

Lecture Notes in Electrical Engineering 678

Mohd Azraai Mohd Razman ·
Jessnor Arif Mat Jizat ·
Nafrizuan Mat Yahya · Hyun Myung ·
Amar Faiz Zainal Abidin ·
Mohamad Shaiful Abdul Karim *Editors*

Embracing Industry 4.0

Selected Articles from MUCET 2019

 Springer

Lecture Notes in Electrical Engineering

Volume 678

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India

Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at <http://www.springer.com/series/7818>

Mohd Azraai Mohd Razman ·
Jessnor Arif Mat Jizat · Nafrizuan Mat Yahya ·
Hyun Myung · Amar Faiz Zainal Abidin ·
Mohamad Shaiful Abdul Karim
Editors

Embracing Industry 4.0

Selected Articles from MUCET 2019

Editors

Mohd Azraai Mohd Razman
Faculty of Manufacturing & Mechatronic
Engineering Technology
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

Jessnor Arif Mat Jizat
Faculty of Manufacturing & Mechatronic
Engineering Technology
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

Nafrizuan Mat Yahya
Faculty of Manufacturing & Mechatronic
Engineering Technology
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

Hyun Myung
School of Electrical Engineering
Korea Advanced Institute of Science
and Technology
Daejeon, Korea (Republic of)

Amar Faiz Zainal Abidin
Technical University of Malaysia Malacca
Malacca, Malaysia

Mohamad Shaiful Abdul Karim
Faculty of Electrical and Electronics
Engineering Technology
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-15-6024-8

ISBN 978-981-15-6025-5 (eBook)

<https://doi.org/10.1007/978-981-15-6025-5>

© Springer Nature Singapore Pte Ltd. 2020

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The 11th editions of Malaysian Technical Universities Conference on Engineering and Technology (MUCET 2019) were held in Kuantan, Malaysia, from November 19, 2019, to November 22, 2019. It was jointly organized by Malaysian Technical Universities Network (MTUN) comprising of four universities, namely Universiti Tun Hussein Onn Malaysia (UTHM), Universiti Teknikal Malaysia Melaka (UTeM), Universiti Malaysia Perlis (UniMAP), and, the 11th edition host, Universiti Malaysia Pahang (UMP). MUCET 2019 aims at serving the researchers and practitioners in related fields with timely dissemination of the recent progress of the innovative research in science, engineering, and technology.

The 11th edition of the conference bears a theme of “Communitising Technology in the Context of Industrial Revolution 4.0”. The advancement of Industrial Revolution 4.0 will be driven by smart, interconnected devices that will be affecting local communities. Bringing the communities to adapt to the pervasive environment is indeed a huge challenge for researchers, innovators, technologists, and scientists.

MUCET 2019 received more than 100 submissions. All submissions were reviewed in a single-blind manner, and the best 20 papers recommended by the reviewers are published in this volume. The editors would like to thank all the authors who submitted their papers as the papers are of good quality and represented good progress: civil engineering, chemical and natural resources engineering, mechanical and manufacturing engineering, information and communication technology, science and mathematics, education, social sciences, and technology management toward Industry 4.0.

We also would like to thank Prof. Seeram Ramakrishna, Adjunct Professor Zulkifli Abd. Rani, and Mr. Mohd Parid Sulaiman for delivering their keynote speeches at the conference. It gave a good insight for academia–industry–community roles toward the success of Industrial Revolution 4.0

We hope that readers find this volume informative. We thank Springer for undertaking the publication of this volume. We also would like to thank the conference organization staff and the members of the International Program Committees for their hard work.

Kuantan, Malaysia
Pekan, Malaysia
Pekan, Malaysia
Daejeon, Korea (Republic of)
Malacca, Malaysia
Pekan, Malaysia

Mohd Azraai Mohd Razman
Jessnor Arif Mat Jizat
Nafrizuan Mat Yahya
Hyun Myung
Amar Faiz Zainal Abidin
Mohamad Shaiful Abdul Karim

Contents

An Adaptive Self-assessment Model for Improving Student Performance in Language Learning Using Massive Open Online Course (MOOC)	1
H. Hashim, S. Salam, S. N. M. Mohamad, K. M. Cheong, and P. E. Tan	
Assessment on Average Participation Versus Bialek’s Methods for Transmission Usage Evaluation Scheme	19
N. S. Ahmad, N. H. Radzi, S. A. Jumaat, M. N. Abdullah, and S. Salimin	
Field-Effect Transistor-based Biosensor Optimization: Single Versus Array Silicon Nanowires Configuration	31
Chee Chung Ong, Mohamad Faris Mohamad Fathil, Mohd Khairuddin Md Arshad, Mohammad Nuzaihan Md Nor, Ruslinda A. Rahim, Uda Hashim, Rafizatul Fitri Abdullah, Mohd Hazmi Mohd Ghazali, and Nurulazlina Tamjis	
Top-Down Fabrication of Silicon Nanogap for Detection of Dengue Virus (DENV)	41
Muhamad Nurshahriza Fitri Zulkiffli, Mohammad Nuzaihan Md Nor, Mohamad Faris Mohamad Fathil, Zarimawaty Zailan, Nor Atiqah Md Isa, Conlathan Ibaun, Wan ’Amirah Basyarah Zainol Abidin, Aidil Shazereen Azlan, and Mohd Khairuddin Md Arshad	
Design and Performance Analysis of IoT Based Sensor System Using LoRa	51
Fakrulradzi Idris, Muhammad Syahmi Anuar, Muhammad Hazim Zin Azman, and Norlezah Hashim	
A PSPT-MAC Mechanism for Congestion Avoidance in Wireless Body Area Network	61
Wan Aida Nadia Wan Abdullah, Naimah Yaakob, R. Badlishah Ahmad, Mohamed Elshaikh Elobaid, S. N. Azemi, and Siti Asilah Yah	

A Planar Slotted RCS Based UWB RFID Tag on a PCB and a Flexible Substrate for Packaging Application	73
A. K. M. Zakir Hossain, M. Z. Abidin Bin Abd Aziz, A. W. Y. Khang, and W. A. Indra	
Design and Analysis of Automated Inspection System for Relays Fault Detection	83
W. H. M. Saad, W. Y. Tai, N. M. Saad, M. S. Karis, and S. A. A. Karim	
Visible Light Communication-Based Indoor Notification System for Blind People	93
J. X. Jerry Chong, S. Saon, A. K. Mahamad, M. B. Othman, Nawir Rasidi, and M. Ikhsan Setiawan	
The Anxiety of Parent to the Prolong Computer Games Interaction: A Survey	105
N. S. Razak and M. Rahmah	
Design and Development of an Optical Sensor System to Measure Visibility of Air	115
S. Nurulain, R. M. Ramli, H. Manap, and Mohd Anwar Zawawi	
The Classification of Skateboarding Tricks by Means of Support Vector Machine: An Evaluation of Significant Time-Domain Features	125
Muhammad Amirul Abdullah, Muhammad Ar Rahim Ibrahim, Muhammad Nur Aiman Shapiee, Anwar P. P. Abdul Majeed, Mohd Azraai Mohd Razman, Rabiuh Muazu Musa, and Muhammad Aizzat Zakaria	
Fuzzy Coordination System for Traffic Light	133
Tang Ren Shyang, Adrian Loh Sheng Pang, Nureize Arbaiy, Mohammad Haris Haikal Othman, Mohd Zaki Mohd Salikon, and Pei Chun Lin	
Metaheuristic Optimization Approaches for Overcurrent Relays Discrimination: A Comparative Study	145
Noor Zaihah Jamal, Mohd Herwan Sulaiman, and Abdul Nasir	
A Mutation Triggering Method for Genetic Algorithm to Solve Traveling Salesman Problem	159
Khandker M. Qaiduzzaman, Sabira Khatun, Maliha Afsa, Sadman Sobhan, Md. Elias Hossain, Syamimi Mardiah Shaharum, and Mostafijur Rahman	
Optimized Weight Selection in Levenberg-Marquardt Back-Propagation Using Wolf Search	171
M. Z. Rehman, Aisha Tanveer, Nazri Mohd Nawi, and Kamal Z. Zamli	

**The Classification of Skateboarding Trick Manoeuvres:
A Frequency-Domain Evaluation 183**
 Muhammad Ar Rahim Ibrahim, Muhammad Nur Aiman Shapiee,
 Muhammad Amirul Abdullah, Mohd Azraai Mohd Razman,
 Rabiū Muazu Musa, and Anwar P. P. Abdul Majeed

**The Identification of RFID Signal Using *k*-Means
for Pallet-Level Tagging 195**
 Chun Sern Choong, Ahmad Fakhri Ab. Nasir, Anwar P. P. Abdul Majeed,
 Muhammad Aizzat Zakaria, and Mohd Azraai Mohd Razman

**Flat Price Prediction Using Linear and Random Forest Regression
Based on Machine Learning Techniques 205**
 Julakha Jahan Jui, M. M. Imran Molla, Bifta Sama Bari, Mamunur Rashid,
 and Md Jahid Hasan

**The Classification of Skateboarding Tricks by Means
of the Integration of Transfer Learning and Machine
Learning Models 219**
 Muhammad Nur Aiman Shapiee, Muhammad Ar Rahim Ibrahim,
 Mohd Azraai Mohd Razman, Muhammad Amirul Abdullah,
 Rabiū Muazu Musa, and Anwar P. P. Abdul Majeed

**$\pi 1a$ and $\pi 1b$ Impedance Matching for Capacitive Power
Transfer System 227**
 Yusmarnita Yusop, Shakir Saat, Khairul Kamarudin, Huzaimah Husin,
 and Sing Kiong Nguang

**Application of Smart Rotatable Solar Powered Electrocoagulation
System for Pollutants Removal from Palm Oil Mill Effluent 251**
 B. K. Zaied, Mohd Nasrullah, L. Singh, and A. W. Zularisam

**Variable Frequency Phase-Disposition Pulse Width Modulation
Technique for Multilevel Inverter 265**
 Auni Syahirah Abu Bakar, Syamimi Mardiah Shaharum,
 and Nur Huda Ramlan

**Discrepancy Resolution: A Review of Missing Tags Detection in RFID
Systems for Inventory Shrinkage 277**
 Yusuf Musa, Hazalila Kamaludin, and Hairulnizam Mahdin

An Adaptive Self-assessment Model for Improving Student Performance in Language Learning Using Massive Open Online Course (MOOC)



H. Hashim, S. Salam, S. N. M. Mohamad, K. M. Cheong, and P. E. Tan

Abstract Massive Open Online Course (MOOC) provides an effective learning platform with various high-quality educational materials accessible to learners from all over the world. On the other hand, assessment plays an important role to improve student performance in MOOC learning. However, issues in assessment designs contribute to the lack of student engagement. Hence, a suitable assessment model should be developed to improve student performance in MOOC learning. This study proposes an adaptive self-assessment model based on learner characteristics to improve student performance in language learning using MOOC. A literature review was performed to identify existing learner characteristics, functional features in assessment and elements of learner characteristics. Four research questions have been constructed to assist the study. The results of the study are then used in formulating a conceptual model for an adaptive self-assessment based on MOOC functional features, and elements of students learning styles & cognitive styles. Based on the conceptual model, an adaptive self-assessment model for language learning was produced to build a complete learning design for Mandarin MOOC. The model was validated by two Subject Matter Experts (SMEs) and two Instructional Design Experts (IDEs) who contributed to the production of the complete learning design. The findings of this study are two folds: (i) a conceptual adaptive self-assessment model based on learner characteristics for improving student performance in MOOC learning, and (ii) an adaptive self-assessment model based on learner characteristics to improve language learning using MOOC. The proposed model is meant to guide MOOC developers in assessment design. A complete learning design for Mandarin MOOC that applies the proposed model has also been developed.

H. Hashim (✉) · S. Salam · S. N. M. Mohamad
Centre for Advanced Computing Technology (C-ACT), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
e-mail: hasmainie76@gmail.com

Fakulti Teknologi Maklumat dan Komunikasi (FTMK), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

K. M. Cheong · P. E. Tan
Pusat Bahasa Dan Pembangunan Insan, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

Keywords Adaptive · Self-assessment · Student performance · Language · MOOC

1 Introduction

Current MOOCs are (i) lacking personalized learning guidance, and (ii) intelligent assessment for individuals [1]. Previous researchers listed three emerging issues for MOOC pedagogy: (i) the role of the teacher and (ii) assessment [2]. Assessment plays a very important role in measuring any learning process or measuring the progress of the learning experience [3]. Researchers mentioned that the assessment can increase the performance of students and at the same time it brings benefits to the teacher [4]. Assessment refers to all activities that provide the information required in making a decision about individuals learning [5]. According to Holmes [6], that assessment is (i) what students regard as important and (ii) how they spend their time. There are few challenges for all students in MOOC assessment and how to apply the learning theory into assessment questions [7, 8]. There are three main problems that have been identified in this study. These include (i) poorly designed assessments [9]; (ii) factors contribute to student performance [10]; and (iii) lack of engagement [11, 12]. According to the previous study, the term assessment is purposefully used to refer to the measurement of learner's achievement and progress in a learning process [13]. Two major forms of assessment exist (i) formative and (ii) summative assessments [7]. In MOOC assessment, the issue is, there was a reduction of interest and activity of students during the session of the course [12]. The authors stated the causes of poor student performance in MOOCs are (i) challenged in interactive practical exercises and (ii) reduced engagement level along the course [14]. Using MOOC, student learning time involves more independent learning time than guided learning time. In MOOC learning, student independent time will involve self-study time which mainly requires self-assessment. Therefore, there is a need for a self-assessment model that can be used to guide MOOC developers in assessment designs to give continuous improvement to students in MOOC learning. This study is aimed at proposing an adaptive self-assessment model for improving student performance in language learning using MOOC. To assist the study, the following research questions were constructed:

RQ1: What are the existing functional features for self-assessment in existing MOOC platforms?

RQ2: What are the elements for each dimension of learner characteristics?

RQ3: How to design MOOC self-assessment activities that incorporate elements for selected dimensions of learner characteristics?

RQ4: How to design MOOC self-assessment activities that incorporate assessment functional features?

This paper is organized as follows: Sect. 2 is a review of related work. Section 3 describes the research methods. In Sect. 4, we review the results, discuss the designs

and propose an adaptive self-assessment model based on elements of learner characteristics and functional features in MOOC assessment. Finally, Sect. 5 gives a conclusion of the main finding in this paper.

2 Related Work

Massive Open Online Course (MOOC) is a new online learning method in education that is currently developing [15]. The researcher highlighted, in MOOC assessment there is a big challenge for a large number of students to get detailed and timely feedback [16]. Assessment is an important aspect for every learning process and it has a strong impact on learning [17, 18]. According to Reinholz [19], assessment can be used to (i) evaluate student outcomes and (ii) support student learning. The limitation of experiential learning and many high-impact practices in higher education is lacking assessment for embedded learning outcomes at the individual learner level [20]. Baleni [17], listed two types of assessment exist; which are (i) formative and (ii) summative assessments. Formative assessment monitor student for learning with learning activities while summative assessment used evidence of learning effectively with the test [18, 21].

The previous researcher mentioned that self-assessment helped to (i) effective learning for a learner and (ii) develop assessor judgment for an instructor [22]. In the year 2010, Taras also highlighted that self-assessment has been shown to support student learning. The author mentioned that the term 'self-assessment' is used to cover all judgments by learners of their work which subsumes terms such as (i) 'self-evaluation' and (ii) 'self-appraisal' [23].

Traditionally, models used for adaptive assessment have been mostly summative: they measure or rank effectively examinees, but do not provide any other feedback [24]. In 2017, Liebetrueth mentioned that assessment model has to cover the (i) aspects structure (framework within which the content has to be filled in), (ii) content (criteria that separate a good and a poor execution of the focal function) and (iii) method of the assessment process (process how a maturity assessment is conducted) [25].

In the year 2016, McKevitt listed three elements in the self-assessment model (i) in which students engage in self-assessment of their work, (ii) receive tutor feedback and (iii) then take corrective action [22]. According to Taras, self-assessment models classified into three classifications with five models: (i) weaker, (ii) median and (iii) stronger [25–27]. Table 1 shows a comparison of the five different self-assessment models. In this study, the stronger self-assessment type 4 models, which is a self-assessment model with integrated peer/tutor feedback, has been chosen as a model to improve student performance in language learning using MOOC platform.

