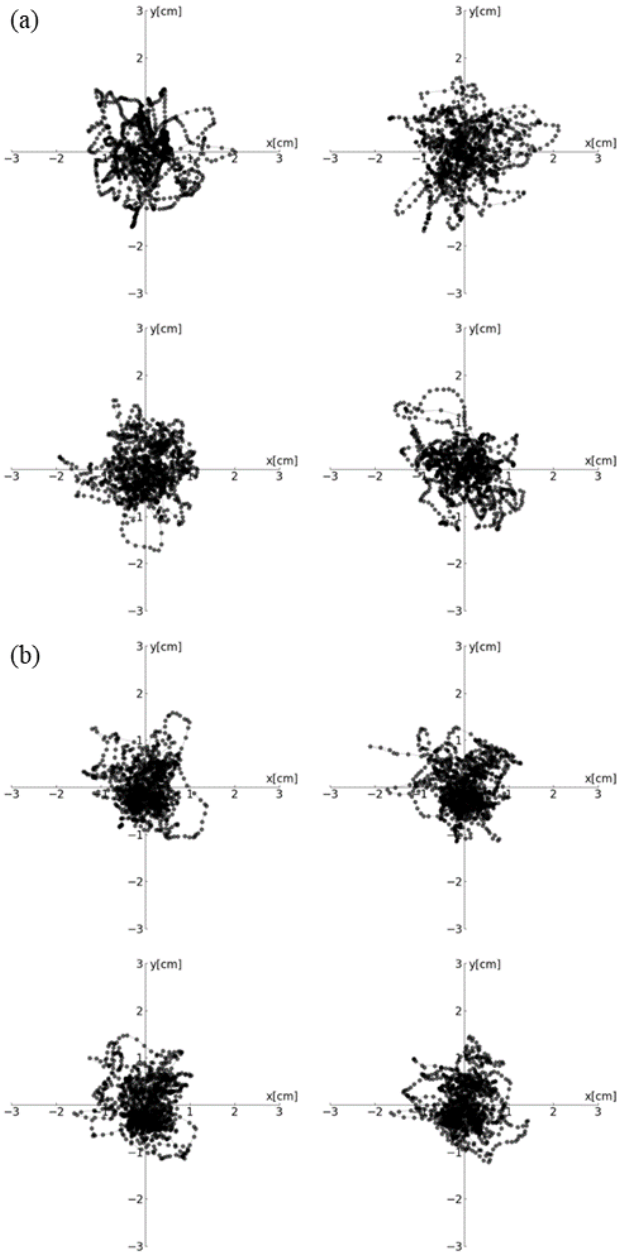


Figure 3: Typical SKGs: (a) measured data, (b) AI simulations

4.1. Measurement

In order to investigate the effects of stereoscopic video clips on the elderly, the body sway was simultaneously measured with the radial motion while/after viewing 2D/3D video clips for 60 s. In this experiment, 204 SKGs were observed from the elderly subjects that stood with Romberg posture on a gravicorder GS3000 (Anima Corp. Ltd., Tokyo) [30]. The SKGs were recorded at 20 Hz sampling in this experiment, which was approved by the Ethics Committee of the Graduate School of Information Science, Nagoya University.

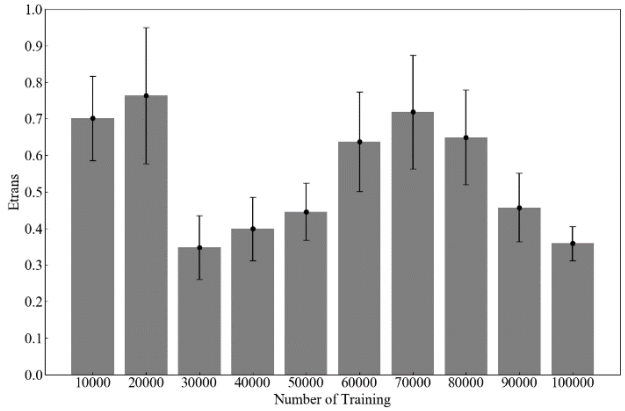


Figure 4: Translation errors for each learning step in the simulation of SKGs.

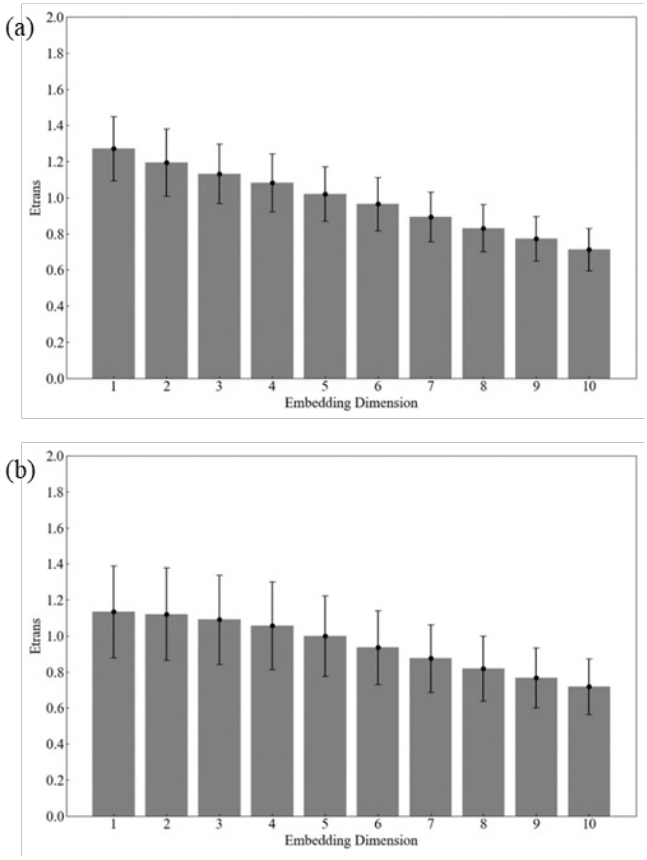


Figure 5: Translation error in SKGs: measured data (a), simulation pattern (b).

The results have shown that, in the elderly, the equilibrium function is affected while/after tracking the visual target in 3D video clips without statistical significance (sickeness-induced) [30]. Based on the previous studies on visually induced motion sickness (VIMS) [18-20], [30], the motion sickness is expected to be induced after/while viewing 3D video clips rather than 2D ones.

4.2. Simulation

In the constructed GAN model, it was confirmed that the machine-learning was not stable due to the small amount of raw data. Therefore, independent two-dimensional noise was generated

from the Winner process, and one million time series were obtained for each component. Using the weights of the GAN model obtained from this learning process, the model was reconstructed so that GAN simulation that imitated the SKGs of 204 cases in the measurements could be generated (Tables 1–2). At this time, the information is also input to "D" to determine whether the input data is a sickness-induced SKGs or a sickness-free SKGs in addition to the condition whether the input data is raw data or fake data and determines. A total of 100,000 steps were learned. Thus, 2D/3D video clips are herein regarded as visual stimulus that do not/do induce the VIMS, respectively. SKGs generated from the GAN every 10,000 steps was evaluated by calculating the translation error using the Wayland algorithm [31–32].

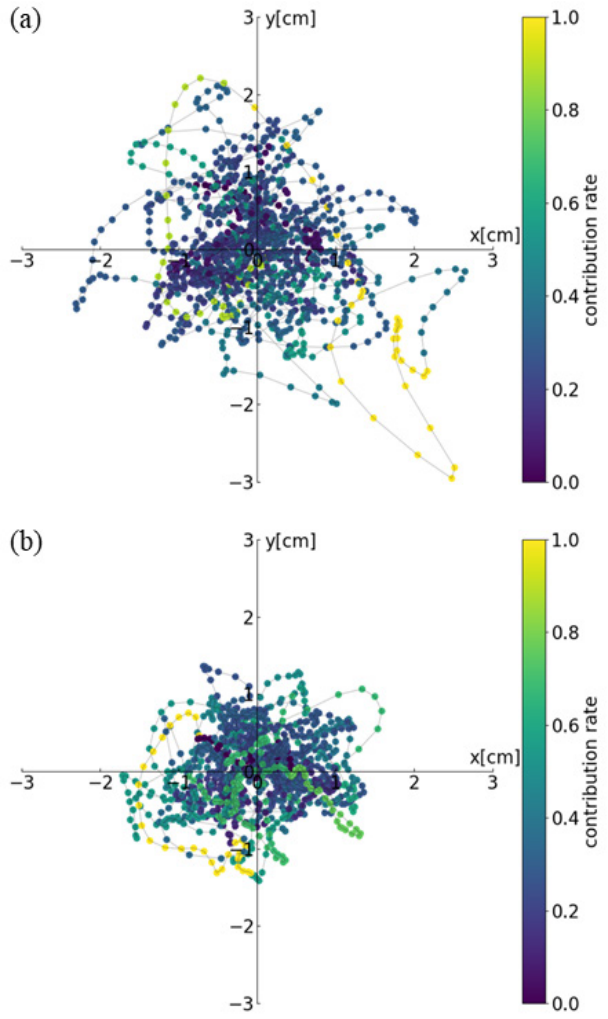


Figure 6: Patterns in SKGs that visualizes the characteristics of the following observation: sickness-induced (a), sickness-free (b)

4.3. Visualization of 3D Sickness Criteria

Using the trained GAN, the data obtained from the subject with the VIMS and the those without sickness were firstly input into the "D", respectively. Secondly, outputs of the intermediate layer close to the output layer were recorded to calculate the contribution rate for the discrimination of the VIMS. Lastly, the input SKGs were a colored with values of the contribution rate.

5. Results and Discussion

For the simulation SKGs (Figure 3b) generated by the GAN, the translation error was calculated every 10,000 steps using the Wayland algorithm (Figure 4). As a result, the value was approximately 0.7 after the machine learning up to 20,000 steps, and it decreased sharply to approximately 0.3 after 30,000 steps. After that, values of translation error increased to 70,000 steps again, and then decreased in learning after 80,000 steps. The GAN simulation might be influenced by the Winner process used for the pre-learning. That is why values of the translation error was high up to 20,000 steps. After that, the influence of the Winner process would be weakened around 30,000 steps. It was considered that the influence of the SKGs observed up to 70,000 steps gradually increased. After that, mode collapse might occur during the GAN-learning process. That is why values of the translation error decreased again after 80,000 steps. Mode collapse is a problem in which the generated data become similar because the variation of the data generated by "G" becomes smaller. This problem may occur in learning after 80,000 steps.

Next, the translation error of the observed SKGs (Figure 3a) are compared with that of the GAN simulation (Figure 3b) after 70,000 steps in the GAN-learning. Both values were about 0.8 in 10-dimensional embedding space, indicating very similar determinism (Figures 5).

Finally, the SKGs was colored using a GAN discriminator to visualize the characteristics of the sickness (Figures 6). It could be seen that the sickness-induced SKGs were widespread with respect to the sickness-free SKGs. It has been also discovered that GAN captures the shape of the cusp as a factor to determine the sickness. The contribution rate to judge the sickness-induced SKGs has been enhanced by the trajectories of the cusp (Figure 6a). We succeeded in extracting partial figure patterns that the AI regards as a property of 3D sickness. Consistency can be seen in the figure patterns that have been also extracted by our previous mathematical method [33]. Compering with the properties resulting from the mathematical analysis, interpretations can be given for the black box processing in the AI.

6. Conclusion

In this study, we proposed a method for the numerical analysis of biological data using AI. So far, there have been many studies that have attempted numerical simulations using SDE, but there is no research that performs simulations using GAN and visualizes/extracts the properties of the time series as shapes. As a result, the possibility of SKGs simulation using GAN was suggested. In the next steps, we would like to accumulate fundamental knowledge so that simulations using GAN can be applied in all fields.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported in part by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) Number 20K12528.

References

[1] I. Matsuba, Statistics of long-term memory processes: Self-similar time series theory and method, Tokyo: Kyoritsu Shuppan, 2007. (In Japanese)

[2] H.E. Hurst, "Long-term storage capacity of reservoirs," Transactions of the American Society of Civil Engineers, **116**(1), 770–799, 1951. Doi: 10.1061/TACEAT.0006518

[3] K. Aihara, T. Ikeguchi, T. Yamada, M. Komuro, Basics and applications of chaos time series analysis, Tokyo: Sangyo Tosho, 2000. (In Japanese)

[4] K. Nakane, H. Takada, T. Hirata, "On Generation of Pseudo Exchange Rate Using GAN," IEICE Technical report, NLP2019-13 (2019-05), 71–76, 2019. (In Japanese)

[5] L .Wang, J. Wittenstein, "Dow Jones Industrial Average average, worst decline in history - drops nearly $1600," Bloomberg, 2018. https://www. bloomberg.co.jp/news/articles/2018-02-06/P3PHW06TTDS601 (2020.3. 26)

[6] S. Ponczek, E. Popina, L. Wang, "Dow Average, who dropped-machine criminal theory emerging," Bloomberg, 2018. Doi: 10.1061/ P3PJGE6JIJVN01

[7] N. Baker, S. Mamudi, "Do 'Circuit Breakers' Calm Markets or Panic Them?" Bloomberg, 2020-3-9. https://www.bloomberg.com/news/articles/2020-03-09/what-are-circuit-breakers-and-do-they-calm-markets-quicktake (2020.3. 26)

[8] C. Mullen, A. Haigh, "Only One Market Circuit-Breaker May Be Left: Controlling Virus," Bloomberg, 2020-3-12. https://www. bloomberg.com /news/articles/2020-03-13/only-one-market-circuit-breaker-may-be-left-controling-virus (2020.3.26)

[9] C. Ballentine, "Stocks at Risk of Tripping Circuit Breakers as Rout Resumes," Bloomberg, 2020-3-16. https://www.bloomberg.com/news/articles/2020-03-16/stocks-at-risk-of-setting-off-circuit-breakers-as-rout-resumes (2020.3. 26)

[10] V. Hajric, "Circuit Breakers for Stocks Triggered a Second Time in 3 Days," Bloomberg, 2020-3-18. https://www.bloomberg.com/news/articles/2020-03-18/circuit-breakers-for-stocks-triggered-a-second-time-in-3-days (2020. 3.26)

[11] F. Black, M. Scholes, "Pricing of options and corporate liabilities, Journal of Political Economy, **81**(3), 637–654, 1973. Doi: 10.1142/9789814759588_0001

[12] T. Tasaki, "Motion equation of for stock prices," Bussei Kenkyu, **81**(4), 518–519, 2004. (In Japanese)

[13] S. Porter, J. Nantel, "Older adults prioritize postural stability in the anterior-posterior direction to regain balance following volitional lateral step," Gait & Posture, **41**(2), 666-669, 2015. Doi: 10.10161/14759588_01156

[14] P.A. Goldie, T.M. Bach, O.M. Evans, "Force platform measures for evaluating postural control: reliability and validity," Arch Phys. Med. Rehabil, **70**, 510–517, 1989.

[15] R.E.A. Emmerrik, R.L.V. Sprague, K.M. Newell, "Assessment of sway dynamics in tardive dyskinesia and developmental disability: sway profile orientation and stereotypy," Mov. Dis., **8**, 305-314, 1993. Doi: 10.1136/naid/10026086168/

[16] J.J. Collins, C. J. De Luca, "Open loop and closed-loop control of posture: a random-walk analysis of center of pressure trajectories," Exp. Brain Res., **95**, 308–318, 1993. Doi: 10.1007%2FBF00229788

[17] K.M. Newell, S.M. Slobounov, E.S. Slobounova, P.C. Molenaar. "Stochastic processes in postural center of pressure profiles," Exp. Brain Res., **113**, 158-164, 1997.

[18] H. Takada, M. Miyao, "Visual fatigue and motion sickness induced by 3D video," Forma **27**(Special Issue), 67–76, 2012.

[19] Y. Mori, Y. Maeda, H. Takada, "Numerical Analysis of Body Sway While Viewing a 3D Video Clip Without Perspective Clues," LNCS, **9176**, 238-245, 2015.

[20] F. Kinoshita, Y. Mori, M. Miyao, H. Takada, "On Mathematical Models of Two-Minute Stereoscopic Viewing on Human Balance Function," Forma, **32**, 11–17, 2017.

[21] K. Wakabayashi, T. Namatame, "Examination of applicability of deep learning to consumers' purchasing behavior," Journal of the Japan Social Data Science Society, **1**(1), 48–57, 2017. (In Japanese)

[22] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv:1506.02640v5 [cs.CV], 2016. https://arxiv.org/abs/1506.02640 (2021.3.15)

[23] H. Wu, J. Zhang, K. Huang, K. Liang, Y. Yu, "Fast FCN: Rethinking Dilated Convolution in the Backbone for Semantic Segmentation," arXiv:1903.11816v1 [cs.CV], 2019. https://arxiv.org/abs/1903.11816 (2021.3.15)

[24] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, "Generative adversarial Nets," 1–9, 2014.

[25] A. Radford, L. Metz, "Unsupervised representation learning with deep convolutional generative adversarial networks," Proceedings of ICLR, 1–16, 2016.

[26] A.L. Maas, A.Y. Hannun, Y. A.Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," Proceedings of ICML, **30**(1), 3, 2013. Doi: 10.1061/ICML.0006518

[27] J. Snoek, H. Larochelle, R.P. Adams, "Practical Bayesian optimization of machine learning algorithms," Proceedings of NIPS, 1–9, 2012.

[28] K. Nakane, R. Ono, S. Yamamoto, M. Takada, F. Kinoshita, A. Sugiura, Y. Matsuura, K. Fujikake, H. Takada, "Numerical Analysis of Body Sway for Evaluation of 3D Sickness," The 15th International Conference on Computer Science & Education (ICCSE 2020), 2020, doi: 10.1109/ICCSE49874.2020. 9201640.

[29] K. Nakane, H. Takada, S. Yamamoto, R. Ono, M. Takada, "Numerical analysis of bio-signal using generative adversarial networks," Lecture Note in Computer Science,12427, 601-613, 2020, doi: 10.1007/978-3-030-60152-2_44.

[30] N. Amano, H. Takada, Y. Johno, T. Tanimura, F. Kinoshita, M. Miyao, M. Takada, "Analysis of the body sway while/after viewing visual target movement synchronized with background motion," LNCS, **10908**, 3-14, 2018. Doi: 10.1007/978-3-319-92052-8_1

[31] R. Wayland, D. Bromley, D. Pickett, A. Passamante, "Recognizing determinism in a time series," Physical Review Letters, **70**, 580–582, 1993.

[32] H. Takada, T. Morimoto, H. Tsunashima, T. Yamazaki, H. Hoshina, M. Miyao, "Applications of Double-Wayland algorithm to detect anomalous signals," Forma, **21**(2), 159–167, 2006.

[33] H. Takada, Y. Kitaoka, M. Ichikawa, M. Miyao, "Physical Meaning on Geometrical Index for Stabilometry," **62**(3), 168-180, 2003. Doi: 10.3757/jser.62.168

Appendix GAN model

The generator could be constructed by the following components. Optimum combination was found as Table 2.

- Kernel sizes: 1–10
- The number of convolutional layers: 1-4
- That of filters in the convolutional layers :16, 32, 64, 128

Table 2: Network configuration in the generator G

Layers	Layer Name	Units	Kernel Size	Filters	Output Shape	Activation
0-1	Noise-Input	-	-	-	-	-
1-1	Dense	100	-	-	100	-
0-2	State-Input	-	-	-	-	-
1-2	Dense	100	-	-	100	-
2	Concatenate	-	-	-	200	-
3	BatchNormalization	-	-	-	200	-
4	Dense	100	-	-	100	LeakyReLU
5	BatchNormalization	-	-	-	100	-
6-1	Dense	150	-	-	150	LeakyReLU
7-1	BatchNormalization	-	-	-	150	-
8-1	Convolution	-	1×3	128	1×300	LeakyReLU
9-1	BatchNormalization	-	-	-	1×300	-
10-1	Convolution	-	1×3	128	1×300	LeakyReLU
11-1	BatchNormalization	-	-	-	1×300	-
12-1	Convolution	-	1×3	64	1×600	LeakyReLU
13-1	BatchNormalization	-	-	-	1×600	-
14-1	Convolution	-	1×3	64	1×600	LeakyReLU
15-1	BatchNormalization	-	-	-	1×600	-
16-1	Convolution	-	1×3	32	1×1200	LeakyReLU
17-1	BatchNormalization	-	-	-	1×1200	-
18-1	Convolution	-	1×3	32	1×1200	LeakyReLU
19-1	BatchNormalization	-	-	-	1×1200	-
20-1	Convolution	-	1×1	1	1×1200	Tanh
6-2	Dense	150	-	-	1×150	LeakyReLU
7-2	BatchNormalization	-	-	-	1×150	-
8-2	Convolution	-	1×3	128	1×300	LeakyReLU
9-2	BatchNormalization	-	-	-	1×300	-
10-2	Convolution	-	1×3	128	1×300	LeakyReLU
11-2	BatchNormalization	-	-	-	1×300	-
12-2	Convolution	-	1×3	64	1×600	LeakyReLU
13-2	BatchNormalization	-	-	-	1×600	-
14-2	Convolution	-	1×3	64	1×600	LeakyReLU
15-2	BatchNormalization	-	-	-	1×600	-
16-2	Convolution	-	1×3	32	1×1200	LeakyReLU
17-2	BatchNormalization	-	-	-	1×1200	-
18-2	Convolution	-	1×3	32	1×1200	LeakyReLU
19-2	BatchNormalization	-	-	-	1×1200	-
20-2	Convolution	-	1×1	1	1×1200	Tanh
21	Concatenate	-	-	-	2×1200	-

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 956-964 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Performance Evaluation of Convolutional Neural Networks (CNNs) And VGG on Real Time Face Recognition System

Showkat Ahmad Dar*, S Palanivel

Department of Computer Science & Engineering, Annamalai University, Annamalainagar, Chidambaram, 608002, India

ARTICLE INFO

Article history:
Received: 18 January, 2021
Accepted: 27 March, 2021
Online: 15 April, 2021

Keywords:
Deep learning
Convolutional Neural Network (CNNs)
VGG16
Face authentication
Real time face images
Classifiers

ABSTRACT

*Face Recognition (FR) is considered as a heavily studied topic in computer vision field. The capability to automatically identify and authenticate human's faces using real-time images is an important aspect in surveillance, security, and other related domains. There are separate applications that help in identifying individuals at specific locations which help in detecting intruders. The real-time recognition is mandatory for surveillance purposes. A number of machine learning methods along with classifiers are used for the recognition of faces. Existing Machine Learning (ML) methods are failed to achieve optimal performance due to their inability to accurately extract the features from the face image, and enhancing system's recognition accuracy system becomes very difficult task. Majority of designed FR systems has two major steps: extraction of feature and classifier. Accurate FR system is still a challenge, primarily due to the higher computational time and separate feature extraction. In general, for various applications using images, deep learning algorithms are mostly recommended for solving these problems because it performs combined feature extraction and classification task. Deep learning algorithm reduces the computation time and enhances the recognition accuracy because of automatic extraction of feature. The major novelty of the work is to design a new VGG-16 with Transfer Learning algorithm for face recognition by varying active layers with three levels (3, 4, and 7). It also designs the Convolutional Neural Network (CNN) for FR system. The proposed work introduced a new Real Time Face Recognition (RTFR) system. The process is broken into three major steps: (1) database collection, (2) FR to identify particular persons and (3) Performance evaluation. For the first step, the system collects 1056 faces in real time for 24 persons using a camera with resolution of 112*92. Second step, efficient RTFR algorithm is then used to recognize faces with a known database. Here two different deep learning algorithms such as CNN and VGG-16 with Transfer Learning are introduced for RTFR system. This proposed system is implemented using Keras. Thirdly the performance of these two classifiers is measured using of precision, recall, F1-score, accuracy and k-fold cross validation. From results it concludes that proposed algorithm produces higher accuracy results of 99.37%, whereas the other existing classifiers such as VGG3, VGG7, and CNN gives the accuracy results of 75.71%, 96.53%, and 69.09% values respectively.*

1. Introduction

In recent years, an active research area is Face Recognition (FR) technology because of technology's potential in commercial use and law enforcement as well as rise in security demands [1]. The issues and developments in face recognition have been alluring a lot of scientists working in computer vision, pattern recognition, and biometrics domain. Various face recognition algorithms are used in diverse applications like indexing and video compression that comes under the domain of biometrics. Face recognition concepts can be used in classifying multimedia content and to help in the quick searching of materials that interests the end user. A comprehensive face recognition mechanism can be of assistance in domains like surveillance and forensic sciences. It can also be used in the areas of law enforcement and to authenticate security and banking systems. In addition to that, it also gives control and preferential access to

*Corresponding Author: Showkat Ahmad Dar, showkatme2009@gmail.com

secured areas and authorized users. The issues in face recognition have garnered even more significance after the spike in terrorism in the recent years. It largely decreases the need for passwords and can offer enhanced security. For this, FR should be used with additional security mechanisms.

In spite of facial recognition's rapid growth and acclaim as a critical authentication mechanism, the algorithms used for facial recognition have not been developed significantly. It has been close to two decades since facial recognition has come to the fore but a comprehensive system that is capable of producing positive results in real-time conditions has not been developed yet. The Face Recognition Vendor Test (FRVT) test carried out by National Institute of Standards and Technology (NIST) is demonstrated that modern face recognition mechanisms will not be able to perform optimally under certain circumstances.

Modern FR systems intended for complex environments has attracted a huge attention in recent years. FR systems that are automated is a developing technology that has garnered a lot of interest. There are a number of conventional algorithms that are used in developing color images and still-face images. The data complexity is increased in color images as the pixels are mapped to a high-dimensional space. This significantly decreases the accuracy and processing efficiency of FR [2].

There are four stages in FR technology namely, classification, representation (extraction of facial feature), alignment and detection [3]. Feature representation technique used for extracting features is the major challenge of FR system. For a specified biometric trait, representations are done using better techniques. In image classification, highly important step is feature extraction. Highly important information is retained in feature extraction, which is used in classification. Feature extraction methods improved FR accuracy slowly.

FR systems are reported with failures or unstable performance often with different false rates in real-world environment due to technical insufficiency. Making it formal and complete use of its performance is being yet a final solution. In the recent years, it has been inferred that deep learning works a lot better for face recognition [4]. For classification and feature extraction, processing unit's multiple layer cascading is used in deep learning techniques. Facial image's very high recognition rate is achieved by this.

The proposed study processes color images to recognize and detect faces with a good deal of accuracy in a real-time scenario. This work aims to apply pre-trained Convolutional Neural Network (CNN) with VGG-16 algorithm for FR and classification accuracy using analysis. These methods have been used for enhancing recognition accuracy. Most relevant challenge for such a system is recognition and feature extraction. A system has been proposed here that makes use of deep learning techniques to extract features. System uses deep learning to recognize faces in an accurate manner. The proposed system will be capable of recognizing more number of faces which can be used for searching suspects as the errors are reduced significantly.

2. Literature Review

FR has become a popular topic of research recently due to increases in demand for security. Highly attractive biometric technology is FR and its accuracy is highly enhanced using recent advancements in technology. According to Deep Learning (DL) and Machine Learning (ML) techniques, FR techniques are performed.

In [5] proposed Haar classifier that made use of a surveillance camera for face recognition. The system had four sequential steps that included (1) real-time image training (2) face recognition with the help of Haar-classifier (3) comparing the real-time images having images that were captured from camera (4) generation of the result as per the comparison. Haar is used to detect the faces in a robust manner in real-time scenarios. Face detection uses an algorithm called as Haar cascading. A system called as Aadhar has been adopted by India for recognizing the citizens. If this is used as a database for the citizens, a local citizen and a foreigner can be easily recognized. This information can be eventually used for identifying if the person is a criminal or not. The major advantage of this work is to apply this system to citizenship database. This has less computational cost, better discriminatory power and high recognition precision. Least processing is required by main features in small images for training Haar wavelets. If number of features becomes more it requires increased computation time for FR system. It is left as scope of the work.

In [6] proposed a Local Binary Pattern (LBP) for identifying faces. This was used along with other image processing methods like Histogram Equalization (HE), Bilateral Filter, Contrast Adjustment and Image Blending to improve overall system accuracy. The LBP codes are improved here due to which the performance of the system is enhanced. The major advantage of this work is that it is reliable, robust and accurate. In real-life setting, this may be used as an attendance management system. But this solution limits only the noise. In facial recognition, mask faces and occlusion issues are not addressed by this system. But, in future, this work can be extended for addressing these issues.

For face feature extraction, combination of Local Binary Pattern (LBP) and Histograms of Oriented Gradients (HOG) descriptors [7]. Low computation power is required by these descriptors. For face classification, Random Forest (RF) classifier based accurate and latest methods are applied. Under a controlled environment, from video broadcast, identification and verification of one or more person can be done accurately using this proposed algorithm. For FR system. Still there is a need to have separate feature extraction technique.

For facial expression recognition, a real-time system is presented [8]. Student's 8 basic facial expressions can be recognized using this proposed system and expressions includes natural, surprise, sad, nervous, happy, feat, disgust and anger inside E-learning environment. Proposed system's efficiency is tested by using Support Vector Machine (SVM), k-Nearest Neighbor (k-NN) classifiers and their results are compared. Techniques used in this study for recognizing facial expressions includes SVM and k-NN classifiers for expression recognition, feature selection based on Principal Component Analysis (PCA), feature extraction based on Gabor Feature approach, Viola-Jones technique based face detection. In a real-time system, for facial expression recognition, it can be used. However the k-NN and

SVM classifier needs more time complete the FR process due to feature extraction and feature selection steps.

For addressing partial face images irrespective of its size, Fully Convolutional Network (FCN) with Sparse Representation Classification (SRC) [9] and it is termed as Dynamic Feature Matching (DFM). For optimizing FCN, introduced a sliding loss based on DFM. Between subject's face image and face patch, an intra-variation is reduced for enhancing DFM performance. For other visual recognition tasks like partial person re-identification, this technique can be extended very easily. This solution limits the noise alone and image count is also restricted.

For suppressing unreliable local features from occluded regions, a fuzzy max-pooling scheme [10]. On every subclass, automatic weighting is done for enhancing robustness in final average-pooling. While dealing with data sufficiency and occlusion problem simultaneously, better performance enhancement is shown by this technique, which is a major advantage of this technique. Every subclass is treated as an individual classifier, where ensemble late fusion framework is used for obtaining final decision. Remarkable enhancement in performance can be achieved as shown in results.

Under various conditions, for face recognition, a framework called Optimized Symmetric Partial Face graph (OSPE) [11]. For instance, light variation, facial expression; occluded faces are used in their experimentation. Partial facial data are introduced for enhancing recognition rate as shown in their experimental results. Local spatial information is not explored fully in these techniques, which is a major drawback of it.

For dealing with FR, a Principal Component Analysis (PCA) technique based on patch [12]. Total scatter is computed for computing divided patches correlation directly. For feature extraction, projection matrix is obtained by optimizing projected samples total scatter. Nearest Neighbor (NN) classifier is used at last. For this large sized covariance matrix, eigenvectors evaluation consumes more time.

For real time face recognition, a LBP [13]. The image of the face is represented by using information about the texture and shape. For representing the face comprehensively, the facial area is divided into different sections. LBP histograms are then extorted which are combined to a single histogram. Facial recognition is then using the Nearest Neighbor (NN) classifier. The validation of the algorithm is carried out by devising a prototype model that makes use of the raspberry Pi single-board computer, Matrix Laboratory (MATLAB). The results indicate that LBP algorithm's recognition rate is relatively higher when compared to other approaches.

In [14], four various algorithms are combined with Discrete Wavelet Transform (DWT). Algorithm includes Convolutional Neural Network (CNN), Linear Discriminant Analysis (LDA) Eigen vector, PCA Eigen vector and Principal Component Analysis (PCA) error vector. Then Fuzzy system and detection probability's entropy are used for combining these four results. Database diversity and image defines the recognition accuracy as indicated in results. For best case, 93.34% recognition rate and for worst case, 89.56% recognition rate are provided by this technique. This technique is better than other techniques, where individual technique is implemented on specific images set. In human face recognition, illumination impact is ignored, which is a only limitation in this work.

All these techniques, reviewed don't completely address issues affecting facial recognition accuracy like feature extraction and noise variation. Unfortunately, the processing time and training period for these algorithms are considerably large. In the recent times, Face Recognition (FR) techniques have been replaced by deep learning. Deep learning has been observed to perform better for large datasets. On the contrary, traditional ML algorithms at an optimum level with comparatively smaller datasets. In conventional ML algorithms, a difficulty needs to be broken down into individual steps. FR based on CNN is trained with large datasets and has attained a high level of accuracy. The increased use of deep learning has accelerated the research involved in FR. Recently, with deep learning emergence, impressive results are achieved in face recognition. In computer vision applications, most popular deep neural network is CNN and it has automatic visual feature extraction which is a major advantage [15].

For recognizing face images, Support Vector Machine (SVM) and Convolutional Neural Network (CNN) [16]. For automatically acquiring remarkable features, CNN is utilized as feature extractor. Using ancillary data, CNN pre-training is proposed at first and updated weights are computed. Target dataset is used for training CNN, which extracts highly hidden facial features. At last, for recognizing all classes, proposed a SVM classifier. With high accuracy, face images are recognized using SVM where facial features extracted from CNN are given as an input. In experimentation, for pre-training, images in Casia-Webfaces database are used and to test and train, used the Facial Recognition Technology (FERET) database. With less training time and high recognition rate, efficiency is demonstrated in experimentation results. Moreover, it is highly difficult to acquire some facial features manually from face images. But, effective facial features are automatically extracted using CNN. With more optimization techniques, deeper CNN based training time and recognition rate's balance point are computed and larger dataset is left as scope of this work.

A modified CNN architecture [17], where two normalization operations are added to two layers. Batch normalization is used as a normalization operation and network is accelerated using this. Distinctive face features are extracted using CNN architecture and in CNN's fully connected layer, faces are classified using Softmax classifier. With better recognition results, face recognition performance is enhanced using this proposed technique as shown in experiment part and it uses Georgia Tech Database.

A novel technique based on CNN [18] which is termed as Deep Coupled ResNet (DCR) model. It consists of two branch networks and trunk network. For face images with various resolutions, discriminative features are extracted using trunk network. High-Resolution (HR) images are transformed using two branch networks and targeted Low Resolution (LR) images are also transformed using this.

Pre-trained CNN's performance [19] with multi-class Support Vector Machine (SVM) classifier and for performing classification, AlexNet model based transfer learning

performance is also proposed. An excellent performance is achieved using hybrid deep learning as shown in results when compared with other techniques of face verification.

For facial recognition, CNN and Multi-Layer Perceptron (MLP) [20]. For performing facial recognition, it is a system based on open code deep-learning. For fiducial point embedding and extraction, deep learning techniques are used. For classification task, SVM is used. In inference and training, it has fast performance. For facial features detection, 0.12103 error rate is achieved using this system, which is pretty close for state of the art algorithms and for face recognition, it has 0.05. In real-time, it can run.

In [21] proposed a CNN for facial recognition that has been implemented on the embedded GPU system. The method uses facial recognition based on CNN along with face tracking and deep CNN facial recognition algorithm. The results have indicated that the system is capable of recognizing different faces. An estimated 8 faces can be recognized simultaneously within a span of 0.23 seconds. The recognition rate has been above 83.67%. As a result, processing time is enhanced and for real-time multiple face recognition, it can be used with acceptable recognition rate.

For face recognition, a Pyramid-Based Scale-Invariant - CNN model (PSI-CNN) [22]. From image, untrained features are extracted using PSI-CNN model and with original feature maps, these features are fused. With respect to matching accuracy, experimentation results are shown and it outperforms model derived from VGG-Face model. Stable performance is maintained using PSI-CNN during the experimentation with low-resolution images which are acquired from CCTV cameras. Robust performance is exhibited with image quality and resolution change.

For regressing facial landmark coordinates, at CNN's intermediate layers [23]. In specific poses and appearances, for regressing facial landmark coordinates, specialized architecture is designed in novel CNN architecture. For every specialized sub-networks, for providing sufficient training examples, designed a data augmentation techniques and it address training data's shortage specifically in extreme profile poses. On true positive detected faces, accuracy is reflected by this. At last, trained and code models are made publicly available using project webpage, for promoting results reproducibility.

Using CNN, real-time face recognition system's evaluation [24]. Standard AT&T datasets is used for performing proposed design's initial evaluation and for designing a real-time system, same can be extended. For enhancing and assessing proposed system's recognition accuracy, CNN parameters tuning are used. For enhancing system performance, systematic technique for tuning parameters is proposed. From the review it concludes that the deep learning algorithms give lesser computation time and more recognition accuracy than the other methods.