Table 1 Classification of different Self-assessment models (adapted from Taras)

Model		Description	Advantages
Weaker	Self-marking	Require an understanding of criteria and standards while making the	Providing students with immediate feedback on their work [23].
	Sound standard	comparison between their work and the models which provide a common ground for dialogic discussions with tutors [26].	Focuses on grading and students learning to understand and appreciate standards and make grading transparent [23].
Median	Standard	Requires students to evaluate their work according to agreed criteria and standards and possibly provide distinctions of stronger and weaker elements, and optionally a grade [27]. Requires learners to use criteria to judge, provide feedback and grade their work prior to submission to tutors [23].	Essentially separate from the tutor assessment and permits learners to consider their work against criteria and standards, and has the advantage of encouraging learner independence [23].
Stronger	Self-assessment model with integrated peer/tutor feedback	Requires integrated peer and tutor feedback prior to self-assessment [26]. Requires learners to integrate tutor and peer feedback before students self-assess [23].	Learners are required to engage with and integrate tutor and peer feedback, and secondly, they are required to self-assess and grade with the benefit of tutors' guild knowledge [23].
	Learning Contract Design	Requires students to make decisions from curriculum choices to assessment design, outcomes and criteria [26].	Controlled by student requirement students, disadvantages: failed to examine the assessment processes in relation to tutors and students within this model [26].

3 Design and Methods

This study is a design-based research that involves qualitative and quantitative methods. The research framework has been divided into four (4) phases which are (i) literature review, (ii) analysis, (iii) design & development and (iv) evaluation.

Literature Review: The study was initiated by conducting a literature review to find answers for RQ1 and RQ2 of the study.

Analysis: Analysis of RQ1 results tabled out the existing functional features for assessment in MOOC platforms. Meanwhile, the results of the RQ2 analysis listed the elements for each dimension of learner characteristics.

Design & Development: Based on findings from the analysis phase, RQ3 and RQ4 are then used to guide the researchers in proposing a conceptual model of self-assessment that is adaptive to learner characteristics to improve student performance. By applying the conceptual model, the researchers built a draft learning design for a Mandarin MOOC.

Evaluation: Two Subject Matter Experts (SMEs) in Mandarin language and two Instructional Design Experts (IDEs) participated actively and played important roles in applying, improving and verifying the adaptive self-assessment model that was used in producing a complete learning design for teaching the Mandarin language via MOOC. Based on the expert feedbacks', a few minor structural changes were made to improve and finalize the model construction.

Figure 1 shows the research design used in this study to develop an adaptive self-assessment model based on learner characteristics to improve student performance using MOOC.

3.1 Existing Functional Features for Assessment in MOOC Platforms

The previous researchers mentioned most of the existing platforms can greatly improve the automatic correction of exercises using artificial intelligence techniques, the adaptation of assessment activities or the gamification features, although some platforms have already done some initial steps in these directions [18]. Authors suggested that in the future, new developments are required to make a perfect solution for assessment in MOOCs and these platforms need to improve in the different aspects and features. Table 2 listed the existing functional features for assessment in various MOOC platforms.

4 Results and Discussion

In this section, the results of the analysis and design phases are presented and discussed based on the research questions.

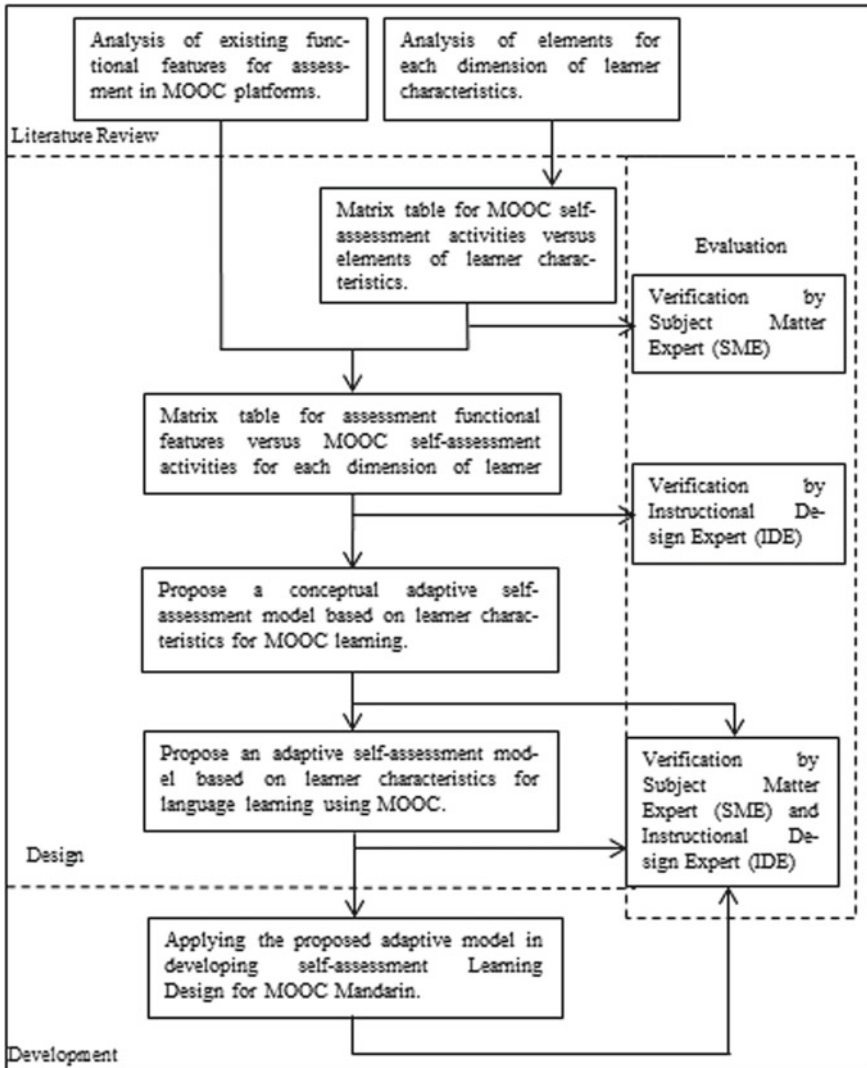




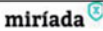































Fig. 1 Research design

4.1 RQ1: What Are the Existing Functional Features for Assessment in Existing MOOC Platforms?

Five popular MOOC platforms were analyzed to find existing functional features for assessment. Figure 2 shows a comparison of the assessment functionality for MOOCs between the MOOC platforms by the indicator (Low ↓, Medium →, and High ↑). Results of the analysis show that Open Learning has the most assessment

Table 2 Existing functional features for assessment in MOOC platforms

Assessment Features MOOC Platform	Features Indicator Availability
Hints, adaptive exercises, etc.	Indicator: Low, Medium, High
Automatic correction of exercises with artificial intelligence	
Interactive exercises	
Peer review	
Gamification	
Adaptation of the learning process and recommendations	
Authoring tools for complex assessment activities	
Learning analytics	
Social interaction functionality	
Setting students own learning objectives and goals	

Assessment features/MOOC Platform					
Hints, adaptive exercises, etc.		-	-		
Automatic correction of exercises with artificial intelligence		-	-	-	
Interactive exercises		-	-		
Peer review				-	
Gamification	-	-	-		
Adaptation of the learning process and recommendations	-	-	-		
Authoring tools for complex assessment activities		-	-	-	
Learning analytics					
Social interaction functionality					
Setting students own learning objectives and goals	-	-	-		-

Legend : (Low ↓ - 1 point, Medium → - 2 points and High ↑ - 3 points)

Fig. 2 Comparison of the assessment functionality in MOOC platforms

features required on a MOOC platform. However, Open Learning is as placed as the second top based on the average score for 10 requirement assessment features on a MOOC platform. For us to do further study, we focused on two assessment features that are provided by our chosen MOOC platform which is Open Learning. The two features are (i) hints and (ii) interactive exercises. The choices are supported by [28] who highlighted that hints are used to support students to learn academic and [29] also highlighted that interactive is the support method to increase students’

understanding of difficult concepts in online education. Meanwhile, many learners also indicated that instructors needed to be more involved, interactive, and responsive in the online course [30].

4.2 RQ2: What Are the Elements for Each Dimension of Learner Characteristics?

In 2015, Park mentioned learner characteristics are important in all forms of online learning [31]. The previous researchers listed three types of learner characteristics that are effective factors for student learning: (i) learning styles, (ii) cognitive styles, and (iii) multiple intelligence. In this study, the researchers have focused on learning styles and cognitive styles only [32].

Elements for the Dimension of Learning Styles. The findings of the first literature study are shown in Table 3. All together, there are 29 existing elements in 8 dimensions of learning styles: Active, Reflective, Sensing, Intuitive, Visual, Verbal, Sequential and Global.

Elements for the Dimension of Cognitive Styles. Table 4 shows our findings of the second literature study. Based on finding, there are 27 existing elements in 8 dimensions of cognitive styles: Extrovert, Introvert, Sensing, Intuitive, Feeling, Judging and Perceiving.

4.3 RO3: How to Design MOOC Self-assessment Activities that Incorporate Elements for the Selected Dimensions of Learner Characteristics?

To answer RQ3, the researchers focused on designing MOOC self-assessment activities for two selected dimensions of learning styles and two selected dimensions of cognitive styles. The chosen dimensions were made based on findings of our preliminary study on 50 students using Felder and Silverman Learning Styles Model and Ancona et al. (1997) cognitive styles questionnaire [41]. Once the dimensions have been determined, the MOOC topic learning outcomes were used as guidance in selecting which types of MOOC self-assessment activities can be incorporated with which learner characteristic dimensions. Table 5 shows the dimensions of learner characteristics that can be adapted to potential MOOC self-assessment activities to improve student performance in language learning using MOOC.

Results of further analysis of each question/item in self-assessment activities for each element in learner characteristic dimensions are presented in Table 6. In this table, a shaded box means that the design of a question for this self-assessment activity (e.g. MC) will incorporate the element of its associated learner characteristic (e.g. See).

Table 3 Existing elements for the dimensions of learning styles

Dimension	Elements	Author				
		[33]	[34]	[35]	[36]	[37]
Active	Trying things	X				
	Impulsive		X			
	Risk-takers		X			
	Do not prefer lectures		X			
	Interpersonal		X			
	Group work	X	X	X	X	X
	Task immediately			X	X	
Reflective	Think before action		X			
	Like writing		X			
	Not inclined to too much note-taking		X			
	Thinking	X	X			X
	Intrapersonal & introspective		X			
	Analytical approach			X	X	
	Working alone	X		X	X	X
Sensing	Concrete material	X	X			X
	Prefer facts			X	X	
	Follow tutors' approaches			X	X	
Intuitive	Abstract material	X	X	X	X	X
	Innovative			X	X	
Visual	See	X				
	Pictorial materials		X	X	X	X
Verbal	Words	X	X			X
	Written and listen			X	X	
Sequential	Step by step	X		X	X	
	Sequenced steps		X			
	Focus on details				X	
	Continual small steps					X
Global	Large leaps, skipping, understand and look at detail	X		X	X	
	Holistically in large jumps		X			X

4.4 RQ4: How to Design MOOC Self-assessment Activities that Incorporate Assessment Functional Features?

To answer RQ4, the researchers focused on designing MOOC self-assessment activities using selected MOOC functional features that are suitable for a particular learner characteristic dimension. As mentioned earlier in sub-section RQ1, two assessment

Table 4 Elements for the dimension of cognitive styles

Dimension	Elements	Author			
		[32]	[38]	[39]	[40]
Extrovert	Outer world	X		X	
	Collaborate with others		X		
	Try things out for himself			X	X
Introvert	Inner world	X		X	
	Independent		X	X	X
	Reflect on thoughts and ideas			X	
Sensing	Real objects and solid facts.	X		X	X
	Physical qualities and affection by other information		X		
	Rely on past experiences			X	X
Intuitive	Possibilities and personal meaning	X			X
	Intuitive types		X		X
	Speculations			X	
Thinking	Analyzing fact	X		X	
	Structure and function		X		
	Logical and rational decisions			X	X
Feeling	Subjective values and views	X		X	
	Initial energetic condition and interactions		X		
	Decisions based on the people and their actions			X	X
Judging	Planned	X			X
	Organized way	X	X	X	X
	Prefers control	X			X
	Seek closure		X		
	Think sequentially		X		X
	Orderly			X	
Perceiving	Flexible	X	X	X	
	Spontaneous way of life	X		X	
	To keep things open ended		X		X

Table 5 Learner characteristics dimensions versus MOOC Self-assessment activities

Dimension	MOOC Self-assessment activities
Visual	Quiz: Multiple Choice (MC), True/False (T/F) Drag & Drop, Matching
Active	Listening Assessment
Thinking	Forum
Intuitive	Mid-Term
	Project: Writing, Presentation

Table 6 Self-assessment activities versus element of learner characteristics in MOOC platform

Dimension	Element	Quiz				Project				
		Multiple Choice	True False	Drag & Drop	Match	Listening Assessment	Forum	Mid-Term	Writing	Presentation
Visual	See Pictorial materials	■	■	■	■					
	Trying things	■	■	■	■	■	■	■	■	■
Active	Impulsive	■	■	■	■	■	■	■	■	■
	Risk-takers	■	■	■	■	■	■	■	■	■
	Do not prefer lectures	■	■	■	■	■	■	■	■	■
Thinking	Interpersonal	■	■	■	■	■	■	■	■	■
	Group work	■	■	■	■	■	■	■	■	■
	Tasks immediately	■	■	■	■	■	■	■	■	■
Intuitive	Analyzing fact	■	■	■	■	■	■	■	■	■
	Structure and function	■	■	■	■	■	■	■	■	■
	Logical and rational decisions	■	■	■	■	■	■	■	■	■
Intuitive	Possibilities and personal meaning	■	■	■	■	■	■	■	■	■
	Existence intuitive types	■	■	■	■	■	■	■	■	■
	Speculations	■	■	■	■	■	■	■	■	■

functional features were selected to be focused in this study: (i) hints and (ii) interactive. Table 7 shows the assessment functional features that can be adapted into potential MOOC self-assessment activities to improve student performance in language learning using a MOOC platform.

Further analysis of each question in self-assessment activities for each functional feature selected has produced results presented in Table 8. In this table, a shaded box means that the design of a question for this self-assessment activity (e.g. MC) will incorporate the feature of its associated learner characteristic (e.g. ‘Hints’ in ‘Graphic’ form).

Development of a Conceptual Model for an Adaptive Self-assessment Model Based on Learner Characteristics. Based on findings from RQ1, RQ2, RQ3, and RQ4, the researchers proposed a conceptual model for an adaptive self-assessment model based on learner characteristics that can be applied by other MOOC developers and MOOC platform developers. Figure 3 shows the proposed adaptive self-assessment model based on learner characteristics to improve student performance

Table 7 Assessment functional features versus activities based on learner characteristics dimension

Dimensions	Features	MOOC Self-assessment activities
Visual	Hints Interactive	Quiz: Multiple Choice (MC), True False (T/F) Drag & Drop, Match
Active		Listening Assessment
Thinking		Forum
Intuitive		Mid-Term
		Project: Writing, Presentation

Table 8 Assessment functional features versus MOOC self-assessment activities

Construct	Features		Quiz					Project			
	Hints	Interactive	Multiple Choice	True False	Drag & Drop	Match	Listening Assessment	Forum	Mid-Term	Writing	Presentation
Visual	Graphic		■	■			■	■	■	■	■
		Animation			■				■		
Active		Drag & Drop			■						
		Match			■						
		Link	■	■			■	■			■
		Upload file					■				
		Down load file	■	■			■				
Thinking	Graphic		■	■			■	■	■	■	■
		Drag & Drop			■						
		Match			■						
Intuitive	Graphic		■	■		■	■	■	■	■	■
		Drag & Drop			■	■					
		Match			■						
		Online Discussion					■				

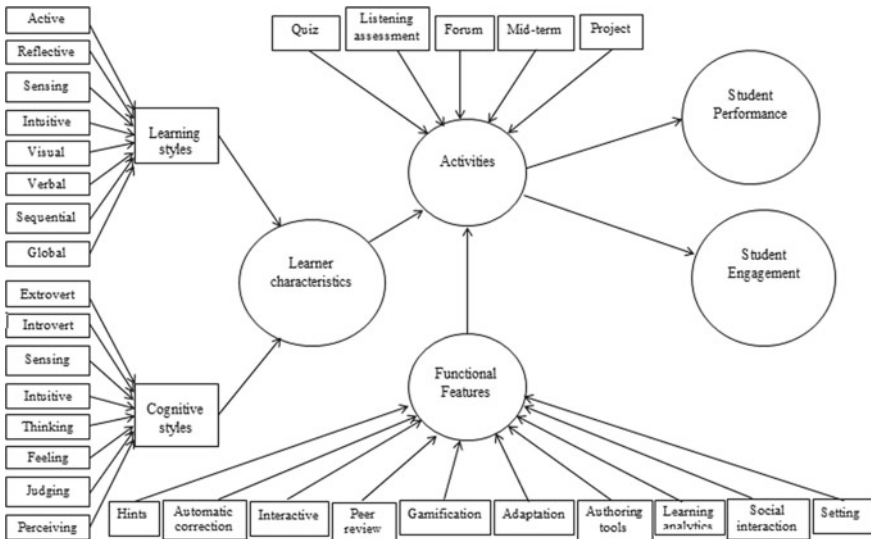


Fig. 3 A Proposed adaptive self-assessment model for improving student performance in Mandarin learning using MOOC



LO	Activities	Features	Item Design
Lesson 1			
LO1	Identify syllable, initials, finals, and tones in Mandarin. Quiz 1 Match the given term with its meaning accordingly.	Hints - Graphic Interactive - Match - Link	<p>1. _____ is a combination of elements such as initial, final, and tone. Hints (Graphic-image)</p>  <p>2. _____ is a tool for people to learn Chinese phonetics. Interactive-link https://www.youtube.com/watch?v=ocgsfnEgqY</p> <p>3. _____ is usually placed at the beginning of a syllable. Hints (keyword) Initial</p> <p>4. _____ is always placed at the end of a syllable. Hints (keyword) Final</p> <p>5. _____ is represented by a different tone mark. Hints (Graphic-image)</p> 

Fig. 4 Example of a storyboard for a question in a Self-assessment activity

in language learning using MOOC. In this model, MOOC self-assessment activities are adaptive to learner characteristics by using functional features in the MOOC platform.