The VGG16 based multi-level information fusion model [25]. On fully connected neural network, enhancement is proposed. Computation time is reduced using this technique. In calculation and propagation process, some useful feature information are lost in CNN model. This work enhances model to be a convolution calculation's multi-level information fusion and discarded feature information are also recovered and it enhances image's recognition rate. Network is divided into five groups using VGG and they are combined as a convolution sequence. The main advantage here is the representational efficiency. Face recognition can be attained using 128 bytes per face.

3. Proposed Methodology

The proposed system is broken into three major steps: (1) database collection, (2) face recognition to identify particular persons, and (3) Performance evaluation. For the first step, the system collects the faces in real time. In the database consists of 24 different persons are in the form of 1056 images having the resolution of 112*92; Olivetti Dataset is also used for implementation. In second step, Convolutional Neural Network (CNN) and VGG-16 Deep Convolutional Neural Network (DCNN) are introduced for improving recognition accuracy. Finally results evaluation of these two classifiers is measured using precision, F1-score, recall and accuracy. These classifiers recognize faces in real-time with high accuracy. The recognition building blocks are shown in the Figure 1.

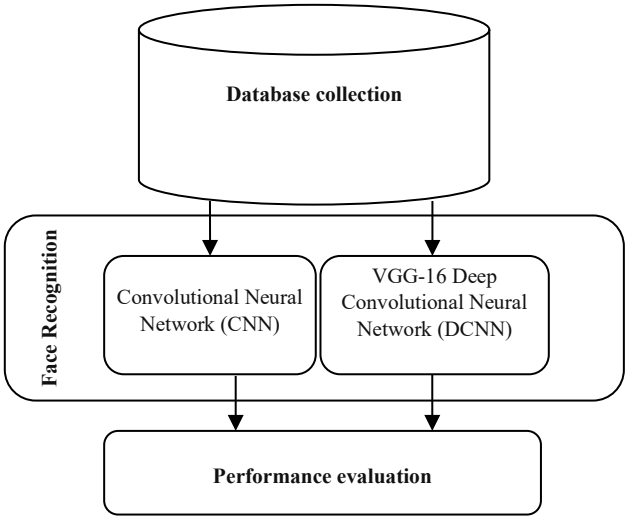


Figure 1: Face Recognition Building Blocks

3.1. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is attracted image processing application's attraction [26, 27, 28] and it is also used for feature extraction capacity from real-time facial images. General CNN model which is used for real time facial recognition has been illustrated in Figure 2. The CNN architecture has three major layers: Convolutional, Fully Connected, and Pooling Layer. The initial layer, facial image samples are given as input. Then in convolution layer, real time face images are converted into facial feature vectors via the kernel with filters=5. It is followed by ReLU activation function for recognition as well as maximum pooling layer. Next in line comes the full connection layer followed by the output layer.

CNN's basic building blocks are explained as follows:

Input layer- This layer has image's raw input having width 112, height 92, depth 3.

Convolution Layer – In CNN**,** a matrix called as the kernel is passed over the input real time face matrix with size of (112*92)

to devise a feature map for subsequent layer. Mathematical operation termed as the convolution is executed via sliding Kernel size (3*3) over input real time face matrix. On each of the locations, real time face matrix multiplication is carried out and adds result set onto final feature map. For example, let use consider a 2-Dimensional kernel filter (K=5, 3*3), and a 2-Dimensional real time facial input image, I. This layer computes output volume via computing dot product between all filters and facial image. Output volume is computed using this layer via computing dot product between facial image and all filters. In this layer, filters are used to produce output volume with 112x 92 x 5 dimension.

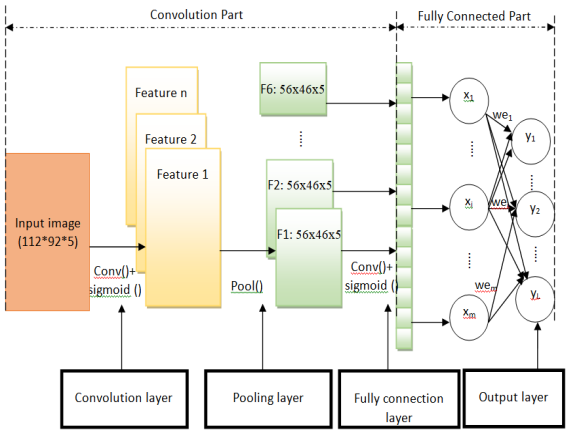


Figure 2: Convolutional Neural Networks Architecture

In layer l, assume In^las neurons input and as neurons output Op^l. Every neuron's activation function and input are used for computing every neurons output. Layer number is represented as l, for instance, for first layer l=1 and for last layer, l=-1. Row number is expressed as i and column number is represented as j. Three-dimensional matrix is produced by CNN' every In^l and Op^l layers data structures, while one dimensional vector is produced by every layer In^l and Op^l in FC.

Assume layer l's weight as we_{ij}^l and bias as ba_j^l. FC last layer's weight is represented as we_{ij}^{-1} and bias is represented as ba_j^{-1}. If layer l is pooling or convolutional layer, pooling windows or convolutional kernel's size is represented as $size^l \times size^l$. If layer l is fully-connected layer, neurons count is given by$size^l$. In layer l, neurons every input value is represented as In_{mn}^l . For convolution computation, convolution(Op^{l-1}, we^l, m, n) function is used. Previous layer's output is given by Op^{l-1}. Between layer l (In^l))'s input and previous layer (Op^{l-1})'s output, weights matrix is given by we^l. Layer l's bias is given by ba^l. The convolutional layer (In^l)'s input is computed as,

$$In_{mn}^l = Convolution(Op^{l-1}, we^l, m, n) + ba^l = \sum_{i=0}^{size^l-1} \sum_{j=0}^{size^l-1} \left(Op_{m+i,n+j}^{l-1} . we_{i,j}^l + ba^l\right) \quad (1)$$

The convolutional layer l (Op_{mn}^l)'s output is calculated as equation (2), where, sigmoid () is activation function

$$Op_{mn}^l = F(net_{mn}^l) = sigmoid(net_{mn}^l) = \frac{1}{1+e^{-In_{mn}^l}} \quad (2)$$

Non-linear activation functions (ReLU) –Once the kernel filter is formed then the next step is to perform FR system. A node that comes next to the convolutional layer is called as activation function. The Rectified Linear Unit (ReLU) can be considered as piecewise linear function which is given as output, if it is positive, or else the output will be given as zero. Total 5 filters for this layer will get output volume 112x 92 x 5 dimension. The expression of ReLu function is R(z)=max(0,z) the function and its derivative image are shown in Figure 3.

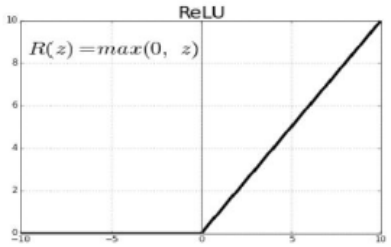


Figure 3: RELU Activation Function

Pooling Layer– In covnets, pooling layer is inserted periodically and its major objective are, minimization of volume size, making fast computation, reducing memory and prevention of overfitting. With stride 2 and 2*2 filters, this work uses a max pool. For face images recognition, resultant volume will have 56x46x5 dimension. Face matrix's average pooling is represented as pool(x) function. Expression (3) gives formula used for computing pool(x). Pooling window size is represented as $size^l$.

$$y_{ij} = pool(x, i, j) = \frac{\sum_{m=1}^{size^l} \sum_{n=1}^{size^l} x_{size^l \times (i-1)+m, size^l \times (j-1)+n}^{l-1}}{size^l \times size^l} \quad (3)$$

As indicated in expression (3), previous layer (Op^{l-1})'s output forms base for pooling layer Op^l's output. In other words, previous layer (Op^{l-1})'s output will be equal topooling layer l (net^l)'s input. As mentioned in above definition, last FC layer's input is represented as t^{-1} , first FC layer's input is represented as net^{-2}, layer before FC's input (pooling layer's last layer) is represented as net^{-3}.

Fully Connected Layer (FC)- Fully Connected (FC) Layer indicates that every node in initial layer is connected with next layer's every node. It computes class scores of facial images and outputs 1-D array of size equal to classes count via softmax function .The CNN model that was proposed was trained with 5 epochs, batch size=32 and using Adam optimizer (adaptive moment estimation). Specify optimizer's learning rate is set at 0.001. The forward propagation's result is$\hat{y}_n$, which is formulated in equation (4) to equation (6).

$$In_j^{-1} = \sum_{m=1}^{size^{-2}} \left(Op_i^{-2} . we_{ij}^{-1} + ba^{-1}\right), j = 1, \ldots size^{-1} \quad (4)$$

$$Op_j^{-1} = F\left(In_j^{-1}\right) = Sigmoid\left(In_j^{-1}\right) = \frac{1}{1+e^{-In_j^{-1}}} \quad (5)$$

$$\hat{y}_n = Op^{-1} \quad (6)$$

The CNN's pseudocodeis listed in Algorithm 1.

Algorithm 1: Training Algorithm of CNN

Input: tr_x, tr_y is considered as the features and labels of training facial images set

te_x, te_y is considered as the features and labels of testing facial images set.

Output: $we_{i,j}^{l}, ba_j^{l}$ weights and bias of convolution and pooling layers respectively

we_{jk}, ba_{jk} weights and bias of Fully Connected (FC) layer respectively

Required parameters: Max_{iter} Maximum number of iterations to complete face recognition task,

Error: when the training error is less than error , the training is finished , η: learning rate of the classifier

Initialization work:

t: t is the current iteration which is initialized as t=1 before training loop, L(t): L(t) is the Mean Square Error (MSE) at iteration t, $L(t)$ is initialized as $L(1) = 1 \geq error$

Begin

Set the required parameters and complete the initialization work

While $t < Max_{iter} \& L(t) > error$

For all tr_x

tr_r (Recognition label of training set tr_x and compute calculation from equation (1-6)

End for

Lt is recalculated as $L(t) = \frac{1}{2}\sum_{n=1}^{N}\left(tr_r(n) - tr_y(n)\right)^2$, N is the total images count in dataset

Increment t=t+1

End

3.2. VGG-16 Deep Convolutional Neural Network (DCNN) with Transfer Learning

VGG16 is considered to be a CNN model that enhances the AlexNet by making replacements of the large kernel-sized filters [29] having various 3×3 kernel-sized filters sequentially. VGG-16 proposed method has four major building blocks namely Softmax classifier, FC-layer, convolution module and attention module.

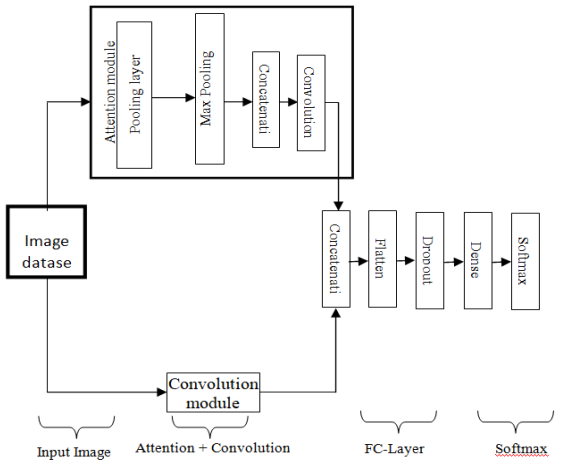


Figure 4: Block Diagram of the Proposed Attention-Based VGG-16 for Face Recognition

3.2.1. Attention Module

In facial image's feature's relationship are captured using this module. On input tensor, average pooling and max pooling are performed and in this technique, VGG-16 technique's 4th pooling layer is formed using this. For performing 7 × 7 filter size (f)'s convolution, max pooled 2D tensor is concatenated with every other layer via sigmoid function (σ). Figure 4 shows attention module's high level diagram. Expression (7) gives concatenated resultant tensor ($M_s(F)$).

$$M_s(F) = \sigma(f^{7\times7}[F^s{}_{max}]) \quad (7)$$

where $F_{avg}^{s} \in \mathbb{R}^{1\times H\times W}$ gives 2D tensors achieved using max pooling operation on input tensor F. Here tensor's height is represented as H and width is represented as W.

3.2.2. Convolutional Module

Convolution module is used which is VGG-16 model's 4th pooling layer. Facial image's features are captured using scale-invariant convolution module. From midlevel layer, extracted the features which are highly needed for real time facial images. For real time facial images, features from other layers like low or high are not appropriate as images are neither more specific nor more general. Thus, attention module is given with 4th pooling layer as first input. Then, that module's result is concatenated using 4th pooling layer.

3.3.3. Fully Connected (FC)-Layers

For representing concatenated features derived from convolution and attention block are converted as one-dimensional (1D) features using fully connected layers. It has three layers namely, dense, dropout and flatten as illustrated in Figure 4. Dropout is fixed as 0.5 and dense layer is limited to 24.

3.3.4. Softmax Classifier

From FC layer, features are extracted for classification and for facial image's final recognition, softmax layer is used. The unit number is defined by categories count in softmax layer, which is a last dense layer. According to classification, probability score's multinomial distribution is produced at softmax layer output. This distribution's output is given by,

$$P(a = c|b) = \frac{e^{b_k}}{\sum_j e^{b_j}} \quad (8)$$

where b probabilities that are retrieved from softmax layer and c represents facial image recognition dataset class used in proposed method. Table 1 gives proposed model's architecture details. Here, units in final dense layer (softmax layer) varies from one dataset to another based on categories count. ReLu activation function is applied in all layers except last one.

Table 1: VGG16 Proposed Model's Architecture

Layer (Type)	Output shape
VGG-16 model	90×110×5
conv2d_6 (Conv2D) layer	90×110×5
max_pooling2d_6(MaxPooling2) layer	45×55×5

conv2d_7 (Conv2D) layer	43×53×5
max_pooling2d_7 (MaxPooling2 (None, 21, 26, 5)	21×26×5
flatten_3 (Flatten)	2730
dense_3 (Dense)	24
Total params: 65,914 Trainable params: 65,914 Non-trainable params: 0	

4. Results and Discussion

This section evaluates facial recognitions method's performance. Real time database consists of 1056 images each of size 112*92.Olivetti database [30], dataset has 400 images with grayscale 64×64 pixels. For every person, there are 10 images so there is 40 persons (target) which make it 40×100 equals 400 rows of data. A confusion matrix needs to be computed for each class $g_i \in G = \{1, \dots, K\}$, in such a way that the ith confusion matrix assumes class g_i as the positive class and the remaining classes g_j with $j \neq i$ as negative class. As each confusion matrix pools together the entire observations labelled with a separate class apart from $g_i$as the negative class, this method increases the number of true negatives. This gives us:

- "**True Positive (TN)**" for event values that are correctly analyzed.
- "**False Positive (FP)**" for event values that are incorrectly analyzed.
- "**True Negative (TN)**" for no-event values that are correctly analyzed.

"**False Negative (FN)**" for no-event values that are incorrectly analyzed

Let usTP_i, TN_i, FP_i and $FN_i$to indicate the true positives respectively, false negatives, true negatives, false positives, in the confusion matrix associated with the ithclass. Let their call here be indicated by R and precision by P.

Micro average pools the performance over the least possible unit (the overall facial images):

$$P_{micro} = \frac{\sum_{i=1}^{|G|} TP_i}{\sum_{i=1}^{|G|} TP_i + FP_i} \quad (9)$$

$$R_{micro} = \frac{\sum_{i=1}^{|G|} TP_i}{\sum_{i=1}^{|G|} TP_i + FN_i} \quad (10)$$

The micro-averaged precision, P_{micro}, and recall, R_{micro}, give rise to the micro F1-score:

$$F1_{micro} = 2.\frac{P_{micro} \cdot R_{micro}}{P_{micro} + R_{micro}} \quad (11)$$

Given that a classifier gets a large$F1_{micro}$, it denotes that it performs exceedingly well. Here, micro-average may not be sensitive to the overall predictive performance. Due to this, the micro-average can be misleading when there is an imbalance in the class distribution.

Macro average averages over bigger groups and over the performance of individual classes than observations:

$$P_{macro} = \frac{1}{|G|} \sum_{i=1}^{|G|} TP_i / TP_i + FP_i \quad (12)$$

$$R_{macro} = \frac{1}{|G|} \sum_{i=1}^{|G|} TP_i / TP_i + FN_i \quad (13)$$

The recall and macro-averaged precision leads to the macro F1-score:

$$F1_{macro} = 2.\frac{P_{macro} \cdot R_{macro}}{P_{macro} + R_{macro}} \quad (14)$$

If $F1_{macro}$ has a bigger value, it points out to the fact that a classifier is able to perform well for each of the individual class.

Multi-class accuracy is termed as the average of the correct predictions:

$$accuracy = \frac{1}{N} \sum_{k=1}^{|G|} \sum_{x:g(x)=k} I(g(x) = \hat{g}(x)) \quad (13)$$

where I is defined as the indicator function, which returns 1 when there is a match between the classes and 0 otherwise. For significance testing, cross-validation is a statistical method used for estimating skill of classifiers. k-fold cross validation is a procedure used for estimating skill of model on new data. The value for k is fixed as 10, value that is found using experimentation to generally result in model skill estimate with low bias modest variance.

Table 2: Benchmark Datasets Results Comparison of Metrics Vs. Classifiers

Metrics	Datasets	kNN	SVM	CNN	VGG3	VGG7	VGG16
Macro Average-Precision (%)	Real Time face	65.00	70.00	72.00	78.00	97.00	99.00
	Olivetti	72.00	76.00	79.00	83.00	85.00	90.00
Macro Average-Recall (%)	Real Time face	67.00	71.00	74.00	80.00	97.00	99.00
	Olivetti	75.00	78.00	82.00	87.00	88.00	92.00
Macro Average-F1-score (%)	Real Time face	66.00	70.50	73.00	79.00	97.00	99.00
	Olivetti	73.50	77.00	80.50	85.00	86.50	91.00
Accuracy (%)	Real Time face	64.00	67.00	69.09	75.71	96.53	99.37
	Olivetti	75.00	80.00	84.00	88.00	90.00	94.00
K-fold cross validation Accuracy (%)	Real Time face	65.23	67.86	70.48	76.94	97.23	99.52
	Olivetti	76.10	81.24	84.90	88.66	91.81	95.18

Figure 5 shows the macro average-precision results comparison of six different classifiers like kNN, SVM, CNN, VGG3, VGG7, and VGG16 with two datasets. The proposed VGG16 classifier gives higher macro-average precision results of 99%, the other methods such as kNN, SVM, CNN, VGG3, VGG7 gives 65%, 70%, 72%, 78%, and 97% in real time dataset. The proposed VGG16 classifier gives higher macro-average precision results of 34%, 29%, 27%, 21% and 2% for kNN, SVM, CNN,

VGG3, and VGG7 methods respectively in real time dataset (See Table 2). It can be concluded that VGG16 gives higher macro average-precision, since the proposed work 16 sequential layers are used for recognition.

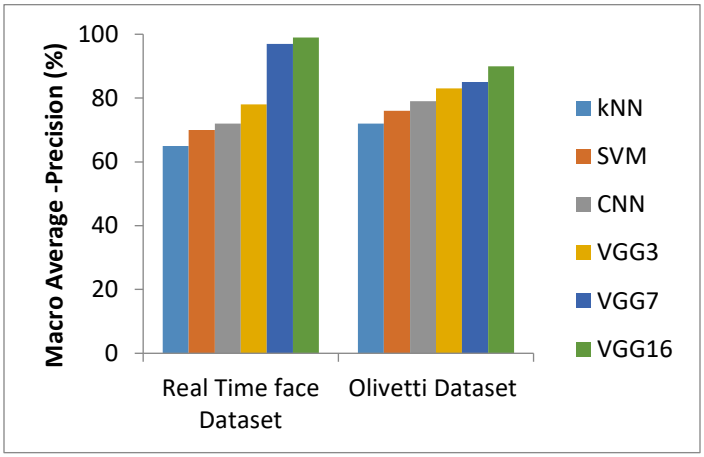


Figure 5: Macro Average-Precision Results Comparison vs. Classifiers

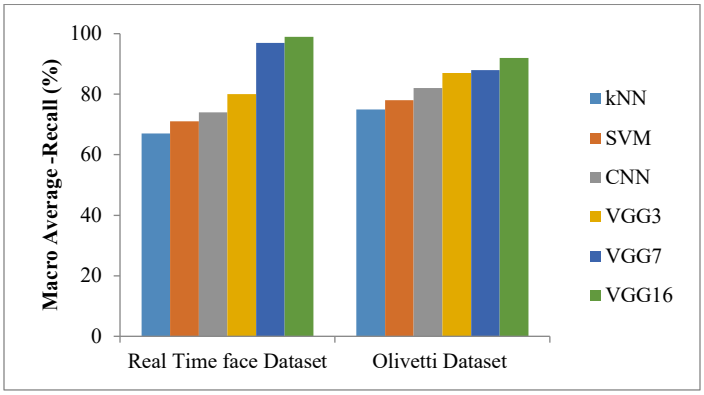


Figure 6: Macro Average-Recall Results Comparison vs. Classifiers

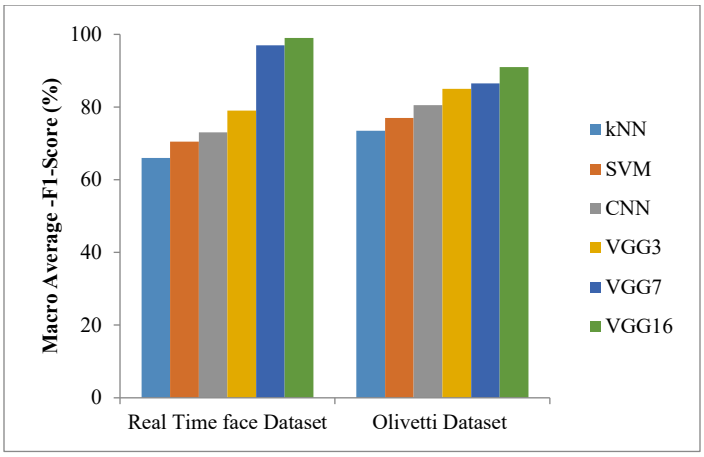


Figure 7 : Macro Average-F1-Score Results Comparison vs. Classifiers

Figure 6 shows macro average-recall results comparison of six different classifiers like kNN, SVM, CNN, VGG3, VGG7, and VGG16 with two datasets. The proposed VGG16 classifier gives higher macro-average recall results of 99%, the other methods such as kNN, SVM, CNN, VGG3, VGG7 gives 67%, 71%, 74%, 80%, and 97% in real time dataset. The proposed VGG16 classifier gives higher macro-average recall results of 32%, 28%, 25%, 19% and 2% for kNN, SVM, CNN, VGG3, and VGG7 methods respectively in real time dataset (See Table 2).

Macro average-F1-score results comparison of four different classifiers such of six different classifiers such as kNN, SVM, CNN, VGG3, VGG7, and VGG16 with two datasets are shown in the figure 7.The proposed VGG16 classifier gives higher macro-average F1-score results of 99%, the other methods such as kNN, SVM, CNN, VGG3, VGG7 gives 66%, 70.5%, 73%, 79%, and 97% in real time dataset. The proposed VGG16 classifier gives higher macro-average F1-score results of 33%, 28.5%, 26%, 20% and 2% for kNN, SVM, CNN, VGG3, and VGG7 methods respectively in real time dataset (See Table 2). On VGG16 based transfer leaning, for getting better results, proposed model is applied with both datasets.

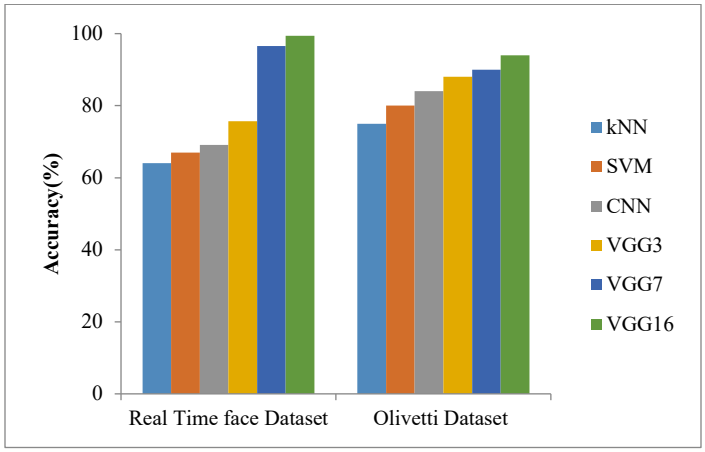


Figure 8: Accuracy Results Comparison VS. Classifiers

Figure 8 shows the accuracy results comparison of six different classifiers such as kNN, SVM, CNN, VGG3, VGG7, and VGG16 with two datasets. The proposed VGG16 classifier gives higher accuracy results of 99.37%, the other methods such as kNN, SVM, CNN, VGG3, VGG7 gives 64%, 67%, 69.09%, 75.71%, and 96.53% in real time dataset. The proposed VGG16 classifier gives higher accuracy results of 35.37%, 32.37%, 30.28%, 23.66% and 2.84% for kNN, SVM, CNN, VGG3, and VGG7 methods respectively in real time dataset (See Table 2). On VGG16 based transfer leaning, for getting better results, proposed model is applied with both datasets. The proposed classifier achieved a very high facial image's recognition rate which approaches human recognition rate.

5. Conclusion and future work

The term Biometrics defines individual's DNA along with other aspects like their facial features, geometry of the hands etc. In addition to that, the behavioral aspects like hand signatures, tone of voice and keystrokes are also taken into consideration. In many circumstances, the recognition of face is becoming more accepted and acclaimed in bio-metric based technologies. This helps in measuring an individual's natural data. This work puts forth a real time face recognition using classification methods. The proposed system contains three major steps that include (1) collection of facial images (2) comparison of trained real time face images via two classifiers such as CNN and VGG16 with transfer learning (3) Results comparison with respect to the metrics like recall, accuracy, precision, F1-score, and precision. The CNN and

VGG16 classifiers recognize faces in real-time with higher accuracy. Both of the classifiers are performed in sequential manner. For VGG16 model is performed based on the transfer learning. Transfer learning intends to extract information from a number of sources tasks and applies it to target task. So VGG16 gives improved accuracy than the CNN classifier. Classifiers are implemented with 1056 face images of24 different persons. The proposed system can successfully recognize 24 different person faces which are which could be useful in searching suspects as its accuracy is much higher than other methods. The proposed VGG16 classifier gives higher values of 99% of macro-average precision, 99% of macro-average recall, 99% of macro-average f1-score and 99.37% accuracy results for real time face images. The major limitation of this work is that it ignores illumination impact in human face recognition which is left as scope of the future work. In future, in various human face sections, this concept can be applied for detecting facial expression for more security.

Conflict of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgment

There is no funding agency supporting our Research.

References

[1] S. Gupta, T. Gandhi, "Identification of Neural Correlates of Face Recognition Using Machine Learning Approach," in Advances in Intelligent Systems and Computing, Springer Verlag: 13–20, 2020, doi:10.1007/978-981-13-8798-2_2.

[2] B.K. Tripathi, "On the complex domain deep machine learning for face recognition," Applied Intelligence, **47**(2), 382–396, 2017, doi:10.1007/s10489-017-0902-7.

[3] G. Hu, Y. Yang, D. Yi, J. Kittler, W. Christmas, S.Z. Li, T. Hospedales, "When Face Recognition Meets with Deep Learning: An Evaluation of Convolutional Neural Networks for Face Recognition," in Proceedings of the IEEE International Conference on Computer Vision, Institute of Electrical and Electronics Engineers Inc.: 384–392, 2016, doi:10.1109/ICCVW.2015.58.

[4] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S.Y. Chang, T. Sainath, "Deep Learning for Audio Signal Processing," IEEE Journal on Selected Topics in Signal Processing, **13**(2), 206–219, 2019, doi:10.1109/JSTSP.2019.2908700.

[5] P. Apoorva, H.C. Impana, S.L. Siri, M.R. Varshitha, B. Ramesh, "Automated criminal identification by face recognition using open computer vision classifiers," in Proceedings of the 3rd International Conference on Computing Methodologies and Communication, ICCMC 2019, Institute of Electrical and Electronics Engineers Inc.: 775–778, 2019, doi:10.1109/ICCMC.2019.8819850.

[6] S.M. Bah, F. Ming, "An improved face recognition algorithm and its application in attendance management system," Array, **5**, 100014, 2020, doi:10.1016/j.array.2019.100014.

[7] H. Mady, S.M.S. Hilles, "Face recognition and detection using Random forest and combination of LBP and HOG features," in 2018 International Conference on Smart Computing and Electronic Enterprise, ICSCEE 2018, Institute of Electrical and Electronics Engineers Inc., 2018, doi:10.1109/ICSCEE.2018.8538377.

[8] H. Ab., A. A., E. E., "A Real-Time System for Facial Expression Recognition using Support Vector Machines and k-Nearest Neighbor Classifier," International Journal of Computer Applications, **159**(8), 23–29, 2017, doi:10.5120/ijca2017913009.

[9] L. He, H. Li, Q. Zhang, Z. Sun, Dynamic Feature Learning for Partial Face Recognition.

[10] Y. Long, F. Zhu, L. Shao, J. Han, "Face recognition with a small occluded training set using spatial and statistical pooling," Information Sciences, **430**–**431**, 634–644, 2018, doi:10.1016/j.ins.2017.10.042.

[11] B. Lahasan, S. Lebai Lutfi, I. Venkat, M.A. Al-Betar, R. San-Segundo, "Optimized symmetric partial facegraphs for face recognition in adverse conditions," Information Sciences, **429**, 194–214, 2018, doi:10.1016/j.ins.2017.11.013.

[12] T.X. Jiang, T.Z. Huang, X. Le Zhao, T.H. Ma, "Patch-Based Principal Component Analysis for Face Recognition," Computational Intelligence and Neuroscience, **2017**, 2017, doi:10.1155/2017/5317850.

[13] P. Shubha, M. Meenakshi, "Human face recognition using local binary pattern algorithm - Real time validation," in Advances in Intelligent Systems and Computing, Springer: 240–246, 2020, doi:10.1007/978-3-030-37218-7_28.

[14] F. Tabassum, M. Imdadul Islam, R. Tasin Khan, M.R. Amin, "Human face recognition with combination of DWT and machine learning," Journal of King Saud University - Computer and Information Sciences, 2020, doi:10.1016/j.jksuci.2020.02.002.

[15] Y. Shin, I. Balasingham, "Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification," in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Institute of Electrical and Electronics Engineers Inc.: 3277–3280, 2017, doi:10.1109/EMBC.2017.8037556.

[16] S. Guo, S. Chen, Y. Li, "Face recognition based on convolutional neural network & support vector machine," in 2016 IEEE International Conference on Information and Automation, IEEE ICIA 2016, Institute of Electrical and Electronics Engineers Inc.: 1787–1792, 2017, doi:10.1109/ICInfA.2016.7832107.

[17] M. Coskun, A. Ucar, O. Yildirim, Y. Demir, "Face recognition based on convolutional neural network," in Proceedings of the International Conference on Modern Electrical and Energy Systems, MEES 2017, Institute of Electrical and Electronics Engineers Inc.: 376–379, 2017, doi:10.1109/MEES.2017.8248937.

[18] Z. Lu, X. Jiang, A. Kot, "Deep Coupled ResNet for Low-Resolution Face Recognition," IEEE Signal Processing Letters, **25**(4), 526–530, 2018, doi:10.1109/LSP.2018.2810121.

[19] S. Almabdy, L. Elrefaei, "Deep Convolutional Neural Network-Based Approaches for Face Recognition," Applied Sciences, **9**(20), 4397, 2019, doi:10.3390/app9204397.

[20] W. Passos, I. Quintanilha, G. Araujo, "Real-Time Deep-Learning-Based System for Facial Recognition," Sociedad Brasileira de Telecomunicacoes, 2018, doi:10.14209/sbrt.2018.321.

[21] S. Saypadith, S. Aramvith, "Real-Time Multiple Face Recognition using Deep Learning on Embedded GPU System," in 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2018 - Proceedings, Institute of Electrical and Electronics Engineers Inc.: 1318–1324, 2019, doi:10.23919/APSIPA.2018.8659751.

[22] G. Nam, H. Choi, J. Cho, I.-J. Kim, "PSI-CNN: A Pyramid-Based Scale-Invariant CNN Architecture for Face Recognition Robust to Various Image Resolutions," Applied Sciences, **8**(9), 1561, 2018, doi:10.3390/app8091561.

[23] Y. Wu, T. Hassner, K. Kim, G. Medioni, P. Natarajan, "Facial Landmark Detection with Tweaked Convolutional Neural Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, **40**(12), 3067–3074, 2018, doi:10.1109/TPAMI.2017.2787130.

[24] K.B. Pranav, J. Manikandan, "Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks," in Procedia Computer Science, Elsevier B.V.: 1651–1659, 2020, doi:10.1016/j.procs.2020.04.177.

[25] G. Lou, H. Shi, "Face image recognition based on convolutional neural network," China Communications, **17**(2), 117–124, 2020, doi:10.23919/JCC.2020.02.010.

[26] L. Wen, X. Li, L. Gao, Y. Zhang, "A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method," IEEE Transactions on Industrial Electronics, **65**(7), 5990–5998, 2018, doi:10.1109/TIE.2017.2774777.

[27] J. Ding, B. Chen, H. Liu, M. Huang, "Convolutional Neural Network with Data Augmentation for SAR Target Recognition," IEEE Geoscience and Remote Sensing Letters, **13**(3), 364–368, 2016, doi:10.1109/LGRS.2015.2513754.

[28] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, R.S. Tan, "A deep convolutional neural network model to classify heartbeats," Computers in Biology and Medicine, **89**, 389–396, 2017, doi:10.1016/j.compbiomed.2017.08.022.

[29] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," International Journal of Scientific and Research Publications (IJSRP), **9**(10), p9420, 2019, doi:10.29322/IJSRP.9.10.2019.p9420.

[30] Face Recognition on Olivetti Dataset | Kaggle, Apr. 2021.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 990-1001 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

A New Video Based Emotions Analysis System (VEMOS): An Efficient Solution Compared to iMotions Affectiva Analysis Software

Nadia Jmour*,[1], Slim Masmoudi[2], Afef Abdelkrim[1]

[1]*Research Laboratory Smart Electricity & ICT, SE&ICT Lab., LR18ES44. National Engineering School of Carthage, University of Carthage, Charguia, 2035, Tunisia*

[2]*Innovation Lab, Naif Arab University for Security Sciences, Riyadh, 14812, KSA*

ARTICLE INFO

Article history:
Received: 10 February, 2021
Accepted: 02 April, 2021
Online: 22 April, 2021

Keywords:
Emotion recognition
Micro-facial expressions
Deep learning
Video analysis
Affectiva
iMotions

ABSTRACT

The Micro-facial expression is the most effective way to display human emotional state. But it needs an expert coder to be decoded. Recently, new computer vision technologies have emerged to automatically extract facial expressions from human faces. In this study, a video-based emotion analysis system is implemented to detect human faces and recognize their emotions from recorded videos. Relevant information is presented on graphs and can be viewed on video to help understanding expressed emotions responses. The system recognizes and analyzes emotions frame by frame. The image-based facial expressions model used deep learning methods. It was tested with two pre-trained models on two different databases. To validate the video-based emotion analysis system, the aim of this study is to challenge it by comparing the performance of the initial implemented model to the iMotions Affectiva AFFDEX emotions analysis software on labeled sequences. These sequences were recorded and performed by a Tunisian actor and validated by an expert psychologist. Emotions to be recognized correspond to the six primary emotions defined by Paul Ekman : anger, disgust, fear, joy, sadness, surprise, and then their possible combinations according to Robert Plutchik's psycho-evolutionary theory of emotions. Results show a progressive increase of the system's performance, achieving a high correlation with Affectiva. Joy, surprise and disgust expressions can reliably be detected with an underprediction of anger from the two systems. The implemented system has shown more efficient results on recognizing sadness, fear and secondary emotions. Contrary to iMotions Affectiva analysis results, VEMOS system has recognized correctly sadness and contempt. It has also successfully recognized surprsie and fear and detect the alarm secondary emotion. iMotions Affectiva has confused surprise and fear. Finally, compared to iMotions the system was also able to detect peak of morbidness and remorse secondary emotions.

1. Introduction

Humans communicate through the expression of their emotions, using various channels (physiological reactions, behaviors, voice, etc.). The emotion is the strong link between cognition and behavior, and between attention, representation and human performance [1]. Facial expression is the most direct and effective way to display human emotional states [2]. In addition, humans use involuntary micro-expressions to unveil their feelings [3]. Nevertheless, they are unable to recognize them accurately. Indeed, only 47% of recognition accuracy is observed at trained adults [4].

Since the 1960's, several facial expressions and emotions classification models have been psychologically studied. Psychologists such as Paul Ekman and Friesen have studied micro-expressions in their research, showing that these micro-expressions are involuntary, incontrollable, and expressed by humans as a very short reaction to a lived situation. Usually, they last less than 1/25 second which makes it hard to find them in real-time [5, 6].

*Corresponding Author: Nadia Jmour, nadiajmour@gmail.com

The authors showed that micro-expressions can be a lie detection tool, In fact, micro-expressions allow recognizing the real emotions expressed by people being in critical situations as the case of criminal interrogation or job recruitment interview [7].

Thus, psychologists established a method to recognize micro expressions by defining criteria of the facial muscles' movements related to each specific micro-expression. Facial movements are defined into 46 action units and each unit presents a contraction or relaxation of a facial muscle. Thereafter, each micro-expression corresponds to the presence of several action units. This allows, for example, recognizing a true involuntary smile, characterized by the movement of the zygomaticus major and of the orbicularis of the eye muscles, [8], from a false smile activating with the movement of the zygomaticus muscle. The developed system is known as the Facial Action Coding System (FACS) and it has become the most used in scientific research [9].

However, intelligent automated recognition of emotional micro-expressions is still a quite new topic in computer vision. Recently, based on machine learning algorithms, and with the availability of facial expressions databases, emotion analysis systems have become able to automatically recognize micro-facial expressions and emotional states, and give an added value technological undertaking [10, 11].

As defined in [12], these systems operate in three steps: Face detection, feature extraction, facial expression and emotion classification.

In traditional methods, feature extraction and feature classification steps are independent. Compared to recent research [13], these machine learning techniques for facial expression recognition are considered as classical techniques.

Nowadays, researchers tend to highlight the advantage of applying deep learning techniques for facial expression recognition in different applications [14, 15]. They consider that applying deep learning reduces human intervention in the analysis process by automating the phase of extracting and learning features [16, 17]. Also, illumination, identity bias and head pose problems are avoided. The need of a large amount of data is also resolved, [18–20].

Several researchers as [21–24], have studied and developed emotion analysis systems for applications in different fields such as human resources, medical field, etc. Authors in [11], employed in his research a deep learning system based on face analysis for monitoring customer behavior specifically for detection of interest. Authors in [25], applied deep learning-based face analysis system in radiotherapy that monitors the patient's facial expressions and predicts patient's advanced movement during the treatment. Authors in [26], considered that facial expression analysis system can be applied on the field of marketing.

Systems such as EmotioNet, iMotions platform [27], using Affectiva Affdex algorithm, [28], Noldus [29], EmoVu [30], and Emotient employ facial expression recognition technologies to create commercial products using information related to their customers. These systems are able to recognize different metrics as head orientation, facial landmarks and basic emotions and most of them apply deep learning [31–33].

The first contribution of this study is to develop an emotion analysis system that recognizes human emotions from recorded videos.

The second contribution is to challenge the system's performance by comparing measures with iMotions Affectiva analysis software measures obtained by analyzing labeled recorded videos.

This study aims also to employ the deep learning-based image analysis model presented in our recent research [18, 20].

Since micro-facial expressions occur during a split second, the proposed system extracts images frame by frame, and every frame is analyzed by the implemented deep learning model. Recorded information is viewed either in an analyzed video and/or presented on the business analysis Power BI Desktop from Microsoft [34], by using visualization graphs to help understanding expressed emotional responses.

Emotions to be recognized correspond to the six facial expressions defined by authors: anger, disgust, fear, joy, sadness, surprise, and then their possible combinations according to Robert Plutchik's psycho-evolutionary theory of emotions.

We expected that the six basic facial expressions can be correctly recognized by the implemented system, with a relatively high accuracy, and can give as an output person's true emotions. As a criterion/external validity of the implemented system, we further expected its significant correlations with iMotions Affectiva measures. Also, secondary emotions identified by Plutchik were expected to be correctly recognized by the proposed new system.

The remainder of this paper is organized into four sections. Section 2 presents the implemented video-based emotion analysis system with an example of a tested scenario. Section 3 presents comparison and challenging results of the implemented system with iMotions Affectiva software. Finally, section 4 summarizes the conclusions

2. Video-based Emotion analysis system (VEMOS) implementation

2.1. Process analysis description

Micro-expressions cannot be controlled, difficult to hide and occur within a split of second. The proposed system extracts first frames from recording video every 1/25 second.

Then, it analyzes and recognizes automatically frame by frame the six basic emotions [32]. Extracted frames are affected as input in the image-based facial expression analysis system, which was implemented in our recent research [18].

This system used deep learning techniques to recognize facial expression images more accurately. It performed by training a deep convolutional neural network (CNN) on a set of face images. This network is composed of two parts:

- The convolutional part for feature extraction: it takes the image as input of the model and then convolution maps are flattened and concatenated into a features vector, called CNN code.

- The fully connected layer part: this part combines the characteristics of the CNN code to classify the image.

To explore the power of these CNN with accessible equipment and a reasonable amount of annotated data, it is possible to explore pre-trained neural networks available publicly [35].

This technique is called transfer learning, it uses knowledge and information acquired from a general classification task and applies it again to a particular new one.

In our recent research, we compared the use of the VGG-16 and the InceptionV3 pre-trained models on the ImageNet dataset, [36], by applying transfer learning technique using the Kaggle (Facial Expression Recognition Challenge) and Extended Cohn-Kanade CK+ facial expression datasets [37].

We used for our implementation Python libraries such OpenCV and TensorFlow on GPU NVIDIA version 391.25, processor type 16xi5-7300HQ. We trained the model on 100 epochs, with a learning rate of 0.0001 [20]. We leveraged the Keras implementation of VGG-16 and InceptionV3.

The data is split into 90% for training and 10% for testing. Experiments had accuracies of 45.83% and 43.22% using the two pre-trained models on the FER-2013. Then, an accuracy of 73.98% was observed by transferring information from the VGG-16, and by training the new architecture on the Extended Cohn-Kanade facial expression dataset using the top frame from every sequence.

Thus, our recent research approved the effect of the distribution and the quality of the data for an efficient training and application of the facial expression recognition task. So, every fraction of a second, this model assigns a score corresponding to each emotion. Output Information is recorded from every analyzed frame.

Based on the recorded facial expressions information, we proposed a new approach for the determination of other emotional states from video data. This approach is inspired from the Robert Plutchik's psycho-evolutionary theory of emotions [38–40].

According to Plutchik, basic emotions can be combined with each other to form different emotions. Table 1 presents the possible combinations.

Table 1: Plutchik's emotions possible combinations.

Primary dyads	Results
Joy and Trust	Love
Trust and fear	Submission
Fear and surprise	Alarm
Surprise and sadness	Disappointment
Sadness and disgust	Remorse
Disgust and anger	Contempt
Anger and anticipation	Aggression
Joy and fear	Guilt
Trust and surprise	Curiosity
Fear and sadness	Despair
surprise and disgust	horror
sadness and anger	Envy
Disgust and anticipation	Cynism
Anger and joy	Pride
Anticipation and trust	Fatalism
Joy and surprise	Delight
Trust and sadness	Sentimentality
Fear and disgust	Shame
Surprise and anger	Outrage
Sadness and anticipation	Pessimism
Disgust and Joy	Morbidness
Anger and trust	Dominance
Anticipation and fear	Anxiety

The Plutchik theory defines emotions as confidence and anticipation. In our case, we explore only combinations based on the six emotions below: anger, disgust, fear, joy, sadness and surprise. Emotions which Plutchik presents as conflict combinations are not considered in our approach.

This approach allows having more information about the emotional state of the person from the video data type. For example, according to Plutchik's theory, the emotion of delight would be a mixture of joy and surprise. Delight's score is defined as the mean of the two basic emotions.

At the end of the process, two types of output are provided. First, the following output information is extracted and visualized through graphs on the Power BI Desktop [41] :

- Dominant basic emotion occurrence: It gives dominant and non-dominant emotion occurrence.
- Maximum basic emotion peak: It gives maximum peak and the corresponding time.
- Emotions Variability: Its records frame by frame the six emotions score over the time.
- Dominant secondary emotion occurrence: It gives dominant and non-dominant secondary emotion occurrence.
- Maximum secondary emotion peak: It gives maximum peak and the corresponding time.

Then, an analyzed video is extracted where face traking and basic and secondary emotions scores over the time are displayed. We leveraged Pandas, and Csv librairies for recording pertinent information from extracted images. Opencv, Mathplotlib and Dlib librairies were used for a better visualization of the resulting analyzed video.

2.2. Tested scenario and results visualization

We chose Carlos Ghosn intervention as a tested video. Carlos Ghosn is the head of Nissan and the French automaker Renault. In November, 2018, Nissan announced that an internal investigation into Ghosn and another executive found Ghosn's compensation had been underreported. There was also evidence that Ghosn committed other misconduct, including personal use of company assets. Japanese authorities arrested Ghosn in Tokyo. He has already spent a hundred and eight days before his second arrestation. On April 9th, 2019, Ghosn posted a structured YouTube video, [42], that he recorded four days before, in which he presented three messages and where he publicly proclaimed that he was innocent of all the accusations that came around these charges that are all biased, taken out of context, twisted in a way to paint a personage of greed, and a personage of dictatorship. Carlos Ghosn spoke calmly, neutral and looked fixedly at the camera in the recorded video.

We analyzed the first message in the video with our implemented video-based emotion analysis system.

Output Extracted information is presented using visualization graphs on the Power BI desktop. Primary and secondary emotions occurrences are shown in Figure 1.

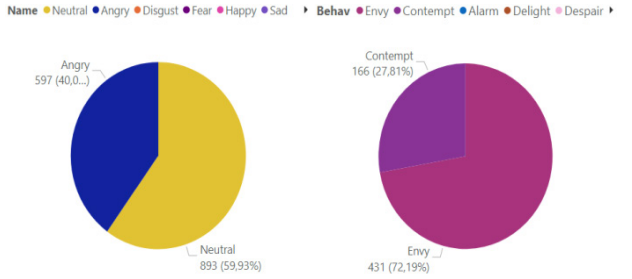


Figure 1: Emotions Occurrence.

Neutral and Anger categories dominated on the whole video with 59.9% and 33.56% respectively. Envy and contempt secondary emotions occurred with 72.1% and 27% respectively.

The variability of each basic emotion over the time is presented in Figure 2. Each color describes an emotion as it is presented on the legend.

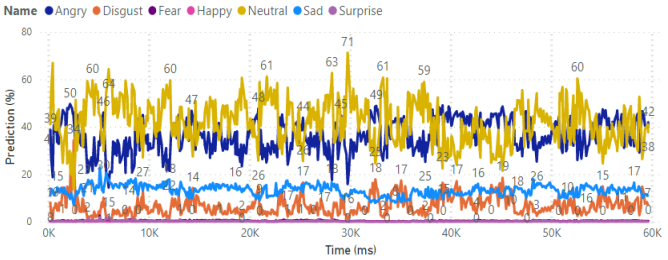
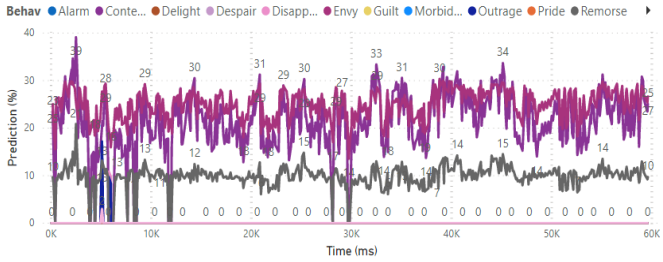


Figure 2: Basic emotions variability.

The figure shows that the two dominant emotions vary with important values. The maximal recorded peak value of the Anger micro-facial expression is 50%. Disgust and Sadness appear with smaller values and they do not exceed 30%.

Secondary emotions are also detected by the system and analysis results are presented in Figure 3.



We noticed the appearance of contempt and envy during the time. Peaks values are 30% and 39% respectively.

The authors tried to be calm and keep his emotions stables. The switch between Anger, Neutrality, on one hand, and contempt, envy, on the other hand, explains also how he maybe tried to manage his emotions.

Figure 4 presents a screenshot from the analyzed video. It shows Carlos while announcing his innocence and expressing envy combining anger and sadness emotions.

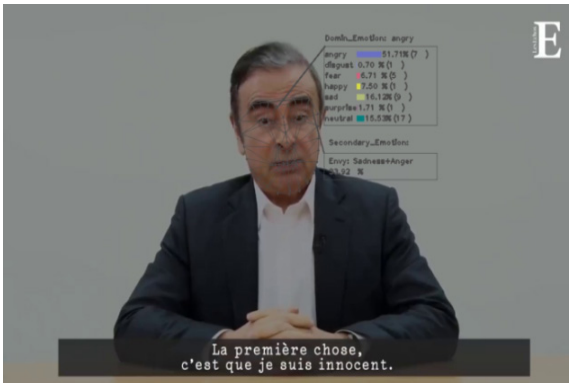


Figure 4: Screenshot of Carlos Ghosn's analyzed video.

Thanks to VEMOS, we unveiled the undetected basic and secondary emotions by a human.

Later, we correlate our results with the 22th April event where the Ex-CEO of Renault-Nissan Carlos Ghosn has been charged again, for breach of trust aggravated by the Tokyo court. This can maybe explain the presence of the detected negative emotions.

3. Comparison with iMotions Affectiva facial expressions analysis software

The performance of the implemented video-based emotions analysis system VEMOS had to be validated. So, we decided to challenge the system by comparing the performance with the iMotions Affectiva AFFDEX emotion analysis software.

3.1. iMotions Biometric Research Platform

iMotions [27, 43], is a software solution helping for data collection, analysis, and quantifying engagement and emotional responses. The iMotions Platform is an emotional recognition software that integrates multiple sensors including facial expression analysis, GSR, eye tracking, EEG, and ECG/EMG along with survey technologies for multimodal human behavior research.

Videos are imported and analyzed in iMotions Platform using the Affectiva Affdex facial expression recognition engine [27, 28].

The Affectiva's Affdex SDK toolkit (in collaboration with iMotions) [28, 44–46], is one of the most known and used emotions analysis tool. It uses the Facial Action Coding system (FACS). The tool detects facial landmark, classify facial action, and recognizes the seven basic emotions.

The system's algorithm uses first a histogram of oriented gradient (HOG) to extract pertinent images regions. Then, an SVM is used for the classification step. The training process is applied on 10000 images and determined a score of 0–100 for each class. The testing process is applied on 10000 images.

3.2. Carlos Ghosn's video correlation results

The video of Carlos Ghosn is analyzed with iMotions Affectiva software. Extracted information are presented in the Table 2.

Table 2: iMotions Affectiva emotions occurence results.

Emotion	Occurrence
Anger	184
Disgust	22
Fear	0
Joy	0
Sadness	29
Surprise	321
Neutral	1429

Through the iMotions Affectiva system, the occurrence of emotions is detected when its corresponding score exceeds the threshold value of 50%.

Taking into account this characteristic, we noticed that results given by iMotions and VEMOS systems lead to the same psychological interpreation of the emotional state of Carlos that we have concluded from the video. Then, we used the Bravais-Pearson correlation coefficient to compare analysis results obtained by the two systems.

Obtained values range between -1.0 and 1.0. A correlation of -1.0 shows a perfect negative correlation, while a correlation of 1.0 shows a perfect positive correlation. A correlation of 0.0 shows no linear relationship between the movement of the two variables. We considered that values less than 0.5 as low values and values greater than 0.5 as high values. Table 3 shows the correlation results.

Table 3: Correlation coefficient with iMotions software

Emotion	Correlation coefficient Initial model	p
Anger	0.051	$p>.05$
Disgust	0.38	$p<.05$
Fear	-0.19	$p<.05$
Joy	0.11	$p<.05$
Sadness	0.06	$p>.05$
Surprise	0.0015	$p>.05$

We obtained low values of correlation for the six basic emotions. As the video is not labeled and is not analyzed by non-verbal professional decoder, results remain open and they do not validate the performance of the system.

To ensure the reliability of our results we proposed to apply our analysis on labeled videos. So, we suggested to prepare two corpus of labeled sequences that were recorded and performed by a Tunisian actor and validated by a professional psychologist. The first corpus contains basic emotions sequences and the second one contains secondary emotions sequences.

3.3. Labeled recorded Sequences

We first collected labeled sequences from YouTube video. The psychologist confirms that emotions expressed in these sequences present muscles' mouvement that correspond to the needed emotions. Then, a Tunisian Actor was hired to perform videos under the monitoring and guidance of a professional psychological doctor.

Before recording the videos, the psychologist gives more details about the muscles to move specifying the facial action coding units per emotion. Then, video recording starts with the actor looking into the camera and showing successively the six facial expressions anger, disgust, fear, joy, sadness, and surprise, repeating the action several times. Thus, many sequences were recorded and only one per emotion was chosen and validated by the psychologist. Table 4 shows the details about the videos characteristics.

Table 4: Basic emotions Sequences characteristics

Attribute	Description
Length of sequences	5000 to 10000 milliseconds
Number of frames	25 fps
Expressions classes annotation	Angry, disgust, fear, happy, sad, surprise
Video format	mp4
Number of subjects	2

Similarly, the actor looked into the camera and expressed the expected secondary emotions. Sequences were recorded and validated by the psychologist. Table 5 below presents characteristics of the prepared sequences.

Table 5: Prepared secondary emotions Sequences characteristics.

Attribute	Description
Length of sequences	5000 to 10000 milliseconds
Number of frames	25 fps
Expressions classes annotation	alarm, contempt, morbidness, despair, remorse
Video format	mp4
Number of subjects	1

3.4. Basic emotions correlation results

To explore the reliability of the Affectiva algorithm for basic emotions recognition, we proposed to take advantages of the constructed corpus of the labeled sequences. So, recorded primary emotion sequences were analyzed by the VEMOS implemented system. They were also analyzed by the iMotions Affectiva Affdex algorithm and temporal information was imported for the 6 sequences.

The analysis results should first correlate with the expected emotion. Then, for a better visualization, the extracted values are presented on graphs using the Power BI Desktop.

By plotting these graphs, facial expression measures are compared to iMotions stored facial expression measures.

Figure 5 shows emotions variability over the time obtained by the VEMOS system for every sequence. Analysis results show that the initial implemented system has difficulties on classifying some facial expressions. Compared to the expected emotions disgust, fear, joy and sadness categories are misclassified. Emotions as angry and surprise are correctly classified and reached maximum values of 87% and 99% respectively.

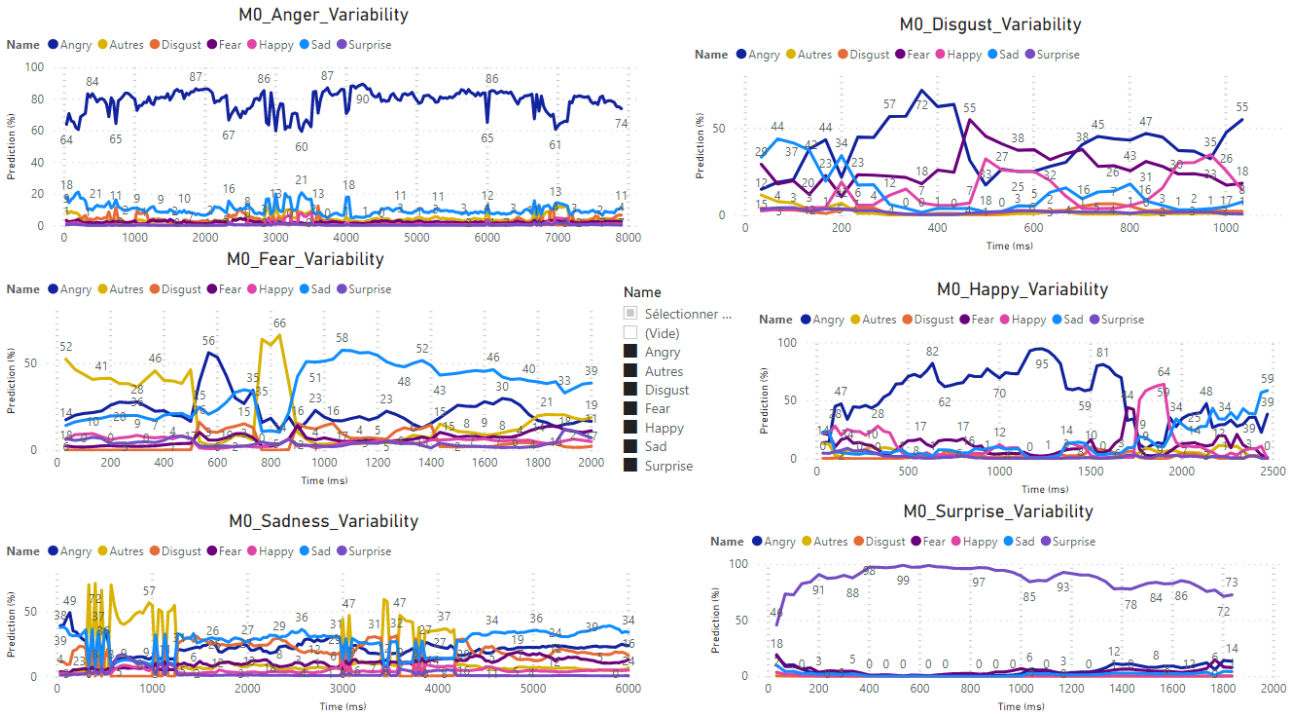


Figure 5: Analysis results with the initial implemented model.

In order to compare results to iMotions platform analysis results, we used the Bravais-Pearson correlation coefficient. Table 6 shows correlation coefficient of each emotion.

Table 6: Correlation coefficient with iMotions software

Emotion	Correlation coefficient Initial model	*p*
Anger	0.01	$p>.05$
Disgust	0.33	$p<.05$
Fear	-0.41	$p<.05$
Joy	0.19	$p<.05$
Sadness	0.03	$p>.05$
Surprise	0.15	$p<.05$

Obtained results show small values and low correlation between the implemented VEMOS system and iMotions Affectiva results.

Consequently, we proposed to challenge the system and improve its performance. The idea was to challenge the number of data and the distribution of emotions categories used for the training of the image-based facial expression system.

Our recent approach, [18], has shown the importance of the quality of the data in the performance of the implemented model. The Fer-2013 dataset presents images that were collected from google and not properly treated. Sometimes, we found images where there is text on the faces. So, compared to the Fer-2013, our best accuracy was obtained by training and testing the model on the CK+ dataset, which was implemented by exploring the top frame of each sequence from the data.

In this approach, we propose to increase the number of input frames from the CK+ dataset by taking in consideration the data distribution. Then, the model is trained by keeping the same proposed architecture and its same characteristics. We propose to construct a new corpus based on the CK+ dataset by taking five frames per sequence and to integrate a new category called “others”. Neutral faces and all other faces movements that are not recognized by the model as one of the basic emotions are put on the new class "others".

The model reached 93.98% test accuracy. Confusion matrix is shown in Figure 6.



Figure 6: Confusion matrix.

To test the reliability and the performance of the challenging system, we used the same labeled sequences. Sequences are already analyzed by the iMotions Affectiva software. So, they were only imported and analyzed by the VEMOS system. The following output information is used for comparison:

- Basic emotions variability of the VEMOS system over the time.

- Basic emotions variability of the iMotions software over the time.

Bravais-Pearson Correlation coefficient is then applied to these measures. Variability over the time of the class anger is presented on Figure 7.

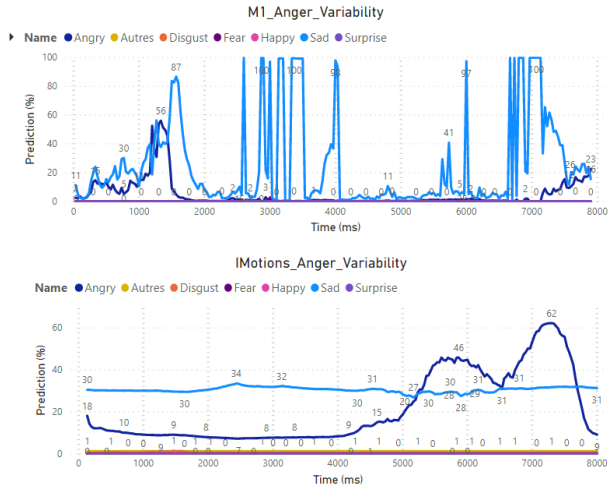


Figure 7: Anger sequence's analysis.

The two systems confirm the presence of the same emotions over the time and recorded correlation coefficient is 0.22 ($p<.05$). Compared to the initial model, correlation has increased with 0.12 ($p<.05$). But compared to the expected emotion, the two systems show confusion between anger and sadness emotions. It is noteworthy that AFFDEX often confused anger with sadness.

So, we can conclude that VEMOS has this same characteristic. It has an underprediction of anger and overprediction of sadness. Figure 8 presents the analysis results for sadness sequence.

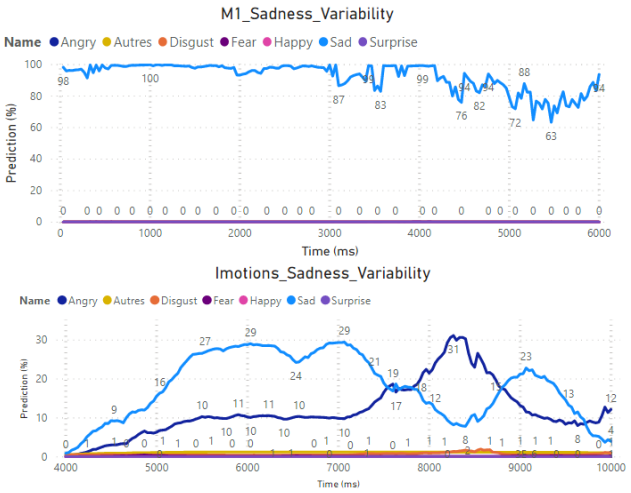


Figure .8: Sadness sequence's analysis.

Sadness is dominating in this sequence and correctly classified by the VEMOS system. Moreover, we recorded satisfying correlation coefficients with values of 0.32 ($p<.05$) with iMotions Affectiva results.

Although the dominance of sadness shown in the analysis results of iMotions Affectiva, we can see again the same confusion of recognizing anger and sadness. VEMOS recognizes sadness emotion with 100% of occurrence.

For disgust sequence, analysis results are presented in Figure 9.

Disgust is correctly recognized by the VEMOS system and reaches a maximum value of 98% over time. Correlation between the two measures is significant 0.21 ($p<.05$).

Happy and surprise sequences analysis are presented on Figures 10 and 11.

The corresponding graphs show that these emotions are correctly classified compared to the expected emotion and reach a maximum peak of 100%. We also recorded high significant correlations with 0.81 ($p<.05$) and 0.54 ($p<.05$).

Figure 12 shows the analysis results for fear sequence. It shows the dominance of the expected category, and measured correlation coefficient is 0.79 ($p<.05$).

Our analysis shows that the implemented VEMOS system recognizes correctly joy, fear, surprise, sadness and disgust. But we can notice a rarely confusion of the anger with sadness.

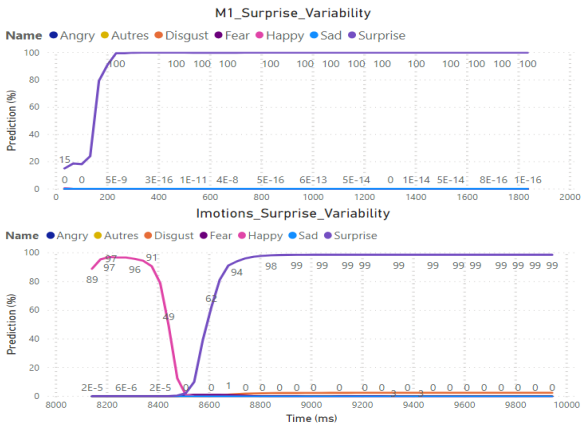


Figure.9: Disgust sequence's analysis.

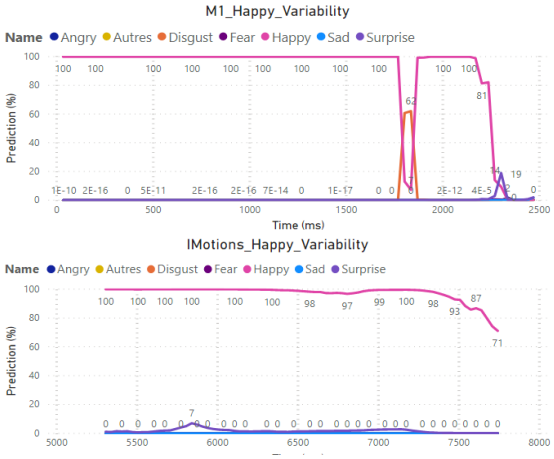


Figure.10: Surprise sequence's analysis

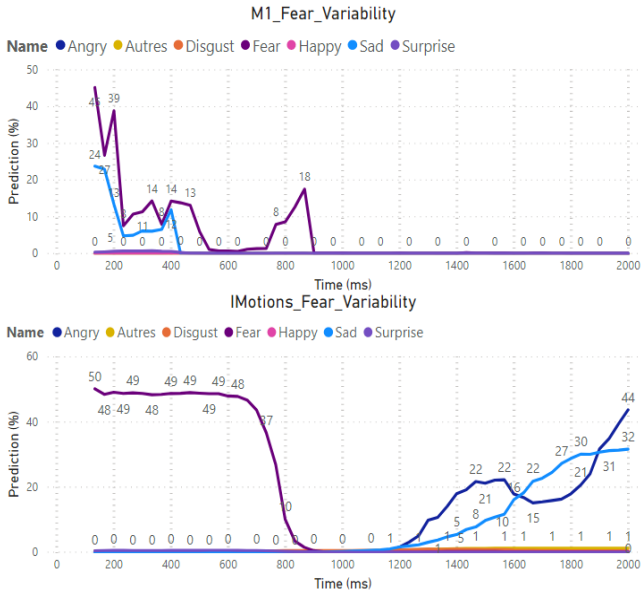


Figure 11: Happy sequence's analysis

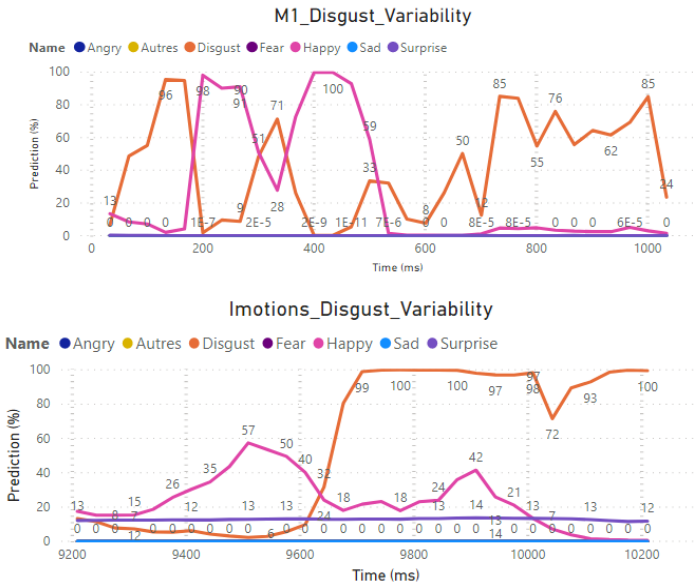


Figure 12: Fear sequence's analysis

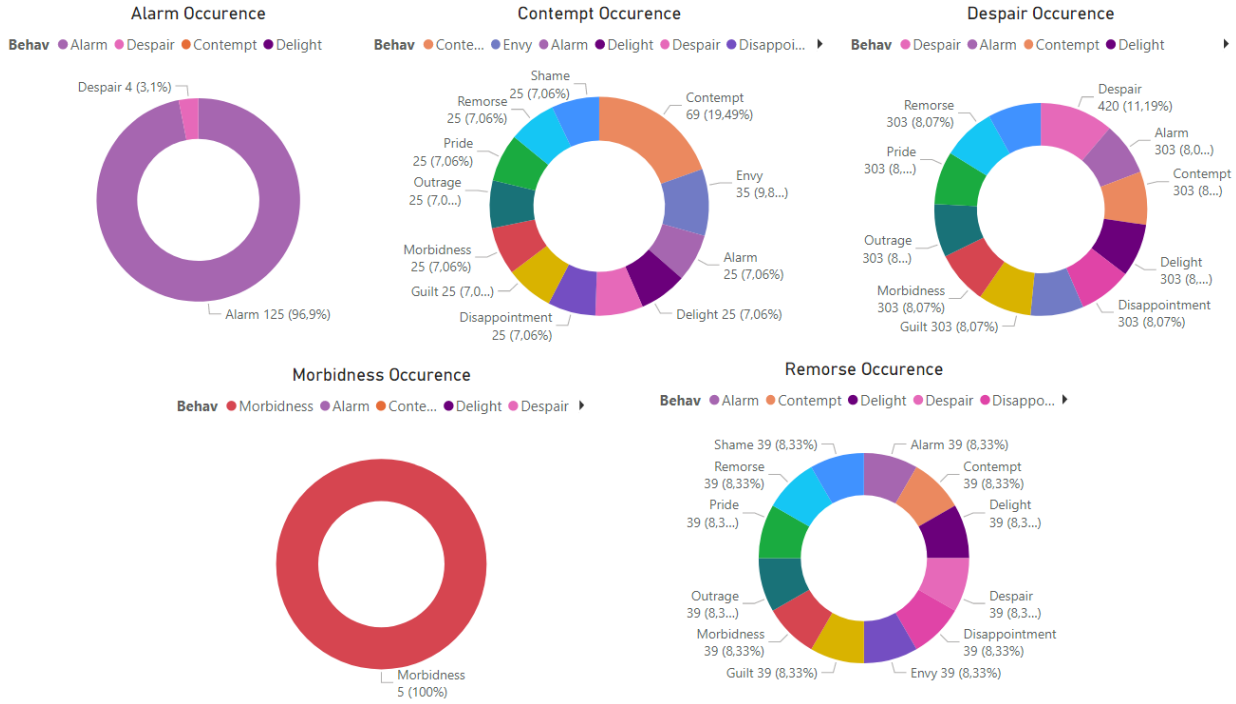


Figure 13: Secondary emotions occurrence

3.5. Secondary Emotions correlation results

The authors defined 13 possible combinations based on the 6 primary emotions: anger, disgust, fear, joy, sadness, surprise. Our system was able to analyze the 13 possible combinations.

Tested emotions are: alarm, contempt, despair, morbidness and remorse. As we have mentioned in section 1, the recorded score for each secondary emotion presents the average of the defined combination. We proposed that at a given moment, if one of the basic emotions is zero, the secondary emotion is assigned to 0. In this case, the detected emotion at this moment will be the primary detected emotion.

To validate and test the performance of the VEMOS system, secondary emotions sequences of the corpus 2 recorded by the actor, are analyzed.

For each sequence, the analysis results are presented on the Power BI Desktop based on the following information:

- Basic emotions variability from the challenging system (VEMOS).

- Basic emotions variability from the iMotions Affectiva software.
- Dominant Secondary emotions occurrence from the challenging system (VEMOS).

Similarly, these sequences were imported and analyzed by the iMotions Affectiva software. The first analysis and validation considered the comparison with the expected emotion.

The graphs describing the dominant secondary emotions for each sequence are presented in Figure 13.

These graphs allow us to compare the dominant emotion with the expected emotion expressed by the actor. Results show relevant information.

We notice that alarm is correctly recognized by the VEMOS system. 96.9% of the extracted frames expressed the expected emotion and only 3% show despair. Contempt is correctly recognized compared to the sequence's label. It dominates with 19.49%.

Figure 14 presents a screenshot of the analyzed video of the actor while expressing contempt emotion. The display of information to be visualized has also been modified and improved. The complete analyzed video of the actor is also available in, [47][1].

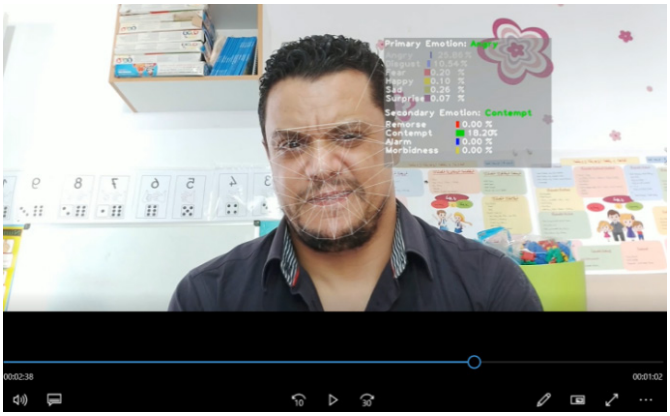


Figure 14: Screenshot of the Actor while expressing Contempt[1].

Analyses of despair and morbidness sequences show that the two emotions are correctly recognized and occurred with 11.19% and 100% respectively.

Figure 13 shows also that the remorse is not correctly recognized. Thus, compared to the expected emotion, alarm, contempt, despair and morbidness are correctly recognized by our system.