Development of a Storyboard for Each Self-assessment Activity. With reference to matrix Tables 6 and 8, MOOC developers can develop a storyboard for each self-assessment activity. In this study, the researchers proposed a storyboard template that consists of 4 elements: (i) assessment types: formative or summative, (ii) domain of learning: affective, cognitive or psychomotor, (iii) self-assessment activities: quiz, listening assessment, forum, project or assignment, and (iv) indicators: marks, no. of views, no. of comments or time spent. Also, for each question in a self-assessment activity, the researchers proposed another 4 elements: (i) learning outcomes, (ii) questions, (iii) functional features and (iv) item design. Figure 4 shows an example of a storyboard for a quiz question with Learning Outcomes 1 in Lesson 1.

A prototype of learning design for Mandarin language learning that applies the proposed adaptive model was developed to verify its usability. The model was verified by two Subject Matter Experts (SMEs) and Instructional Design Experts (IDEs). SMEs were consulted in verifying the application of Table 6 to each MOOC self-assessment question while IDEs were consulted in verifying the application of Table 8 to each MOOC self-assessment question. Figure 5 shows the Overall Mandarin Learning Design developed in the case study. Table 9 describes the operational definition for each layer in the Overall Mandarin Learning Design in Fig. 5.

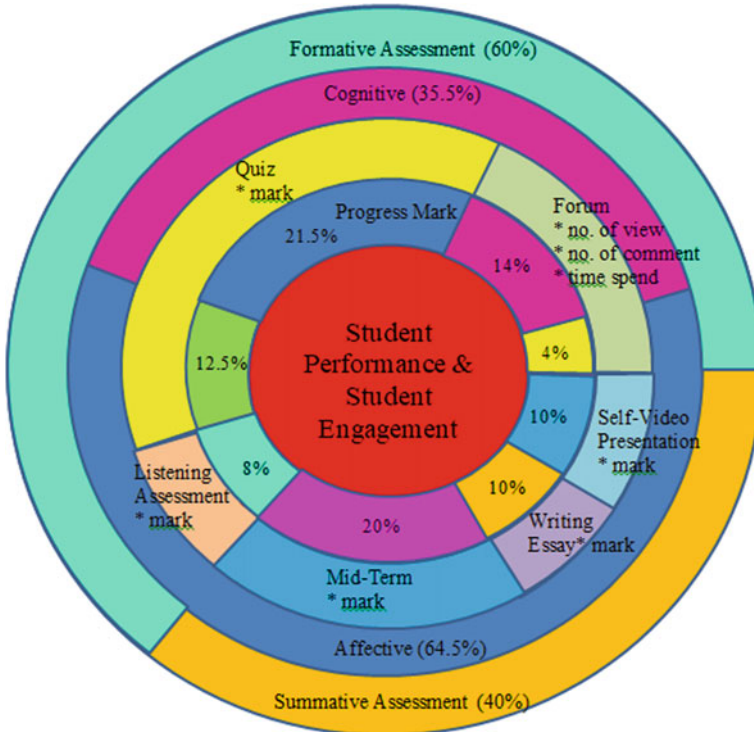


Fig. 5 The overall Mandarin learning design on the proposed adaptive MOOC Self-assessment model

Table 9 The operational definition for each layer in MOOC self-assessment activities design

Layer No.	Label	Operational Definition
5	Student Performance	• Student performance refers to the student progress for each activity, student perception, and self-efficacy.
5	Student Engagement	• Student engagement refers to the degree of attention and interest that students show when they are learning, which extends to the level of motivation they have to learn and progress in their education.
4	Progress Mark	• Progress mark refers to the percentage of student completion in assessment activities using MOOCs.
3	Activities	• In this case, the MOOC self-assessment design consists of 5 types of self-assessment activities with online quizzes, listening assessment, forum, mid-term, and project.
2	Domain	• In this case study, the Course Learning Outcome (CLO) covers 2 types of learning domains only: (i) cognitive and (ii) affective.
1	Assessment	• In this case study, the assessment types used are: (i) formative assessment and (ii) summative assessment.

5 Conclusion

This study presents the findings on the development of an adaptive self-assessment model based on learner characteristics for improving student performance in language learning using MOOC. The formulation takes account of learning styles and cognitive styles to be adapted into the learning design of Self-assessment activities in language learning using MOOC. These activities based on learner characteristics (learning styles and cognitive styles) dimensions and assessment feature (hints and interactive) elements. The framework also incorporates the roles of SMEs and IDEs to ensure the proposed model is practical and applicable. The proposed model is meant to guide MOOC developers in assessment design. This model consists of type assessment, taxonomy bloom, online activities method, % progress mark and student (performance and engagement). In the future, student performance and engagement will be investigated using that adaptive MOOC self-assessment model.

References

1. Wang Z (2016) Structured knowledge tracing models for student assessment on Coursera. In: Proceedings of the third (2016) ACM conference on learning @ scale, pp 209–212
2. Leppisaari I (2017) Searching for effective peer assessment models for improving online learning in HE—Do-It-Yourself (DIY) case. *E-Learn* 2017:53–65
3. Moreno-Ger P, Burgos D, Martínez-Ortiz I, Sierra JL, Fernández-Manjón B (2008) Educational game design for online education. *Comput Human Behav* 24(6):2530–2540
4. Tenorio T, Bittencourt II, Isotani S, Silva AP (2016) Does peer assessment in on-line learning environments work? A systematic review of the literature. *Comput Human Behav* 64:94–107
5. Orooji F, Taghiyareh F (2018) Peer assessment and self-assessment in social learning environments through a new crowd-sourced mechanism. *IEEE Access* 6:7321–7339
6. Holmes N (2017) Engaging with assessment: increasing student engagement through continuous assessment. *Act Learn High Educ*, p 146978741772323
7. Admiraal W, Huisman B, Van De Ven M (2014) Self- and peer assessment in massive open online courses. *Int J High Educ* 3(3):119–128
8. Admiraal W, Huisman B, Pilli O (2015) Assessment in massive open online courses. *Electron J e-Learning* 13(4):207–216
9. Aldahdouh AA, Osório AJ (2016) Planning to design MOOC? Think first! *Online J Distance Educ e-Learning* 4(2):47–57
10. Bozkurt A, Akgün-özbek E, Zawacki-Richter O (2017) Trends and patterns in massive open online courses: review and content analysis of research on MOOCs (2008–2015). *Int Rev Res Open Distance Learn* 18(5):120–147
11. W. W. Goh, S. Y. Wong, and E. Ayub, “The Effectiveness of MOOC Among Learners Based on Kirkpatrick’s Model,” in *Redesigning Learning for Greater Social Impact*, 2018, pp. 313–323
12. Lisitsyna LS, Evgenii AE (2017) Making MOOCs more effective and adaptive on basis of SAT and game mechanics. *Smart Educ e-Learning 2017, Smart Innov Syst Technol* 75
13. Gikandi JW, Morrow D, Davis NE (2011) Online formative assessment in higher education: a review of the literature. *Comput Educ* 57(4):2333–2351
14. Barreto D, Vasconcelos L, Orey M (2017) Motivation and learning engagement through playing math video games. *Malaysian J Learn Instr* 14(2):1–21
15. Xiao F, Pardamean B (2016) MOOC model: Dimensions and model design to develop learning. *New Educ Rev* 43(1):28–40

16. Wong BT (2016) Factors leading to effective teaching of MOOCs. *Asian Assoc Open Univ J* 11(1):105–118
17. Baleni ZG (2014) Online Formative Assessment in Higher Education: Its Pros and Cons. *Proc. Eur. Conf. e-Learning* 13(4):43–47
18. Muñoz-Merino PJ, Rui Pérez-Valiente JA, Sanz JL, Delgado Kloos C (2015) Assessment activities in MOOCs
19. Reinholz D (2015) The assessment cycle: a model for learning through peer assessment. *Assess Eval High Educ* 2938:1–15
20. Heinrich WF, Rivera JE (2016) A framework for PLA in traditional higher education: experiential learning assessment for embedded outcomes. *Prior Learn Assess Inside Out* 5(5):1–5
21. Louhab FE, Bahnasse A, Talea M (2018) Towards an adaptive formative assessment in context-aware mobile learning. *Proc Comput Sci* 135:441–448
22. McKevitt C (2016) Engaging students with self-assessment and tutor feedback to improve performance and support assessment capacity. *J Univ Teach Learn Pract* 13(1):1–20
23. Taras M (2010) Student self-assessment: processes and consequences. *Teach High Educ* 15(2):199–209
24. Vie J-J, Popineau H, Bruillard É, Bourda Y, Jill-Jënn (2017) A review of recent advances in adaptive assessment. *Learn Anal Fundam Appl Trends* 94:113–142
25. Liebethuth T (2017) Sustainability in Performance Measurement and Management Systems for Supply Chains. *Procedia Eng.* 192:539–544
26. Taras M (2016) Situating power potentials and dynamics of learners and tutors within self-assessment models. *J Furth High Educ* 40(6):846–863
27. Taras M (2015) Student self-assessment: what have we learned and what are the challenges? *Reli—Rev Electron Investig Eval Educ* 21(1):1–16
28. Chen SY, Yeh C-C (2017) The effects of cognitive styles on the use of hints in academic english: a learning analytics approach. *Educ Technol Soc* 20(2):251–264
29. Zhong SH, Li Y, Liu Y, Wang Z (2016) A computational investigation of learning behaviors in MOOCs. *Comput Appl Eng Educ.* 1–13
30. Howell SL, Hricko M, Ebrary I (2006) *Online assessment and measurement.* Information Science Publishing, USA
31. Park Y, Jung I, Reeves TC (2015) Learning from MOOCs: a qualitative case study from the learners' perspectives. *EMI Educ Media Int* 52(2):72–87
32. Lever-Duffy J, McDonald JB (2009) *Teaching and learning with technology,* 3rd edn. Pearson, Boston
33. Hmedna B, El Mezouary A, Baz O, Mammass D (2016) A machine learning approach to identify and track learning styles in MOOCs. In: *International conference on multimedia computing and systems (ICMCS' 16)*
34. Sadhasivam J, Babu R (2017) MOOC: a framework for learners using learning style. *Int Educ Res J* 3(2):21–24
35. Al-Azawei A, Al-bermani A, Lundqvist K (2016) Evaluating the effect of Arabic engineering students' learning styles in blended programming courses. *J Inf Technol Educ Res* 15:109–130
36. Al-azawei A, Parslow P, Lundqvist K (2017) Investigating the effect of learning styles in a blended e-learning system: an extension of the technology acceptance model (TAM). *Australas J Educ Technol* 33(2):1–23
37. Fasihuddin H, Skinner G, Athauda R (2017) Towards adaptive open learning environments: evaluating the precision of identifying learning styles by tracking learners' behaviours. *Educ Inf Technol* 22(3):807–825
38. Mukherjee S (2016) Learning style of humanities, commerce and science students: a study on higher secondary students from West Bengal. *Int J Indian Psychol* 3(3):15–22
39. Hashim H, Salam S, Nurul S, Syafiatun N (2018) The designing of adaptive self-assessment activities in second language learning using Massive Open Online Courses (MOOCs). *Int J Adv Comput Sci Appl* 9(9):276–282

40. Hooda M, Devi R (2017) Significance of cognitive style for academic achievement in mathematics. *Sch Res J Humanit Sci English Lang* 4(22):5521–5527
41. Hashim H, Salam S, Mohamad SNM (2017) Investigating learning styles for adaptive Massive Open Online Course (MOOC) learning. *J Adv Humanities Soc Sci* 3(5):282–292

Assessment on Average Participation Versus Bialek's Methods for Transmission Usage Evaluation Scheme



N. S. Ahmad, N. H. Radzi, S. A. Jumaat, M. N. Abdullah, and S. Salimin

Abstract Nowadays, most developing countries have been moved toward the energy restructuring market. There have several primaries in energy elements within this deregulation scheme which are power generation, transmission, and distribution firms. One of the problems that arise with respect to transmission utility is the development of a fair and equitable methodology for the price of transmission that already used by consumers. An effective method of transmission price assessment is an important framework that needs to be built in the transmission pricing methodology in order to solve the problem that arises in the transmission pricing. Next, a direct current load flow (DCLF) approach needs to be used to define the network power flow for each line. By using the Power World Simulator software, the results using the DCLF approach that was used in this research was been specified accurately and correctly. the power contribution for each transmission line that used it either generation or distribution line will be identifiable to ease and solve problems that arise in this transmission pricing. The method that was used to investigate this problem using in this research is Average Participation and Bialek's methods. In order to investigate the problem that arises in the transmission pricing issues, Average Participation and Bialek's methods were used in this research to solve the problem. For the case study, used IEEE 14 bus systems and the simulation has been carried out using the Matlab software. Then, the results for Average Participation and Bialek's methods were compared for identifying and determining the best approach in this transmission usage evaluation scheme.

Keywords Deregulation scheme · Transmission pricing · Average participation method · Bialek's method

N. S. Ahmad · N. H. Radzi (✉) · S. A. Jumaat · M. N. Abdullah · S. Salimin
Green and Sustainable Energy Focus Group (GSEnergy),
Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia,
Parit Raja, Batu Pahat, 86400 Johor, Malaysia
e-mail: nurhanis@uthm.edu.my

1 Introduction

In Malaysia, three entities which are generation, transmission and distribution are controlled by Tenaga Nasional Berhad (TNB). Hence, all the electrical charge from the generation to distribution should be pay to TNB. The electricity industry in other countries like the United States indeed has long been privately owned. By owning a private company this will increase competition and decreased regulation [1, 2]. However, it is different for developing countries which the electricity business is managed by different companies such as generator companies (GENCOs), transmission companies (TRANSCOs) and distribution companies (DISCOs). The new scheme established free access to the transmission lines and boosting competition among generator and customer. In this situation, the transmission line is the key factor of the electricity markets. To deliver energy from generation to load, transmission is an important role as a separate business. There have many issues arise in transmission pricing. The transmission pricing methodology is a connection charge, interconnection charge, service charge that imposed on power generation and distribution. There have six principles of transmission pricing that should be followed when designing electricity transmission prices. This principle was organized by the Energy Modelling Forum of Stanford University [3].

One of the transmission pricing issues is determining the efficient transmission usage allocation method. Transmission system usage (TSU) cost allocation problem is a complex problem [4]. The purpose of this scheme is to identify the power usage of transmission users to the networks. Hence, in this paper, two methods were explored which are the Average Participation and Bialek's methods in order to determine the best method for transmission usage evaluation scheme. The Average Participation method deals when the net power flow is determined. The idea is to determine the participation share of the generators and loads in the flow of all elements [5]. In this case, the flows of electricity trace from a generator to its consumption buses while the Bialek's method is an algorithm which works only on lossless flows when the flows at the beginning and end of each line are the same [6–8].