Then, second analysis and validation considers the comparison with iMotions Affectiva software results. Basically, this comparison concernes primary emotion defined in the combination of each secondary emotion. So, the variability of these primary emotions over the time was recorded for each sequence by the two systems and compared by measuring the correlation coefficient.

Consequently, we can see how close the VEMOS results are to that of iMotions Affectiva Software and we can take advantages to confirm again the performance of our system on recognizing primary emotions.

Alarm emotion is a combination of the fear and surprise. Compared to the meaning of fear emotion, Alarm is unexpected and apprehended. It can concern a doubt or lack of self-confidence. Alarm basic emotions variability over the time is analyzed by the two systems. Graphs are presented in the Figure 15.

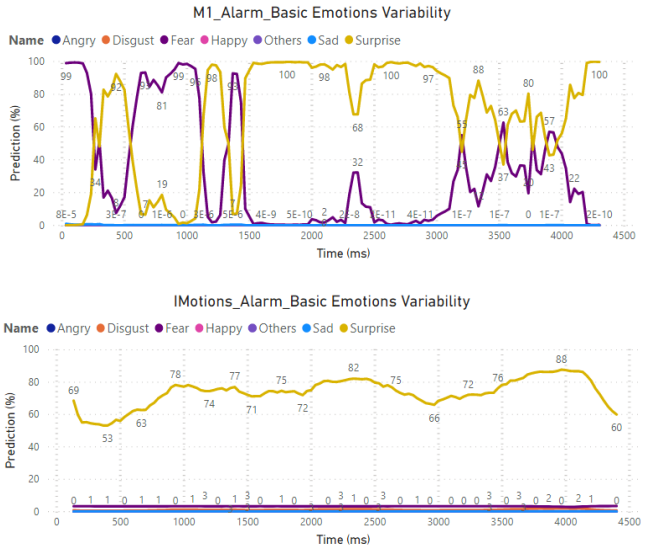


Figure 15: Alarm's sequence analysis.

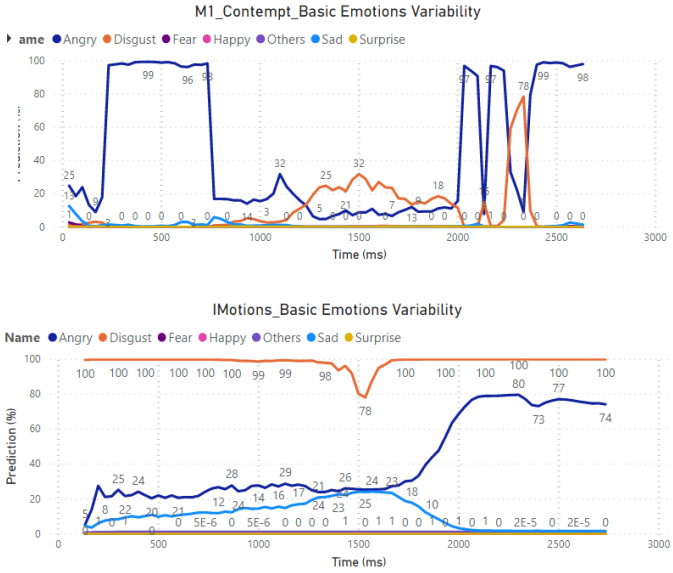


Figure 16: Contempt's correlation analysis.

Analysis present relevant results compared to iMotions Affectiva. The system shows an appearance of fear and surprise over the time. Obtained correlation coefficients of fear and

[1]https://youtu.be/W_0-CuAwDFE

surprise are -0.03 and 0.14 respectively. But in this case, iMotions shows only surprise.

Another peculiarity to note is that Affectiva AFFDEX usually confused fear with surprise. This can probably explain the obtained results.

Contempt has the characteristics of anger and disgust. Analyses results are presented in Figure 16.

The two graphs show an appearance of anger and disgust over the time. iMotions analysis also shows the appearance of small variations of sadness. Obtained correlation coefficients of anger and disgust are 0.24 and -0.25 respectively.

Despair emotion combines sadness and fear. This emotion expresses a complete despondency and a loss of hope. We analyzed the sequence with the two systems. Basic emotions variability is recorded and presented on graphs in Figure 17.

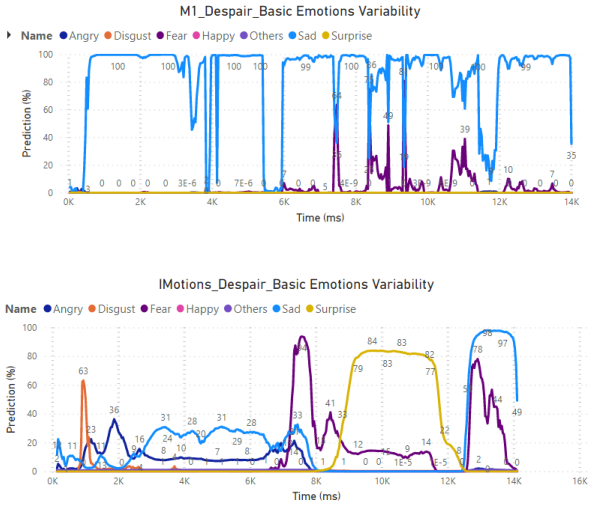


Figure 17: Despair's emotion analysis.

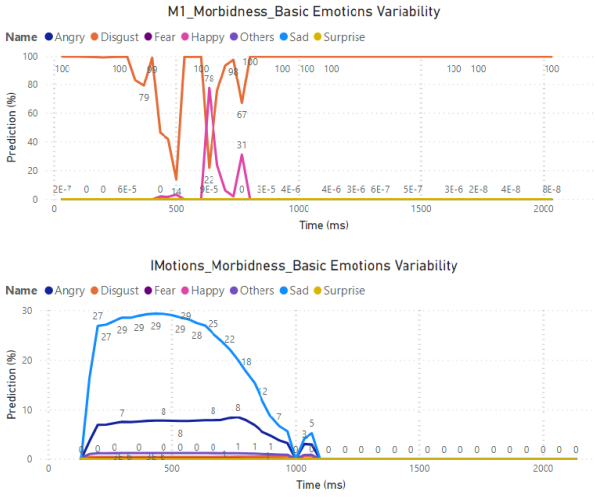


Figure 18: Morbidness's correlation analysis

Our system showed performing results. Sadness and fear appear over the time as it is defined on the Plutchik's emotion combination. Obtained values of correlation coefficient for fear and sadness with iMotions Affectiva analysis results are 0.65 and 0.11, respectively. The second graph in Figure 17 shows iMotions Affectiva analysis. Results show dominance of sadness and fear. Graph also shows an appearance of the surprise with high values. This appearance can be explained by the fact that iMotions Affectiva confuses surprise and fear. Anger is shown with small variations at the recorded resulting variability of emotions. This can be explained by the fact that iMotions confuses between anger and sadness.

For morbidness, analyzes are presented in Figure 18. According to Plutchik, morbidness combines disgust and joy emotions.

The first graph shows that the two dominant emotions throughout the video are happiness and disgust. It shows that the secondary emotion appears as a peak between 500 and 1000 milliseconds.

iMotions Affectiva analysis shows that the person has expressed sadness and anger. Therefore, the iMotions Affectiva system did not show the expected result. Obtained values of correlation coefficient for disgust and happiness are -0.35 and -0.02 respectively. Figure 19 presents analysis results for remorse emotion.

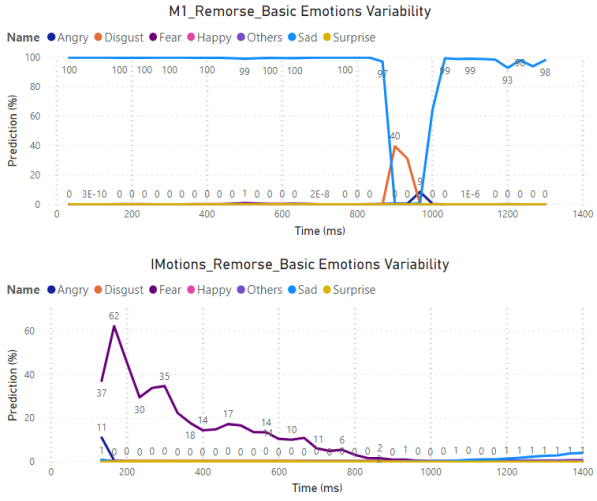


Figure 19: Remorse's correlation analysis.

In the case of the remorse sequence, only sadness has occurred with high scores. The second emotion is disgust, as it is defined in the remorse combination. But disgust appears with one peak all over the video and with low score of 39.53%. iMotions Affectiva system assigned the sequence to the fear category. The correlation values of the basic emotions for remorse combination are 0.05 for disgust and 0.016 for sadness.

We can conclude that our system has successfully recognized secondary emotions. First, results were validated by comparing them to the expected emotion expressed by the actor and validated by the psychologist. These results were also validated according to Plutchik's definition of the basic emotions combination.

Analysis results were also compared to those of iMotions Affectiva results. Joy, disgust and surprise were correctly recognized by the two systems.

More efficient performance was obtained by our implemented system VEMOS for the other emotions.

We notice that iMotions Affectiva confuses anger and sadness emotions. Analysis results of the primary emotions sequences have shown that the system has an underprediction of sadness and overprediction of anger. Because of this confusion, the system was not able to recognize correctly despair and contempt emotions that combine one on the confused primary emotion.

VEMOS system presents an underprediction of anger. But, the system recognizes correctly the sadness emotion. Then, results have shown that the implemented system recognizes correctly despair and contempt emotions.

In addition, VEMOS recognizes more efficiently fear and surprise emotions. Besides, Alarm was correctly recognized. But, we noticed that iMotions Affectiva system has an overprediction of surprise and underprediction of fear. This was shown in the analysis results of Alarm sequence that combine fear and surprise emotions. Only surprise was detected.

Finally, morbidness and remorse secondary emotions were not recognized by iMotions Affectiva system. But, VEMOS was able to detect a maximum peak of the corresponding combination.

4. Conclusion

In this study, a video-based emotion analysis system, VEMOS, was implemented. It detected frame by frame images from video. Then, it employed deep learning techniques for basic emotions recognition. Information about emotions occurrence, emotions variability during the time and maximum peak of each emotions are detected and presented in Power BI Report. Also, an analyzed video showing the variation of pertinent information during the time is obtained.

Then, the system was validated and challenged by comparing correlation with the iMotions Affectiva software on labeled sequences. These sequences were recorded by an actor and validated by a psychologist.

Our validation studies reveal that the implemented system has the potential to recognize basic emotions expressed by individuals. Then, we obtained significant correlations between the implemented system and Affectiva measures. More efficiently than Affectiva, VEMOS system recognizes successfully 5 secondary emotions defined by Plutchik.

With this efficient performance, VEMOS can be exploited in different application. It can help customs officers, known as detection officer, to unveil travelers with illegal items or goods in the customs airport. It is also able to unveil wanted and terrorist people in airport. So, detection officer can stop them from entering and leaving the country. Furthermore, judge can resort to VEMOS to reveal criminal people. Thus, it can assist to know the truth and change the verdict, from a criminal to an innocent.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was supported by Biware Consulting company (https://biware-consulting.com/) from Tunisia and by the National Engineering School of Carthage.

Authors thank Biware Consulting Groups for their confidence and help. Special thanks to Amine Boussarsar and Walid Kaâbachi for spirited discussion and helpful comments on an earlier draft. Thanks also to the Tunisian actor Oussema Kochkar for providing videos and accepting to perform the needed emotions.

References

[1] S. Masmoudi, D. Yun Dai, N. Abdelmajid, Attention, Representation, and Human Performance : Integration of Cognition, Emotion, and Motivation, Psychology Press, 2012, doi:10.4324/9780203325988.

[2] A. Freitas-Magalhaes, The psychology of emotions: The allure of human face., Uni. Fernando Pessoa Press, 2007.

[3] P. Ekman, Lie Catching and Microexpressions, Oxford University Press, Jan. 2021.

[4] Q. Wu, X. Shen, X. Fu, "The Machine Knows What You Are Hiding: An Automatic Micro-expression Recognition System," in: D'Mello, S., Graesser, A., Schuller, B., and Martin, J.-C., eds., in Affective Computing and Intelligent Interaction, Springer, Berlin, Heidelberg: 152–162, 2011, doi:10.1007/978-3-642-24571-8_16.

[5] P Ekman, WV Friesen, W. Friesen, J Hager, Facial action coding system: A technique for the measurement of facial movement – ScienceOpen, 1978.

[6] P. Ekman, "Facial expression and emotion," American Psychologist, **48**(4), 384–392, 1993, doi:10.1037/0003-066X.48.4.384.

[7] D. Matsumoto, H.C. Hwang, "Microexpressions Differentiate Truths From Lies About Future Malicious Intent," Frontiers in Psychology, **9**, 2018, doi:10.3389/fpsyg.2018.02545.

[8] J.-W. Tan, S. Walter, A. Scheck, D. Hrabal, H. Hoffmann, H. Kessler, H.C. Traue, "Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions," Journal of Ambient Intelligence and Humanized Computing, **3**(1), 3–10, 2012, doi:10.1007/s12652-011-0084-9.

[9] P. Ekman, J.C. Hager, W.V. Friesen, Facial action coding system: the manual, Research Nexus, Salt Lake City, 2002.

[10] M.S. Hossain, G. Muhammad, "Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data," Information Fusion, **49**, 2018, doi:10.1016/j.inffus.2018.09.008.

[11] G. Yolcu, I. Oztel, S. Kazan, C. Oz, F. Bunyak, "Deep learning-based face analysis system for monitoring customer interest," J. Ambient Intell. Humaniz. Comput., 2020, doi:10.1007/S12652-019-01310-5.

[12] A. Krause, Picture Processing System by Computer Complex and Recognition of Human Faces, PhD, Kyoto University, Kyoto, 1973.

[13] B. Reddy, Y.-H. Kim, S. Yun, J. Jang, S. Hong, K. Nasrollahi, C. Distante, G. Hua, A. Cavallaro, T.B. Moeslund, S. Battiato, "End to End Deep Learning for Single Step Real-Time Facial Expression Recognition," in Video Analytics. Face and Facial Expression Recognition and Audience Measurement, Springer International Publishing, Cham: 88–97, 2017, doi:10.1007/978-3-319-56687-0_8.

[14] A. Sun, Y.J. Li, Y.M. Huang, Q. Li, "Using facial expression to detect emotion in e-learning system: A deep learning method," in Emerging Technologies for Education - 2nd International Symposium, SETE 2017, Held in Conjunction with ICWL 2017, Revised Selected Papers, Springer Verlag: 446–455, 2017, doi:10.1007/978-3-319-71084-6_52.

[15] H. Sikkandar, R. Thiyagarajan, "Deep learning based facial expression recognition using improved Cat Swarm Optimization," Journal of Ambient Intelligence and Humanized Computing, 2020, doi:10.1007/s12652-020-02463-4.
[16] R. Ranjith, K. Mala, S. Nidhyananthan, "3D Facial Expression Recognition Using Multi-channel Deep Learning Framework," Circuits, Systems, and Signal Processing, **39**, 2020, doi:10.1007/s00034-019-01144-8.
[17] S. Loussaief, A. Abdelkrim, "Deep learning vs. bag of features in machine learning for image classification," in 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 6–10, 2018, doi:10.1109/ASET.2018.8379825.
[18] J. Nadia, Z. Sehla, A. Afef, "Deep Neural Networks for Facial expressions recognition System," International Conference on Smart Sensing and Artifical Intelligence (ICS2'AI), 2020.
[19] Customer Satisfaction Recognition Based on Facial Expression and Machine Learning Techniques, Journal, Jan. 2021.
[20] N. Jmour, S. Zayen, A. Abdelkrim, "Convolutional neural networks for image classification," in 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 397–402, 2018, doi:10.1109/ASET.2018.8379889.
[21] Deep Learning Model for A Driver Assistance System to Increase Visibility on A Foggy Road, Journal, Jan. 2021.
[22] M. Edwards, E. Stewart, R. Palermo, S. Lah, "Facial emotion perception in patients with epilepsy: A systematic review with meta-analysis," Neuroscience and Biobehavioral Reviews, **83**, 212–225, 2017, doi:10.1016/j.neubiorev.2017.10.013.
[23] S.H. Lee, "Facial Data Visualization for Improved Deep Learning Based Emotion Recognition," Journal of Information Science Theory and Practice, **7**(2), 32–39, 2019, doi:10.1633/JISTAP.2019.7.2.3.
[24] W. Medhat, A. Hassan, H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, **5**(4), 1093–1113, 2014, doi:10.1016/j.asej.2014.04.011.
[25] Kim, "Facial expression monitoring system for predicting patient's sudden movement during radiotherapy using deep learning," Journal of Applied Clinical Medical Physics, 2020.
[26] A.M. Barreto, Application of Facial Expression Studies on the Field of Marketing, Jan. 2021.
[27] iMotions Biosensor Software Platform - Unpack Human Behavior, Imotions, Jan. 2021.
[28] D. McDuff, R. el Kaliouby, T. Senechal, M. Amr, J.F. Cohn, R. Picard, "Affectiva-MIT Facial Expression Dataset (AM-FED): Naturalistic and Spontaneous Facial Expressions Collected 'In-the-Wild,'" in 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 881–888, 2013, doi:10.1109/CVPRW.2013.130.
[29] Noldus | Innovative solutions for behavioral research., Noldus | Innovative Solutions for Behavioral Research., Jan. 2021.
[30] pwtempuser, EmoVu, ProgrammableWeb, 2014.
[31] H.M. Fayek, M. Lech, L. Cavedon, "Evaluating deep learning architectures for Speech Emotion Recognition," Neural Networks: The Official Journal of the International Neural Network Society, **92**, 60–68, 2017, doi:10.1016/j.neunet.2017.02.013.
[32] P. Barros, D. Jirak, C. Weber, S. Wermter, "Multimodal emotional state recognition using sequence-dependent deep hierarchical features," Neural Networks, **72**, 140–151, 2015, doi:10.1016/j.neunet.2015.09.009.
[33] L. Bozhkov, P. Koprinkova-Hristova, P. Georgieva, "Learning to decode human emotions with Echo State Networks," Neural Networks, **78**, 112–119, 2016, doi:10.1016/j.neunet.2015.07.005.
[34] J. Kim, B. Kim, P.P. Roy, D. Jeong, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure," IEEE Access, **7**, 41273–41285, 2019, doi:10.1109/ACCESS.2019.2907327.
[35] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, "A Comprehensive Survey on Transfer Learning," Proceedings of the IEEE, **109**(1), 43–76, 2021, doi:10.1109/JPROC.2020.3004555.
[36] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012.
[37] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 94–101, 2010, doi:10.1109/CVPRW.2010.5543262.
[38] R. Plutchik, Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION, Academic Press: 3–33, 1980, doi:10.1016/B978-0-12-558701-3.50007-7.
[39] R. Plutchik, Chapter 1 - MEASURING EMOTIONS AND THEIR DERIVATIVES, Academic Press: 1–35, 1989, doi:10.1016/B978-0-12-558704-4.50007-4.
[40] R. Plutchik, "The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," American Scientist, **89**(4), 344–350, 2001.
[41] Visualisation des données | Microsoft Power BI, Jan. 2021.
[42] CBS News, Carlos Ghosn clame son "innocence" dans une vidéo, 2019.
[43] Facial Expression Analysis: The Complete Pocket Guide, Imotions, Jan. 2021.
[44] In Lab Biometric Solution, Affectiva, Jan. 2021.
[45] L. Kulke, D. Feyerabend, A. Schacht, "A Comparison of the Affectiva iMotions Facial Expression Analysis Software With EMG for Identifying Facial Expressions of Emotion," Frontiers in Psychology, **11**, 2020, doi:10.3389/fpsyg.2020.00329.
[46] S. Stöckli, M. Schulte-Mecklenbeck, S. Borer, A.C. Samson, "Facial expression analysis with AFFDEX and FACET: A validation study," Behavior Research Methods, **50**(4), 1446–1460, 2018, doi:10.3758/s13428-017-0996-1.
[47] Nadia Jmour, VEMOS the new video based emotion analysis system: Actor while expressing Alarm and Contempt, 2021.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 1091-1100 (2021)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Enhanced Data Transportation in Remote Locations Using UAV Aided Edge Computing

Niranjan Ravi*, Mohamed El-Sharkawy

Indiana University Purdue University Indianapolis, Department of Electrical and Computer Engineering, Institute, Indianapolis, IN, 46202, USA

ARTICLE INFO

Article history:
Received: 16 February, 2021
Accepted: 18 March, 2021
Online: 28 April, 2021

Keywords:
Unmanned Aerial Systems (UAV)
IoT
UAVCAN
ECC
IBM Cloud
Thread Mesh Network
Ble Mesh

ABSTRACT

In recent years, the applications in the field of Unmanned Aerial Vehicle (UAV) systems has procured research interests among various communities. One of the primary factors being, thinking beyond the box of what could UAV system bring to the table other than military applications? Evidence to any answer for this question is the current day scenarios. We could see numerous applications of UAV starting from commercial applications of delivering consumer goods to life saving medical applications such as delievery of medical products. Using UAVs in for data transportation in remote locations or locations with no internet is a trivial challenge. In-order to perform the tasks and satisfy the requirement, the UAVs should be equipped with sensors and transmitters. Addition of hardware devices increases the number of connections in hardware design, leading to exposure during flight operation. This research proposes an advanced UAV system enabling wireless data transfer ability and secure data transmission with reduced wiring in comparison to a traditional design of UAV. The applications of this research idea targets using edge computing devices to acquire data in areas where internet connectivity is poor and regions where secured data transmission can be used along with UAV system for secure data transport.

1 Introduction

In current day scenarios, internet is playing a significant role in everyone's daily life. There has been a expeditious technological developments in various fields for betterment of human life. One of such developments is discovering various new modes of data exchange like heterogeneous communication between humans and things compared to a long standing traditional homogeneous communication. Wireless Sensor Networks (WSN) has amplified the applications of embedded systems and Internet of Things (IoT). WSN, a network of randomly dispersed nodes which can have communication with every other node in the network with help of radio signals.

WSN is made up of independent micro-sensor nodes which can sense and communicate the information further to other nodes. A single micro-sensor node can be embedded with a sensor, a microcontroller or a microprocessor, depending upon the target applications [1], [2]. The micro-sensor nodes comes with an advantage of inbuilt power source and a wireless transmitter to enable wireless data transfer facility both online and offline environments. A classic example of WSN is using them to sense atmospheric conditions like temperature and humidity. The usefulness of WSN is tremendous in industrial sectors with hazardous nature where it is difficult to deploy humans for data acquirement.

If the WSN are enhanced with wireless technologies such as ble, thread and zigbee it would enable them to send the data without the need for internet or physical wired connections. These technologies can operate on coin cell battery for years which makes this WSN a low-power architecture. By deploying of edge devices, wireless mesh networks such as ble mesh, thread mesh network would enable a single node to transfer data for longer distance to avoid any failures and ensure no data is lost during transmission [3]. Owing to above factors, the WSN is used in industrial and non industrial applications. At the same time, there has been an increasing demand to use WSN at remote locations to acquire sensitive data but the problem arises how can the data be transmitted for longer distance when it exceeds the limitations of mesh technologies? On the other had, how can we secure the data when it is being transmitted and make it readily available around the globe at the same time? Is it possible to carry out the operation using IoT edge devices?

Edge computing devices helped us to understand and answer

*Corresponding Author: Niranjan Ravi, ECE Department, IUPUI, IN, USA & ravin@iu.edu

the questions. These unanswered questions triggered the motivation to use UAV to transport the acquired data from remote WSN when it is difficult for any human to enter the region. UAVs also has a significant quality of quicker transportation owing to less air traffic.

To ensure a secure data transmission,a ready to use solution, NXP's A71ch, a secure element which can store provision credentials and aid in connecting to private/public clouds [4] was utilized. This process would require new hardware installation in UAV design. The increment of hardware comes with a cost of increased wiring which is addressed by using light-weight protocol called Unmanned Aerial Vehicle Controller Area Network (UAVCAN).

2 Literature Review

In 1999, when term IoT was first coined, it was believed that in future every object would have a globally unique Radio Frequency Identification tag (RFID). IoT provides a new path to explore the data which has been gathered by the embedded systems using its sensors. IoT provides a guiding hand to devices which are resource constrained and would require additional computational and communication capabilities. They help in monitoring the devices at remote locations through the internet in event of any failures. Embedded system devices used in IoT applications are referred as smart objects because they act according to environment changes. But the question of storage arises. How much data can these devices store? Would it be efficient to store the data in the stand-alone device? Is there any alternative?

Cloud platform answers these questions. Cloud computing plays a vital role in enhancing IoT sector. The smart objects are resource-constrained and performing advanced computational would require re-designing the hardware which would result in increased production and operating costs. Also, it would increase the size of the component. Cloud provides alternate approach like data storage, handling and unlimited distributions. This brings up new features like scalability and flexibility for the researchers to deploy the products in various new applications. By uploading the data to cloud complex machine learning algorithms can be deployed to observe patterns in the data on real-time. This cloud infrastructure solves the problem of storage and real-time data visualization. But how secure is the data transmission to the cloud?

Crytographic solutions answers the above questions by offering end-to-end security to minimize data breach. The idea of choosing a right encryption technique in IoT field is widely debated. Because, the data generated by IoT devices are of vast magnitude ranging from a temperature sensor value to medical records which are exchanged with cloud platforms. A pressing need for security measures that should be compatible enough to be deployed in IoT devices is required.

Elliptic Curve Cryptography (ECC), a type of asymmetric cryptography is focused in this research. Asymmetric cryptography knows as public key cryptography algorithm involves a pair of keys: a public key which can be shared/exchanged and a private key, kept as a secret. This provides higher security level at Transport Layer Security (TLS) for devices which are resource constrained, as shown in this research. TLS handshake is for authentication and key exchange in order to establish a secured connection. TLS protocol version and usage of asymmetric encryption algorithm is managed by TLS handshake protocol.

On the other hand, taking advantage of some of the above developments in the field of sensors and electronics, UAV system has evolved its reach into civilian and military applications [5]. They are currently used in surveillance of a location and even launch drone strike on enemy targets. One of the primary reasons for advancement in the field of UAV systems are

1. Environmental safety and protection.
2. Military drone strikes.
3. Drone surveillance during natural disasters.
4. Geological study.

A flight controller module, GPS for navigation purpose, a telemetry device to transmit and receive data from ground control station are used in design and construction of a traditional UAV. It is also equipped with 9 axis sensors to get real time flight dimensions in motion [6]. A lot of new physical interfaces have been introduced in the flight controller design. One such development is introduction of CAN port to reduce the wiring and reduce the noise. A thorough background study on UAVCAN principles was carried out with help of this research [7] where the author has experimented a proof of concept of establishing a one way CAN bus communication between flight controller with an arduino module.

The basic understanding of arduino was helpful but further questions arose like how would this interface work when flight controller is paired with a resource constrained microcontroller? What are the parameters needed to be considered when integrating various operating systems with different clock frequencies and operating speed? How much of the payload(data) can be transmitted for the usage? How many devices can be connected in a single UAVCAN bus? What are the applications it would bring on this successful completion? How to transmit data from a drone system securely to a data centre/cloud networks ? The results of the question are discussed on this research.

3 Background Study

3.1 802.15.4 Standard

This standard defines an effective Medium Access Control (MAC) layer with numerous physical layers (PHY) which can be optional. IEEE 802.15.4 outlines Information elements (IEs) to help in better utilization of the standard. The header IE and payload IE. The header can be useful in authentication and payload can be encrypted. IE is advantageous such as providing backwards-compatibility for future versions of the standard.

3.2 6LowPAN

6LowPAN is abbreviated as IPv6 over Low Power Wireless Personal Networks. As a successor of IPv4, IPv6 has increased the address size of 128 bits as compared to 32 bits. 6LowPAN transmits and receives IPv6 packets over IEEE 802.15.4 link. Enabling the

maximum frame size to be sent transmitted. It is aimed at low power and low cost applications. In places of Ethernet, a packet of Maximum Transmission Unit (MTU) size can be transmitted in a single frame. Between 802.15.4 link layer and IPv6 networking layer, 6LowPAN exists as an intermediate layer. This aids in the process of fragmentation of sender's packet and re-fragmenting it on the receiver node. The transmission load is reduced when fewer bits of data is transmitted. Thread network inherits this feature to transmit data packets. Mesh header supports more efficient methods of compression of messages inside a thread network .

3.3 Constrained Application Protocol

Most of the resources deployed in the field of IoT are resource-constrained in nature. In order to support the data transmission, a specialized web transfer method/protocol called constrained application protocol (CoAP) is deployed. CoAP contains RESTful features to run on any resource constrained embedded devices. CoAP works on User Datagram Protocol (UDP) protocol than Transfer Control Protocol (TCP). A feature of getting response from destination nodes to parent node called Confirmable requests (CON) is also present.

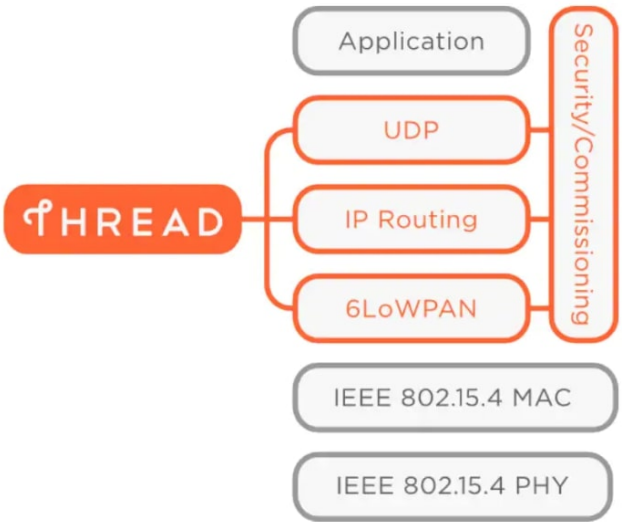


Figure 1: Thread network stack [8]

3.4 Thread Protocol

Thread is an Internet Protocol (IP) based on wireless networking protocol aimed at low powered applications in the field of embedded systems and IoT. Thread is developed based on many standards resulting in providing many advantages like reliability, cost-effective and low power consumption. Thread used UDP on Network layer and uses IEEE 802.15.4 PHY and MAC layers wireless specifications supporting up to 250 devices in a mesh network by operating at 2.4 GHz band as shown in Figure 1. With help of Wi-Fi on their Home Area Network (HAN), users can exchange and communicate with other thread devices within the thread network by using smartphones.

3.4.1 Thread and Border Router

Thread router offer connectivity for the devices inside the network and offers security services to new devices trying to join the network. Border router carries out connectivity between one network and its adjacent. A thread network would consists of at least one border router and it can have many as well depending upon the application. A leader in the thread network would receive requests from end devices to become router. The leader also contains the registry of router IDs in the network. In any instance of the application, if the leader fails, another router is elected to be the leader immediately without any user intervention .

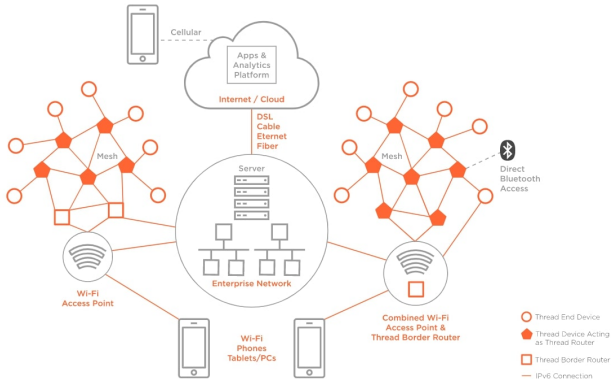


Figure 2: Commercial thread network topology

3.4.2 Router Eligible End Devices (REEDS)

As the name suggest REEDs are devices which primarily acts as end devices. The REEDs are eligible to become active routers. This action is carried out by the network topology without any user intervention. The REEDs can also act as sleepy end devices which has the function to gather data and transmit the information further to the parent node. Sleepy end devices are highly efficient since they are helpful in extending the battery life .

3.4.3 Ensuring No Failure

The thread system has the advantage of working to full efficiency despite network failure or nodes losing connectivity by autonomously re configuring itself in-case of any faults. For instance if a end devices requires a parent node to transmit and the parent node fails due to connectivity problems, then the end device would choose another parent to transmit the data. Thread network topology is shown in Figure 2. A similar procedure is applied in case of failure of border router as well without any user intervention .

3.5 Real-Time Operating Systems (RTOS)

RTOS can be visualized as a software component that allows the system to rapidly switch between various tasks simultaneously on a single processing core and it gives an impression to the user that the system is working on various operations at a same time. Because in reality, the system can run only a single thread and RTOS decides the priority of the task to be handled to execute an efficient operation [9].

3.5.1 FreeRTOS

FreeRTOS is designed to run even on a small microcontroller applications. Microcontroller is an important component of a deeply embedded system and it would require a freeRTOS system to maximize the efficiency during critical applications. The critical features of RTOS are

1. Interrupt latency.
2. Resource-constrained processors.
3. Priority scheduling.
4. Improved efficiency.

3.6 Fight Controller and QGround Control

The flight controller acts like a heart to the drone system. The usage of fibre optics provide considerable advantages to the design. The flight movements are converted into electric signals and are transmitted to ground control station to monitor the flight activity. Pixhawk and Ardupilot are predominantly used flight controllers [10]. QGround control provides complete UI setup and configuration for PX4 software. Real-time screenshot of QGround system is shown Figure 3. Currently, PX4 is predominantly used in developing UAV applications. QGC provides flight map, way points and flight track which are essential to plan a flight mission.

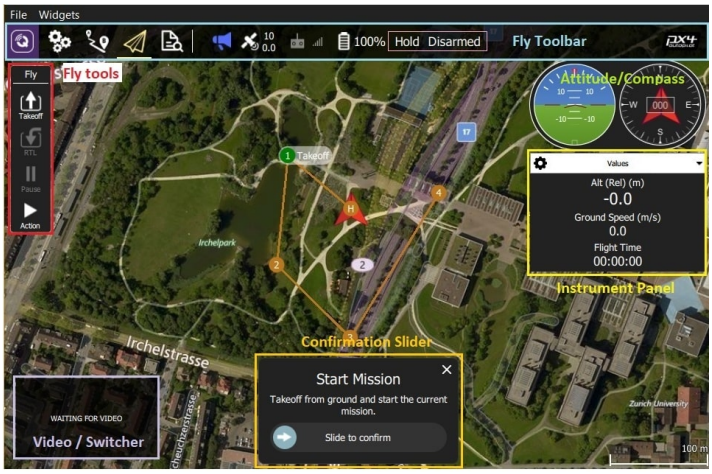


Figure 3: QGroundControl - Ground control station

3.7 Controller Area Network

CAN standard was introduced to harness complex wiring with two wires. It has more immunity to electrical noises and became a popular standard in automobile industries and other sectors. CAN bus works on the principle of multi-master and multi-slave system with message broadcast and maximum signalling rate of about 1 Mbps. The message/payload is Figure 4. ISO specification For CAN Bus broadcasted to all other nodes in the system. Types of CAN,

1. Standard CAN.
2. Extended CAN.

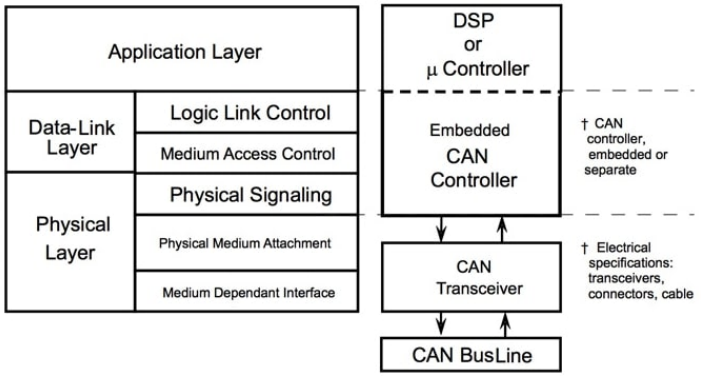


Figure 4: ISO specification for CAN Bus

3.8 UAVCAN

Unmanned Aerial Vehicle Controller Area Network (UAVCAN) protocol was initially designed for robotic and aerospace communications over the robust networks such as CAN bus . The design goals of UAVCAN are

1. Democratic network.
2. Nodes can exchange payloads.
3. Support for redundant interfaces and redundant nodes.
4. High throughput, low latency communication.
5. Simple logic, low computational requirements.
6. Common high-level functions to be defined.

In UAVCAN, each node is equipped with unique numeric identifier known as node ID as shown in Figure 5. 8 bytes of data can be sent in a single frame. In the event of huge payload, payload is divided and transmitted in multi CAN frames. CAN payload structure is displayed in Figure 6.