1.1 Transmission Usage Evaluation Scheme

Accurate erudition in transmission usage is basically important in the performance of usage-based expense allocation methods. In theory is immensely difficult to decompose the network flows into components linked with individual customers based on the nonlinear nature of power flow equations, On the contrary, from an engineering point of opinion, it is a likelihood and tolerable to request approximate models or sensitivity indices to define the contributions to the network flow from individual user [9]. The DC power flow is considered as this is the simplest way to calculate the actual power flow that flows for each line which is the net power flow at the transmission lines rather than AC power flow method. The net power flow for each

line can be calculated by using Eqs. (1)–(7).

$$P_{ij} = -X_{ij}^{-1} \theta_{ij} \quad (1)$$

where,

$$\begin{aligned} P_{ij} &= \text{circuit flow (p.u)} \\ X_{ij} &= \text{circuit reactance (p.u)} \\ \theta_{ij} &= \text{angle between the buses } i \text{ and } j \text{ (rad)} \end{aligned}$$

$$P_i = P_{Gi} - P_{Li} \quad (2)$$

where,

$$\begin{aligned} P_i &= \text{net injection} \\ P_{Gi} &= \text{MW of generator at bus } i \\ P_{Li} &= \text{MW of load at bus } i \end{aligned}$$

In matrix form:

$$[P_i] = [B_{ij}] \times [\theta_{ij}] \quad (3)$$

$$B_{ij} = -X_{ij}^{-1} \quad (4)$$

$$B_{ii} = \sum X_{ij}^{-1} \quad (5)$$

where,

[B] = the bus susceptance matrix

Matrix B:

$$B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

Let B = 0, matrix is singular, no inverse (delete row 1 and column 1-reference bus)

$$[P] = [B'] \times [\theta] \quad (6)$$

From (3) the value of angle can be calculated as below:

$$[\theta] = [B']^{-1} \times [P] \tag{7}$$

After identified the net power flow of each line, the power contribution can be determined by using the transmission usage evaluation schemes. In this paper, the Average Participation and Bialek’s methods were used in order to identify the best method for transmission usage evaluation scheme. The flowchart in Fig. 1 shows the

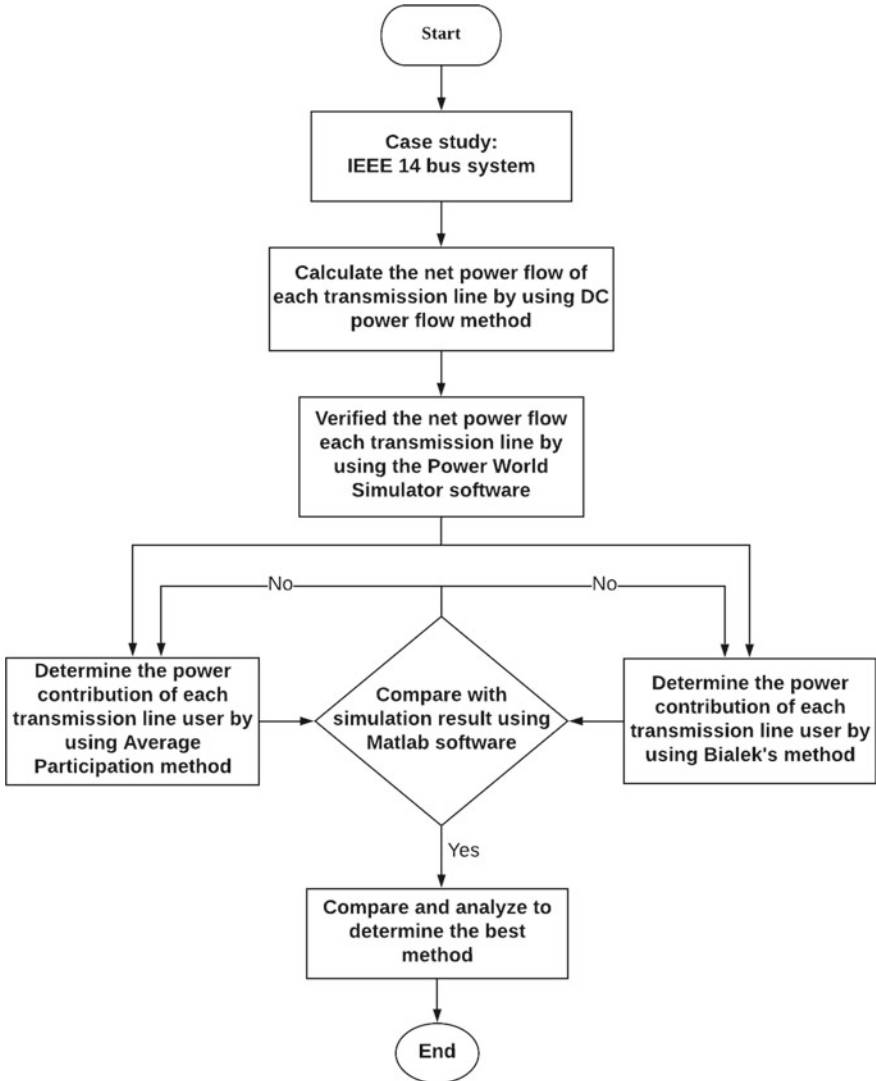


Fig. 1 Flowchart of research methodology

summarization of the transmission usage evaluation process for both methods that has been used in this research.

1.2 Average Participation Method

Complete power flow is needed as an input replying to the required system condition of importance in the Average Participation method. Based on an algorithm, the assumption that electricity supply flows can be detected or the responsibility to produce can be tasked by supposing at any network node and the inflows being distributed proportionally between the outflows [9]. Each generator is injecting power into the network then power flow starting at the generator into the grid until that reach certain loads. By using Average Participation method, the calculation uses the nodal method to calculate the power flow usage each line at the transmission line. The calculation of Average Participation method will use Eq. (8) and Eq. (9).

For generator:

$$L_{ij} = \frac{G_x}{G_T} \times l_{ij} \quad (8)$$

where,

$G_x =$ Generation at bus x (MW)

$G_T =$ Total power of generation (MW)

$L_{ij} =$ Power contribute to line $i - j$

$l_{ij} =$ Power flow at line $i - j$

For load:

$$L_{ij} = \frac{L_x}{L_T} \times l_{ij} \quad (9)$$

where,

$L_x =$ Load at bus x (MW)

$L_T =$ Total power load (MW)

1.3 Bialek's Method

Bialek's method is used to determine the power usage at of each transmission line for transmission usage evaluation method that used for transmission pricing and also to know the fixed costs at the transmission line [10]. Generally, this method is the proportional sharing principle and uses a topological approach to determine the contribution of individual generators or loads to every line flow. By using Bialek's method, which will assume the nodal inflows is share proportionally between nodal outflows. To calculate the power flow for each line at the transmission by using the Bialek's method, there have two algorithms which are the downstream-looking and upstream-looking method. The calculation of Bialek's method will use Eqs. (10)–(17).

If using the upstream-looking method, the power injection in each bus to the generator is given by [10]:

$$P_i^g = \sum_{j \in \alpha_i^u} |P_{ij}^g| + P_{Gi} \quad \text{for } i = 1, 2, \dots, n \quad (10)$$

where,

P_i^g = an unknown gross nodal power flow through node i

P_{ij}^g = an unknown gross line flow in line $i - j$

α_i^u = set of buses supplying directly bus i

P_{Gi} = generation in node i

For the downstream-looking method, the power passing through each bus to the loads is given by [6]:

$$P_i^g = \sum_{j \in \alpha_i^d} |P_{ij}^g| + P_{Li} \quad \text{for } i = 1, 2, \dots, n \quad (11)$$

where,

P_{Li} = load in node i

After that, eliminating the losses it will be, $|P_{ij}^g| = |P_{ij}^g|$. Then, the line flow $|P_{ij}^g| = |P_{ij}^g|$ can be related to the nodal flow at the node j by substituted the $|P_{ij}^g| = c_{ij} P_j^g$. So, simplify the equation that will become:

$$P_i^g = \sum_{j \in \alpha_i^d} c_{ij} P_j^g + P_{Gi} \quad (12)$$

where,

c_{ij} = An unknown gross line flow in line $i - j$ divide
by an unknown gross nodal power flow through
node i

Then can simplify the equation:

$$P_g = A_u P_G \tag{13}$$

$$[A_u]_{ij} = \begin{cases} 1 & i = j \\ -c_{ij} = -\frac{|P_{ij}^g|}{P_j^g} & j \in \alpha_i^u \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where,

A_u = upsteamdistributionmatrix

Invert matrix A_u :

$$P_g = A_u^{-1} P_G \tag{15}$$

The gross power at node i :

$$P_i^g = \sum_{k=1}^n [A_u^{-1}]_{ik} P_{Gk} \tag{16}$$

where,

P_{Gk} = generationinnodek

Then, the gross outflow of line i - j by using the proportional sharing principle is given as:

$$P_{ij}^g = \frac{P_{ij}^g}{P_i^g} \sum_{k=1}^n [A_u^{-1}]_{ik} P_{Gk} \quad \text{for } i = 1, 2, \dots, n \tag{17}$$

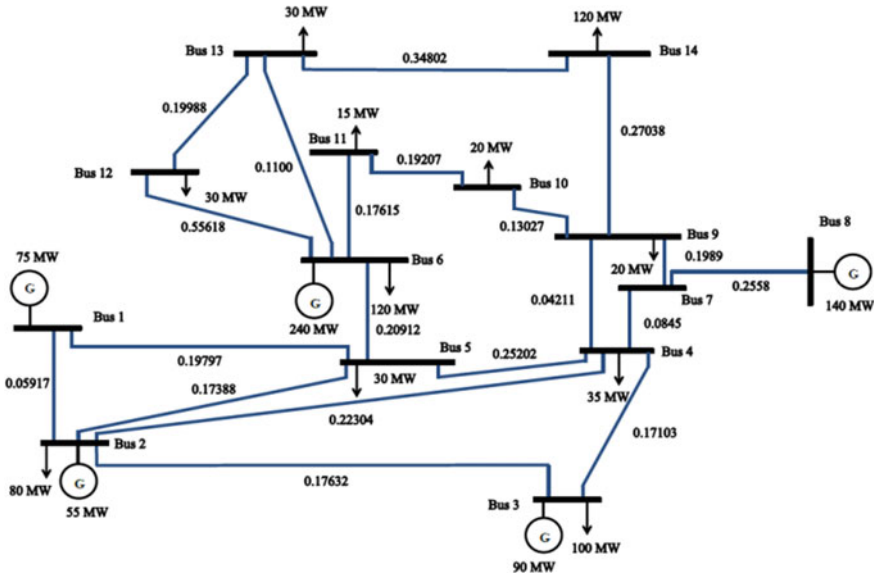


Fig. 2 IEEE 14-bus system [11]

1.4 Case Study

The proposed approach has been tested on IEEE 14-bus system. This case study was based on DC power flow and tested by using Power World Simulator and Matlab programming system. The IEEE 14-bus system is shown in Fig. 2.

Figures 3 and 4 show the result of power flow contribution for each generator and load to networks using Average Participation method for IEEE 14 bus system.

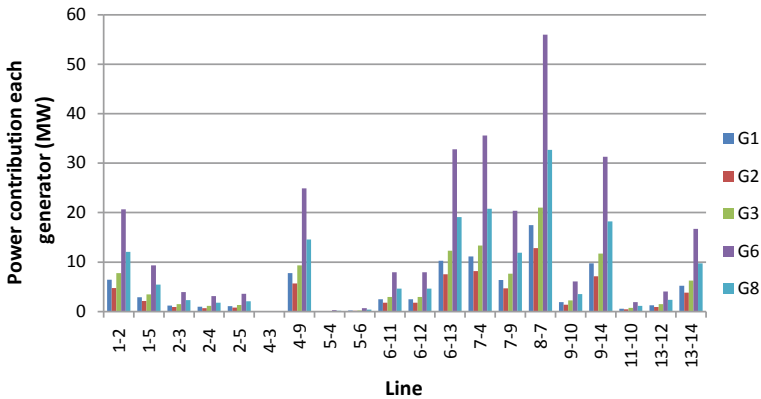


Fig. 3 Power contribution for each generator to networks by using average participation method

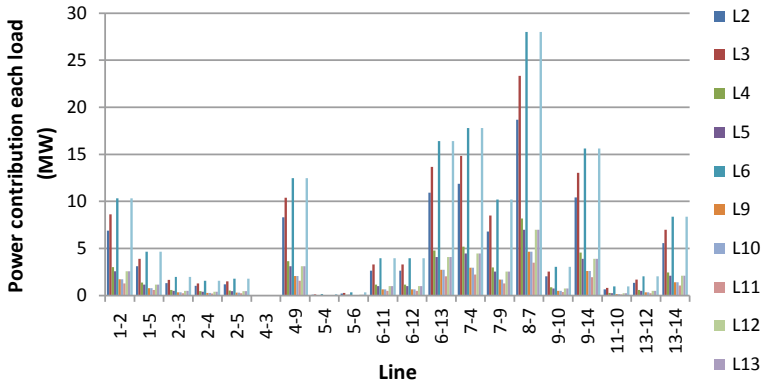


Fig. 4 Power contribution of IEEE 14 bus system for each load to networks by using average participation method

From these both figures, it can be seen that the generators and loads for this method have contributed power flow for each line and follows the nature of power injection. Based on that, all elements have contributed power in each transmission line when the power was injected. It is because the assumption that electricity supply flows can be detected or the responsibility to produce can be tasked by supposing at any network node and the inflows being distributed proportionally between the outflows.

Meanwhile, Figs. 5 and 6 show the result of power flow contribution for each transmission users to the transmission line using Bialek's method. Based on these both figures, it can be seen that for Generator 3, Load 4 and Load 13 have no power contribution for all transmission lines where the value is 0 MW. Load 14 only contribute the power flow at line 11-10 otherwise no power contributes at other lines.

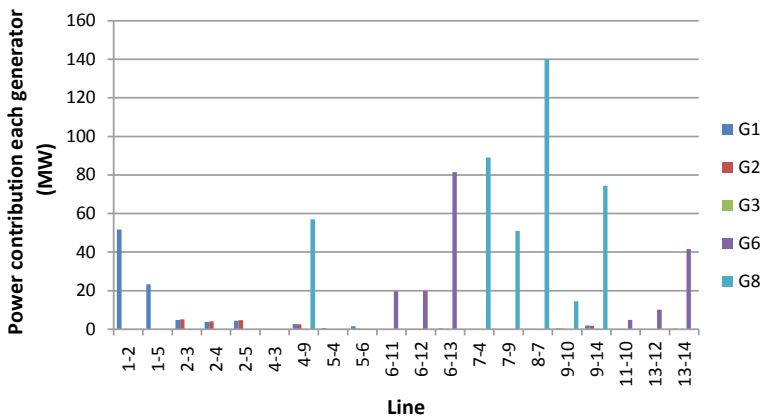


Fig. 5 Power contribution of IEEE 14-bus system for each generator to networks by using Bialek's method

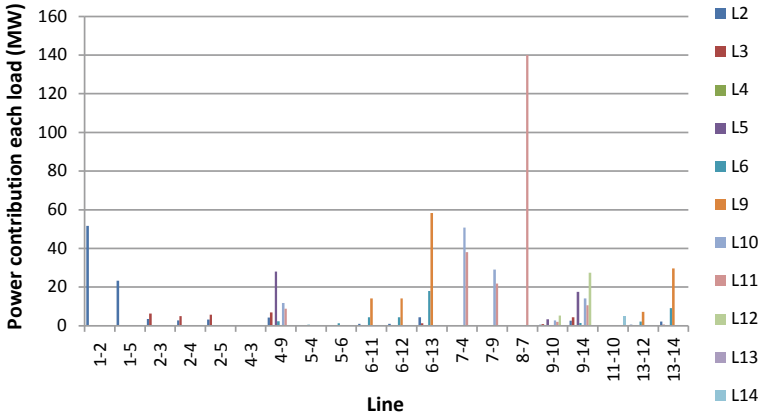


Fig. 6 Power contribution of IEEE 14 bus system for each load to networks by using Bialek’s method

This is because of the Bialek’s method applied the proportional sharing principle and uses a topological approach to determine the contribution of individual generators or loads to every line flow.

By comparing both of the methods that used in this research, for Average Participation method, each network have the power flow contribution from each user. Meanwhile, for Bialek’s method there has no power contribution from few generators and loads to the networks. Based on the nature power injection, all users will be using the transmission services.