Message frame

Field name	Priority				Message type ID														Service not message	Source node ID						
CAN ID bits	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Allowed values																						0	1...127						
CAN ID bytes	3					2																1	0						

Figure 5: CAN frame format

Each published message would consists of an unique data type ID and each node in the network would consists of a node ID. This helps in preventing the redundancy in the network. Data Structure Description Language (DSDL) defines data structures format of data which has to be transmitted. Each DSDL file would consists of an unique identifier and a data type name. Few data types in the system are reserved for Vendor specific data types. DSDL helps in optimization which is predominant in deeply embedded systems where dynamic memory allocations may not be permitted. Two methods of data exchange are observed in UAVCAN. They are

- Message broadcasting.
- Service invocation.

Field name	Transfer payload	Start of transfer / End of transfer / Toggle / Transfer ID							
Payload byte	Up to 7 bytes	Tail byte							
Bit position		7	6	5	4	3	2	1	0

Figure 6: CAN payload

Message transfer and service transfer are popularly used modes of transfer. CAN 2.0A has 11 bit identifier and CAN 2.0B has 29 bit identifiers. The payload consists of 7 bytes followed by a single tail byte to indicate start of frame, end of frame, toggle and transfer ID.

3.9 Security Concepts

IoT devices should follow as secured design approach which would enable them for a safe storage of credentials, device authentication and secured data exchange with remote servers. Design criteria of IoT devices is crucial to aid in a safe storage of credentials, device authentication and secured data exchange with remote servers.

RSA (River, Shamir and Adleman) and ECC are recent developments in asymmetric encryption algorithms. ECC is based on algebraic structures of elliptic curves over a finite field which generate a certain public and a private key pair. A digital signature is used in guaranteeing the integrity and authenticity of the device. Digital certificates are like electronic documents which contain details like sender public key, identity of the owner and signature from a trusted entity, Certification Authority (CA).They are used in proving authenticity of the shared public keys.

Elliptic Curve Digital Signature Algorithm (ECDSA) algorithm utilizes ECC to authenticate the devices. The key pair (public and private key generated from ECC) is used in signing/verification a message between sender and receiver. A signed message using the private key and a digital certificate is transmitted from sender side. The receiver device retrieves sender public key from digital certificate to validate the signature of the received message.

4 Hardware And Software Used

1. NXPs rapid IoT prototyping kit.
2. NXP's kinetic platform kits.
3. Pixhawk/PX4.
4. UAV system.
5. NXPs A71ch - Plug and trust for secure IoT applications.
6. Wireless sensor node.
7. IBM Bluemix - IBM cloud platform.
8. MCUXpresso IDE.

4.1 Rapid IoT Prototyping Kit

It is a power optimized IoT solution for rapid prototyping of ideas into POC. It includes pre-programmed applications that enables users to have a quick understanding about hardware and software architectures. It uses FreeRTOS. Web based IDE and GUI tools help in prototyping without any help from programming languages. Figure 7 displays rapid IoT prototyping steps. Wireless protocols such as ble, thread are supported in which thread protocol was used in this application to create a mesh network to transmit data from a remote end to master/parent node. It also provides A1006 authentication which is offered as a turnkey solution for developers to integrate security solutions into the host platforms. The complete authentication time is lesser than 50 ms. It uses FlexCAN architecture [11].



Figure 7: Rapid IoT - Prototyping

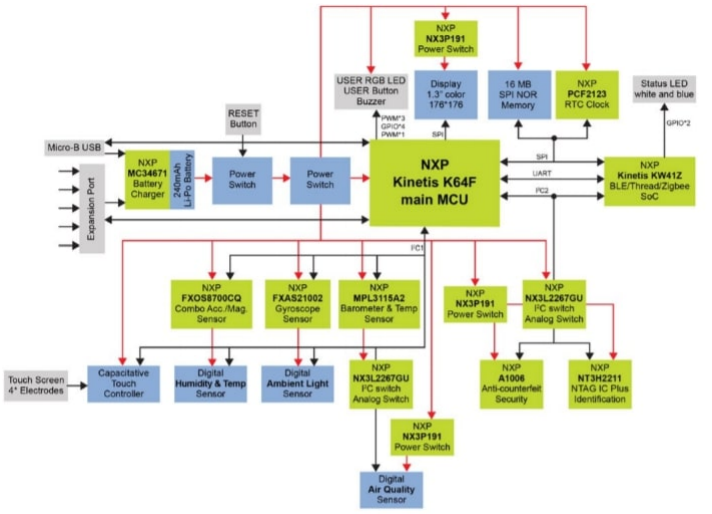


Figure 8: Rapid IoT architecture

The features of rapid IoT utilized in this research are

1. ARM Cortex M4-Core for application processing, Figure 8.
2. Thread connection establishment to a gateway inside the UAV system.
3. A1006 authentication for enhanced security features.

4. Software enablement including RTOS drivers, middleware and cloud platforms.

5. System enablement with advanced sensors like temperature, humidity, light sensors and digital air quality. Temperature sensor was utilized to get real time temperature values to cloud.

Temperature sensors values generated from rapid IoT at predefined intervals are utilized in this research testing.

4.2 Pixhawk/PX4

Pixhawk works on light-weight and energy efficient real-time operating system called NuttX. In order to access Ground Control Station, a command line interface called NuttShell (nsh) is utilized. It also provides micro Object Request Broker (uORB). uORB is used for asynchronous mode of communication to switch between tasks. Data transfer is done by using a simple publish-subscribe pattern. The transmitted and received data is visualized using ground control station. There are multiple feature advancement in recent years and key features utilized in this research are

1. 32bit STM32F427 Cortex M7.

2. 2 MB of memory.

3. 512 KB RAM.

4. ST Micro L3GD20H - 16 bit gyroscope, ST Micro LSM303D - 14 bit accelerometer / magnetometer.

5. Interfaces such as CAN, SPI, UART, I2C, ADC, PWM.

4.3 A71CH - Plug and Trust for secure IoT applications

The data generated by IoT devices is increasing everyday and ways to secure the data is a popular question among research communities nowadays.

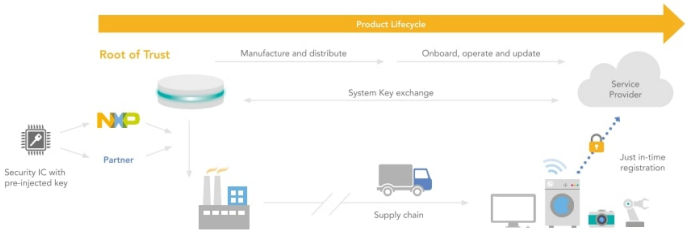


Figure 9: A71ch - Root Of trust

A71ch is a ready-to use solution since it contains necessary private and public keys for the device which enables the device to securely connect to private/public clouds. Figure 9 shows A71ch product features. This security IC can be easily interfaced with various hardware and operating systems. This security IC comes in two product variants:

1. Customer programmable Type.

2. Provisioned and programmable type.

The first type is delivered without any credentials and used in designing new products, to test and evaluate the designs. The second type is ready-to-use IC which is already provisioned by NXP trust provisioning support to ensure a secured TLS connection with IBM Bluemix. A71ch is interfaced to Micro-controller Unit (MCU) using Inter-Integrated Circuit (I2C) protocol. A71ch provides root of trust at the IC level. It delivers proven, chip to the cloud security so connection can be established with cloud platforms such as IBM, AWS and google cloud without the need to write security keys or exposing certificates/credentials. But when the data transmission occurs for the first time, proof of origin of the device and anti-counterfeit protection are required to protect server and device authenticity and avoiding un-trusted servers attacking the system. The complete authentication procedure is composed of below steps:

1. Server certificate verification.

2. Server authenticity verification.

1. IoT device certificate verification.

2. IoT device authenticity verification.

A71ch is used for anti-counterfeit protection of the device against many physical and logical attacks by operating autonomously based on integration of Javacard operating systems and applet firmware. Physical attacks on the system is narrowed down by allowing direct memory access only by fixed functionalities of applet. This isolates the content of memory from host system. The server stores a unique key pair and a digital certificate signed by trusted CA. On the other hand, the IoT device would also have a key pair and a signed digital certificate. The digital certificate's are issued by same trusted CA. These credentials are stored inside the A71ch chip. Server's certificate and authenticity is verified by IoT device using the CA certificate and a random challenge as a sequential process. The challenge involving generating and sending random message to validate the server. The same steps are followed to verify the IoT device's certificate and authenticity.It can be deployed across a broad range of systems to establish a secure connection. Since the communication between A71ch and the Host is based on I2C protocol, it can be carried out through a Secure Channel Protocol (SCP03). This would protect the channel against external attacks such as replay attacks. Also ensuring only the authenticated MCU can exchange information with our security IC by use of session keys. The features of A71ch include

1. Systematic enforced authentication.

2. ECDSA.

3. An unique chip ID.

4. Freezing of credentials.

5. Possibility to lock A71ch as transport lock mechanism.

6. HMAC SHA256 calculation in one shot.

The applications of A71ch include home gateways, home appliances, sensor networks gathering confidential data, connected industrial devices , Figure 10.

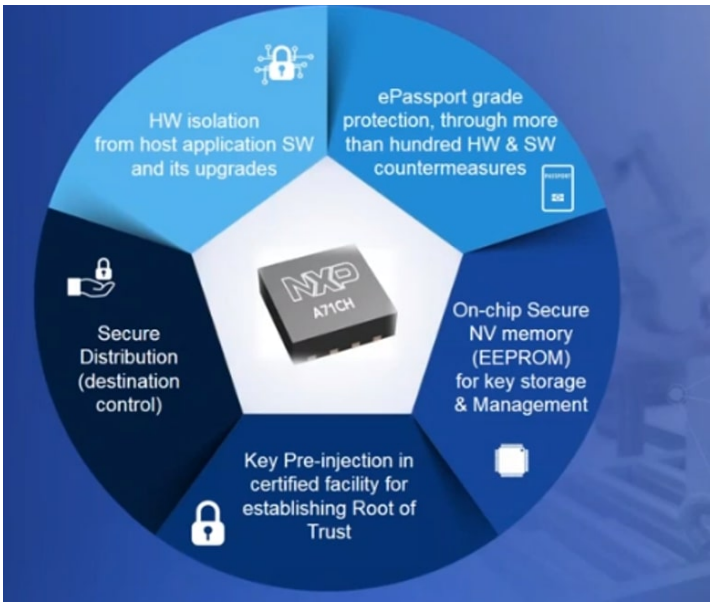


Figure 10: A71ch - Plug and trust for secure IoT [4]

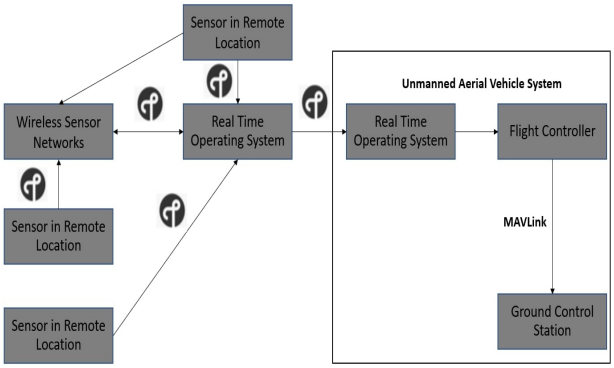


Figure 11: Data transfer using thread protocol from WSN to UAV system

5 Proposed Model and Implementation

This sections explains the initial design phase construction and the modifications in hardware platforms applied for this research work.

5.1 Design Phase

With the help of hardware and software information detailed in section 4, the design phase includes construction of an UAV system using Pixhawk board as flight controller. The interfaces in Pixhawk are helpful in connecting external peripherals like motor, telemetry and GPS module to the UAV system. In order to make the UAV system to adapt for this research, the system was modified by utilizing the CAN bus interface present in flight controller. Two new systems such as Rapid IoT and A71ch are connected to flight controller module.

In order to facilitate the data exchange and visualization aspect, inside the PX4 firmware, two new UAVCAN messages were created each having unique .c and .h files structures. One message to receive data from external input sources and other message to send the information to external devices using a simple uORB pub-sub methodology. In order to encrypt and decode the CAN frame, UAVCAN GUI tool was adopted in testing and result phases.

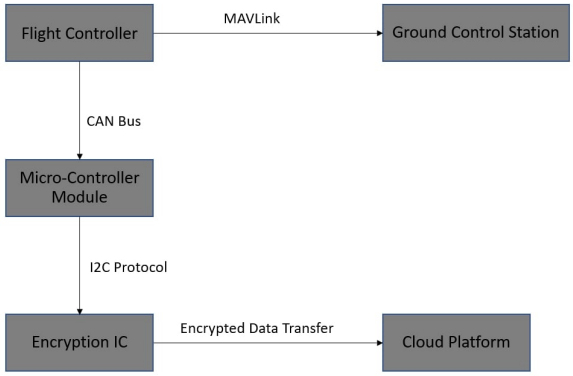


Figure 12: Data transfer to cloud platforms from UAV system

5.2 Integration of Rapid IoT with Pixhawk

Owing to advantages of rapid IoT system, documented in section 4.1, it is embedded with the UAV design using UAVCAN (CAN bus interface) and with help of a docking station. External CAN transreceivers are attached to the docking station. They convert CAN TX and RX signals into CAN High and Low signals which in-turn will be used for transmission. NXP's FRDM KW41Z is assigned to be remote wireless sensor node and it is placed at a distance of 70 - 100 metre. Both the Rapid IoT and remote node are pre-programmed with thread for this research. The remote node would gather atmospheric temperature value for every second. Any sensor data/ raw data can be used here. Once the UAV system is armed, the data from remote node is transmitted wirelessly using thread to rapid IoT in the UAV system, Figure 11.

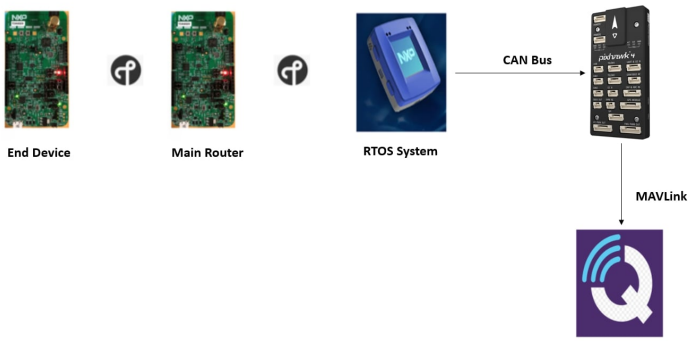


Figure 13: Integration of rapid IoT with pixhawk

The rapid IoT would in-turn send the data as CAN signals to the flight module/controller. The flight controller would re-transmit the data to ground Control station using telemetry and readings are visualized in QGroundControl. This testing phase was carried out at initial step. The design setup was increased by increasing the number of wireless sensor nodes and establishing a mesh network connectivity among them. This enabled UAV system to gather data from the nodes which are about 350 metres. Figure 13 shows the hardware design of UAV system with rapid IoT.

5.3 Establishing a secure data transmission

Using telemetry, the data from flight controller is sent to ground station as radio signals. In order to secure the data, A71ch is connected to Pixhawk using CAN bus protocol. Provisioned and programmable type of A71ch was used in this research. But A71ch is a security IC and it has to be used along with a base board like NXP's LPC54018 which has an inbuilt WiFi module. Establishing a secure connection between A71ch and IBM cloud require three steps. In the first step, A71ch is interfaced with host MCU using I2C protocol. Using the host MCU's debugger port, keys and certificates are inserted into A71ch. Device ID (Identification number - unique number) and generated CA certificate from the device will be utilized in next step. Later, a new user account was created in IBM Bluemix with an organization ID (name of your organization) and registration is completed by entering the device ID and uploading CA certificate from previous step. This portion of initial setup is important to establish a secured connection between cloud and security IC. The final step would be to acquire API key, authentication token from cloud and use it in our code base to initiate the connection from our IC. If the API key is lost or keys are refreshed, the connection setup will not succeed. The above process has to be repeated again in this case. Every provisioned A71ch has to carry-out this one step procedure to function without any failures. The generated data from flight controller is transmitted as CAN frames to LPC module, where security IC receives the data through I2C communication. It would encrypt the data and upload it to cloud as shown in Figure 12. This process would secure the data to prevent data leakage and restrict unwanted access to the data. Hardware design is shown in Figure 14.

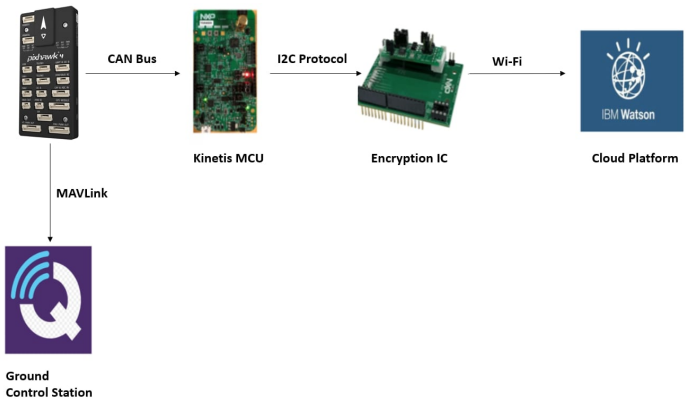


Figure 14: Integration of security IC with pixhawk

6 Testing Phase And Deployment Design

At first, the hardware connections were tested and telemetry was checked if a proper communication channel has been established with the UAV system from the ground control station. The CAN communication was tested between NXP's Kinetis boards using Tera term.



Figure 15: Deployment system of UAV to acquire data from WSN

Later, the interfaced connections were also tested. In various phases of testing, an oscilloscope was used to check for CAN frames. As a next step, UAVCAN functionality in Pixhawk was tested in the lab environment. To ensure whether the CAN frames are being transmitted from the hardware. A publish-subscribe pattern is tested initially to understand the uORB functionality. The status of UAVCAN frames transmitted is witnessed with the help of UAVCAN frame status command in nsh shell.

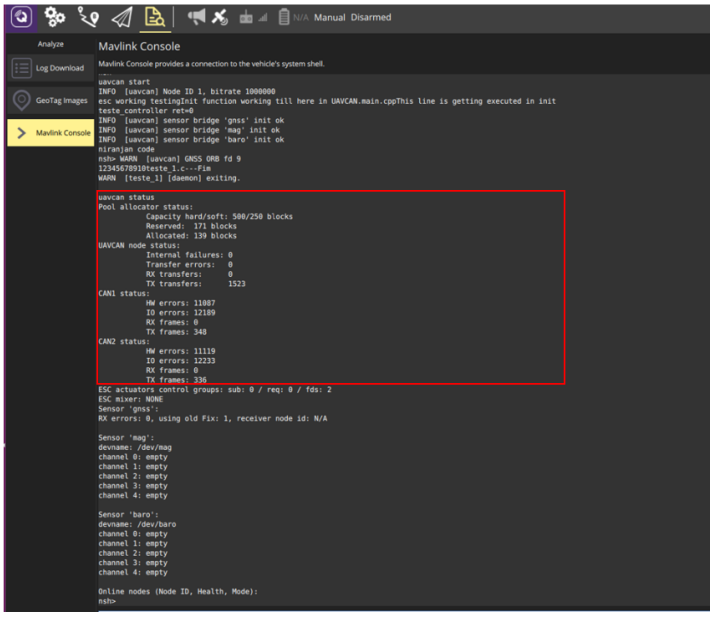


Figure 16: QGroundControl - UAVCAN frame status

An alternative way of testing and visualizing the CAN frames was UAVCAN GUI tool. It shows the sender ID of the CAN frame and the payload. The other alternative was to store the log data and understand the payload and its format during any transmission.

After the testing, two advanced design were implemented using above specified hardware designs.



Figure 17: Deployment design of UAV to transmit data to IBM Bluemix

The designs are shown separate for a better understanding, Figure 15, 17. They can be integrated together and deployed as a single design as well. It had also been tested in this research.

7 Results And Demonstration

Extending the range of data transmission using wireless technologies like thread and incorporating security IC to secure the data transport is tested with the help of NXPhlite,Pixhawk, NXPs rapid IoT and NXP's security IC in real-time environment is tested. Before performing an actual field testing, an initial testing was carried out in two steps. The first step is to ensure the connectivity of hardware to check if there are any wiring issues and second step ensuring all the network devices were working perfectly.

As a first step, Pixhawk module was interfaced with four devices using a CAN network. 8 bytes of data for each CAN frame was transmitted and it was received successfully by other devices as well. This data transmission rate of 8 bytes/second is significantly higher compared to traditional UART (Universal Asynchronous Receiver/Transmitter). The successful completion of this step ensured a fully functional CAN network was established between all the hardware with varied CAN protocols.

In the second step, thread network connectivity was tested between the network devices and connection establishment to IBM cloud was also verified. As mentioned in section 5.3, the unique device ID of A71ch IC is cross-checked with ID displayed in IBM cloud, Figure 19. After successful completion, field testing was carried out.

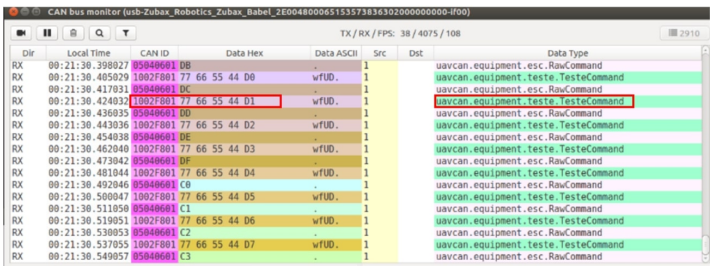


Figure 18: CAN frames - UAVCAN GUI tool

As proposed in section 5.2, with the help of QGround control, Figure 16 shows status of CAN frames both received/transmitted. Standard CAN of payload 8 bytes is tested in this practice. UAVCAN GUI tool displays the values displaying the CAN frames, 8 individual bytes which are being sent from rapid IoT system and which are received by the flight controller, Figure 18.

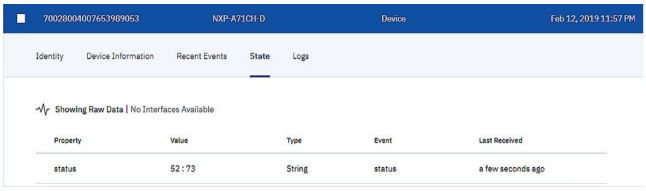


Figure 19: IBM Bluemix - Output data visualization

To establish a secure data transmission, as proposed in section 5.3, the generated values from the flight controller is transmitted to security module which then encrypts and shows the value in IBM Bluemix for each second/whenever the data is published. Figure 19, shows the device ID of the security IC and value column displays the actual value in a string datatype. This is give access across the globe to see the results instantly regarding the flight data. Figure 19 displays instant data transfer into the IBM cloud.

8 Design Challenges

Various design challenges were faced during the research. NuttX operating system had very little documentation with regards to UAVCAN. GitHub repository of PX4 is also being upgraded from time to time but the documentation did not reflect all the changes and new features. UAVCAN is a developing research area and it is difficult for developers to re-design/upgrade the existing design. Only few articles/documentations were available on how to interface flight controller with external hardware devices with UAVCAN [12]. Next challenge was, NXP's rapid IoT used FlexCAN and Pixhawk had UAVCAN. The CAN frame structures were different and without suitable debugging tool, it was difficult to proceed further. To facilitate communication in CAN network, external CAN transmitters were used.

Another challenge was to place the new hardware in UAV system. 3D printed models were very helpful in overcoming above issue. With regards to security IC, understanding the background of security concepts was mandatory. Initial setup and key exchange between IBM cloud and security IC had to be carried out in sequential steps. Any error in the steps will lead to failure and start the process all over again.

9 Conclusion

Technological development and growth in various sectors of life plays a crucial role in making human lives better each day. Internet, a recent trend acts as a stimuli to the process. But there are places in globe, where technological advancements are not fully available and such places remains isolated from the rest of the world. The ideal aim of the research is to provide connectivity/access to areas where

internet connectivity is not good and the places where it is difficult to reach because of terrain. Also, ensuring secure data transport in regions where limited internet connectivity is available.

To achieve that, the design system of an UAV system was proposed. The background knowledge of software technologies and hardware components were detailed in sections 3 and 4. The system was equipped with remote wireless devices which are connected with thread network, the interfacing of flight controller module with other MCUs were carried out using UAVCAN, security of data transmission during over the air transfer (OTA) was addressed with the help of security IC, A71ch. Using the proposed UAV system with wireless data transfer facility, would foster growth quick data gathering and transmission compared to traditional methods. By combining these two ideas, it would result in building an Eco-friendly and a safe environment all over the globe.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgments We would like to thank NXP, UAVCAN forums for guidance and supporting in completion of this research.

References

[1] L. Mainetti, L. Patrono, A. Vilei, "Evolution of wireless sensor networks towards the Internet of Things: A survey," 2011 International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2011, **16**(8), 16–21, 2011.

[2] C. Rotariu, R. G. Bozomitu, V. Cehan, A. Pasarica, H. Costin, "A wireless sensor network for remote monitoring of bioimpedance," Proceedings of the International Spring Seminar on Electronics Technology, **2015-Septe**, 487–490, 2015, doi:10.1109/ISSE.2015.7248046.

[3] R. Chitanvis, N. Ravi, T. Zantye, M. El-Sharkawy, "Collision avoidance and Drone surveillance using Thread protocol in V2V and V2I communications," Proceedings of the IEEE National Aerospace Electronics Conference, NAECON, **2019-July**, 406–411, 2019, doi:10.1109/NAECON46414.2019.9058170.

[4] N. Ravi, M. El-Sharkawy, "Integration of UAVs with Real Time Operating Systems using UAVCAN and Establishing a Secure Data Transmission," Ph. D Thesis, IUPUI, 2019.

[5] N. Ravi, R. Chitanvis, M. El-Sharkawy, "Applications of Drones using Wireless Sensor Networks," Proceedings of the IEEE National Aerospace Electronics Conference, NAECON, **2019-July**, 513–518, 2019, doi:10.1109/NAECON46414.2019.9057846.

[6] S. R. Vangimalla, M. El-Sharkawy, "Remote wireless sensor network range extension using UAVs with thread protocol," Proceedings - 2018 International Conference on Computational Science and Computational Intelligence, CSCI 2018, 902–906, 2018, doi:10.1109/CSCI46756.2018.00178.

[7] P. Bernardo, "Development of technology and procedures for health monitoring of UAV subsystems Examination Committee," M.S Thesis, Instituto Superior Técnico, (November), 2015.

[8] L. Zimmermann, N. Mars, M. Schappacher, A. Sikora, "Development of Thread-compatible Open Source Stack," Journal of Physics: Conference Series, **870**(1), 2017, doi:10.1088/1742-6596/870/1/012001.

[9] A. Gerstlauer, H. Yu, D. D. Gajski, "RTOS modeling for system level design," Design, Automation, and Test in Europe: The Most Influential Papers of 10 Years Date, 47–58, 2008, doi:10.1007/978-1-4020-6488-3_4.

[10] L. Meier, P. Tanskanen, F. Fraundorfer, M. Pollefeys, "PIXHAWK: A system for autonomous flight using onboard computer vision," Proceedings - IEEE International Conference on Robotics and Automation, 2992–2997, 2011, doi:10.1109/ICRA.2011.5980229.

[11] J. Pimentel, J. A. Fonseca, "FlexCAN: A Flexible Architecture for highly dependable embedded applications," Int. Workshop on Real-Time Networks, **1**(3), 2004.

[12] N. Ravi, M. El-Sharkawy, "Integration of UAVs with Real Time Operating Systems using UAVCAN," in 2019 IEEE 10th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON), 0600–0605, 2019, doi:10.1109/UEMCON47517.2019.8993011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1199-1207 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Performance Evaluation Reprogrammable Hybrid Fiber-Wireless Router Testbed for Educational Module

Muhammad Haqeem bin Mohd Nasir[1], Wan Siti Halimatul Munirah binti Wan Ahmad[1], Nurul Asyikin binti Mohamed Radzi[*,1,2], Fairuz Abdullah[1,2]

[1]*Institute of Power Engineering (IPE), Universiti Tenaga Nasional, Kajang Selangor, 43000, Malaysia*

[2]*Department of Electrical and Electronic Engineering, College of Engineering, Universiti Tenaga Nasional, Kajang Selangor, 43000, Malaysia*

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 02 February, 2021
Online: 25 February, 2021

Keywords:
Raspberry Pi
Fiber-Wireless Network
FiWi Router
Educational Module

ABSTRACT

Fiber-Wireless (FiWi) network is an integration of fiber optic and wireless connections in the same network. It is one of the best solutions to overcome rapid increment of Internet users and bandwidth-hungry services. To facilitate fundamental knowledge and further understanding on FiWi for students and researchers at the university level, this article proposes the development of a fast integration and scalable FiWi router testbed using Raspberry Pis as the embedded system-based hardware for lab-scale experiments. The performance of the router testbed in terms of end-to-end delay and throughput for upstream and downstream are evaluated. The delay values comply with IEEE 802.15.4 routing scheme. The performance of the router testbed is compared with the industrial grade off-the-shelf router in terms of throughput for each network. A testbed stress test is conducted by sending two data traffics simultaneously, and the performance test is repeated for Wireless-Fiber-Wireless and Fiber-Wireless-Fiber network architecture. The results show the proposed router testbed is scalable, flexible, and capable of fast integration.

1. Introduction

This paper is an extension of work originally presented in 2020 IEEE 8th International Conference on Photonics (ICP) [1]. However, in that work, fiber connection alone would not be able to accommodate all connections, especially to mobile and smart devices that are rapidly increasing. Furthermore, the demand for Internet and leased line bandwidth are growing continuously at more than 20% per year due to more video streaming, cloud computing, social media and mobile data delivery [2]. By 2020, the bandwidth demand will continue to grow due to an enhancement of video quality such as 8K Ultra High Definition (UHD) and increasing number of user subscriptions. Because of that, the estimated traffic in terms of mobile data and fixed systems will be 2600 times more than the traffic in 2010 [2]. This growth is accelerated by the new type of communication services such as device-to-device (D2D) and machine-to-machine (M2M) [3]. Therefore, to ensure users experience are at the same level of Quality of Service (QoS) regardless the time and location while the demand is increasing, Fiber-Wireless (FiWi) deployment has become a necessity because it can cover a large area and support large number of users. According to [4], FiWi is still an ongoing study and there are plenty of rooms for improvement.

It is important to equip fundamental knowledge and further understanding on FiWi for students and researchers at the university level. Therefore, to facilitate this, it is essential to have a scalable FiWi testbed with fast integration capability, which is proposed in this article. FiWi network is a combination of fiber and wireless connection in the same network [5, 6] to provide fixed and mobile services [7-11]. A typical FiWi network consists of high-capacity passive optical network (PON) as backhaul comprising of Optical Line Terminal (OLT) and multiple Optical Network Units (ONU), connected to a cluster of wireless front end routers [12]. ONU is integrated with a wireless gateway that provides wide-area connectivity to users. At the optical backhaul, the OLT at the central office (CO) forms a root connected to the optical backbone via fiber link to provide cloud computing services. ONU is connected to the OLT via 1 : N (N = 32, 64) splitter to form a leaf-shaped nodes. The CO is responsible to manage the information

*Corresponding Author: Nurul Asyikin Mohamed Radzi, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia, +603-89212020 ext. 3286, email: asyikin@uniten.edu.my

https://dx.doi.org/10.25046/aj0601136

transmission between mobile devices with ONUs and acts as a gateway to other networks [13]. In the wireless end of FiWi network, mobile devices such as smart phones [14, 15], virtual reality (VR) glasses [16], smart watches and other Internet of Things (IoT) devices [16-18] have access to the Internet either via ONU or multi-hop wireless mesh network. Based on a study by [19], one of the advantages of FiWi is that it provides high speed optical backhaul for a wireless mesh network. In [20] stated that the network has a broadband access where it uses wide range of frequencies enabling a large number of data to communicate simultaneously. Whereas, mobile backhaul combines reliability and capacity of Ethernet-based optical backhaul with the wide range of coverage and flexibility of Ethernet-based wireless devices. By utilizing the efficacy of both optical fiber and wireless, a fast speed and low cost of service areas can be achieved [21]. It provides cost effectiveness, robustness, flexibility, high capacity, reliability and self-organization compared to fiber or wireless connection alone [22-24] Another advantage of having FiWi network is the combination of fiber optic and radio access technologies in multi-tier Radio Access Network (RAN) [2].

Radio access technologies will be used to deliver wireless services with high capacity, high link speed, and low latency [25]. The multi-tier RAN will improve the cell edge performance for mobile fronthaul and backhaul, resource sharing, and centralization of multiple standards with different frequency bands and modulation formats. FiWi network is a promising technology to support high demand of bandwidth and large number of users in different types of topologies and geographical environments.

Over the years, research communities have been working hard to improve the performance of FiWi network. Hence, in order to study and perform an extensive performance evaluation, testbeds are needed. Testbed is known as a prototype for proof-of-concept of any technology features where further studies can still be conducted [26]. There are various types of testbed in various platforms such as lab-scale testbed and field testbed. According to [27], built in lab-scale testbed is portable, which means, it can easily be packed away, securely stored and safely transported. The testbed can also be reused and assembled by other researchers in the future. For field testbeds, they consist of high-end setup which require installation by trained professionals and usually require large spaces. For example, in [28], an experiment was conducted consisting of 64 massive multiple-input multiple-output (MIMO) antennas for the FiWi testbed setup. The setup is comparing with collocated massive MIMO to simultaneously serves many users using the same time-frequency resource. For educational module, a simpler setup is preferable because fast integration capability is an important criterion for an educational testbed. This will speed up the installation of the testbed and the lecturers can have more time on explaining rather than focusing on testbed installation.

A testbed can have flexible architecture to be connected to any devices such as multiple sensors [29]. Despite of its compactness, it has powerful local computation unit. A testbed can also be reconfigured to different topologies as required. A review on two types of router testbeds (commercial-based and embedded system-based) were presented in [30]. The commercial-based are store-bought off-the-shelf routers [31, 32]. Some examples of embedded system-based routers are using netFPGA [33], Banana Pi [34] and Raspberry Pi [35-37]. From the study in [30], the embedded system-based routers are reconfigurable, reprogrammable, scalable and cost less than the commercial routers. These advantages will be a suitable choice for the lab-scale router testbed in research and academic field.

Developing a lab-scale testbed is also one of the practical ways of studying and experimenting a particular technology before an actual implementation due to its portability and flexibility. This article describes the development of a FiWi router testbed with the said advantages, with several additional features: fast integration capability, scalable and flexible. The proposed router testbed uses Raspberry Pi for its programmable function (coded in Python language) to be a router. The details of the router structure will be described in the following section.

2. Hybrid Fiber-Wireless Router Testbed

Figure 1 shows a flowchart that reflects the overview of technical development for the proposed router testbed. Firstly, after the proposed routers in the testbed are connected with one another in FiWi architecture, connection reliability test is done by sending a packet from one client to another client. If the packet is not received by the client, an additional processing delay is increased in the proposed router. This delay is an additional time for the router to process the packets including label injecting, label checking, and packet forwarding. The process is repeated until the clients are able to receive the packet without fail. Afterwards, the performance of the proposed router is tested in FiWi architecture in terms of throughput and end-to-end delay in order to validate its correctness and comply with the current communication standards. After it is validated, a stress test is done by sending two traffics at a time to test the stability of the proposed router. Then, the architecture is changed to Fi-WiFi and Wi-FiWi in order to test its flexibility and scalability.

The proposed FiWi router testbed is using Raspberry Pi 3B+ to support data transmission for research and education purposes. It consists of four routers that are connected in tree topology and may be varied accordingly, as shown in Figure 2. It is a typical topology for FiWi network [12, 13]. Each router is an integration of four Raspberry Pi 3B+; one Header Pi and three Forwarding Pis. The three Forwarding Pis are to emulate the real router, with each of them having a unique static IP address. To form a router, the Raspberry Pis will be connected to two Ethernet switches via CAT5e Ethernet cable, as illustrated in Figure 3. Ethernet Switch 1 is to represent the internal connection of the routers (between Header Pi and Forwarding Pis). While Ethernet Switch 2 representing the external connection between routers. An Access Point (AP) is added as an element to the router structure to support wireless transmission.

2.1. Topology

The router testbed is set up in tree topology consisting of four proposed routers for the default FiWi architecture, as shown in Figure 2. The connection between routers is using single mode fiber optic patch cord. Since Raspberry Pi has only one Ethernet port, a Fiber Media Converter (FMC) is needed as an adapter. Then, the architecture is expanded to Fi-WiFi and Wi-FiWi as shown in Figure 4 and Figure 5, respectively. Compared to other topologies such as ring and mesh, the flexibility and stability of the proposed router can be observed after going through a number of

medium changes from wireless to fiber and vice versa in tree topology. Moreover, the purpose of these setups is to test the scalability of the testbed.