2 Conclusion

This paper has presented the assessment on the Average Participation method versus Bialek’s method for transmission usage evaluation scheme. In this research work, the Average Participation method is the best method for transmission usage evaluation scheme in order to determine the power flow contribution of each transmission users to each network. The results show that this method has followed the nature of power injection where all the elements of the transmission lines used the networks when power was injected. In addition, with the implementation of this method will leads to fair and equitable transmission service charges scheme.

Acknowledgements This work was supported by Universiti Tun Hussein Onn Malaysia through TIER 1 grant vot H156.

References

1. Hakvoort R (2000) Technology and restructuring the electricity market, April, pp 4–7
2. Hirst E, Kirby B (2001) Transmission planning for a restructuring U.S. electricity industry, June
3. Green R (1997) Electricity transmission pricing: an international comparison. *Util Policy* 6(3):177–184
4. Rao MSS, Soman SA, Chitkara P, Gajbhiye RK, Hemachandra N, Menezes BL (2010) Min-max fair power flow tracing for transmission system usage cost allocation: a large system perspective. *IEEE Trans Power Syst* 25(3):1457–1468
5. Junqueira M, da Costa LC, Barroso LA, Oliveira GC, Thomé LM, Pereira MV (2007) An Aumann-Shapley approach to allocate transmission service cost among network users in electricity markets. *IEEE Trans Power Syst* 22(4):1532–1546
6. Murali M, Kumari MS, Sydulu M (2013) An overview of transmission pricing methods in a pool based power market. *Int J Adv Sci Eng Technol* 6–11. ISSN 2321-9009
7. Orfanos GA, Tziasiou GT, Georgilakis PS, Hatzigryriou ND (2011) Evaluation of transmission pricing methodologies for pool based electricity markets. In: 2011 IEEE PES Trondheim PowerTech Power Technol a Sustain Soc POWERTECH 2011, vol 218903, pp 1–8
8. Hogan WW (2014) Markets and electricity restructuring transmission expansion
9. Pan J, Teklu Y, Rahman S, Jun K (2000) Review of usage-based transmission cost allocation methods under open access, vol 15, no 4, pp 1218–1224
10. Anjaneyulu V, NarasimhaRao PV, Prakash KNSD (2013) Fixed transmission cost allocation using power flow tracing methods. *Int J Adv Res Electr Electron Instrum Eng* 2(8):3895–3904
11. Srikanth P, Rajendra O, Yesuraj A, Tilak M, Raja K (2013) Load flow analysis of IEEE 14 bus system using MATLAB. *Int J Eng Res Technol* 2(5):149–155

Field-Effect Transistor-based Biosensor Optimization: Single Versus Array Silicon Nanowires Configuration



Chee Chung Ong, Mohamad Faris Mohamad Fathil, Mohd Khairuddin Md Arshad, Mohammad Nuzaihan Md Nor, Ruslinda A. Rahim, Uda Hashim, Rafizatul Fitri Abdullah, Mohd Hazmi Mohd Ghazali, and Nurulazlina Tamjis

Abstract This paper demonstrated the effect of different number of silicon nanowire transducer channels, in other word single, double, and triple channels towards the performance of field-effect transistor-based biosensor through simulation tool. These silicon nanowire field-effect transistor biosensors were designed and simulated in device simulation modelling tool, Silvaco ATLAS with fixed length, width, and height of the silicon nanowire. Different negatively interface charge density values were applied on the transducer channels' surface of the biosensors to represent as detected target biomolecules that will bind onto the surface of the transducer regions. Based on the results, more negatively interface charges density values presented on the sensing channels had reduced the electron carrier accumulation at the channel's interface, therefore, reduced drain current flow between the source and drain terminal. With the increase number of the transducer channels, significant change in drain current for every applied negatively interface charges became more apparent and increased the sensitivity of the biosensor. The triple transducer channels silicon nanowire field-effect transistor biosensor had demonstrated highest sensitivity, that is $2.83 \mu\text{A}/e\bullet\text{cm}^2$, which indicates it has better response for the detection of interface charges, thus increases chances for transducer channels reaction to the target biomolecules during testing or diagnosis.

Keywords Device simulation · Field-effect transistor biosensor · Silicon nanowire · Single nanowire · Array nanowire

C. C. Ong · M. K. Md Arshad · R. F. Abdullah · M. H. Mohd Ghazali · N. Tamjis
School of Microelectronic Engineering, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia

M. F. M. Fathil (✉) · M. K. Md Arshad · M. N. Md Nor · R. A. Rahim · U. Hashim
Institute of Nano Electronic Engineering, Universiti Malaysia Perlis, 01000 Kangar, Perlis, Malaysia
e-mail: mohamadfaris@unimap.edu.my

1 Introduction

A biosensor is an analytical device which converts a biological response into a quantifiable output signal. Among the available biosensors, optical- and electrical-based biosensors are the most famous among the researchers. Optical-based biosensor has higher sensitivity and wider range of detection, but it lacks real-time and label-free detection. These problems can be fulfilled by electrical-based biosensors due to real-time, low fabrication cost, ease of miniaturization, and label-free detection. Therefore, electrical-based biosensor has the potential to become a good point-of-care and illness detection sensor [1]. One example of electrical-based biosensor is the field-effect transistor (FET) device that is used in medical field due to its good platform that can detect the biological and chemical species directly [2]. FET-based biosensors have high scalability, mass reproducibility, and compatibility with the on-chip integration on both the sensor and measurement system [3]. In the ion-sensitive FET (ISFET) biosensor, the metal gate of the FET is replaced by a biofilm layer material. The bio-modulated gate surface changes the drain current (I_D) of the device. This phenomenon is known as the “gating effect” which influences the channel conductivity of the FET. FET biosensor has been widely used due to its high sensitive measurements, portable instrumentation, easy operation with a small amount of sample, low cost with mass reproducible, and high speed detection of target biomarker [4]. Silicon nanowire (SiNW) is a novel FET-based biosensor because it is highly sensitive and direct target detection for large spectrum of species due to high surface-to-volume ratio, rapid response, low detection limit, and 1-dimensional structure [5].

In recent years, SiNW biosensor has been widely applied in many areas, e.g. food industry, environmental monitoring, and medical diagnosis [7]. Effective illness diagnosis is essential for illness identification and subsequent proper treatment. The SiNW has been intensely studied for biosensor due to its surface can be sensitive to charge species combined with their high surface-to-volume ratio [6], hence preferable for highly sensitive disease detection to prevent late or misleading detection. The sensitive detection of the SiNW biosensor can be improved through various nanowire diameters and pattern designs [7]. Therefore, the nanowire diameter and pattern design are the critical aspects to enhance the detection sensitivity. The output signal from detection of target biomolecule can be effectively transduced due to large surface-to-volume ratio. Larger surface-to-volume ratio has larger electrical property, such as conductance to be dominated by surface contributions. Hence, the biosensor allows detection of the biomarker with ease [8].

In addition, by increasing the number of transducer channels between the source and drain region using array SiNW formation may further improve the performance of FET-based biosensor. In this work, several SiNW-FETs based on single and array configurations were designed and simulated to elucidate their performance by using Silvaco ATLAS device simulator. These biosensor functionalities were validated by comparing and analysing the produced current-voltage ($I-V$) characteristics. Moreover, this study also demonstrated SiNW-FET as biosensor by introducing interface

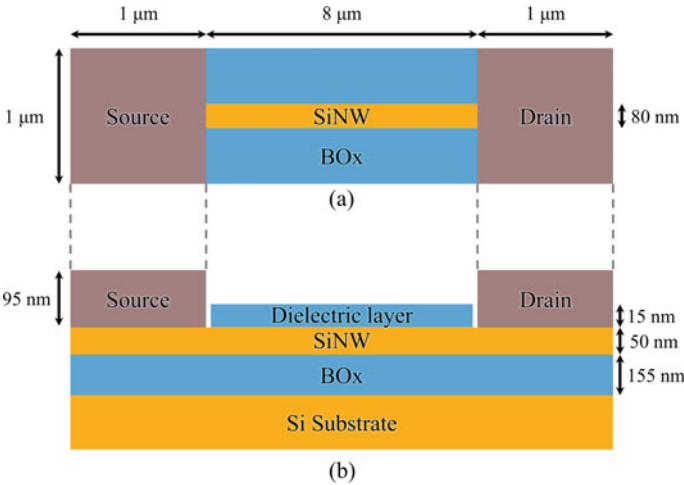


Fig. 1 Structure of single SiNW-FET biosensor from **a** top and **b** side view

charge density (Q_F) on the surface of the SiNW channel as a representation of target biomolecule followed by analysis of the result to determine the device sensitivity (S).

2 Research Methodology

2.1 Structural Specification

Silvaco ATLAS device simulation software was used to model the desired pattern of electronic device. The parameters of the device modelling was referred according to [9]. Silicon-on-insulator (SOI) wafer was used as substrate due to its good electrostatic control on the transistor device channel and its ability to combine device on SOI wafer with other device on silicon bulk by etching the device layer and buried oxide (BOx) layer. Figure 1 shows the top and side view of simulated device with the presence of single SiNW-FET in between the source and the drain region.

2.2 Device Simulation

To simulate the effect of different number of transducer channels on the performance of the biosensor, several models were used, including Lombardi transverse field dependent mobility (CVT), Shockley-Read-Hall recombination using fixed lifetimes (SRH), band-gap narrowing (BGN), Fermi-Dirac carrier statistics (FERMIDIRAC),

Selberherr's (IMPACT Selb). The CVT model is suitable for non-planar devices, while the SRH model that utilized fixed minority carrier lifetime and is recommended to be used in most simulation. The structure of the SiNW-FET biosensor has heavily doped regions ($> 1 \times 10^{17} \text{ cm}^{-3}$) [10, 11], especially at the channel region, therefore it is important to select the BGN model in order to decrease the band gap by lowering the conduction band and raising the valence band by the same amount [12].

For the representation of the availability of charged target biomolecule analytes on the SiNW-FET biosensor during detection, the surface of the transducer channel was applied with negative Q_F of 1×10^{11} , 5×10^{11} , and $1 \times 10^{12} \text{ e/cm}^{-2}$ [13, 14]. Based on the simulation, the effect of Q_F on the electrical behavior in term of the I - V characteristic of the SiNW-FET biosensor with different number of transducer channels was observed and analyzed. The performance of these biosensors was tested by applying a direct current (DC) voltage swept from 0 to 1 V. From the output and measure properties, the sensitivities of the biosensor were determined and comparison of their performances in terms of different number of transducer channels based on device under test was made in order to recognize the which design can be used for the actual biosensor fabrication in the future.

3 Results and Discussion

3.1 Device Structure

The 3D and 2D cross-sectional views of the simulated SiNW-FET biosensors structures are shown in Figs. 2 and 3, respectively for (a) single, (b), and (c) triple transducer channels. The biosensors were made on BOx layer with thickness of 155 nm as insulator layer on a silicon substrate. On top of the BOx layer is another silicon layer with thickness of 50 nm that formed the drain and source regions with a transducer channel region in between them. Both source and drain regions were doped at concentration of $1 \times 10^{20} \text{ cm}^{-3}$ while the SiNW transducer channel was doped at concentration of $1 \times 10^{18} \text{ cm}^{-3}$. Phosphorous was used as n -type dopant for source, drain, and transducer channel regions.

In order to investigate the effect of number of transducer channels toward the performance of the biosensors, another two different structures were formed i.e. double, and triple SiNW transducer channels in between the source and drain regions. For double and triple SiNW, the distances that separate between the transducer channels are 380 nm and 150 nm, respectively. The surface of the SiNW channels was covered with a uniform SiO₂ dielectric layer with thickness of 15 nm as binding site for target biomolecules. Aluminum with thickness of 95 nm was deposited and formed as electrode pads on source and drain regions for the biosensor. Table 1 listed the parameters of the device.

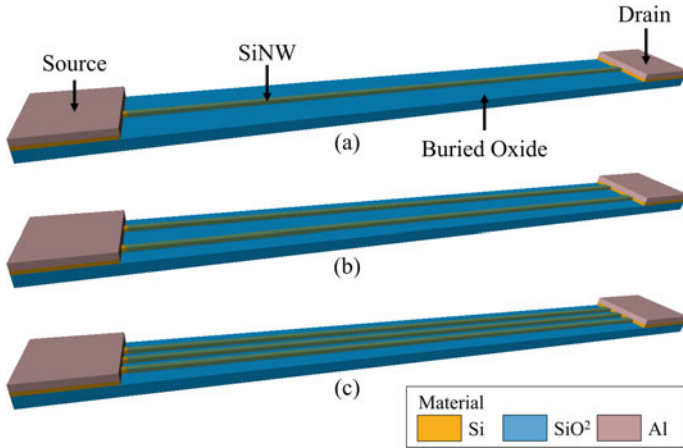


Fig. 2 3D SiNW-FET biosensor structures for **a** single, **b** double, and **c** triple transducer channels with channel length, width, and thickness of 8 μm , 80 nm, and 50 nm, respectively illustrated in TonyPlot

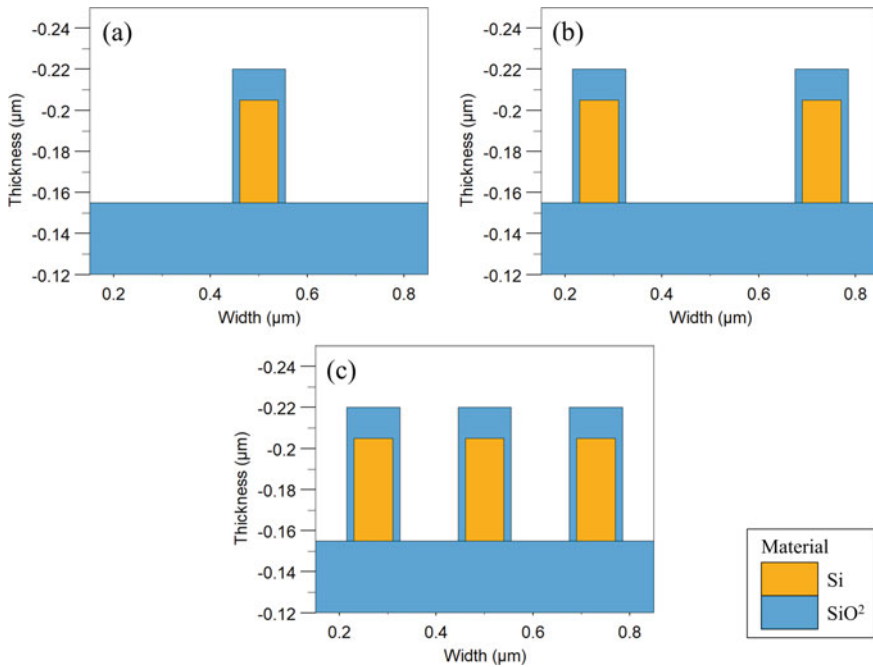


Fig. 3 2D cross-sectional views of SiNW-FET biosensor structures along z -axis of for **a** single, **b** double, and **c** triple transducer channels illustrated in TonyPlot

Table 1 Parameters of SiNW-FET biosensor structure

Region	Material	Dopant type / concentration (cm^{-3})	Length (μm)	Width (μm)	Thickness (nm)
BOx	SiO_2	–	10	1.000	155
Source and drain	Si	n -type / 1×10^{20}	1	1.000	50
SiNW	Si	n -type / 1×10^{18}	8	0.080	50
Electrode	Al	–	1	1.000	95

3.2 Electrical Characteristics

Figure 4a shows ohmic behaviour of the SiNW-FET biosensor for all designs of SiNW transducer channels (i.e. single, double, and triple). The I_D is linearly increased with the increase of drain voltage (V_D). This is related to the high n -type dopant concentration at the channel (i.e. $1 \times 10^{18} \text{ cm}^{-3}$), which enables more current flow from drain to source region before reaching saturation. The significant different