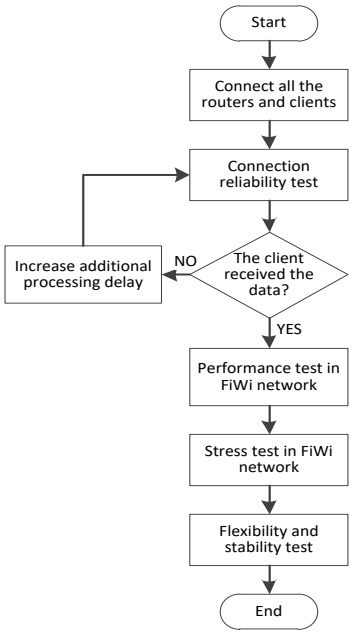


Figure 1: The overall flowchart of this work

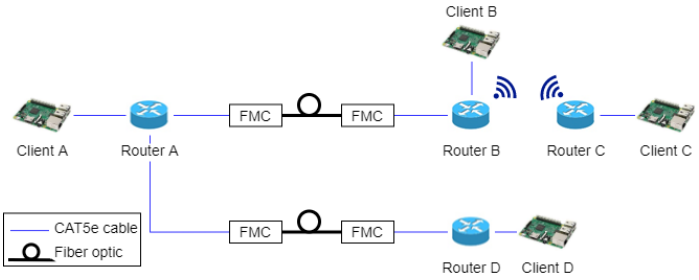


Figure 2: The architecture of FiWi router testbed

The architecture of this testbed can be scaled easily to Fi-WiFi and Wi-FiWi without any tedious hardware rearrangement. The routing mechanism for FiWi, Fi-WiFi and Wi-FiWi can be achieved by changing the label for each client, Header Pi and Forwarding Pi in the router program code. The labels for routers and clients in any of the three architectures (FiWi, Fi-WiFi and Wi-FiWi) must be unique to prevent misinterpretation of data destination by the router.

2.2. Routing Mechanism

The flowcharts of the testbed routing mechanism are shown in Figure 6, 7 and 8, for the client, Header Pi and Forwarding Pi, respectively. When a client sends a packet of data to the destination, a pre-label is added to the payload of the packet to indicate where the data should end. The client sends the data to Header Pi. Then, the Header Pi checks the label to know the beginning (source) and end (destination) of the data. After the Header Pi has identified the destination of the data, it replaces the old label with a new label onto the payload and then broadcast the data to each Forwarding Pis. Each of the Forwarding Pi has its own unique label because each identity label corresponds to one destination only. When the Forwarding Pi receives data from Header Pi, it will check the label on the payload of the packet with its own identity label. If both labels are the same, the Forwarding Pi will continue to forward the data to desired destination set by the client. Otherwise, if both labels are not the same, the Forwarding Pi turns the data to zero and drops the packet.

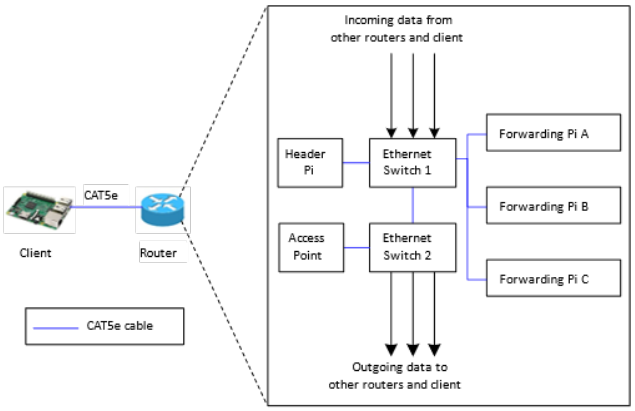


Figure 3: The structure of proposed router

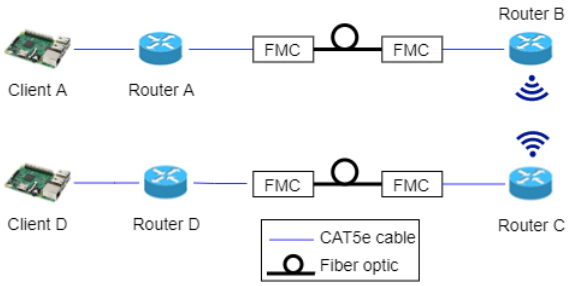


Figure 4: The architecture for Fi-WiFi router testbed

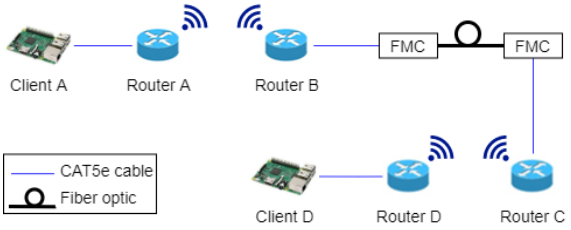


Figure 5: The architecture for Wi-FiWi router testbed

2.3. Experimental Setup

2.3.1. FiWi Router Testbed Performance Test

For this test, Raspberry Pi client is used to send a packet to another client via a combination of fiber and wireless transmission as in Figure 2. The initial data size transmitted is at 100 bytes for proof of concept [38]. Then, the data size is increased by 100 bytes each time until 1445 bytes. The performance of the experiment in terms of end-to-end delay and throughput are plotted on the graph. The proposed router is compared with industrial grade router to check its functionality and to observe its correctness. The setup for the industrial grade router is using the same topology as the proposed testbed. The upstream and downstream transmissions are done by using two computers. The throughput of the proposed router testbed is validated with industrial grade router. The bandwidth of the industrial grade router is set to 1 Gbps. However,

the initial data size is set to 10 kB and increased by 10 kB for each transmission until 100 kB. This is because the industrial grade router is connected to an access point with similar specification to the proposed testbed setup, limiting the actual bandwidth to 1 Mbps. Then, the throughput of the experiment for upstream and downstream are captured using Wireshark.

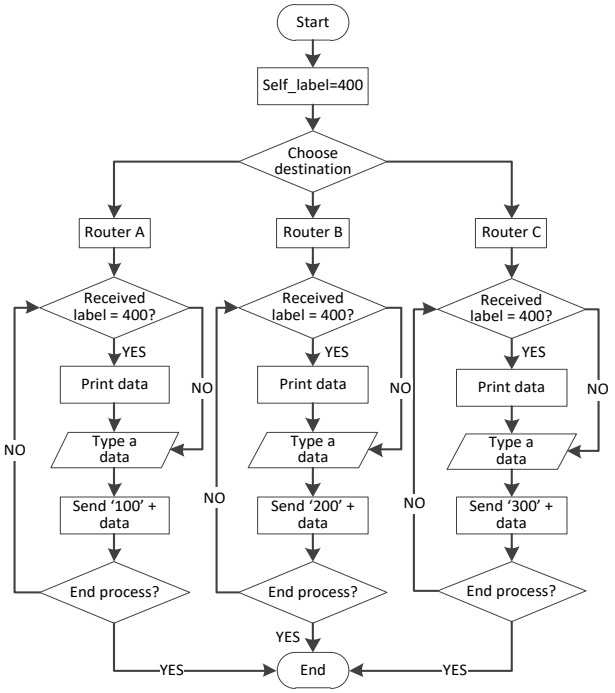


Figure 6: Client flowchart

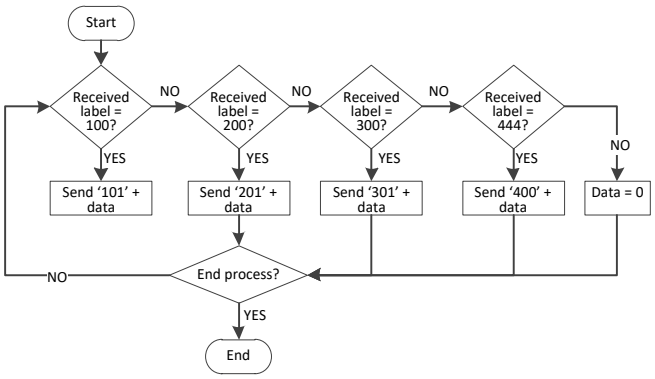


Figure 7: Header Pi flowchart

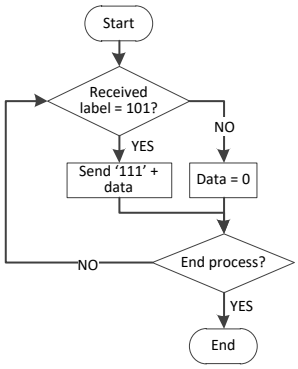


Figure 8: Forwarding Pi flowchart (label example for A)

2.3.2. FiWi Router Testbed Stress Test

The limit of the FiWi router testbed is tested by sending two traffics simultaneously from one router to another. This stress test is done by connecting two clients for each proposed router. The performance will be analysed in terms of the end-to-end delay and throughput for downstream and upstream for two simultaneous traffic transmission. The purpose of this test is to evaluate the stability of the proposed testbed.

2.3.3. Scalability Test

Scalability is defined as the ability for the testbed to adapt to a new architecture or arrangement easily and its ability to expand the number of new components. In order to add or remove a component in the testbed, minor changes need to be done to its arrangement. In this case, the testbed architecture is changed to Fi-WiFi and Wi-FiWi to test its scalability performance and to prove the testbed is flexible to work in various architecture. The scalability test presents the performance for Fi-WiFi and Wi-FiWi in terms of end-to-end delay and throughput. Hence, the purpose of this section is to analyse the scalability performance and to compare which architecture is better for the testbed.

2.4. Parameters

2.4.1. Design

Design parameters are defined as the input for the testbed; the data size and the end-to-end delay. The data size is the length of data in bytes (B). It can be varied easily by the user at the client side. In this experiment, the increment of the transmitted data size is 100 bytes for each transmission. The summary of data size and corresponding throughput scale used in the experiment are listed in Table 1. The data size affects the end-to-end delay where greater data size may result in greater end-to-end delay. The end-to-end delay is used to study the throughput of the testbed.

Table 1: Summary of data size and throughput scale used in the testbed experiment

	Start	Data size increment	End	Throughput scale
Proposed testbed	100 bytes	100 bytes	1445 bytes	linearly
Industrial grade router	10 kB	10 kB	100 kB	1

2.4.2. Performance

Performance parameters are the output of the testbed which indicate its performance. Two performance parameters are taken into consideration; end-to-end delay and throughput. End-to-end delay is the time taken when the user sends the data from a client to another client. Throughput is the overall performance of the testbed in bit per second (bps) when the data size in bit (b) is divided by end-to-end delay, L, in second (s). The throughput, Th_{TL} calculation is summarized in (1) [39], where D_T is data transmitted and D_L is data loss.

$$Th_{TL} = \frac{D_T - D_L}{L} \quad (1)$$

3. Results and Discussion

3.1. Fiber-Wireless Performance Test

Figure 9 shows the throughput for downstream and upstream transmissions. The purpose of this validation is to observe the correctness of the router testbed. The throughput of the testbed is scaled up to match the throughput industrial grade router. Based on the figure, the throughput of the proposed testbed has similar increasing trend as the industrial grade router, thus validates the correctness. The downstream throughput for proposed testbed increases from 0 bps at 0% offered load to 719 kbps at 100% offered load. The throughput of industrial grade router increases from 0 bps at 0% offered load to 697 kbps at 90%, then decreases at 100% offered load when reaching its bandwidth limit. This situation is due to the QoS algorithm in the industrial grade router limits the maximum throughput.

Similar trend is observed for upstream transmission, where the throughput of the proposed testbed increases as the offered load increases from 0 bps to 770.351 kbps at 0% to 100% offered load. The throughput for industrial grade router is also increasing from 0 bps to 774 kbps at 0% to 100% offered load.

For both downstream and upstream transmissions, industrial grade router is able to transmit data up to 700 kbps and 774 kbps respectively at 100% offered with 1445 bytes of in each packet. But the proposed testbed can only transmit for one packet of 1445 bytes, up to 52.097 kbps for downstream transmission and 54.335 kbps for upstream transmission. Hence, if the proposed testbed sends as many packets as the industrial grade router, the transmitted data in the proposed testbed is vertically scaled up by factor of 13.436 (700 kbps/59.097 kbps = 13.436) for downstream and 12.883 for upstream (774 kbps/59.097 kbps = 13.436). Vertical scaling is a method to translate a graph without losing the original properties where all y-values in the graph are multiplied by a specific factor. Therefore, for the proposed testbed, the throughput for each offered load is multiplied with 13.436 and 12.883 for downstream and upstream respectively.

Overall, the throughput of the proposed testbed is slightly lesser than the industrial grade router due to its lower technical specification, bandwidth and processing power but is adequate enough for lab-scale experiment. From the graphs, we can conclude that the trend of the proposed testbed is correct.

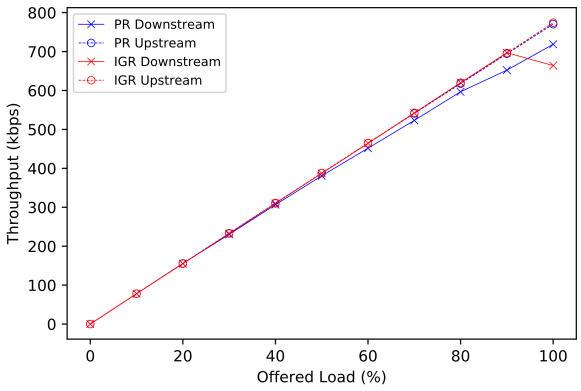


Figure 9: FiWi downstream and upstream throughput for the proposed router (PR) testbed and the industrial grade router (IGR)

Figure 10 shows the end-to-end delay for downstream and upstream transmission in FiWi architecture. For downstream, the minimum and maximum end-to-end delay are 0.1246 s and 0.14484 s at 100 bytes and 1445 bytes, respectively. For upstream, the minimum end-to-end delay is 0.12193 s at 100 bytes, and the maximum is 0.12430 s at 1445 bytes. The increased delay is due to testbed takes longer time to process the data. The subtle differences in the graphs caused by very small data size for a large bandwidth of fiber (1 Gbps). Furthermore, the data is transmitted by using light pulses in fiber. Hence, it transmits faster compared to electrical pulses in copper. For example, during downstream transmission, the end-to-end delay at 100 bytes is 0.1246 s, whereas at 200 bytes the end-to-end delay is 0.1250 s. The end-to-end delay is expected to be higher in the proposed router testbed compared to a typical fiber-supported router like industrial grade router, also because of hardware limitation. This is acceptable as part of the aim of this article is to create a reconfigurable router testbed that supports fiber-wireless transmission for educational module.

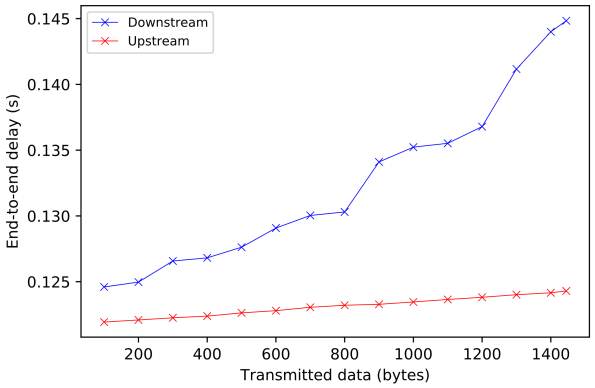


Figure 10: FiWi downstream and upstream end-to-end delay

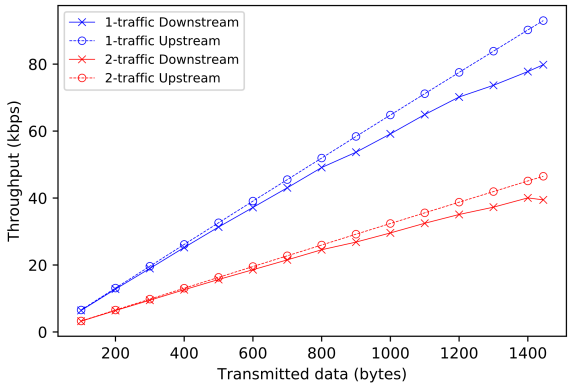


Figure 11: FiWi stress test downstream and upstream throughput for 1-traffic and 2-traffics transmissions

3.2. Fiber-Wireless Stress Test

Figure 11 shows the throughput for downstream and upstream transmissions. The throughput is still increasing gradually with the data size. The minimum and maximum throughput are 3.210 kbps at 100 bytes and 40.047 kbps at 1400 bytes. For downstream

with 2 traffics, the throughput decreases at 1445 bytes from 40 kbps to 39.5 kbps due to the end-limit of the testbed. Compared with the throughput of a single traffic, the throughput of two traffics is about half of the throughput of single traffics. The reason of this occurrence is two traffics causing the proposed testbed to process the data twice. For example, at 100 bytes in downstream transmission, the throughput for a single traffic is 6.42 kbps whereas the throughput for two traffics is 3.21 kbps. This statement is supported by [40] where more traffics contribute to less throughput.

Figure 12 shows the downstream and upstream end-to-end delay graphs. For downstream, the trend of the graph is increasing with the data size, from 0.249 s at 100 bytes to 0.293 s at 1445 bytes. At 1445 bytes, there is a sudden increase of end-to-end delay, due to the testbed's end-limit. Theoretically, when number of traffics increases, the end-to-end delay also increases [41]. Therefore, comparing the plots for downstream, the end-to-end delay of two traffics is twice as big as single traffic, because testbed needs to process the data twice. For example, the end-to-end delay at 1445 bytes for single traffic and two traffics are 144.838 ms and 293.01 ms, respectively.

The trend for upstream end-to-end delay is also similar where the delay increases from 243.863 ms at 100 bytes to 248.591 ms at 1445 bytes. At 1445 bytes, the end-to-end delay for two traffics and single traffic are 248.591 ms and 124.296 ms, respectively. Even though the trends of the graphs are almost constant, it is actually increasing as the data increases. The constant trends of the graphs are due to the presence of fiber which make the difference in end-to-end delay between two data such as 100 bytes and 200 bytes are not significant.

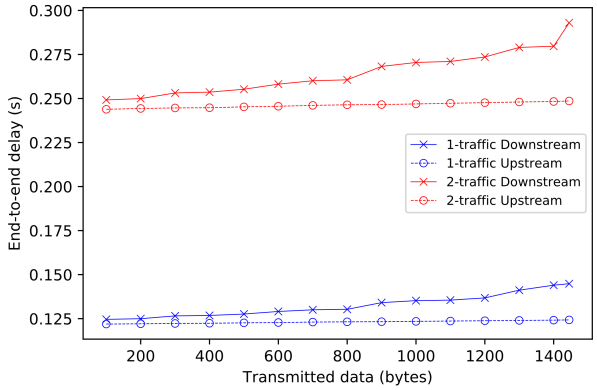


Figure 12: FiWi stress test downstream and upstream end-to-end delay for 1-traffic and 2-traffics transmissions

3.3. Scalability Test

3.3.1. Fiber-Wireless-Fiber

The plot for Fi-WiFi downstream and upstream throughput is shown in Figure 13. The downstream throughput increases with the data size from 2.626 kbps at 100 bytes to 22.403 kbps at 1445 bytes. For upstream, the throughput is increasing up until 1100 bytes only, unstable at 1200 bytes to 1300 bytes, then continue to decrease from 1400 bytes to 1455 bytes. Even though the downstream throughput is increasing, but the graph gradient is gradually decreasing starting at 700 bytes. Whereas the upstream throughput becomes unstable at 1200 bytes to 1300 bytes because the proposed testbed is almost at its limit, resulting in greater graph gradient. Then, the throughput decreases at 1400 bytes and 1445 bytes due to the sudden increase of upstream end-to-end delay at 1400 bytes in and the delay continues to increase until 1445 bytes.

The downstream and upstream end-to-end delay for Fi-WiFi architecture is shown in Figure 14. The downstream delay increases with the data size from 0.609 s at 100 bytes to 1.032 s at 1445 bytes. For upstream, the increases from 0.505 s at 100 bytes to 0.948 s at 1445 bytes. The end-to-end delay for this setup is higher than the FiWi architecture because the data has to undergo multiple routers and medium conversions from fiber to wireless and then wireless to fiber [42]. For downstream transmission, there is a sudden increase in end-to-end delay at 1400 bytes due to the end-limit of the proposed testbed, thus the longer processing time.

3.3.2. Wireless-Fiber-Wireless

The downstream and upstream throughput for Wi-FiWi architecture can be seen in Figure 13. The trend of the downstream throughput increases with the transmitted data size from 2.833 kbps at 100 bytes to 38.09 kbps at 1445 bytes. For upstream, the throughput increases from 2.81 kbps at 100 bytes to 38.348 kbps at 1445 bytes. At 100 bytes, the throughput for upstream transmission is slightly lower than downstream. However, as the transmitted data increases, the throughput for upstream transmission is higher than downstream: the throughput at 1100 bytes is 29.754 kbps (upstream) and 29.941 kbps (downstream). However, starting 1200 bytes, the throughput for upstream transmission is 32.276 kbps, while for downstream is 32.087 kbps. At 1445 bytes, the performance for both transmissions are slightly decreased due to the end-limit of the testbed, referring to the slope of the graph at this point. The slight differences in the throughput values proving that the proposed testbed in Wi-FiWi architecture is stable for both transmissions. This proposed router testbed is therefore suitable to be a scalable testbed.

Figure 14 shows the end-to-end delay of downstream and upstream transmissions respectively for Wi-FiWi architecture. The downstream end-to-end delay increases with the transmitted data size from 0.565 s at 100 bytes to 0.607 s at 1445 bytes. There is a sudden increase to the downstream end-to-end delay at 1200 bytes, because the proposed testbed is approaching its limit. Hence, the data processing time takes longer at this point. For upstream, the end-to-end delay is also increasing with the transmitted data, from 0.569 s at 100 bytes to 0.603 s at 1445 bytes. It can be observed that the downstream and upstream end-to-end delay for Wi-FiWi architecture is also higher than the FiWi because there are more medium changes in between clients. Therefore, the process of medium conversion contributes to higher end-to-end delay.

3.3.3. Fi-WiFi versus Wi-FiWi

Figure 13 clearly shows the comparisons in terms of throughput between Fi-WiFi and Wi-FiWi architectures for downstream and upstream transmissions from 100 bytes to 1445 bytes. The performance for Wi-FiWi is more stable than Fi-WiFi based on the behaviour of the graphs. This is because Fi-WiFi has more conversion from electrical pulses to light pulses and vice versa compared to Wi-FiWi causing the instability in Fi-WiFi.

Thus, the overall performance in Wi-FiWi is more stable because it has better observable graphs compared to Fi-WiFi.

Figure 14 shows the end-to-end delay comparisons between Fi-WiFi and Wi-FiWi architectures for downstream and upstream transmissions. Both have increasing end-to-end delay when the transmitted data size increased from 100 bytes to 1445 bytes. However, Fi-WiFi architecture has higher delay than Wi-FiWi because of the higher number of FMCs in Fi-WiFi. Hence, the medium conversions from light pulses to electrical pulses or vice versa is more than Wi-FiWi, thus contributes to more delay. From the graphs, we can conclude that Wi-FiWi has more stable transmission because throughout the transmissions, the end-to-end delay changes subtly. The summary of the throughput and end-to-end delay results at full load (1445 bytes) for all experiments using the proposed router testbed are listed in Table 2. Throughput results for FiWi architecture are not scaled, to fairly compare with other experiments.

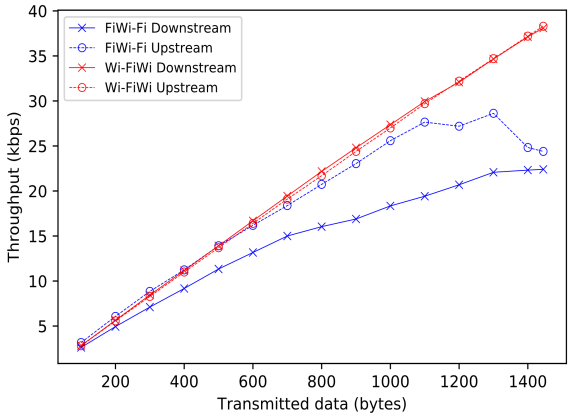


Figure 13: Fi-WiFi vs Wi-FiWi throughput

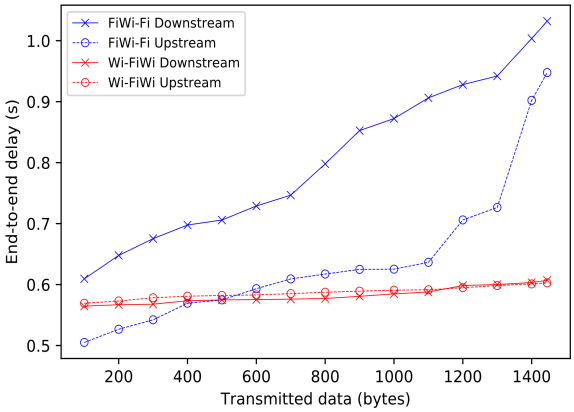


Figure 14: Fi-WiFi vs Wi-FiWi end-to-end delay

4. Conclusion

FiWi is seen to be one of the best technologies for future global data communication network architecture due to the advantages of providing robustness and mobility to the consumers by deploying FIber and wireless in one network. To equip fundamental knowledge and further understanding on FiWi for students and researchers at the university level, this article proposed a FiWi router testbed. It is a fast integration and scalable, which is essential for an educational module testbed. The focus of this research is on lab-scale environment router testbed in FiWi architecture to provide a proof-of-concept solution on reconfigurable router testbed in FiWi network. Raspberry Pi is chosen to be the embedded system hardware in this research for the development of scalable and reconfigurable FiWi router testbed. It has a socket module which enables the user to program Raspberry Pi to communicate with each other using IP addresses.

The proposed FiWi testbed throughput performance has been validated with industrial grade router to prove the correctness. As the proposed testbed approaches its limit, the throughput is maintained with no further increment. For the end-to-end delay, it increases with the data size with acceptable measure and its behaviour satisfies the end-to-end delay of IEEE 802.15.4 routing scheme. Hence, the proposed testbed is suitable to serve as an educational module FiWi testbed due to its simplicity, reconfigurable and fast implementation. A stress test shows that the end-to-end delay for two traffics is twice as high as single traffic due to the twice data processing. This causes the throughput for two traffics is halved of the single traffic throughput. For scalability performance test, the end-to-end delay in Fi- WiFi is higher than Wi-FiWi architecture due to more electrical pulses to light pulses conversion and vice versa from FMC in Fi-WiFi. Furthermore, Wi-FiWi is more stable compared to Fi-WiFi, thus making the testbed flexible and scalable for various possible setups. The proposed router testbed also shows a promising stability, a huge advantage to test various architectures.

Table 2: Summary of throughput and end-to-end delay results at full load (1445 bytes) for all experiments

Experiment	Parameters			
	Throughput (kbps)		End-to-end delay (s)	
	Up stream	Down stream	Up stream	Down stream
FiWi	54.34	59.10	0.124	0.144
FiWi stress test	46.50	39.45	0.250	0.290
Fi-WiFi	24.40	22.40	0.948	1.030
Wi-FiWi	38.35	38.09	0.600	0.610

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The author gratefully acknowledges the funding of Yayasan Canselor UNITEN (201901001YCU/10), and UNITEN R&D Sdn. Bhd. for TNB Seed Fund (U-TD-RD-20-03).

References

[1] M. H. M. Nasir, W. S. H. M. W. Ahmad, N. A. M. Radzi, F. Abdullah, A. Abdullah, A. Ismail, and M. Z. Jamaludin, "Raspberry Pi as Reconfigurable Fiber-based Router and Its Performance Analysis," in 2020 IEEE 8th International Conference on Photonics (ICP), 30-31, 2020, doi:10.1109/ICP46580.2020.9206497.

[2] M. Tornatore, G.-K. Chang, and G. Ellinas, Fiber-wireless convergence in next-generation communication networks, Springer, 2017.

[3] E. A. Kosmatos and N. D. Tselikas, "Dominant handover algorithms for vehicular radio-over-fiber networks at 60 GHz: a performance Evaluation study" Performance Evaluation, **102**, 21-33, 2016, doi: 10.1016/j.peva.2016.05.001.

[4] M. Ridwan, N. A. M. Radzi, F. Abdullah, N. M. Din, and C. Rashidi, "Feasibility study of a reconfigurable fiber-wireless testbed using universal software radio peripheral" International Journal of Engineering and Technology Innovation, **8**(4), 274-283, 2018,

[5] N. Ghazisaidi and M. Maier, "Fiber-wireless (FiWi) access networks: challenges and opportunities" IEEE Network, **25**(1), 36-42, 2011, doi:10.1109/MNET.2011.5687951.

[6] M. Maier and N. Ghazisaidi, FiWi access networks, Cambridge University Press, 2011.

[7] B. P. Rimal, M. Maier, and M. Satyanarayanan, "Experimental testbed for edge computing in fiber-wireless broadband access networks" IEEE Communications Magazine, **56**(8), 160-167, 2018, doi:10.1109/MCOM.2018.1700793.

[8] V. Mishra, R. Upadhyay, and U. R. Bhatt, Progress in Advanced Computing and Intelligent Engineering, Springer, 2018.

[9] J. Liu, H. Guo, H. Nishiyama, H. Ujikawa, K. Suzuki, and N. Kato, "New perspectives on future smart FiWi networks: scalability, reliability, and energy efficiency" IEEE Communications Surveys & Tutorials, **18**(2), 1045-1072, 2015, doi:10.1109/COMST.2015.2500960.

[10] Y. Liu, L. Guo, B. Gong, R. Ma, X. Gong, L. Zhang, and J. Yang, "Green survivability in fiber-wireless (FiWi) broadband access network" Optical Fiber Technology, **18**(2), 68-80, 2012, doi: 10.1016/j.yofte.2011.12.002.

[11] Q. Dai, G. Shou, Y. Hu, and Z. Guo, "A general model for hybrid fiber-wireless (FiWi) access network virtualization," in 2013 IEEE International Conference on Communications Workshops (ICC), 858-862, 2013, doi:10.1109/ICCW.2013.6649354.

[12] Y. Yu, C. Ranaweera, C. Lim, L. Guo, Y. Liu, A. Nirmalathas, and E. Wong, "Hybrid fiber-wireless network: An optimization framework for survivable deployment" Journal of Optical Communications and Networking, **9**(6), 466-478, 2017, doi: 10.1364/JOCN.9.000466.

[13] H. Guo and J. Liu, "Collaborative computation offloading for multi-access edge computing over fiber–wireless networks" IEEE Transactions on Vehicular Technology, **67**(5), 4514-4526, 2018, doi:10.1109/TVT.2018.2790421.

[14] Z. Zhang, J. Kong, C. Huang, Q. Wu, J. Wu, and J. Li, "Virtual network embedding algorithm considering resource fragmentation in virtualized industrial fiber-wireless (FiWi) access network," in Proceedings of the International Conference on Imaging, Signal Processing and Communication, 148-152, 2017, doi: 10.1145/3132300.3132328.

[15] H. Beyranvand, W. Lim, M. Maier, C. Verikoukis, and J. A. Salehi, "Backhaul-aware user association in FiWi enhanced LTE-A heterogeneous networks" IEEE Transactions on Wireless Communications, **14**(6), 2992-3003, 2015, doi:10.1109/TWC.2015.2399308.

[16] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila, and T. Taleb, "Survey on multi-access edge computing for internet of things realization" IEEE Communications Surveys & Tutorials, **20**(4), 2961-2991, 2018, doi:10.1109/COMST.2018.2849509.

[17] W. Sun, J. Liu, and H. Zhang, "When smart wearables meet intelligent vehicles: Challenges and future directions" IEEE wireless communications, **24**(3), 58-65, 2017, doi:10.1109/MWC.2017.1600423.

[18] B. P. Rimal, D. P. Van, and M. Maier, "Mobile edge computing versus centralized cloud computing over a converged FiWi access network" IEEE Transactions on Network and Service Management, **14**(3), 498-513, 2017, doi:10.1109/TNSM.2017.2706085.

[19] P.-Y. Chen and M. Reisslein, "FiWi network throughput-delay modeling with traffic intensity control and local bandwidth allocation" Optical Switching and Networking, **28**, 8-22, 2018, doi: 10.1016/j.osn.2017.11.001.

[20] B. P. Rimal, D. P. Van, and M. Maier, "Mobile edge computing empowered fiber-wireless access networks in the 5G era" IEEE Communications Magazine, **55**(2), 192-200, 2017, doi:10.1109/MCOM.2017.1600156CM.

[21] H.-H. Lu, C.-Y. Li, T.-C. Lu, C.-J. Wu, C.-A. Chu, A. Shiva, and T. Mochii, "Bidirectional fiber-wireless and fiber-VLLC transmission system based on an OEO-based BLS and a RSOA" Optics Letters, **41**(3), 476-479, 2016, doi: 10.1364/OL.41.000476.

[22] P. Singh and S. Prakash, "Optical network unit placement in Fiber-Wireless (FiWi) access network by Moth-Flame optimization algorithm" Optical Fiber Technology, **36**, 403-411, 2017, doi: 10.1016/j.yofte.2017.05.018.

[23] B. Kantarci, N. Naas, and H. T. Mouftah, "Energy-efficient DBA and QoS in FiWi networks constrained to metro-access convergence," in 2012 14th International Conference on Transparent Optical Networks (ICTON), 1-4, 2012, doi:10.1109/ICTON.2012.6253908.

[24] M. Lévesque, M. Maier, F. Aurzada, and M. Reisslein, "Analytical framework for the capacity and delay evaluation of next-generation FiWi network routing algorithms," in 2013 IEEE Wireless Communications and Networking Conference (WCNC), 1926-1931, 2013, doi:10.1109/WCNC.2013.6554859.

[25] J. Liu, H. Guo, Z. M. Fadlullah, and N. Kato, "Energy consumption minimization for FiWi enhanced LTE-A HetNets with UE connection constraint" IEEE Communications Magazine, **54**(11), 56-62, 2016, doi:10.1109/MCOM.2016.1600169CM.

[26] N. Choosri, Y. Park, S. Grudpan, P. Chuarjedton, and A. Ongvisesphaiboon, "IoT-RFID testbed for supporting traffic light control" International Journal of Information and Electronics Engineering, **5**(2), 102-106, 2015, doi:10.7763/IJIEE.2015.V5.511.

[27] W. Hurst, N. Shone, A. El Rhalibi, A. Happe, B. Kotze, and B. Duncan, "Advancing the micro-CI testbed for IoT cyber security research and education," in The Eighth International Conference on Cloud Computing, GRIDs, and Virtualization, 129-134, 2017.

[28] S. T. Abraha, D. F. Castellana, X. Liang, A. Ng'oma, and A. Kobyakov, "Experimental study of distributed massive MIMO (DM-MIMO) in in-building fiber-wireless networks," in 2018 Optical Fiber Communications Conference and Exposition (OFC), 1-3, 2018,

[29] Z. Gong, W. Xue, Z. Liu, Y. Zhao, R. Miao, R. Ying, and P. Liu, "Design of a reconfigurable multi-sensor testbed for autonomous vehicles and ground robots," in 2019 IEEE International Symposium on Circuits and Systems (ISCAS), 1-5, 2019, doi:10.1109/ISCAS.2019.8702610.

[30] M. H. M. Nasir, N. A. M. Radzi, W. S. H. M. W. Ahmad, F. Abdullah, and M. Z. Jamaludin, "Comparison of router testbeds: embedded system-based, software-based, and multiprotocol label switching (MPLS)" Indonesian Journal of Electrical Engineering and Computer Science, **15**(3), 1250-1256, 2019, doi:http://doi.org/10.11591/ijeecs.v15.i3.pp1250-1256.

[31] S. M. Blair, F. Coffele, C. Booth, B. De Valck, and D. Verhulst, "Demonstration and analysis of IP/MPLS communications for delivering power system protection solutions using IEEE C37. 94, IEC 61850 sampled values, and IEC 61850 GOOSE protocols," in 2014 CIGRE Session, 1-8, 2014.

[32] K. Tantayakul, R. Dhaou, B. Paillassa, and W. Panichpattanakul, "Experimental analysis in SDN open source environment," in 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 334-337, 2017, doi:10.1109/ECTICon.2017.8096241.

[33] V. Sivaraman, A. Vishwanath, D. Ostry, and M. Thottan, "Greening router line-cards via dynamic management of packet memory" IEEE Journal on Selected Areas in Communications, **34**(12), 3843-3853, 2016, doi:10.1109/JSAC.2016.2600412.

[34] D. Posch, B. Rainer, S. Theuermann, A. Leibetseder, and H. Hellwagner, "Emulating NDN-based multimedia delivery," in Proceedings of the 7th International Conference on Multimedia Systems, 1-4, 2016. doi: 10.1145/2910017.2910626.

[35] S. Y. Jang, B. H. Shin, and D. Lee, "Implementing a dynamically reconfigurable wireless mesh network testbed for Multi-Faceted QoS support," in Proceedings of the 11th International Conference on Future Internet Technologies, 95-98, 2016, doi: 10.1145/2935663.2935678.

[36] X. Piao, L. Huang, K. Yuan, J. Yuan, and K. Lei, "The real implementation of NDN forwarding strategy on android smartphone," in 2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 1-6, 2016, doi:10.1109/UEMCON.2016.7777909.

[37] V. Gupta, K. Kaur, and S. Kaur, "Developing small size low-cost software-defined networking switch using Raspberry Pi," in 50th Annual Convention of Computer Society of India : Next-Generation Networks, 147-152, 2018, doi: 10.1007/978-981-10-6005-2_16.

[38] S. Brown, "An analysis of loss-free data aggregation for high data reliability in wireless sensor networks," in 2017 28th Irish Signals and Systems Conference (ISSC), 1-6, 2017, doi:10.1109/ISSC.2017.7983622.

[39] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: a simple model and its empirical validation," in Proceedings of the ACM SIGCOMM'98 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, 303-314, 1998, doi: 10.1145/285237.285291.

[40] H. Xu, X.-Y. Li, L. Huang, H. Deng, H. Huang, and H. Wang, "Incremental deployment and throughput maximization routing for a hybrid SDN" IEEE/ACM Transactions on Networking, **25**(3), 1861-1875, 2017, doi:10.1109/TNET.2017.2657643.

[41] Q. T. Minh, T. K. Dang, T. Nam, and T. Kitahara, "Flow aggregation for SDN-based delay-insensitive traffic control in mobile core networks" IET Communications, **13**(8), 1051-1060, 2019, doi:10.1049/iet-com.2018.5194.

[42] M. A. Abidini, O. Boxma, C. Hurkens, T. Koonen, and J. Resing, "Revenue maximization in optical router nodes" Performance Evaluation, **140**, 102-108, 2020, doi: 10.1016/j.peva.2020.102108.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1208-1214 (2021)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Factors Impacting Digital Payment Adoption: An Empirical Evidence from Smart City of Dubai

Anas Najdawi[*,1], Zakariya Chabani[2], Raed Said[2]

[1]*Faculty of Management, Amity University, Dubai, 12346, UAE*

[2]*Faculty of Management, Canadian University of Dubai, Dubai, 12346, UAE*

ARTICLE INFO

Article history:
Received: 02 January, 2021
Accepted: 03 February, 2021
Online: 25 February, 2021

Keywords:
E-Payment Adoption
Digital Payment Technologies
Digital Transformation
Generations XYZ
Factor Analysis

ABSTRACT

The emergence of new digital payment technologies has introduced both opportunities and challenges across all industries. This research aims to examine the significant factors that influence the adoption of new e-payment technologies, specifically in smart cities, as in Dubai. A comprehensive theoretical framework based on several previous studies included the following factors; Perceived Usefulness, Perceived Trust, Perceived Personal Innovativeness, Perceived Ease of Use, Perceived Risk, and Generation Cohort. The results of this research confirm that all proposed factors significantly affect the adoption of e-payment in Dubai as a case of a smart city; however, the perceived usefulness is not as significant as the other factors. Moreover, comparative analysis across the three generations showed almost similar patterns of adopting e-payment systems.

1. Introduction

E-payment is among the revolutionary changes caused by technology in the field of financial services that tremendously changed humans' life by making them simpler and easier than before. Currently, e-payments have taken over the entire world of financial transactions whether these are meant for personal or commercial use [1] because it offers numerous benefits to the consumers including cost-effectiveness and conservation of time and energy. The UAE in general, and Dubai in particular, has extensively adopted new innovative e-payment systems with over 25 e-payment service providers offering more than 250 payment services for traffic fines, electricity bills, parking fines, university fees, etc... [2]. Moreover, the e-payment transactions in the UAE increased by 21% between 2015 and 2016, which represents $12.4 billion of goods and services [3] due to the Dubai government's vision of creating the perfect smart city.

Despite their rapid growth, e-payment systems are not perfect. Some disfavors easily noticed, such as hacking risks, the high cost of setting-up, and upgrading the system. These disadvantages prevent some individuals from adopting e-payment systems. This paper investigates the key factors that influence the adoption of new e-payment technologies across several generations in smart cities such as Dubai. Consequently, help the organizations in Dubai, facilitating the adoption of innovative e-payment services through smart and effective design that is more conscious about the essential factors in the adoption criteria. The expansion of new e-payment adoption research allows companies and organizations to spend less effort and time, which can be invested in developing new e-payment services with more significant opportunities in the market.

A large number of studies employed the TAM or UTAUT models to investigate the factors affecting e-payment adoption. However, the framework proposed in [4] is considered more suitable for the goals of the current research. The study investigated the Malaysian consumers' perception of e-payment and recommended the Benefits, Trust, Self-Efficacy, Ease of Use, and Security as the most critical factors affecting the consumers' perception of e-payment. This research extended the previous frameworks by proposing the following factors: Perceived Usefulness, Perceived Trust, Perceived Personal Innovativeness, Perceived Ease of Use, and Perceived Risk. Moreover, the effect of age cohorts (X, Y, and Z Generations) on the previous factors and the users' perception to adopt the e-payment system in Dubai is to be examined.

The research in [5], is the sole study among past studies that shed light on how Age influences the factors of e-payment adoption. However, only some of the factors (Perceived Ease of Use and Perceived Usefulness) were considered by the researchers, overlooking the impact of other significant factors.

[*]Corresponding Author: Anas Najdawi, Amity University, Dubai, UAE, anajdawi@amityuniversity.ae

In [6], the authors explored the factors affecting the adoption of e-payment systems among university students in the UAE. In this study, the researchers employed a modified TAM model in which the factor of trust was also incorporated. However, the population studied in this paper is limited to the university's students only. Therefore, the age factor was not being taken into consideration. The authors addressed students residing in all the emirates of UAE.

In [7], the authors adds to the previous work the demographic variables and e-payment channel adoption in the UAE. The current paper is organized as follows. Section 2 is dedicated to reviewing the previous works. Section 3 will represent our framework. The next section will be the description of the methodology used in order to accept or reject the proposed hypotheses. After that, the results are listed and discussed in Sections 5 and 6. The conclusions and different implications are provided in Section 7. Finally, the research limitations and recommendations for future studies are stated in Section 8.

2. Literature Review

2.1. E-payment systems

Similarly to all e-services, an electronic platform is essential for the functioning of e-payment systems. The e-payment supports various kinds of monetary transactions [1], including regular transactions like paying for grocery or utility bills as well as B2B transactions besides others, and thus, it enables the buyer and seller to reap the benefits of a faster and more convenient payments system [2].

Various modes of payment, including e-money, debit or credit cards, and internet banking, may be considered as E-payment [3]. Moreover, another important revolution in this context is the introduction of mobile money by various organizations. For instance, two of the most successful companies in terms of mobile money are Apple pay and Samsung pay [4].

The e-payment systems have numerous advantages. For instance, they offer faster monetary transactions waiving the need for long queues and waiting periods for payments and receipt of money. People can remotely perform the payment and receipt of money through their cell phones or another mode of payment. Additionally, the fees related to financial transactions when using e-payment systems are far less than transaction fees when using traditional ways in most countries, especially in Dubai. Even in the absence of direct fees, the consumer may have to bear indirect fees, such as transportation fees, to go to the bank to withdraw money or go to the store to buy goods and services.

However, e-payment systems also have some disadvantages, for instance, the security threat that users face when using the e-payment system, the Risk of hacking the system, which causes a loss of money and personal information. Another disadvantage would be the cost; although the cost will be less for users, some costs related to securing the system and installing or upgrading it could be relatively high.

2.2. Generations X Y Z characteristics

The ICTs are changing and transforming individuals and society in general. Various generations are characterized by different traits, preferences, behaviors, principles, and concerns [5]. Therefore, all people with the same characteristics are grouped in what is called a cohort.

The author serves as the bridge that connects the previous and current generations [8]. One of the prominent attributes of Gen X is its reliance on reviews of other people. They make purchase decisions only after thorough research and review of feedback posted by others to make better decisions regarding the purchase [9]. However, they witnessed ICT development and then information society; hence, they became more familiarized with technology [10]. They are adapting to technology very fast.

The author evolved amidst technological advancements and internet growth, allowing them to acquire tons of information without hassle. Gen Y has been introduced to the internet and technology since childhood; hence, they are thoroughly familiar with technology [11], [10], and Media is part and parcel of their lives since they grew in the era of information [12]; therefore, they are comfortable with technology and would prefer to communicate with e-mail or text than talking face-to-face [13] and focus significantly on technical information [14].

Using computers and the internet is part and parcel of Gen Z's lives, who deem social media platforms as the basis of communication and interaction [10]. Therefore, they are excellent with all the digital devices, and they are almost unable to exist without these devices [15].

2.3. Related frameworks and theories

TAM stands for Technology Acceptance Model, which was developed by Davis with the aim to determine and forecast the factors that affect a system before implementing it. TAM was employed by the majority of researchers to develop the framework in order to investigate the factors that influence e-payment adoption. The TAM model identified two major factors; Perceived Usefulness and Perceived Ease of Use, which can be affected by external variables [16]. Later, it was adapted to include more factors; some researchers included Perceived risk, Security, Perceived advantage, Trust, Web assurance seals and Usability [17]. Others, adopted the Perceived Risk and Information on e-payment as additional factors. Adopted Trust, Self-efficacy, and Security as additional factors [18].

In [19], the author performing research on the e-payment adoption in the context of Iran, the factors of Usability (Perceived Ease of use) and Technological and transaction system, as well as access to protection rules (Perceived Usefulness), were used by the authors. Moreover, Technical security, Security reports, Transaction system, and Personal experience pertaining to e-payment systems were identified in [20] as external variables. While the Age was added as a moderating factor in [21].

Other researchers adapted the UTAUT model to incorporate the factor of trust while discarding the factor of user behavior. Another change made in this model was the substitution of Behavioral intention with Continuance Intention [22]

3. Research model and hypotheses

The theoretical framework employed for the current research is shown in Figure 1. This study is carried out to investigate the

factors affecting the Emirate of Dubai residents' perception of using e-payment. The relationships between the factors affecting the residents' perception of using e-payment are used to develop the hypotheses in the following subsections. The factors we used in our conceptual model in light of prior findings from the literature are the following:

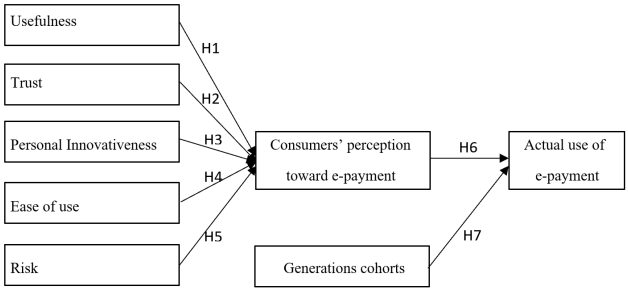


Figure 1: The Extended and Adopted Conceptual Framework

3.1. Perceived usefulness

Perceived usefulness implies the consumer's perception regarding the effectiveness and convenience offered by e-payment systems in simplifying his daily transactions [23]. In other words, it is the benefits that a person can get from e-payment systems. E-payment systems provide the user with a convenient mean of payment and the ability to store and transport a currency value [24].

Moreover, e-payment systems can provide other benefits such as time and cost-saving. Usefulness is one of the most important factors for e-payment systems use [17], [18], [25], [26]. However, as per some researchers, considerable time and effort may have to be invested by the consumer to learn how to use the e-payment system [27]; therefore, e-payment systems may be costly rather than cost-saving systems. Thus, it is hypothesized that:

H1 - There is a significant relationship between perceived usefulness and consumers' perception toward e-payments.

3.2. Perceived Trust

Perceived Trust in the context of e-payment systems implies the confidence bestowed by the user in the system regarding the reliability of the system in protecting the user's personal data and money from misuse and theft [28] , [29].

Many researchers stated that a high level of trust is essential for an e-payment system adoption [29], [22]. Similarly, users' concern over data privacy, internet security, etc., leads to a low level of trust, and thus, it will impact the adoption of e-payment systems negatively [32].

However, some researchers argue that trust alone is not sufficient to attract individuals adopting e-payment systems; for instance, if a reputed organization offers e-payment functions, the consumer will willingly use them; However, the user will hesitate in using the e-payment services provided by an untrustworthy organization [17]. Others found that trust is not significantly associated with individuals' perception toward e-payment [20] [31]. Thus, it is hypothesized that:

H2 - There is a significant relationship between consumers' perceived trust and perception toward e-payments.

3.3. Perceived personal innovation

Innovativeness is defined as the extent to which individuals consider themselves as technology leaders and pioneers [33]. This factor has been considered as an important factor influencing user's adoption of e-payment systems [18], [34], [35]. Thus, it is hypothesized that:

H3 - There is a significant relationship between perceived personal innovativeness and consumers' perception toward e-payments.

3.4. Perceived ease of use

Ease of use is defined as to which extent a person will not spend effort while using an e-payment system [23]. It can be attained via several factors, such as the content, design [29], information management, and speed [36]. E-payments systems can be affected significantly by these factors [37]. Previous researchers found that ease of use has a positive effect on the adoption of e-payment systems [18], [21]. Thus, it is hypothesized that:

H4: There is a significant relationship between perceived ease of use and consumers' perception toward e-payments.

3.5. Perceived Risk

Risk is defined as the individual's uncertainty about e-payment systems [38], due to their concern about fraud, theft, hacking, misuse of personal information, etc... Individuals may feel hesitant about the e-payment system due to Risk as an important factor influencing the adoption of the e-payment systems [39]. And the individuals' concern about the system may be a barrier to its adoption [40].

Contrary to most studies, some researchers argue that there is no relation between Risk and intention to adopt e-payment systems; although the Risk affects the payment system's adoption, it has no direct link; thus, individuals may use the system despite their concern [17]. Therefore, it is hypothesized that:

H5: There is a significant relationship between the perceived Risk and consumers' perception toward e-payments.

3.6. Consumer's perception towards the actual use of e-payment

User's attitudes toward IT systems represent a paramount factor that influences a particular system's actual use by individuals [16]. Thus, it is hypothesized that:

H6: Consumers' perception toward e-payment has a positive effect on the actual use of e-payments

3.7. Generations XYZ

In the current study, the generation cohorts will be defined based on the results from the literature review as follow:

- Generation X (Gen X), born between 1965 and 1979

- Generation Y (Gen Y), born between 1980 and 1994, this generation is also commonly known as 'Millennials' and sometimes as 'the children of Globalization' [41].

- Generation Z (Gen Z), born between 1995 and Nowadays, this generation is also known as the "silent generation," the iGeneration, generation quiet, and the next generation" [13].

The researchers believe that people from different generations do things differently when it comes to adopting e-payment systems in Dubai. Thus, it is hypothesized that:

H7: There are no significant differences among different generations toward the actual use of e-payments.

4. Methodology

The participants in this study were using e-payment in the emirate of Dubai. In this research, a nonprobability sample technique is used for collecting responses. Convenience sampling is a type of non-probability sampling where population elements are selected for inclusion in the sample based on access ease. In some cases, judgment sampling is used where the researchers' judgment is used for selecting participants who are considered as representative of the population.

The survey was carried out using a self-administered online questionnaire, which consisted of two major sections. The first section comprises 10 questions intended to collect demographic information and some information about the use of e-payment. The second section contains 28 statements meant to examine those factors that influence the perception of using the e-payment. This section's items essentially required the respondents to choose to what extent they agreed with each statement. The items were close-ended questions based on a 5-point Likert scale that each spectrum includes strongly disagree, disagree, no idea, agree, and strongly agree.

The proposed factors in accordance with the above hypotheses and the conceptual research framework constructed in section 3 are; Perceived Usefulness (abbreviated as PU), Perceived Trust (PT), Perceived Personal Innovation (PPI), Perceived Ease of Use (PEOU), Perceived Risk (PR), and Consumers' Perception toward e-payment (CP). Notably, items of the questionnaire for each one of these factors were cautiously adapted from prior studies. However, the wording was kept as similar as possible across studies.

The pilot study was done on 20 consumers prior to its dissemination to satisfy face validity. Cronbach Alpha, which is a measure of reliability based on the internal consistency of the constructs, was calculated for each factor. The reliability for each factor was higher than the acceptable 0.75 limit that [42] suggests. The Cronbach Alpha for the data instrument in total was 0.913, indicating that the questionnaire has attained a rather high level of reliability. Hence, all items are retained.

Before the regression analysis, 7 variables were created. The items relating to each factor were added together for each respondent, and the average was found. The mean scores ranged from a possible score of 1 to 5, with 1 representing total disagreement with all stated items, and 5 representing total agreement with all stated items—two types of variables available in the research - dependent variable and independent variables. The main objective of this research is to investigate the predictors/variables that would determine the consumers' perception of the use of e-payment in Dubai. In this study the independent variables are the PU, PEOU, PT, PR, PPI, and CP. These six independent variables would be tested to identify whether these variables possess an influence on the dependent variable, which is the actual use of e-payments (UEP).

5. Analyses and Results

5.1. Descriptive findings

We conducted an online survey to investigate factors affecting consumers' perception of e-payment differences via generations based on our proposed conceptual framework. A total of 379 respondents visited the survey's link, and 242 respondents (63.8%) completed the survey. Of the 242 respondents, 207 respondents (85.5%) were considered eligible for this study after we removed respondents who answered that they are not currently using any of the e-payment methods.

Table 1. Shows the demographic characteristics of the e-payment users in the Emirate of Dubai who use e-payment. From the respondents to the questionnaire, 52.9 % were male, and 47.1 % were female. 12.6 % of the participants have a year of birth less than 1980 (Generation X), 33.3% between 1980 and 1994 (Generation Y), and 54.1% born in 1995 or above (Generation Z). Twenty-two percent of the participants were married, and in terms of the current professional position, 12.1 % were Executive, 10.1% Non-executive, 2.4% Housewife, 61.4% students, 5.3% Self-employed, and 8.7% were in other positions. Also, 30 % of the participants had High school or lower, 12.6 Diploma, 37.7% Bachelor's degree, 10.1% Masters and 9.7% Ph.D. holders. In term of the monthly income, the study found that 55.2% of the respondents having income below 10000 Dirhams, 14.3% between 10000 and 20000, 18.2% between 20000 and 30000, 5.9% between 30000 and 40000 and 6.3% having a monthly income above 40000 Dirhams

Table 1: Demographic characteristics of the sample (n=207)

Criterion	Category	Frequency	%
Gender	Male	109	52.9
	Female	97	47.1
Generation	X	26	12.6
	Y	69	33.3
	Z	112	54.1
Marital Status	Single	161	77.8
	Married	46	22.2
Current Professional Position	Executive	25	12.1
	Non-executive	21	10.1
	Housewife	5	2.4
	Student	127	61.4
	Self-employed	11	5.3
	Others	178	8.7
Education	High school or lower	62	30.0

	College (2 years program)	26	12.6
	Bachelor's degree	78	37.7
	Masters	21	10.1
	Ph.D.	20	9.7
Monthly Income	Below 10000	113	55.2
	10000 up to less than 20000	29	14.3
	20000 up to less than 30000	38	18.2
	30000 up to less than 40000	12	5.9
	40000 or above	14	6.4
Total:		207	100

5.2. Relationships between the factors affecting the Consumers' perception towards e-payment

Table 2 shows the results of the regression model we have done so far using SPSS after we checked the assumptions of regression analysis. From these outputs, we conclude that the factors under investigation explained 78.2% of the variance of the consumers' perception to use e-payments in Dubai. The remaining 21.8% of the variance can be explained by other factors not listed in our conceptual framework. As shown in the ANOVA table (Table 3), the minimal significance F indicates that our regression model is valid at a 0.05 significant level.

Table 2. Regression Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.782[a]	.611	.602	.40269
a. Predictors: (Constant), PR, PEOU, PPI, PT, PU				

Table 3. ANOVA Table

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	51.303	5	10.261	63.274	.000[b]
	Residual	32.595	201	.162		
	Total	83.898	206			
a. Dependent Variable: CP						
b. Predictors: (Constant), PR, PEOU, PPI, PT, PU						

Table 4 shows the estimations of regression coefficients along with their t-ratios and probability values. It is seen that PEOU, PT, and PPI have a positive and statistically significant effect on the CP. In contrast, the PR has a negative and statistically significant effect on the CP. However, the results of Table 4 indicate that there is no significant relationship between PU and the CP (p-value 0.057). Therefore, hypotheses H2, H3, H4, and H5 of the present research have been validated, but there is no strong evidence to accept H1 at a 5% level of significance.

The sixth hypothesis of the study assumes that the consumers' perception toward e-payment has a positive effect on the actual use of e-payment. To test this hypothesis, we used a correlation analysis. The result of the Pearson's correlation coefficient between the consumers' perception toward e-payment and the actual use of e-payments is significant at 0.01 level (r = 0.452, p-value=0.000). The findings support the stated hypothesis H6 and confirm that consumers' perception toward e-payment has a positive effect on the actual use of e-payments

Table 4: Regression coefficients of the factors affecting consumers' perception of e-payment

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	Constant	.656	.239		2.742	.007
	PU	.144	.076	.143	1.911	.057
	PT	.116	.056	.110	2.073	.039
	PEOU	.446	.077	.452	5.782	.000
	PPI	.219	.048	.238	4.529	.000
	PR	-.096	.034	-.127	-2.797	.006
a. Dependent Variable: CP						

5.3. Differences among generations towards the use of e-payments

This section discusses the differences among generations towards the use of e-payments. As indicated in Table 5, the findings show that the respondents of generation X have a better agreement in using e-payments (average 4.10) than that of generation Y (with an average of 3.88) and of generation Z (with an average of 3.99). However, the value of the standard deviation (SD), reveals no significant differences between the respondents of generation X (SD=0.69) and the respondents of generation Z (SD=0.65). This contrasts with the respondents of generation Y (SD=0.97). This means there is a more significant difference among generation Y than the two other generations in regards to their perception to use e-payments.

Moreover, we used ANOVA test to verify the seventh and final hypothesis (H7). The hypothesis assumes a statistically significant difference in the consumers' perception of respondents from different generations toward the use of e-payment. As illustrated in Table 6, the p-value is above 0.05 (Sig. = 0.425). This indicates that there are no significant differences in the consumers' perception of respondents from different generations toward the use of e-payment. Therefore, the seventh hypothesis H7 is supported.

Table 5: Averages and Std. deviations of respondents from different generations toward the use of e-payment

Gen.	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min.	Max.
					Lower Bound	Upper Bound		
X	26	4.1026	.69134	.13558	3.8233	4.3818	2.33	5.00
Y	69	3.8841	.97991	.11797	3.6487	4.1195	1.00	5.00
Z	112	3.9970	.65070	.06149	3.8752	4.1189	2.33	5.00
Tot.	207	3.9726	.77987	.05420	3.8658	4.0795	1.00	5.00

Table 6: ANOVA Table of the respondents from different generations toward the use of e-payment

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.047	2	.523	.859	.425
Within Groups	124.242	204	.609		
Total	125.289	206			

6. Discussion

In the first hypothesis H1, Perceived Usefulness was found not to have a significant relationship with the consumers' actual use of

e-payment in Dubai. The finding was consistent with previous studies [27]. The result implied that regardless of the persons' belief that using an e-payment system will enhance the performance of his day to day activities; consumers are more likely to use e-payment.

In the second hypothesis, H2, Perceived Trust was found to have a significant relationship with the consumers' actual use of e-payment in Dubai. Similar results were demonstrated from previous studies [22], [29]-[32]. And this implied that a high level of trust is essential for an e-payment adoption.

Perceived personal innovation was found to have a significant relationship with the consumers' actual use of e-payment in Dubai, which supported the third hypothesis H3, which was in line with many past studies. [18], [34], [35]. And this implied that the extent to which individuals consider themselves as technology leaders and pioneers is considered an important factor influencing the user's adoption of e-payment.

Perceived Ease of Use was found to have a significant relationship with the consumers' actual use of e-payment in Dubai, which supported the fourth hypothesis H4, which was in line with many past studies [18], [21], [37]. And this implies that ease to use of handling e-payment, such as convenience, speed, flexibility, simplicity, accessibility, and availability, is also considered a vital factor influencing the user's adoption of e-payment.

Perceived Risk was found to have a significant negative impact found to have a significant relationship with the consumers' actual use of e-payment in Dubai, which supported the fifth hypothesis H5. This was in line with many past studies [39], [40]. This implies that individuals' concern about the system may be a barrier to its adoption.

In the sixth hypothesis, H6, the Consumers' Perception toward e-payment was found to have a significant positive relationship with the consumers' actual use of e-payment in Dubai. And this implied that the user's attitude toward IT systems represents a paramount factor that influences the actual use of e-payment.

Finally, this study's results did not show significant differences among the different generations toward the use of e-payment, which supported our seventh hypothesis, H7. This might be indicated that most consumers in Dubai, as a smart city, are sufficiently familiar with using e-payment and accepting any new technologies easily.

7. Conclusion and implications

In this paper, the researchers studied the effect of some factors that might affect the adoption of new e-payment services and technologies. Based on an extended framework that captures the unique status of smart cities such as Dubai that move faster in digital transformation rather than less developed ones. The extended framework included all of Perceived Usefulness, Perceived Trust, Perceived Personal Innovativeness, Perceived Ease of Use and Perceived Risk on the adoption of e-payment systems in Dubai while considering the age cohorts (X, Y, and Z generations).

The results show that most of the factors used in the study are indeed critical and significant to the users' adoption of e-payment systems in Dubai. However, perceived usefulness is not as significant as the other factors; we think it is because some users care more about the system being easy to use than to be useful.

Unexpectedly, we found no serious differences between the generation cohorts regarding the adoption of e-payment in Dubai. We think that this due to the government of Dubai's efforts to in digital transformation of all of its E-Government services, and the adoption of digital payments toward a cashless economy might be a reason. The results imply that the government's strategies encourage consumers to use technologies in general as e-payment systems are working fine. This research will help future studies that focus on more recent technologies in the region based, such as adopting cryptocurrencies by individuals and businesses in the Middle East's smart city context.

8. Study limitations and recommendation for future work

This study is not free from some limitations. And to begin with, the sample employed in our empirical analysis is collected by nonprobability sampling techniques. The use of a convenient sample limits the research results to the specific sample, which means that the findings cannot be generalized to a larger population [43]. Other factors and dimensions, including the individual differences, system characteristics, social influence, and facilitating conditions missing in this study, will be considered in future research.

Additionally, the exact e-payment channels and factors toward adoption in Dubai and the preferred e-payment methods should be conducted to understand how Dubai as a smart city is moving toward the digital transformation of e-payment channels and their diffusion among the consumers; this will compared to other smart cities in the region including Saudia Arabia and Singapore [44], [45] and my be focused on specific sector such as banking [46]. The expected results will be useful to financial institutions, government, and e-payment technology providers in this region and marketers and, will encourage the innovative design of e-payment services that are most likely will be adopted by individuals, businesses, and governments.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Z. J. Zuopeng and J. M. Sajjad, "Knowledge market in organizations: incentive alignment and IT support," Industrial Management & Data Systems, 1101 - 1122, 2012. doi:10.2991/ijcis.d.191025.002

[2] T. Pikkarainen, K. Pikkarainen, H. Karjaluoto and S. Pahnila, "Consumer acceptance of online banking: An extension of the technology acceptance model," Internet Research , 224-235, 2004. DOI: 10.1108/10662240410542652

[3] W. Chaiyasoonthorn and W. Suksa-ngiam, "The diffusion and adoption of electronic payment systems in bangkok," International Journal of e-Business Research, **15**(2), 102-115, 2019.doi:10.1086/261933

[4] J. M. Gray, "How apple pay coincides with the consumer financial protection act: will apple become a regulated entity," Journal of High Technology Law, 170-194, 2015.

[5] S. Lissitsa and O. Kol, "Generation X vs. Generation Y – A decade of online shopping," Journal of Retailing and Consumer Services, **31**, 304-312, 2016. DOI: 10.1016/j.jretconser.2016.04.015

[6] S. A. Salloum and M. Al-Emran, "Factors affecting the adoption of e-payment systems by university students: extending the TAM with trust,"

International Journal of Electronic Business, 371-390, 2018. DOI: 10.1504/IJEB.2018.098130

[7] A. Najdawi, . Z. Chabani, . S. Raed and O. Starkova, "Analyzing the Adoption of E-Payment Technologies in UAE Based on Demographic Variables," in 2019 International Conference on Digitization (ICD), Sharjah, 2019. DOI: 10.1109/ICD47981.2019.9105908

[8] J.-G. Heaney, "Generations X and Y's Internet banking usage in Australia," Journal of Financial Services Marketing, 196-210, 2007.doi: 10.1057/palgrave.fsm.4760052

[9] E. Peralta, "Generation X: The Small But Financially Powerful Generation," 17 September 2015. [Online]. Available: https://www.centro.net/blog/generation-x-the-small-but-mighty-generation/.

[10] E. Csobanka, "The Z Generation," Acta Technologica Dubnicae, 6-10, 2016. doi: 10.1515/atd-2016-0012

[11] I. BAKANAUSKIENĖ, R. BENDARAVIČIENĖ and I. BUČINSKAITĖ, "EMPLOYER'S ATTRACTIVENESS: GENERATION Y EMPLOYMENT," Human Resources Management & Ergonomics, **X**(1), 2016.

[12] M. Muda, R. Mohd and S. Hassan, "Online Purchase Behavior of Generation Y in Malaysia," Procedia Economics and Finance, **37**, 292 – 298 , 2016.doi: 10.1016/S2212-5671(16)30127-7

[13] J. G. Harber, "Generations in the Workplace: Similarities and Differences.," East Tennessee State University, Johnson City, 2011.

[14] M. Rahulan, O. Troynikov, C. Watson, M. Janta and V. Senner, "Consumer behavior of generational cohorts for compression sportswear," Journal of Fashion Marketing and Management: An International Journal, 87–104, 2015. doi: 10.1108/JFMM-05-2013-0072

[15] A. Kolnhofer-Derecskei, R. Zs. Reicher and A. Szeghegyi, "The X and Y Generations' Characteristics Comparison," Acta Polytechnica Hungarica, **14**(8), 2017. doi: 10.12700/APH.14.8.2017.8.6

[16] F. D. Davis , "Perceived usefulness, perceived ease of use, and user acceptance of information technology," MIS Quarterly, 19-340, 1989. doi: 10.2307/249008

[17] S. Ozkan , G. Bindusara and R. Hackney, "Facilitating the adoption of e-payment systems: theoretical constructs and empirical analysis," Journal of Enterprise Information Management, 305 - 325, 2010. doi: 10.1108/17410391011036085

[18] C. Lin and C. Nguyen, "Exploring E-Payment Adoption in Vietnam and Taiwan," Journal of Computer Information Systems, **51**(4), 41-52, 2011.doi: 10.1080/08874417.2011.11645500

[19] M. Barkhordari, Z. Nourollah, H. Mashayekhi, Y. Mashayekhi and M. S. Ahangar, "Factors influencing adoption of e-payment systems: an empirical study on Iranian customers," Information Systems and e-Business Management, **15**(1), 89–116, 2017.doi:10.1007/s10257-016-0311-1

[20] E. Oney, G. O. Guven and W. H. Rizvi, "The determinants of electronic payment systems usage from consumers' perspective," Economic Research-Ekonomska Istraživanja, **30**(1), 394-415, 2017. doi: 10.1080/1331677X.2017.1305791

[21] A. Riskinanto, B. Kelana and D. R. Hilmawan, "The Moderation Effect of Age on Adopting E-Payment Technology," Procedia Computer Science, 536–543, 2018. doi: 10.1016/j.procs.2017.12.187

[22] Indrawati and D. A. Putri, "Analyzing Factors Influencing Continuance Intention of E-Payment Adoption Using Modified UTAUT 2 Model," in 6th International Conference on Information and Communication Technology, 2018. doi: 10.1109/ICoICT.2018.8528748

[23] S. Roy and I. Sinha , "Determinants of Customers' Acceptance of Electronic Payment System in Indian Banking Sector – A Study," International Journal of Scientific & Engineering Research, 177- 187, 2014.

[24] S. Chakravorti, "Theory of Credit Card Networks: A Survey of the Literature," Review of Network Economics, 50-68, 2003. doi: 10.2139/ssrn.419944

[25] S. San Martín , B. López-Catalán and M. A. Ramón-Jerónimo , "Factors determining firms' perceived performance of mobile commerce," Industrial Management & Data Systems, 946-963, 2012. doi: 10.1108/02635571211238536

[26] Y. Chou, C. Lee and J. Chung, "Understanding m-commerce payment systems through," Journal of Business Research, 1423 – 1430, 2004. doi:10.1016/S0148-2963(02)00432-0

[27] C. Kim, W. Tao, N. Shin and K.-S. Kim, "An empirical study of customers' perceptions of security and trust," Electronic Commerce Research and Applications, 84–95, 2009. doi:10.1016/j.elerap.2009.04.014

[28] T. Kongprapunt and N. Pupat, "FACTORS INFLUENCING GENERATION Y'S ONLINE PURCHASE INTENTION TOWARD XYZ ONLINE STORE IN THAILAND," Assumption University GSB e-journal, 94-106, 2018.

[29] D. Abrazhevich, Electronic Payment Systems: a User-Centered Perspective and Interaction Design, Eindhoven: Eindhoven University of Technology, 2004.doi: 10.6100/IR575913

[30] S. Kurnia and B. Lim, "Exploring the Reasons for a Failure of Electronic Payment," Journal of Research and Practice in Information Technology, 231-243, 2007.doi: 10.1007/s10660-011-9083-3

[31] W. M. Teoh, S. C. Chong, B. Lin and J. W. Chua, "Factors affecting consumers' perception of electronic payment: An empirical analysis," Internet Research, **23**(4), 465-485, 2013.doi: 10.1108/IJBM-05-2013-0048

[32] N. B. Tasin, "Factors Influecing Customer's Trust in Online Shopping Among Executives in a Bank," Malaysian Journal of Social Sciences and Humanities, 47-60, 2017.doi: 10.47405/mjssh.v2i3.47

[33] A. Parasuraman and C. Colby, Techno-Ready Marketing: How and Why Your Customers Adopt Technolog, New York: The Free Pres, 2002.

[34] E. M. Rogers, Diffusion of Innovations, 5th edition, New York: Free Press, 2003.

[35] C. Kim , K. Takashima and S. Newell , "How do retailers increase the benefits of buyer innovativeness?," Asia Pacific Journal of Marketing and Logistics, 571-586, 2018.

[36] M. Jun and S. Cai, "The key determinants of Internet banking service quality: a content analysis," International Journal of Bank Marketing, 276-291, 2001.doi: 10.1108/02652320110409825

[37] P. Guriting and N. O. Ndubisi , "Borneo online banking: evaluating customer perceptions and behavioural intention," Management Research News, 6-15, 2006. doi: 10.1108/01409170610645402

[38] S. Ram and J. .. sheth, "Consumer resistance to innovations: the marketing problem and its solution," Journal of Consumer Marketing, 5-14, 1989. doi:10.5281/zenodo.1094243

[39] S. L. Jarvenpaa, N. Tractinsky and M. Vitale, "Consumer trust in an Internet store," Information Technology and Management, 45-71, 2000. doi: 10.1023/A:1019104520776

[40] S. Sarin, T. Sego and N. Chanvarasuth, "Strategic Use of Bundling for Reducing Consumers' Perceived Risk Associated with the Purchase of New High-Tech Products," Journal of Marketing Theory & Practice Theory & Practice, 71-83, 2003. doi:10.1080/10696679.2003.11658502

[41] S. B. Berkup, "Working With Generations X And Y In Generation Z Period: Management Of Different Generations In Business Life," Mediterranean Journal of Social Sciences, **5**(19), 218-229, 2014. doi: 10.5901/mjss.2014.v5n19p218

[42] J. C. Nunnally, Assessment of Realibilty, IN: Pschometric Theory, New York: McGraw-Hill, 1978.

[43] J. Hair, R. Anderson, R. Tatham and W. Black, Multivariate data analysis, NY: Macmillan, 1998.

[44] K. Baskaran, "The impact of digital transformation in Singapore e-tail market," International Journal of Innovative Technology and Exploring Engineering (IJITEE), **8**(11), 2320-2324, 2019.

[45] A. R. Chaudhry, B. Rajput and R. Mishra, "Influence of IoT &AI in place making and creating Smart Cities," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, 2019.