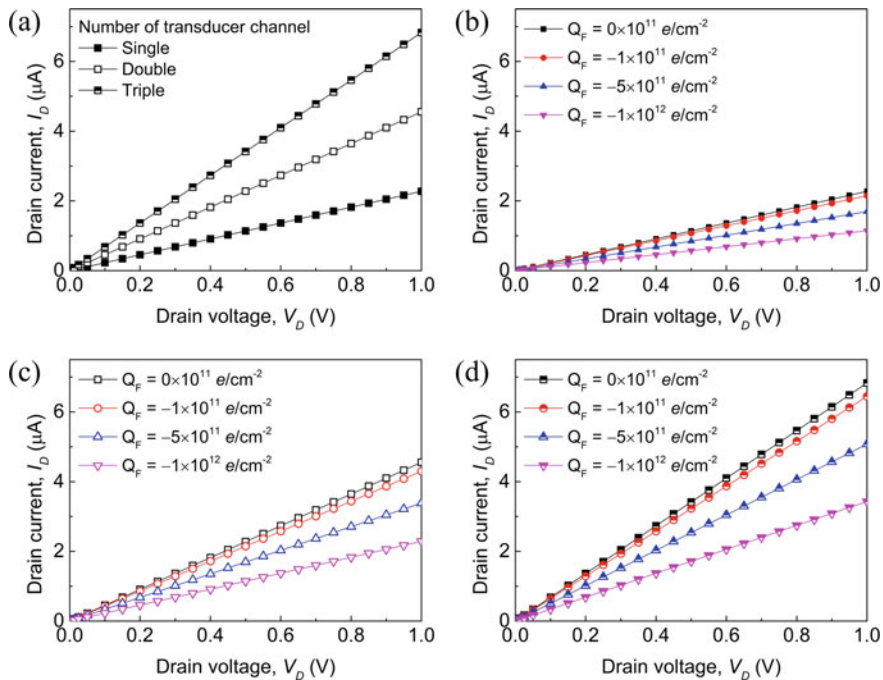


Fig. 4 Electrical characteristics of SiNW-FET biosensor with **a** different number of transducer channels on the electrical and detection of negative Q_F values on SiNW-FET biosensors for **b** single, **c** double, **d** triple SiNW transducer channels

between these three designs is the amount of I_D , which are 2.28, 4.55, and 6.83 μA at $V_D = 1\text{ V}$ for single, double, and triple SiNW transducer channels regions, respectively. The increase of number of transducer channels caused the electrical behaviour of SiNW channel region to change from series to parallel circuit, hence, opened up more pathways for the I_D to flow from drain to source region. The total resistance (R_{Total}) of the channels is not equal to the sum the channels' resistance (as in a series circuit) anymore, however can be determined as in Eq. (1):

$$R_{Total} = 1 / [(1 / R_1) + (1 / R_2) + (1 / R_3)] \quad (1)$$

Based on the Eq. (1), the R_{Total} for channel region can be reduced with the increase of number of transducer channels. According to Ohm's law as in Eq. (2):

$$I_D = V_D / R_{Total} \quad (2)$$

where I_D is indirectly proportional to the R_{Total} . Therefore, when the number of SiNW transducer channels increased from single to double and triple, it reduced the R_{Total} and increased the I_D of the transducer channels region.

Next, several negative Q_F values (i.e. -1×10^{11} , -5×10^{11} , and $-1 \times 10^{12} \text{ e/cm}^{-2}$) were applied on the dielectric layer of the SiNW-FET biosensor to represents the negative charge biomolecules at different concentration as in Fig. 4b–d for single, double, and triple SiNW transducer channels, respectively. When negatively Q_F were applied on the device surface, majority carriers in n -type doped SiNW channels called electrons were repulsed. Lack of electrons accumulation along the SiNW channels due to repulsion caused by the negatively Q_F on the surface had reduced the current flow from the drain to source region. The higher the negative Q_F applied on dielectric layer of the SiNW-FET, the lower I_D across the SiNW channel. The trend is the same for all designs of single, double, and triple SiNW transducer channels. Nevertheless, the changes in I_D values become more significant with the increase number of transducer channels. Table 2 shows the effect of various Q_F on the I_D of the biosensor for different number of SiNW transducer channels based on Fig. 4 for easier observation.

Table 2 The impact of Q_F on I_D for different number of SiNW transducer channels on SiNW-FET at $V_D = 1\text{ V}$

Interface charge density, Q_F (e/cm^{-2})	Drain current, I_D for different number of SiNW transducer channels (μA)		
	Single	Double	Triple
0×10^{11}	2.28	4.55	6.83
-1×10^{11}	2.15	4.30	6.45
-5×10^{11}	1.69	3.38	5.08
-1×10^{12}	1.14	2.29	3.43

3.3 Device Performance Analysis

The sensitivity, S of the SiNW-FET biosensor with different number of SiNW transducer channels is the ratio of between the output signal and the measured property, as in Eq. (3):

$$S = \Delta I_D / \Delta Q_F \tag{3}$$

In this case, the output signal is the difference in I_D (ΔI_D), while the measured property is difference in Q_F (ΔQ_F). S can also be determined as the slope of the linear regression curve, m shown in Fig. 5 using Eq. (4):

$$m = (Y_2 - Y_1) / (X_2 - X_1) \tag{4}$$

High m value of the linear regression curve signifies the high S value. Table 3 tabulates the sensitivities for the biosensor with different number of transducer channels. The SiNW-FET biosensor with triple transducer channel reached the highest S value of $2.83 \mu A/e\text{-cm}^2$, followed by double and single with sensitivities of 1.89 and $0.94 \mu A/e\text{-cm}^2$, respectively. The biosensor with highest S value indicates that with the same amount of negatively Q_F applied on SiNW-FET may allow more significant output I_D can be obtained.

Fig. 5 Determination of sensitivities of SiNW-FET biosensor based on the m values for different number of SiNW transducer channels values at $V_D = 1$ V

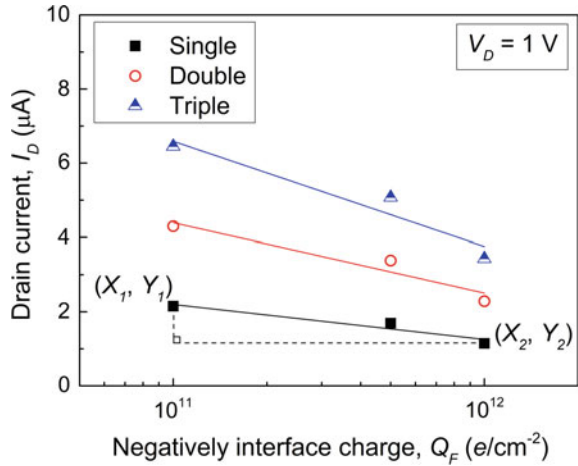


Table 3 Sensitivities of the SiNW-FET Biosensor with different number of SiNW transducer channels at $V_D = 1$ V

Number of transducer channel	Sensitivity ($\mu A/e\text{-cm}^2$)
Single	0.94
Double	1.89
Triple	2.83

4 Conclusion

The investigation to optimize the SiNW-FET biosensor electrical characteristic and performance by increasing the number of SiNW transducer channels in between the source and drain regions had been performed through Silvaco ATLAS device simulation tool. The higher number of transducer channels had increased the output I_D of the biosensor for detection of applied Q_F on SiNW surface. The biosensor with triple SiNW transducer channels had the highest value of I_D . Performance improvement analysis of the biosensor with different number of transducer channels was also conducted in terms of sensitivity. Again, the biosensor with triple SiNW transducer channels achieved the highest sensitivity of $2.83 \mu A/e\text{-cm}^2$. Therefore, the SiNW-FET biosensor with array channels configuration like the high sensitivity triple SiNW transducer channels should be able to improve the performance for detection of target biomolecule at low concentration.

Acknowledgements The authors would like to acknowledge all the team members in Institute of Nano Electronic Engineering (INEE) and School of Microelectronic Engineering, Universiti Malaysia Perlis (UniMAP) for their guidance and help related to this study.

References

1. Kötz B, Schaepman M, Morsdorf F, Bowyer P, Itten K, Allgöwer B (2003) A semiconductor nanobridge biosensor for electrical detection of DNA hybridization. In: Geosci Remote Sens Symp 2003. IGARSS'03. Proceedings. 2003 IEEE Int, vol 4, pp 2869–2871
2. Kim JY, Choi K, Moon DI, Ahn JH, Park TJ, Lee SY, Choi YK (2013) Surface engineering for enhancement of sensitivity in an underlap-FET biosensor by control of wettability. *Biosens Bioelectron* 41:867–870
3. Kim JY, Ahn JH, Choi SJ, Im M, Kim S, Duarte JP, Kim CH, Park TJ, Lee SY, Choi YK (2012) An underlap channel-embedded field-effect transistor for biosensor application in watery and dry environment. *IEEE Trans Nanotechnol* 11:390–394
4. Park J, Hiep H, Woubit A, Kim M (2014) Applications of Field-Effect Transistor (FET)-type. *Biosensors* 23:61–71
5. Zhang G, Zhang L, Huang MJ, Hong Z, Luo H, Kai G, Tay I, Lim EA, Kang TG, Chen Y (2010) Sensors and actuators B: chemical silicon nanowire biosensor for highly sensitive and rapid detection of dengue virus. *Sens Actuators B Chem* 146:138–144
6. Ambhorkar P, Wang Z, Ko H, Lee S, Koo K, Kim K, Cho D (2018) Nanowire-based biosensors: from growth to applications. *Micromachines* 9:679
7. Tran D, Pham T, Wolfrum B, Offenhäusser A, Thierry B (2018) CMOS-compatible silicon nanowire field-effect transistor biosensor: technology development toward commercialization. *Materials (Basel)* 11:785
8. Gao A, Dai P, Lu N, Li T, Wang Y (2013) Optimization of silicon nanowire based field-effect pH sensor with back gate control. In: 8th annual IEEE international conference on nano/micro engineered and molecular systems, IEEE NEMS 2013, pp 116–119
9. Gong X, Zhao R, Yu X (2015) High sensitive detections of Norovirus DNA and IgG by using multi-SiNW-FET biosensors. In: 2015 Transducers—2015 18th international conference on solid-state sensors, actuators and microsystems (Transducers), pp 1537–1540

10. Asthana PK, Ghosh B, Mukund Rahi SB, Goswami Y (2014) Optimal design for a high performance H-JLTFET using HfO₂ as a gate dielectric for ultra low power applications. *RSC Adv* 4:22803–22807
11. Schenk A (1998) Finite-temperature full random-phase approximation model of band gap narrowing for silicon device simulation. *J Appl Phys* 84:3684–3695
12. Boucart K (2010) Simulation of double-gate silicon tunnel FETs with a high-k gate dielectric, 136
13. Fathil MFM, Ghazali MHM, Arshad MKM, Nadzirah S, Ayub RM, Ruslinda AR, Hashim U, Abdullah RF, Ong CC, Tamjis N (2018) Numerical simulation of different silicon nanowire field-effect transistor channel lengths for biosensing application. In: AIP conference proceedings, p 020007
14. Fathil MFM, Tamjis N, Arshad MKM, Rahman SFA, Ayub RM, Ruslinda AR, Hashim U, Ong CC, Abdullah RF, Ghazali MHM (2018) The impact of different channel doping concentrations on the performance of polycrystalline silicon nanowire field-effect transistor biosensor. In: AIP conference proceedings, p 020006

Top-Down Fabrication of Silicon Nanogap for Detection of Dengue Virus (DENV)



Muhamad Nurshahriza Fitri Zulkiffli, Mohammad Nuzaihan Md Nor, Mohamad Faris Mohamad Fathil, Zarimawaty Zailan, Nor Atiqah Md Isa, Conlathan Ibau, Wan 'Amirah Basyarah Zainol Abidin, Aidil Shazereen Azlan, and Mohd Khairuddin Md Arshad

Abstract In this work, a highly sensitive Silicon nanogap biosensor was demonstrated for Deoxyribonucleic acid (DNA) detection related to Dengue virus (DENV). The Silicon nanogap was fabricated using the top-down conventional lithography approach followed by reactive ion etching (RIE) to further thin down the nanogap. The inspections of Silicon nanogap structures were carried out using the scanning electron microscope (SEM) and atomic force microscopy (AFM). The surface of the fabricated Silicon nanogap was functionalized by means of a three-steps procedure involving surface modification, immobilization and hybridization. This procedure acts as a liquid gate control to establish the electrical detection targets of the dengue virus. The electrical detection is based on the changes in the current of the sensor due to the accumulation of the negative charges by the immobilized probe and hybridized target Deoxyribonucleic acid. The limit of detection (LOD) achieved was recorded at 10 pM with a 207 nm of fabricated Silicon nanogap and its sensitivity at 1.5×10^{-10} AM⁻¹. The demonstrated results show that the Silicon nanogap has the excellent properties for detection of dengue virus as biosensor devices.

Keywords Deoxyribonucleic acid (DNA) · Conventional lithography · Silicon nanogap biosensor · Electrical detection · Dengue virus (DENV)

M. N. Md Nor (✉) · M. F. M. Fathil · C. Ibau · W. 'A. B. Zainol Abidin · A. S. Azlan · M. K. Md Arshad

Institute of Nano Electronic Engineering, Universiti Malaysia Perlis, 01000 Kangar, Perlis, Malaysia

e-mail: m.nuzaihan@unimap.edu.my

M. N. F. Zulkiffli · N. A. M. Isa · M. K. Md Arshad

School of Microelectronic Engineering, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia

Z. Zailan

Faculty of Engineering Technology, Universiti Malaysia Perlis, Aras 1, Blok S2, Kampus UniCITI Alam, 02100 Sungai Chuchuh, Padang Besar, Perlis, Malaysia

© Springer Nature Singapore Pte Ltd. 2020

M. A. Mohd Razman et al. (eds.), *Embracing Industry 4.0*,

Lecture Notes in Electrical Engineering 678,

https://doi.org/10.1007/978-981-15-6025-5_4

1 Introduction

1.1 Dengue Fever

Dengue fever (DF) percentage has risen productively over the world in recent decades. DF is a significant illness that kills many people in the worldwide every year. During a premature infection, a patient shows a sign of high temperature without other signs. Subsequently, grave harm to internal organs can happen if no respectable treatment and then can cause death [1]. The infection is caused by the DENV. DENV are restricted into four serotypes, which are DENV-1, DENV-2, DENV-3, and DENV-4 [1]. Each of these serotypes of DENV can bring mild febrile diseases at first primary infection and potentially developed into more severe dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) that can lead to fatality [2]. One recent estimate indicates 390 million dengue infections per year. Another study, of the prevalence of dengue, estimates that 3.9 billion people, in 128 countries, are at risk of infection with dengue viruses [3]. A total of 32,435 cases of dengue fever with 53 deaths were reported nationwide from January to June 2019, which is a decrease of 34.8% or 17,291 cases and 110 deaths reported in the same period in 2018 [4]. However, these numbers still consider as huge cases that we must give attention to prevent it. Currently, in Malaysia a total of 1638 DF cases with one death were reported during week 16, 2019, a decrease of 5.0 percent (%) compared to the previous week, bringing the cumulative number to 41,443 cases and 65 deaths as of April 20, 2019. This is higher compared to 16,971 cases with 36 deaths reported in 2018 during the same period [5]. The trend can be clearly seen in Fig. 1.

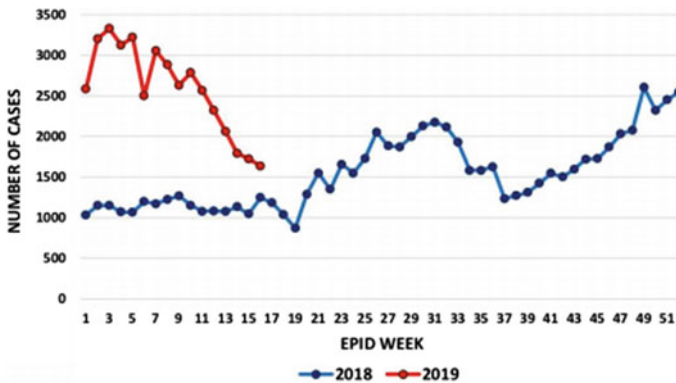


Fig. 1 Dengue cases reported weekly in 2018 and 2019, Department of Health, Malaysia [5]

1.2 Silicon Nanogap Biosensor

In addition, methods for diagnosing DF at a very early stage are urgently needed to prevent the spreading of the outbreaks. One of the best ways to resolve the issue is introducing a Silicon nanogap biosensor. Silicon nanogap biosensor is a device having two electrodes with a nanometer scale of gap and a molecule, that place in between the two terminals will be recognized by analysing the electrical parameters for current, resistance, conductance, capacitance, permittivity, and impedance [6–8] as shown in Fig. 2. There are two basic requirements have favor the researchers to predict the changes of nanogap-based electrical biosensor devices for the high sensitive area of Deoxyribonucleic acid (DNA). At first, DNA can transport charges as being charges carrier and can associate terminals separated by a nanometer of gap, and furthermore charges transport is affected if the molecule encounters immobilization from two fold stranded to single-stranded adjustment [9]. Frequently, a single strand deoxyribonucleic acid (ssDNA) test is immobilized on nanogap area, and target for hybridization is distinguished by planning estimation of an electrical signs or dielectric field affect without the need of a labeling mark [6].

To the best of our knowledge, a Silicon nanogap biosensor for the detection of DENV has not yet been investigated. Because of the inherent high sensitivity and potentials of nanogap-based electrical biosensor devices to reduce and diagnose large number of victims who are suffering from the outbreak of DF, herein we demonstrate a top-down fabrication of Silicon nanogap and an electrical detection of DENV using this fabricated Silicon nanogap. The electrical detection is based on the changes in current, resistance and conductance of the device due to accumulation of negative charges added by the immobilized and hybridized DNA. This work is useful in making electrical detection devices that can be commercialized for DNA detection.

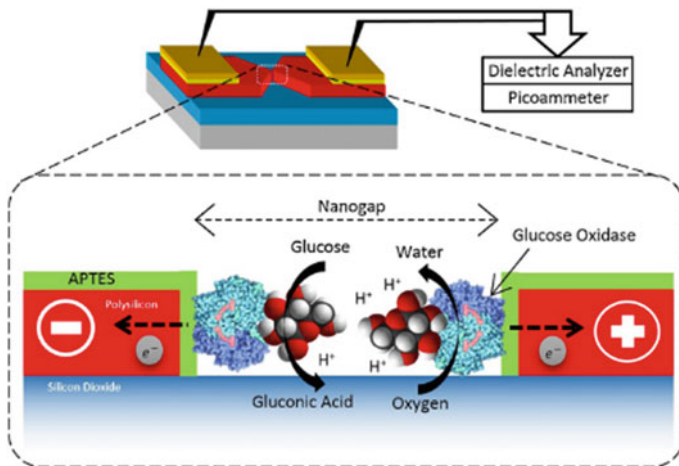


Fig. 2 Molecule detection mechanism on Silicon Nanogap biosensor [8]

Thus, we expect this Silicon nanogap biosensor to be beneficial for point-of-care diagnostic applications.

2 Methodology

2.1 Fabrication of Silicon Nanogap

The Silicon nanogap was fabricated on a p-type silicon wafer with thickness of $500 \pm 25 \mu\text{m}$. Fabrication of Silicon nanogap is shown as in Fig. 3 and consisted of 8 steps processes.

The p-type silicon wafer was used as the substrate. Firstly, the silicon wafer was cleaned using a standard RCA 1 solution. The solution was prepared by mixing de-ionized (DI) water, ammonium hydroxide, NH_3 (27%), and hydrogen peroxide, H_2O_2 (30%). The ratio for RCA 1 solution is 5:1. The purpose of the cleaning process is to remove both organic contaminants. Next, the wafer was soaked in hydro-fluoric acid (HF) to remove the native oxide from the wafer surface. Then, the wafer was transferred to a container and the deionize (DI) water was used to rinse the wafer. This step is to remove the solution from the sample. After the cleaning process, the p-type silicon wafer was cut into small pieces measuring 2 cm by 2 cm. Next, silicon dioxide (SiO_2) was grown on top of the silicon layer through wet oxidation process as insulating layer with average thickness of 180 nm, followed by deposition of polysilicon on the SiO_2 layer. Deposition process was done using low pressure chemical vapor deposition (LPCVD) at 620°C by H_2 ambient pyrolysis of silane with a deposition rate of $80 \text{ \AA}/\text{minute}$. The process continued with the conventional photolithography. This process required several steps, including photoresist coating at 5000 rpm for 25 s using spin coater, followed by soft bake using hot plate for 60 s at 90°C . Next, nanogap mask was aligned on the photoresist coated on the polysilicon. The process continued with UV light exposure for 10 s, followed by development

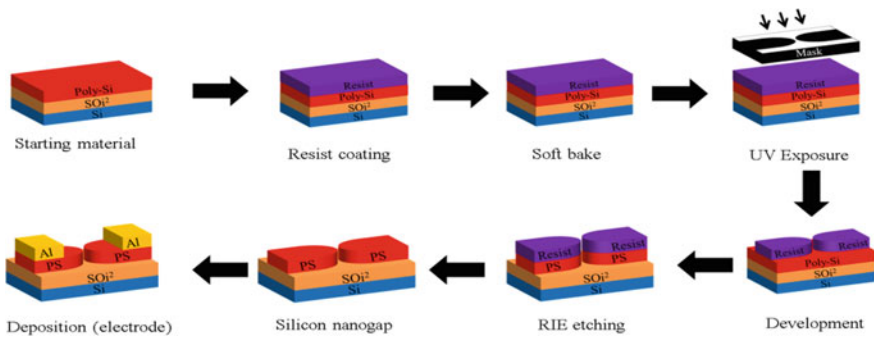


Fig. 3 Top-down fabrication of Silicon nanogap biosensor

process using RD-6 developer solution for 35 s and cleaned with deionized (DI) water, then completed with hard bake for 120 s at 90 °C. The developed sample was then loaded into a reactive ion etching (RIE) for anisotropic etch profile of polysilicon nanogap, followed by the resist pattern was stripped using acetone, which revealed Silicon nanogap with a good anisotropic profile. The process continued with deposition of aluminum (Al) thin film on top of the Silicon nanogap layer. To develop Al contact pads, second conventional lithography was conducted using contact pads. Using the same method of lithography, design of contact pads was transferred on the resist, followed by wet etching process of Al thin film inside Al etchant solution. The rest of the Silicon nanogap device was fabricated after the development of Al contact pads.

2.2 Surface Functionalization of Silicon Nanogap

The three steps involved in the surface functionalization of Silicon nanogap are surface modification, immobilization and hybridization. To activate the Silicon nanogap through surface modification, the Silicon nanogap device was immersed in 2% 3-aminopropyltriethoxysilane (APTES) in a combination of 95% phosphate buffer solution (PBS) and 5% APTES for 1 day at room temperature. Silanization oxygen atom of hydroxyl-terminated groups has performed a covalent bond with the Silicon nanogap area in form of the molecule of APTES. Subsequently, the APTES-functionalized Silicon nanogap surface was immersed in 2.5% of glutaraldehyde. Then the surface was cleaned using PBS and DI water to rinse off excess of glutaraldehyde. Glutaraldehyde was presented as a linker to ensure a chemical bond with groups of amine-based and the aldehyde groups (COH) on the sensing area. Aldehyde groups can subsequently used for DNA immobilization step. For DNA immobilize, a amine-terminated probe was linked to the aldehyde-terminated groups. A 10 μM DNA probe solution diluted with PBS solution was dropped onto the Silicon nanogap sensing area followed by incubation process at room temperature for 1 day. Then, unbound probe was rinse with PBS solution. After this immobilize step, a drop with complementary DNA targets are injected into the sensing are, which was located in the center of the device. The sample was left overnight for hybridization takes place at nanogap area. Noted that in order to ensure the sample is completely in dry condition and not touch between two pads as possible to avoid a short circuit. After the first hybridization process, sample need to de-hybridize before hybridize with different concentration (10 μM to 10 pM) to evaluate the sensitivity of DENV DNA detection. For de-hybridization process, the sample was washed with hot DI water at 80 °C for 5 min to de-hybridize the complementary DNA pairs on the nanogap.

2.3 Characterization of Silicon Nanogap

The morphology of the fabricated Silicon nanogap was investigated using the scanning electron microscopy (SEM) and atomic force microscopy (AFM), as they provide direct visualization of the quality of the shape, size and uniformity of the nanogap. The direct current (DC) electrical measurements were performed using KEITHLEY 6487 picoammeter/voltage source, which is the voltage swept from 0 V to 1 V to quantify the electrical contact behavior with the two terminals of the Al electrodes. Furthermore, all the measurements were carried out at ambient temperature and sample was tested without any vibration or noisy environment to avoid interference results.

3 Results and Discussion

3.1 Morphology of the Fabricated Silicon Nanogap

The morphology of the Silicon nanogap was investigated using SEM and AFM. The images were shown in Fig. 4. The gap of the structure was approximately 207 nm. The images show that gap is formed with normal development profile with desirable resolution, good pattern placement and good uniformity [10].

3.2 Electrical Responses of the Fabricated Silicon Nanogap

To ensure a successful fabrication and functionalization of the silicon nanogap biosensor, an electrical test was conducted using picoammeter/voltage source. The electrical measurement was tested with bare sample (air), APTES, DNA immobilization, and DNA hybridization. Figure 5 clearly depicts the I-V characteristic, which exhibited significant differences in measure I_{ds} values for each step of the surface modification. The biosensor with APTES surface modification on the nanowire surface resulted in $I_{ds} = 4.0 \times 10^{-10}$ A which increased by 1.0×10^{-10} A compared with the $I_{ds} = 3.5 \times 10^{-10}$ observed for the bare sample at $V_{ds} = 1$ V. Meanwhile, the biosensor with the target complementary DNA (10 μ M) resulted in $I_{ds} = 1.6 \times 10^{-9}$ A which is 0.9×10^{-9} A higher than the DNA probe (10 μ M), $I_{ds} = 0.7 \times 10^{-9}$ A at $V_{ds} = 1$ V. The increment of the measured I_{ds} values were originated from the increased positive charges carrier current density and the p-type silicon nanogap encouraged by the negativity charged probe and target DNA [11–13]. Higher concentration of target DNA gives more negative charge, which lead to higher current values over voltage. Furthermore, this observation was also in consistent with previously reported results [6, 8, 9]. The limit of detection (LOD) is defined as the lowest detectable molar concentration of target DNA. In this work, the LOD was 10 pM

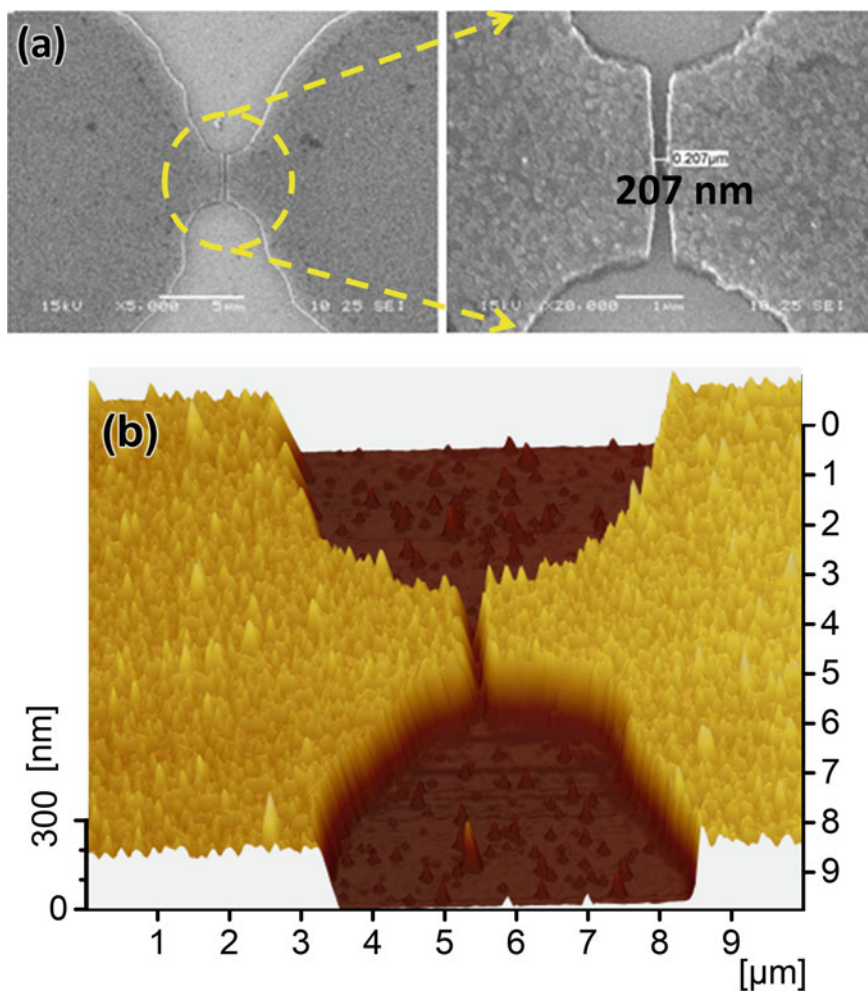


Fig. 4 a SEM and b AFM images of the Silicon Nanogap

(Fig. 5) and the sensitivity of the Silicon nanogaps biosensor attained was $1.5 \times 10^{-10} \text{ AM}^{-1}$ as shown in Fig. 6. These results verify that Silicon nanogap biosensor was feasible as a biosensor to detect low concentration of target DNA related to DENV.

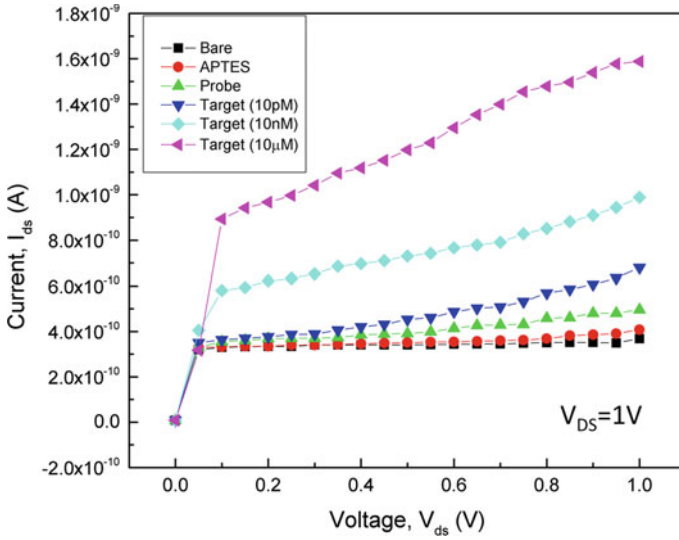


Fig. 5 $I_{ds}-V_{ds}$ characteristic by different steps of the surface functionalization

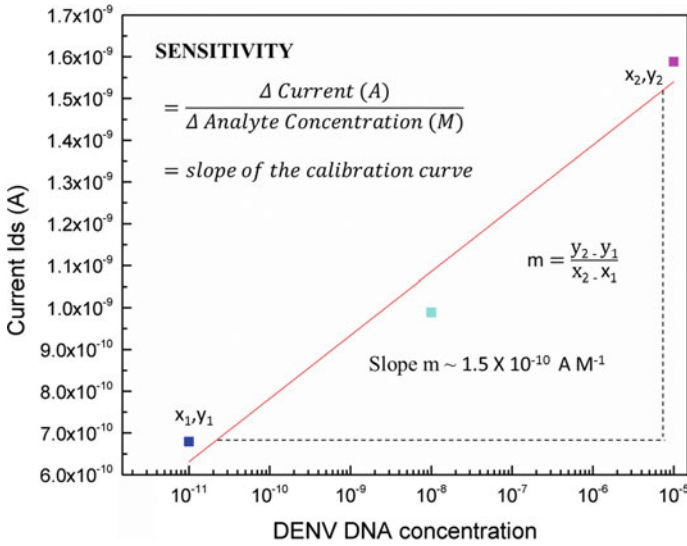


Fig. 6 I_{ds} response curve of Silicon nanogap biosensor

4 Conclusion

The in-house Silicon nanogap was fully fabricated with 207 nm gap using the conventional lithography approach (top-down fabrication). The responsiveness of this biosensor was thoroughly investigated by observing the electrical detection in response to the concentrations of DENV DNA. The use of this approach offers a good sensitivity of $1.5 \times 10^{-10} \text{ AM}^{-1}$ and the LOD of 10 pM for target DENV DNA. Thus, this Silicon nanogap can be used as an improved devices and serves as biosensor for future biomedical diagnostic applications.