[46] A. Mehrotra, "Artificial Intelligence in Financial Services – Need to Blend Automation with Human Touch," in 2019 International Conference on Automation, Computational and Technology Management (ICACTM), London, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1690-1697 (2020)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Analysis and Evaluation of Competitiveness in Medical Tourism Industry in Taiwan

Yen-Hung Chen*, Tin-Chang Chang

Department of Business Administration, Asia University, 500, Lioufeng Rd., Wufeng, Taichung, 41354, Taiwan

ARTICLE INFO

Article history:
Received: 09 November, 2020
Accepted: 20 December, 2020
Online: 28 December, 2020

Keywords:
Medical Tourism
Hospital Management
VIKOR
Entropy
SWOT
COVID-19

ABSTRACT

Tourism is relatively high profit, low-cost and high work opportunity industry. Medical tourism is special kind of business type in tourism industry. The characteristic of medical tourism is high threshold and high profit for hospital and government. But, fewer literatures research about competitiveness of medical tourism industry until now. Especially, medical tourism industry faces big challenge when COVID-19 occurs which stop patient to go abroad for treating in the world. This research arranges relative literatures and invite some expert's opinion to design criteria for evaluating competitiveness of medical tourism industry in each area. After that, linguistic VIKOR and Entropy will be integrated to analyze and evaluate performance of medical tourism industry among China, Taiwan, Japan and South Korea. Finally, some reasonable suggestion will be provided for Taiwan government and hospital.

1. Introduction

Compare with other area in Asia, medical service in Taiwan is relatively good whose price–performance ratio is high in the world. Based on health care index from CEOWORLD magazine, the rank of Taiwan is in top levels among eighty regions in the world (Please refer to Figure 1). This magazine investigates and believe that Taiwan's health care standard ranks first in medical equipment, medical personnel technology and medical expenses fields. Based on this information, Taiwan government can consider to apply this strength of Taiwan medical service to develop economics.

Business volume of tourism relative industry such as accommodation industry, transportation industry, catering industry and amusement industry will increase when government develop tourism industry [2]. Besides, a huge amount of work opportunity can be generated by tourism industry. Based on above reason, medical tourism become the main development business activity in Taiwan. However, medical tourism industry faces big challenge when COVID-19 occurs which stop tourist to travel in the world and foreign patient can't go to Taiwan for treatment [3]. It needs a novel criteria and framework to evaluate competitiveness of medical tourism industry for satisfying the sustainable development goal. The goal (research objectives) of this research is to develop a framework to evaluate and understand competitiveness of medical tourism industry under the sustainability development condition in Taiwan.

The content of this research is organized as follows. Some literatures about medical tourism evaluation will be arranged and discussed in chapter 2. After that, some operation content about linguistic variable and proposed method will be explained in chapter 3. The competitiveness of medical tourism industry among China, Taiwan, Japan and South Korea will be analyzed and evaluated based on proposed method and SWOT analysis about Taiwan medical tourism industry will be executed in chapter 4. Some management implementation will be token over in chapter 5. Finally, conclusion and future research will be discussed as ending.

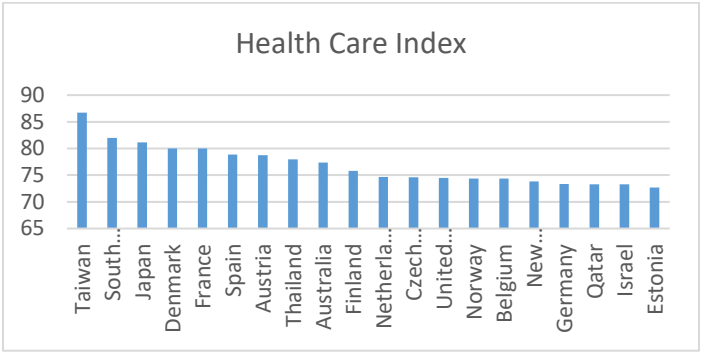


Figure 1: Rank of medical service in the world [1]

*Corresponding Author: Yen-Hung Chen, chenyenhung1234@gmail.com

2. Literature Review

In 2005, the business volume of global medical market and healthcare market reached US$200-300 billion. In the same time point, the tourist volume of the global medical tourism industry market exceeds 19 million person per year and business volume of medical tourism industry approaches US$20 billion [4]. Enterprise in medical tourism industry should manage and evaluate its performance. So, there are various literatures research about medical tourism evaluation has been executed under this development trend. Below literatures are relative research about medical tourism evaluation.

In [5], the author collected relative literatures and surveyed relative experts' opinions to build 33 evaluation criteria. And then, fuzzy VIKOR were applied to evaluate service quality of public medical tourism hospital. In [6], the author developed a comprehensive framework to assess the factors affecting health tourism development in Yazd province in China. Their method integrated interpretive structural modeling (ISM) with fuzzy TOPSIS to evaluate medical tourism competitive ability. Research result showed that patient satisfaction is the most important factors for developing medical tourism. In [7], the author tried to validate the measurement model for medical tourism research. In their research, the research dimension about human capital, physical infrastructure and patient satisfaction had been validated by confirmatory factor analysis (CFA). Future scholar can use their questionnaire to test performance of medical tourism industry. In [8], the author used strengths, weaknesses, opportunities and threats (SWOT) analysis to evaluate business status of medical tourism industry in Turkey. After that, analytical hierarchy process (AHP) were applied to analyze and rank the importance of each factor which will influence the development of tourism industry in Turkey.

In [9], the author took Indian medical tourism hospital as research target. This scholar applied SWOT analysis to evaluate the development status of each organization. And then, technique for order preference by similarity to ideal solution (TOPSIS) were combined with AHP to rank the performance of each hospital. In [10], author considered that medical tourism industry is normal business phenomenon because some country face long waiting queues, high treatment costs or lack of insurance in some country. They applied SWOT model and best-worst techniques to develop and analyze medical tourism industry development strategy in Yazd province of Iran. In [11], the author collected relative information of different city in India and used rough number to integrate experts' opinions. Analytic hierarchy process was applied to evaluate weigh of each medical tourism criteria and multi‐attributive border approximation area comparison methods were used to rank performance of medical tourism site. According to analysis result, we can know that "quality of infrastructure of healthcare institutions", "supply of skilled human resources" and "new job creations" were three most important factor to select medical tourism destination. In [12], the author collected relative information to build medical tourism evaluation criteria and investigated fuzzy TOPSIS with decision making trial and evaluation laboratory (DEMATEL) to analyze the key factor for developing medical tourism industry in Malaysia. According to their research result, human and technological factors are most important key factors for medical tourism industry. In [13], author used SERVQUAL methodology to evaluate service quality of hospital and designed twenty-eight criteria to build hospital service quality. Fuzzy AHP and Elimination and Choice Translating Reality (ELECTRE-I) were employed to evaluate and rank performance of each hospitals. According to experiment result, trustworthiness and empathy were the most important dimension for patient to select hospital.

According to above information, we can know that a lot of literatures discuss about medical tourism in each dimension. They used each kind of multi criteria decision making technology to handle each kind medical tourism evaluation problem (such as importance factor analysis, medical service evaluation, hospital performance evaluation and hospital destination decision etc). However, past literatures investigative about medical tourism evaluation problem who usually lack to consider the dimension about sustainable development of medical tourism industry when some special event will stop development of medical tourism such as COVID-19. This study will fill this research gap.

3. Preliminary and Entropy-VIKOR Method

In this study, we will introduce the research tool – SWOT analysis, linguistic variable, entropy method and Entropy-VIKOR Method.

3.1. SWOT analysis

Strengths, Weaknesses, Opportunities and Threats Analysis (SWOT Analysis) are the most useful analysis tool. This tool has been applied in each kind of fields such as hospitality industry evaluation [14], tourism development strategy [15], medical tourism assessment [8] etc. SWOT is the comprehensive analysis tool which think over relative conditions of relevant competitors and external environmental conditions of research target. This method allows research target to analyze, examine and investigate external opportunities and potential threats. After that, research target will understand and recognize their own advantages and disadvantages for analyzing its competitiveness (Refer to Table 1).

Table 1: SWOT analysis concept

		Impact on Research Target	
		Good for Research Target	Harmful to Research Target
Sources of Influencing Factors	Within the Research Target	Strengths	Weaknesses
	External Environment	Opportunities	Threats

3.2. Linguistic Variable

Linguistic variable is useful tool. This tool can let expert to express their opinion under the same evaluation scale and integrate experts' opinion easily [16]. Below content are the definition, notation and operation about linguistic variable.

Definition 1. The η scale linguistic term set S is the set of linguistic variables. S can be described by following equation [17] and [18].

$$S=\{s_0^{\eta}, s_1^{\eta}, \dots, s_{\eta-1}^{\eta}\} \quad (1)$$

where $s_{\mathfrak{z}}^{\eta}$ ($\mathfrak{z} = 0,1, \dots, \eta-1$) is η scale linguistic variable and η is the total volume of linguistic variable in S.

Definition 2. △ is the linguistic transfer function. This function can transfer linguistic variable into crisp value n (n is between 0 and 1). This function can be described by following equation [17] and [18].

$$\Delta(s_z^{\eta}) = \frac{z}{\eta-1}=n \quad (2)$$

Definition 3. $\triangle^{-1}$ is the linguistic transfer inverse function. This function can transfer crisp value n (n is between 0 and 1) into linguistic variable. This function can be described by following equation [17] and [18].

$$\triangle^{-1}(n)=s_{n*(\eta-1)}^{\eta} \quad (3)$$

3.3. Entropy Method

Entropy Method is one kind of weight evaluation technology. Under this method, importance of each criterion will be decided according to information volume in each criterion. So, multi criteria decision making (MCDM) technology integrated with entropy method can distinguish the performance of each alternative easily [19]. Below content are notation and operation about entropy method.

Definition 4. Suppose that φ_q is entropy value for criterion q. If information in criterion q is low, then φ_q is low. Entropy function should be described by following equation [20] and [21].

$$\varphi_q=-1 * \frac{1}{lnM}\sum_{i=1}^{M} f_{iq} ln(f_{iq}) \quad (4)$$

where $f_{gq}=\frac{x_{gq}}{\sum_{i=1}^{M} x_{iq}}$.

Definition 5. Weight (importance) of each criterion should be decided by entropy value. The weight of criterion q will be acquired by following equation [20] and [21].

$$w_q=\frac{1-\varphi_q}{\sum_{z=1}^{N}(1-\varphi_z)} \quad (5)$$

where N is the total volume of criteria.

3.3. Entropy-VIKOR Method

In this research, Entropy-VIKOR Method are selected to cope with medical tourism competitiveness evaluation problem. In Entropy-VIKOR Method, some notation will be applied and can be described as Table 2.

Table 2: Notation in Entropy-VIKOR Method

Set	Notation	Detail Description
Research target	$T=\{T_1, T_2, \dots, T_m\}$	m represents volume of research target
Criteria	$C=\{C_1, C_2, \dots, C_n\}$	n represents volume of criteria
Weight criteria	$W=\{w_1, w_2, \dots, w_n\}$	n represents volume of criteria
Decision matrix	$X = \begin{matrix} & C_1 & C_2 & \dots & C_n \\ T_1 \\ T_2 \\ \dots \\ T_m \end{matrix} \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \dots & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \dots & \tilde{x}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{x}_{m1} & \tilde{x}_{m2} & \dots & \tilde{x}_{mn} \end{bmatrix}$	$\tilde{x}_{ij}$ represents performance of research target T_i with respect to C_j.
Expert	$E=\{E_1, E_2, \dots, E_K\}$	K represents volume of experts

At first, research targets should be decided by organization. Organization should organize the evaluation project team. After that, some experts will be invited into this team. After experts discuss with each other, criteria for evaluating competitiveness of medical tourism industry can be arranged and selected. And then, experts' opinion will be gathered for evaluating performance of each research target respect to each criterion. At last, Entropy-VIKOR method will be used to evaluate competitiveness of medical tourism industry among each research target in Asia. SWOT analysis will be executed for detailly evaluating competitiveness of medical tourism industry among each research target. Research result will be provided to organization for reference. Relative execution process can refer to Figure 2.

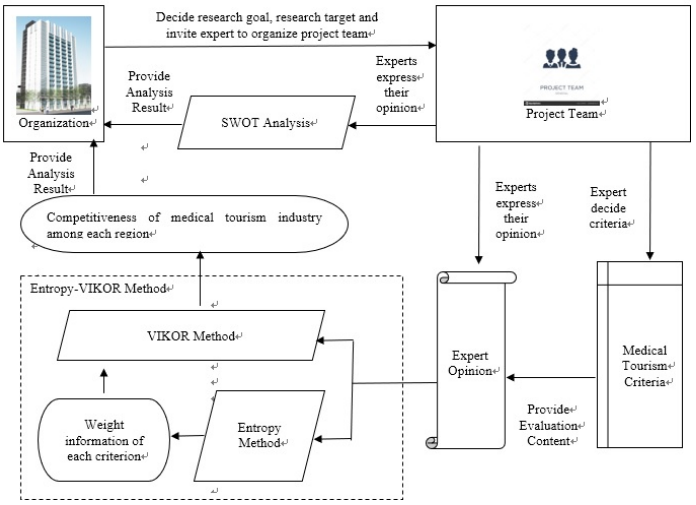


Figure 2: Execution process of proposed method

Let $\tilde{x}_{ijk}$ be expert k's opinion for performance of region i with respect to criterion j. Experts' opinion will be integrated by following equation.

$$\tilde{x}_{ij}=\triangle^{-1}\left(\sum_{k=1}^{K}\frac{\Delta(\tilde{x}_{ijk})}{K}\right) \quad (6)$$

In this research, weight of criteria will be calculated and acquired by entropy method. Weight evaluation function can be describing according to following equation.

$$w_q=\frac{1-\varphi_q}{\sum_{z=1}^{n}(1-\varphi_z)} \quad (7)$$

where $L_{iq}=\frac{\Delta(\tilde{x}_{iq})}{\sum_{i=1}^{m}\Delta(\tilde{x}_{iq})}$, $\varphi_q=-1 * \frac{1}{\ln(m)}\sum_{i=1}^{m} L_{iq} ln(L_{iq})$.

The advantage of using entropy method to acquire the weight of each criteria is that entropy method can decide the weight of each criteria automatically for reducing the evaluation burden of experts.

Below content is relative execution step in VIKOR method.

Positive ideal solution (PIS) is the importance index in VIKOR method. PIS represents the ideal performance of virtual best research target (alternative). PIS can be calculated according to following equation.

$$f_j^*=max_i \triangle^{-1}(\tilde{x}_{ij}) \quad (8)$$

Negative ideal solution (NIS) is another importance index in VIKOR. NIS represents the poorest performance of virtual worst research target (alternative). NIS can be calculated according to following equation.

$$f_j^-=min_i \triangle^{-1}(\tilde{x}_{ij}) \quad (9)$$

Total performance of research target (the largest group utility) S_i represents the overall performance of research target T_i. S_i can be calculated according to following equation.

$$S_i=\sum_{j=1}^{n} w_j * \frac{f_j^*-\triangle^{-1}(\tilde{x}_{ij})}{f_j^*-f_j^-} \quad (10)$$

The individual regret value of research target (the smallest individual regret) R_i represents the maximum regret degree of decision maker if he/she select research target T_i. R_i can be calculated according to following equation.

$$R_i=max_j\left[w_j * \frac{f_j^*-\triangle^{-1}(\tilde{x}_{ij})}{f_j^*-f_j^-}\right] \quad (11)$$

Total performance of research target S_i and individual regret value of research target R_i respectively means different consideration for making decision. So, expert can rank the performance of research target according to largest group utility S_i and smallest individual regret R_i simultaneously. Based above view point, it generates the compromise result Q_i. Q_i should be calculated for ranking performance of each research target based on following equation.

$$Q_i=v*\frac{S_i-S^*}{S^--S^*}+(1-v)*\frac{R_i-R^*}{R^--R^*} \quad (12)$$

where v represents decision parameter. This parameter is used to make decision for adjusting the importance between S_i and R_i. $S^*=min_i S_i$, $S^-=max_i S_i$, $R^*=min_i R_i$, $R^-=max_i R_i$.
Research target T_i will be better if Q_i is fewer.

4. Competitiveness Analysis of Medical Tourism Industry Among Asia Region

Competitiveness of medical tourism industry in Taiwan will be evaluated by an organization. After discussing and considering, four regions have been chosen as research target to evaluate their competitiveness. Those research targets are South Korea (T_1), Japan (T_2), Taiwan (T_3), China (T_4). Above research target are regions in Asia which are main medical tourism competitor for Taiwan.

The execution process of proposed method is as follows:
Step 1. Organize evaluation project team

In the beginning, Taiwan government organize evaluation project team and invite nine experts to execute this evaluation project. Experts in this team include government officer, employees in traditional hotel, employees in travel agency, nurse leader, doctor (hospital dean) and university professors. The investigation target covered already include tourism industry, medical industry, government and university. So, respondents in this survey are representative (Refer to Table 3).

Table 3: Interviewee related information

no	attribute	Content
1	Industry	Hospital dean
2	Industry	Nurse leader
3	Industry	High level employees in travel agency
4	Industry	High level employees in hotel
5	Government	Staff in the Ministry of Economy
6	Government	Staff in Tourism Bureau
7	Government	Staff in the Ministry of Health and Welfare
8	Academia	Professor in department of business management
9	Academia	Professor in department of nursing management

Step 2. Execute evaluation criteria decision

After collecting literatures and experts discuss with each other, eight criteria are employed to evaluate performance of research target. The criteria include medical industry image (C_1), medical ability (C_2), telemedicine ability (C_3), medical cost (C_4), financial status in hospital and travel agency (C_5), volume of medical tourism traveler per year (C_6), medical service speed in hospital (C_7) and maximum service volume in region (C_8). Above criteria are suitable to evaluate competitiveness of medical tourism industry because of below reason.

(a) medical industry image

Patient's medical industry image in this region will influence his/her intention to go to the region [22] and [23], foreign hospital does not have opportunity to service this patient if he/her do not go to this region.

(b) medical ability

Some patients want to receive high quality medical service, so medical ability of medical industry in this region is the main consideration when patients select medical tourism destination region [24] and [25].

(c) telemedicine ability

Hospitals need to treat their foreign patients and maintain relationship with those patients. However, COVID-19 causes that medical tourism patient cannot go abroad to receive face-to -face medical service. So, telemedicine ability is more and more important for region to develop medical tourism industry in order to maintain relationship with foreign patients and continuously treat those patients.

(d) medical cost

In some region, medical cost is very expensive for citizen [26] and [27]. So, patient has the intention to cure his/her disease in foreign region. Medical cost is the main consideration for choosing medical tourism destination region.

(e) financial status in hospital and travel agency

Some hospital's business volume decreases because COVID-19 occurs and foreign patient cannot go abroad to receive medical service. So, robust financial status in hospital and travel agency is key factor if medical tourism industry wants to sustainable development.

(f) volume of medical tourism traveler per year

Volume of medical tourism traveler in this region is the important index to evaluate medical tourism development in this region. High volume of medical tourism traveler not only can reduce medical cost but also can increase potential foreign patient because medical tourism traveler will execute "buzz marketing" if this traveler satisfies medical service in this region.

(g) medical service speed in hospital

Foreign patients go abroad for receiving medical service will request service quality. In some literatures, it justified that medical service time will negative influence patient safety and service quality perception [28]. So, controlling medical service speed in the reasonable range is the important competitiveness of hospital and it will influence the development of medical tourism industry in this region.

(h) maximum service volume in the region

The maximum medical service volume in the region is important index because this index will decide maximum business volume of medical tourism industry in this region. According to scale economics view point, high business volume of medical tourism industry can reduce medical cost in this region because hospital can purchase a huge amount of medicine and medicine price will reduce by bulk purchase [29].

Step 3. Collect information

In this study, nine scale linguistic variables are employed for experts to express their opinions [30] and [31]. Expert's opinion about competitiveness of each research target (region) in medical tourism industry respect to each criterion can refer to Table 4.

Table 4: Performance of each research target

	Expert 1's opinion				Expert 2's opinion			
	South Korea	Japan	Taiwan	China	South Korea	Japan	Taiwan	China
C_1	EG	EG	P	MG	MG	F	EG	VG
C_2	MP	MG	MP	EG	MG	EG	EG	EG
C_3	MP	MG	MG	EG	MG	F	EP	EG
C_4	G	EG	EG	MG	MP	MG	VG	MG
C_5	MG	F	MP	P	F	MG	EG	EP
C_6	EG	F	MP	EP	EG	F	EP	EP
C_7	EG	F	VG	F	G	MP	MP	EG
C_8	MG	G	MG	MP	MG	EG	EG	EP
	Expert 3's opinion				Expert 4's opinion			
	South Korea	Japan	Taiwan	China	South Korea	Japan	Taiwan	China
C_1	VG	MP	F	EP	MP	F	VG	F
C_2	EG	MG	MG	EP	F	G	VG	EP
C_3	MG	MG	G	G	MP	G	VG	MG
C_4	EP	MG	G	MG	F	G	G	F
C_5	MG	MP	EP	F	MG	MP	EG	VG
C_6	VG	F	VP	MP	G	MG	F	MP
C_7	MP	EG	MG	VG	EP	VG	F	P
C_8	EG	MG	F	MG	G	G	MG	G
	Expert 5's opinion				Expert 6's opinion			
	South Korea	Japan	Taiwan	China	South Korea	Japan	Taiwan	China
C_1	F	MG	F	MP	P	VG	VG	F
C_2	EG	VG	MP	F	EG	EG	MP	EG
C_3	MG	F	F	MG	MP	MG	EP	MG
C_4	EP	P	EG	VG	VG	VG	MP	MG
C_5	EG	VG	EP	EP	MP	P	MP	EG
C_6	MG	MG	P	EP	EP	MG	P	EP
C_7	F	EG	MG	F	F	MP	F	P
C_8	MP	F	MG	VG	VG	F	P	MG
	Expert 7's opinion				Expert 8's opinion			
	South Korea	Japan	Taiwan	China	South Korea	Japan	Taiwan	China
C_1	MP	F	VG	P	VG	G	P	MP
C_2	MG	VG	MP	MP	F	EG	MG	MP
C_3	EG	F	MG	EG	MG	VG	MG	VG
C_4	EP	P	F	F	F	F	EG	VG
C_5	EG	MG	MP	EP	MG	EG	F	VG
C_6	EP	F	EP	P	VG	MG	EP	MP
C_7	EP	MP	EG	P	MP	EG	VG	MG
C_8	EP	VG	P	VG	P	VG	F	P
	Expert 9's opinion							
	South Korea	Japan	Taiwan	China				
C_1	EG	G	MG	MG				
C_2	F	EG	MP	MP				
C_3	F	MG	F	VG				
C_4	MP	MP	VP	MG				
C_5	G	MP	MP	P				
C_6	VG	EG	EP	EP				
C_7	P	P	F	F				
C_8	MG	VG	VG	EG				

Step 4. Calculate weight (importance) of each criterion

In this work, we use entropy method (Refer to equation 7) to calculated weight of each criteria. Analysis result can refer to Table 5.

Table 5: Weight of each criterion

Criteria	C_1	C_2	C_3	C_4
Weight	0.1246	0.1902	0.1600	0.1190
Criteria	C_5	C_6	C_7	C_8
Weight	0.0656	0.1043	0.0906	0.1457

In this study, the rank of each research target (region) in medical tourism industry is decided by VIKOR method. The advantage of VIKOR is that VIKOR can solve decision problems with conflicting criteria. In this research, medical ability (C_2) and medical service speed in hospital (C_7) are conflicting criteria. Industry image (C_1) and medical cost (C_4) are conflicting criteria.

Step 5. Calculate PIS and NIS

Positive ideal solution (PIS) is the ideal status for medical tourism industry in Asia. On the opposite direction, negative ideal solution the worst status for medical tourism industry in Asia. PIS and NIS are reference points used to evaluate performance of each medical tourism industry (research target). Equation 8 and equation 9 are employed to calculate PIS and NIS. Analysis result can refer to Table 6.

Table 6: PIS and NIS

PIS	C_1	C_2	C_3	C_4
	0.7344	0.9688	0.9219	0.7969
	C_5	C_6	C_7	C_8
	0.7656	0.7500	0.7344	0.8438
NIS	C_1	C_2	C_3	C_4
	0.5156	0.5781	0.5625	0.4219
	C_5	C_6	C_7	C_8
	0.4688	0.1719	0.4688	0.6406

Step 6. Calculate total performance of research target S_i and the individual regret value of research target R_i .

Total performance of research target S_i means the entire performance of research target T_i. Individual regret value of research target R_i means the maximum regret degree of experts if he/she select research target T_i. S_i and R_i separately represents performance analysis result of research target based on different analysis direction. Equation 10 and equation 11 are employed to calculate S_i and R_i. Analysis result can refer to Table 7.

Step 7. Calculate Q value

In this research, decision parameter v is set up as 0.5. We consider that the evaluation index S_i and R_i is important equally. Equation 12 are employed to calculated Q value. Analysis result can refer to Table 7. According to analysis result, the rank of each region is T_2>T_1>T_3>T_4. The competitiveness rank sequence of medical tourism industry in Asia is Japan, South Korea, Taiwan and China.

Table 7: Total performance, individual regret value, Q value and rank among each medical tourism industry in Asia

	South Korea (T_1)	Japan (T_2)	Taiwan (T_3)	China (T_4)
Total performance of research target S_i	0.5795	0.1897	0.6310	0.6758
Individual regret value of research target R_i	0.1457	0.0974	0.1674	0.1902
Q value	0.6612	0.0000	0.8310	1.0000
Rank	2	1	3	4

Step 8. SWOT Analysis

According to experts' opinion (Refer to Figure 3), Taiwan medical industry possesses some strength is the dimension of medical cost (C_4) and medical service speed in hospital (C_7). However, Taiwan medical industry should take effort in the dimension of medical industry image (C_1), medical ability (C_2) and telemedicine ability (C_3). Medical ability and telemedicine ability of Taiwan medical industry will influence foreigners' medical industry image for Taiwan. Because volume of medical tourism traveler per year (C_6) in Taiwan is relatively small, Taiwan possesses high growth space to develop in the medical tourism market. Compare with south Korea, Japan and China, Taiwan lack scale economics in medical industry because financial status in hospital and travel agency (C_5), maximum service volume in region (C_8) is relatively small. The disadvantage of scale economics in Taiwan will cause that it is hard to reduce Taiwan's overall medical costs and is not easy to enhance Taiwan's overall medical industry image (Refer to Table 8).

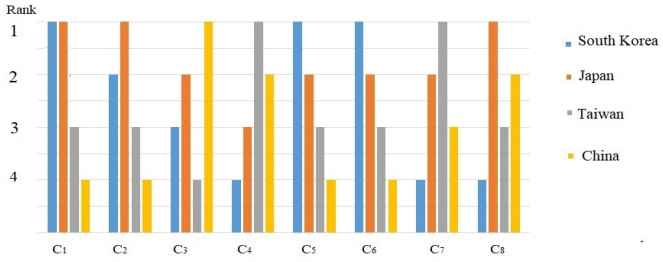


Figure 3: Rank of each region among different dimension

Table 8: SWOT analysis for Taiwan medical tourism industry

Strengths (1) Medical cost is relatively cheap. (2) Medical service speed is relatively fast.	Weaknesses (1) Taiwan medical industry image from tourism traveler is relatively less well-known. (2) Medical ability is weak. (3) Telemedicine ability is weak.
Opportunities (1) Volume of medical tourism traveler is relatively small.	Threats (1) Financial status in hospital and travel agency is relatively small (2) Maximum service volume in region is relatively small

5. Management Implementation

According to experts' opinion, Taiwan possesses high growth space to develop in the medical tourism market because medical cost in Taiwan is relatively cheap which can attract some foreigner whose economics ability is not good enough to receive high price medical service in Japan or South Korea. There are some policies which can promote the competitiveness for Taiwan tourism medical industry.

5.1. Increase volume of medical students each year

There are only 1,300 new students who can enter to study department of medicine in Taiwan each year. Under this total control policy, although Taiwan can maintain high medical standards, it cannot improve Taiwan's medical energy to meet the needs of additional medical services for foreign patients.

5.2. Build overseas medical zone

Taiwan has implemented universal health insurance. Therefore, Taiwan citizen can use low-cost medical services. However, the purpose of developing medical tourism industry is to increase Taiwan's overall GDP. It is unfair to Taiwan citizens if charging standards for foreign patients and Taiwan citizens are the same. Doctor and nurse should use English to serve foreign patients.

This is also a language threshold to develop medical tourism industry.

In order to compete with foreign medical institutions, this study suggests that Taiwan should set up an overseas medical zone. The hospitals in the zone are dedicated to treating foreign patients. It is beneficial to the development for Taiwan's medical tourism industry.

5.3. Development of telemedicine technology

The occurrence of COVID-19 in 2019 has caused the stagnation of the tourism industry in various countries around the world. This event has also affected the development of medical tourism. The medical tourism industry needs to maintain relationship with foreign patients. Telemedicine technology is an important means for hospitals to maintain relationship with foreign patients when international tourism activity cannot be implemented during 2019-2020. Taiwan's medical tourism industry should strengthen and improve telemedicine technology.

5.4. Develop overseas free clinic activities

Taiwan's medical industry is less well-known than south Korea and Japan. Taiwan should go to various countries in southeast Asia and central Asia for free consultations. This will not only enhance Taiwan's medical industry image but also increase the potential patience source for Taiwan's medical tourism industry. The world's low-end industrial manufacturing order may move from China to various countries in southeast Asia and central Asia. Therefore, people from countries in southeast asia and central asia will gradually have the ability to travel to Taiwan for receiving medical service.

6. Conclusion

Medical tourism is very important industry in many countries. A novel framework for evaluating competitiveness of medical tourism industry has been designed in this research. South Korea, Japan, Taiwan and China are selected as research target to evaluate their competitive ability. According to analysis result, we find out that Taiwan's medical tourism competitiveness falls in backward group. So, Taiwan government should provide more resource to improve its medical tourism competitive ability for developing medical tourism industry especially some citizen in other country such as America has high motivation to go abroad for medical treatment because medical expenses is too expensive in America. So, this research provides four suggestion to Taiwan government.

In the future, relative scholar can use this model to evaluate competitiveness of medical tourism industry among Asian countries and European countries. Scholar can analyze how the cultural difference can influence the rank of medical tourism competitiveness.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

I state that the data in this research collect from experts in each field and experts' occupation includes government officer, employees in traditional hotel, employees in travel agency, nurse leader, doctor (hospital dean) and university professors. Other scholars can use the data to execute extended research.

References

[1] I. Sophie, "Revealed: Countries With The Best Health Care Systems. CEOWORLD Manazine," https://ceoworld.biz/2019/08/05/revealed-countries-with-the-best-health-care-systems-2019/, 2019.
[2] H. M. Chen, C. H. Tseng, "The performance of marketing alliances between the tourism industry and credit card issuing banks in Taiwan," Tourism Management, **26**, 15-24, 2005.
[3] S. Gössling, D., Scott, C. M. Hall, "Pandemics, tourism and global change: a rapid assessment of COVID-19," Journal of Sustainable Tourism, 1-20, 2020. https://doi.org/10.1080/09669582.2020.1758708
[4] L. Qi, Y. X. Zhuo, S. Cai, "Taiwan's international medical tourism development strategy," Taiwan Medical Circle, **55**(8), 48-51, 2012. https://doi.org/10.11634/216837861403126
[5] T. H. Chang, "Fuzzy VIKOR method: a case study of the hospital service evaluation in Taiwan," Information Sciences, **271**, 196-212, 2014. https://doi.org/10.1016/j.ins.2014.02.118
[6] A. M. Sharifabadi, F. A. Ardakani, "A model for health tourism development using fuzzy TOPSIS and interpretive structural modeling in Yazd province," Journal of Health Administration (JHA), **17**(55), 73-88, 2014.
[7] Z. Awang, A. Afthanorhan, M. Mohamad, M. A. M. Asri, "An evaluation of measurement model for medical tourism research: the confirmatory factor analysis approach," International Journal of Tourism Policy, **6**, 29-45, 2015.
[8] A. Görener, "A. swot-ahp approach for assessment of medical tourism sector in Turkey," Alphanumeric Journal, **4**, 159-170, 2016.
[9] P. Ajmera, "Ranking the strategies for Indian medical tourism sector through the integration of SWOT analysis and TOPSIS method," International Journal of Health Care Quality Assurance, **30**, 668-679, 2017.
[10] F. Abadi, I. Sahebi, A. Arab, A. Alavi, H. Karachi, "Application of best-worst method in evaluation of medical tourism development strategy," Decision Science Letters, **7**(1), 77-86, 2018.
[11] J. Roy, K. Chatterjee, A. Bandyopadhyay, S. Kar, "Evaluation and selection of medical tourism sites: A rough analytic hierarchy process based multi-attributive border approximation area comparison approach," Expert Systems, **35**, 12232, 2018.
[12] M. Nilashi, S. Samad, A. A. Manaf, H. Ahmadi, T. A. Rashid, A. Munshi, W. Almukadi, O. Othman Ibrahim, O. H. Ahmed, "Factors influencing medical tourism adoption in Malaysia: A DEMATEL-Fuzzy TOPSIS approach," Computers & Industrial Engineering, **137**, 106005, 2019.
[13] A. Singh, "An integrated approach towards ranking hospitals using Fuzzy AHP and ELECTRE-I technique," International Journal of Healthcare Management, 1-10, 2019. https://doi.org/10.1016/j.ins.2014.02.118
[14] J. Takacs, S. A. Vaduva, "A swot analysis of the global hospitality industry," Revista Economica, **69**(6), 105-119, 2017.
[15] A. Reihanian, N. Z. B. Mahmood, E. Kahrom, T. W. Hin, "Sustainable tourism development strategy by SWOT analysis: Boujagh National Park," Iran. tourism management Perspectives, **4**, 223-228, 2012.
[16] P. Liu, F. Teng, "An extended TODIM method for multiple attribute group decision-making based on 2-dimension uncertain linguistic Variable," Complexity, **21**, 20-30, 2016.
[17] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," Learning systems and intelligent robots. Springer, Boston, MA, 1-10, 1974.
[18] F. Herrera, L. Martinez, "The 2-tuple linguistic computational model: Advantages of its linguistic description, accuracy and consistency," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, **9**, 33-48, 2001.
[19] A. I. A1 Maghsoodi, G. Abouhamzeh, M. Khalilzadeh, E. K. Zavadskas, "Ranking and selecting the best performance appraisal method using the MULTIMOORA approach integrated Shannon's entropy," Frontiers of Business Research in China, **12**(1), 2018.
[20] R. Levy, J. P. F. LeBlanc, E. Gull, "Implementation of the maximum entropy method for analytic continuation," Computer Physics Communications, **215**, 149-155, 2017. https://doi.org/10.1016/j.cpc.2017.01.018
[21] W. Huang, B. Shuai, Y. Sun, Y. Wang, E. Antwi, "Using entropy-TOPSIS method to evaluate urban rail transit system operation performance: The China case," Transportation Research Part A: Policy and Practice, **111**, 292-303, 2018.
[22] A. De Nisco, G. Mainolfi, V. Marino, M. R. Napolitano, "Tourism

satisfaction effect on general country image, destination image, and post-visit intentions," Journal of Vacation Marketing, **21**(4), 305-317, 2015.
[23] S. Chaulagain, J. Wiitala, X. Fu, X. "The impact of country image and destination image on US tourists' travel intention," Journal of Destination Marketing & Management, **12**, 1-11, 2019.
[24] C. R. Peters, K. M. Sauer, "A survey of medical tourism service providers," Journal of Marketing Development and Competitiveness, **5**(3), 117-126, 2011.
[25] J. Connell, "Contemporary medical tourism: Conceptualisation, culture and commodification," Tourism management, **34**, 1-13, 2013.
[26] W. Bies, L., Zacharia, "Medical tourism: outsourcing surgery," Mathematical and Computer Modelling, **46**(7-8), 1144-1159, 2007.
[27] D. York, "Medical tourism: the trend toward outsourcing medical procedures to foreign countries," Journal of Continuing Education in the Health Professions, **28**(2), 99-102, 2008.
[28] D. S. Kc, C. Terwiesch, "Impact of workload on service time and patient safety: An econometric analysis of hospital operations," Management Science, **55**(9), 1486-1498, 2009.
[29] H. E. Frech III, L. R. Mobley, "Resolving the impasse on hospital scale economics: a new approach," Applied Economics, **27**(3), 286-296, 1995.
[30] L. Martı, F. Herrera, "An overview on the 2-tuple linguistic model for computing with words in decision making: Extensions, applications and challenges," Information Sciences, **207**, 1-18, 2012.
[31] J. H. Wang, J. Hao, "A new version of 2-tuple fuzzy linguistic representation model for computing with words," IEEE transactions on Fuzzy Systems, **14**, 435-445, 2006.