References

1. Navakul K, Warakulwit C, Thai Yenchitsomanus P, Panya A, Lieberzeit PA, Sangma C (2017) A novel method for dengue virus detection and antibody screening using a graphene-polymer based electrochemical biosensor. *Nanomed Nanotechnol Biol Med* 13(2):549–557
2. Rashid JIA, Yusof NA, Abdullah J, Hashim U, Hajian R (2015) A novel disposable biosensor based on SiNWs/AuNPs modified-screen printed electrode for dengue virus DNA oligomer detection. *IEEE Sens J* 15(8):4420–4421
3. Brady OJ et al (2012) Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl. Trop Dis* 6(8):e1760
4. BorneoPost Online | Borneo, Malaysia, Sarawak Daily News | Largest English Daily In Borneo. 32,435 dengue cases recorded in Jan–June 2018. Retrieved from <http://www.theborneopost.com/2018/07/08/32435-dengue-cases-recorded>, Jan–June 2018. 21 Oct 2018
5. WHO (2014) Update on the dengue situation in the Western Pacific Region. *Dengue Situat Updat* 453:1–5
6. Ali E, Dhahi TS, Das R, Hashim U (2013) DNA hybridization detection using less than 10-nm gap silicon nanogap structure. *Sensors Actuators A Phys* 199:304–309
7. Rahman SFA et al (2016) Enhanced sensing of dengue virus DNA detection using O₂ plasma treated-silicon nanowire based electrical biosensor. *Anal Chim Acta* 942:74–85
8. Nuzaihan M et al (2016) Top-down nanofabrication and characterization of 20 nm silicon nanowires for biosensing applications. *PLoS One* 11(3):e0152318
9. Parkash O, Shueb RH (2015) Diagnosis of dengue infection using conventional and biosensor based techniques. *Viruses* 7(10):5410–5427
10. Dhahi TS, HashimU, Ahmed NM (2011) Reactive Ion Etching (RIE) for Micro and Nanogap Fabrication. *J Basrah Res* 37(2)
11. Balakrishnan SR et al (2014) Development of highly sensitive polysilicon nanogap with APTES/GOx based lab-on-chip biosensor to determine low levels of salivary glucose. *Sensors Actuators A Phys* 220:101–111
12. Humaira MSN, Hashim U, Nazwa T, Ten ST, Ahmad S, Yusof NA (2014) Process development of 40 nm silicon nanogap for sensor application. *IEEE Int Conf Semicond Electron Proceedings ICSE* 96–99
13. Balakrishnan SR et al (2016) Polysilicon nanogap lab-on-chip facilitates multiplex analyses with single analyte. *Biosens Bioelectron* 84:44–52

Design and Performance Analysis of IoT Based Sensor System Using LoRa



Fakrulradzi Idris , Muhammad Syahmi Anuar,
Muhammad Hazim Zin Azman, and Norlezhah Hashim

Abstract The fourth industrial revolution (IR 4.0) can be characterized by extensive use of Internet of Things (IoT) technology. Conventional sensor systems rely heavily on existing wireless technology such as ZigBee, Bluetooth and WiFi. However, these technologies have some limitations which include higher power consumption and lower coverage. LoRa can improve these limitations by enhancing the coverage area while using much lower power consumption. This paper describes the design of a sensor system which combines both IoT and LoRa technologies. The system consists of a LoRa sensor node and a gateway node. Arduino Uno, LoRa module and temperature sensor are used to design the sensor node. The real-time value of the temperature will be forwarded to the gateway node which consists a Raspberry Pi and another LoRa module. The performances of the system are evaluated both in line of sight and non-line of sight conditions. Specifically, the received signal strength indicator (RSSI) and signal to noise ratio (SNR) are studied for different distances between the sensor and gateway nodes. The results showed that RSSI of -117 dBm and SNR of -17 dB can be achieved for a distance of 300 m in non-line of sight condition. On the other hand, RSSI as high as -79 dBm and SNR of 3 dB are recorded for the same distance but in line of sight condition.

Keywords Internet of things · LoRa · Arduino · Raspberry pi · RSSI · SNR

F. Idris (✉) · M. S. Anuar · M. H. Z. Azman

Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia

e-mail: fakrulradzi@utem.edu.my

N. Hashim

Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik (FTKEE), Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia

1 Introduction

The use of wireless sensor system could play a central role in the Internet of Things (IoT) [1–3], especially in the low-power long-range network. The low power communication had evolved until reach multi-kilometer ranges. In addition, this long-range with low-power sensor system could provide better coverage and energy efficiency improvement over existing system based on WiFi or ZigBee technologies.

LoRa [4, 5] is a spread spectrum modulation technique derived from chirp spread spectrum (CSS) technology. This technology was developed by Semtech which is a founding member of the LoRa Alliance. Unlike cellular and Device-to-Device (D2D) communication [6], LoRa uses license-free radio frequency bands such as 433, 868 and 915 MHz. LoRa technology enables a variety of smart IoT applications in many sectors such as agriculture, healthcare, industrial control, metering and logistic. In addition, LoRa technology offers an efficient, flexible and economical solution to many problems where other technologies such as cellular, WiFi, Bluetooth and ZigBee are ineffective [7–10].

The aim of this work is to design an IoT based sensor system using LoRa technology and study the performance of the system. Firstly, a LoRa sensor node is designed using an Arduino board and LoRa module. Secondly, Raspberry Pi which are widely used in IoT applications, together with another LoRa module are used to construct a LoRa gateway. In order to show the usefulness of monitoring temperature information, a test bed is performed with the implemented LoRa system.

2 Methodology

This section presents the detail steps for hardware development and software configuration of the sensor system. The hardware development consists of the design of sensor and gateway nodes. Similarly, the software configuration of both the sensor and gateway nodes are provided.

2.1 Sensor Node Design

Arduino UNO and Dragino LoRa Bee [11] module are the key components of the sensor node, while LM35 is used as the temperature sensor. Sensor node acts as a transmitter to send temperature reading to a remote gateway node. In this work, LoRa Bee with 915 MHz frequency band is used.

As in [12], the LoRa module need to be supplied with 3.3 V input voltage. Therefore, the pin 1 (VCC) of the module is connected to the 3.3 V on the Arduino Uno. Pin 4 (MISO) of the module is connected to the ICSP pin that located on the Arduino board. The ICSP pin that connects to the pin 4 is also designated as MISO. The same

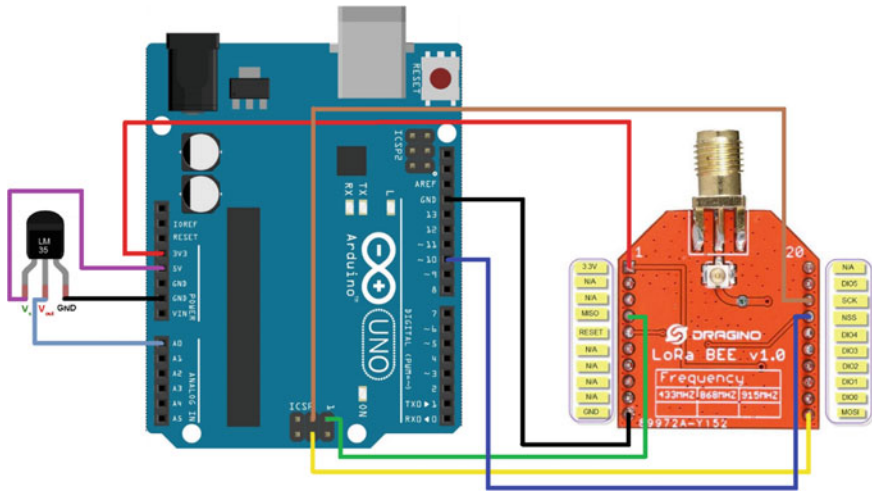


Fig. 1 Sensor node wiring diagram

goes to the pin 11 (MOSI) of LoRa Bee that connects to the ICSP pin (MOSI) of the board. Next, pin 10 (GND) of the module is connected to the respective GND pin on the UNO board. Then, pin 17 (NSS) and pin 18 (SCK) are connected to digital pin 10 and pin ICSP (SCK) respectively. The temperature sensor (LM35) has 3 pins which are VCC, output, and GND. The VCC pin is attached to the 5 V supply of the board, the GND is connected to one of the board’s GND and the output pin is connected to the analog pin A0. The complete wiring diagram of the sensor node is depicted in Fig. 1.

The software implementation of the sensor node is designed using Arduino IDE. Figure 2 shows part of the sketch for the sensor node.

2.2 Gateway Design

For the gateway design, another LoRa module is attached to a Raspberry Pi’s GPIO header. Based on the work in [13], we need to connect the corresponding SPI pin (MOSI, MISO, CLK, CS) of the LoRa module and the Raspberry Pi’s GPIO header. In addition, the GND and 3.3 V pin of the LoRa module are connected to the respective pin of the GPIO header. Figure 3 shows the wiring diagram of the gateway node.

In addition, the software configuration for the gateway need to be performed. Specifically, only the radio configuration for the gateway are modified according to the LoRa parameters as shown in Table 1.

```

#ifdef BAND868
#ifdef SENEGAL_REGULATION
const uint32_t DEFAULT_CHANNEL=CH_04_868;
#else
const uint32_t DEFAULT_CHANNEL=CH_10_868;
#endif
#elif defined BAND900
const uint32_t DEFAULT_CHANNEL=CH_05_900;
// For HongKong, Japan, Malaysia, Singapore, Thailand, Vietnam: 920.36MHz
//const uint32_t DEFAULT_CHANNEL=CH_08_900;
////////////////////////////////////
// CHANGE HERE THE LORA MODE, NODE ADDRESS
#define LORAMODE 1
#define node_addr 6
////////////////////////////////////

////////////////////////////////////
// CHANGE HERE THE THINGSPEAK FIELD BETWEEN 1 AND 8
#define field_index 3
////////////////////////////////////

////////////////////////////////////
// CHANGE HERE THE READ PIN AND THE POWER PIN FOR THE TEMP. SENSOR
#define TEMP_PIN_READ A0
temp = 0.0;
int value;

for (int i=0; i<5; i++) {
// change here how the temperature should be computed depending c
//
value = analogRead(TEMP_PIN_READ);

//LM35DZ
//the LM35D2 needs at least 4v as supply voltage
//can be used on 5v board
temp += (value*TEMP_SCALE/1024.0)/10;

```

Fig. 2 Arduino sketch for sensor node

3 Results and Discussion

This section presents the system prototype which consists of LoRa sensor and gateway nodes. The values of temperature reading from the sensor node are logged using a ThinkSpeak channel. The measurement of RSSI and SNR during real testbed implementation in non-line of sight (NLOS) and line of sight (LOS) scenarios are provided.

3.1 System Prototype

The prototype of both sensor and gateway nodes are shown in Fig. 4. It should be noted that printed circuit boards are designed to simplify the connection between LoRa module to Arduino Uno and Raspberry Pi.

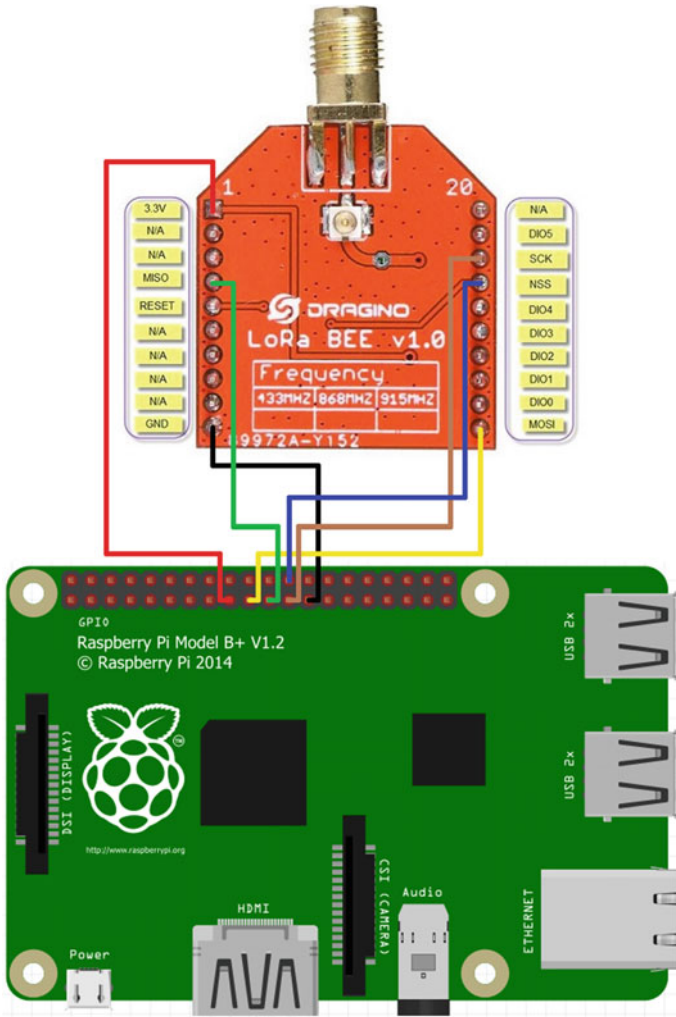


Fig. 3 Gateway node wiring diagram

Table 1 LoRa parameters

Parameter	Value
Bandwidth (BW)	125 kHz
Coding rate (CR)	4/5
Spreading factor (SF)	12
Channel number	5
Central frequency	913.88 MHz

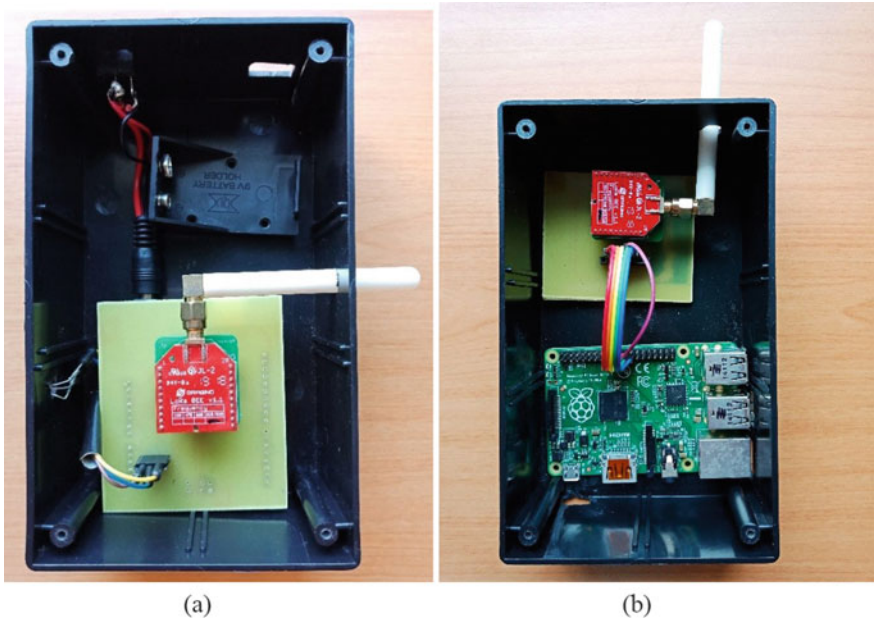


Fig. 4 System prototype. a Sensor node, b Gateway node

3.2 ThinkSpeak Channel

The sensor node consists of an LM35 sensor which measures the temperature of surroundings. The value of temperature reading will be collected and sent wirelessly to gateway node. Once received, the gateway will forward the data to a ThingSpeak channel. For example, the ThingSpeak webpage logged the value of temperature as shown in Fig. 5.

Fig. 5 Sensor values in ThinkSpeak channel

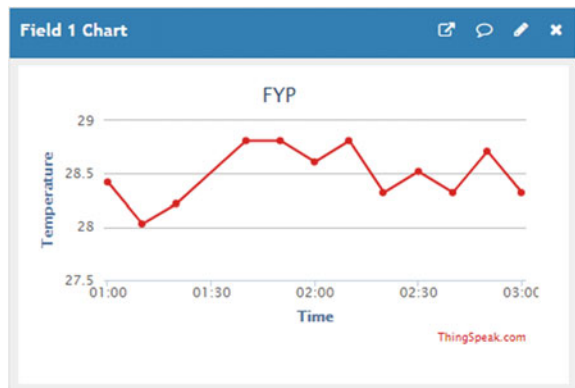




Fig. 6 Location for NLOS measurement. a 100 m, b 200 m, c 300 m

3.3 Testbed Measurement

Non-line of Sight (NLOS)

For NLOS scenario, the system is deployed in a housing area around Taman Tasik Utama, Ayer Keroh, Melaka. The range of the sensor node from the gateway are set to 100 m, 200 m, and 300 m respectively as shown in Fig. 6. These locations are NLOS scenarios because there are a few buildings and trees blocking the signal path between sensor node and gateway node.

Line of Sight (LOS)

Other areas are chosen to deploy and test the performance of the system for LOS condition. Specifically, three different locations around Universiti Teknikal Malaysia Melaka (UTeM) as shown in Fig. 7 are selected for system deployment.

By logging into the gateway node, we can capture the current temperature reading and important parameters such as received signal strength indicator (RSSI) and signal to noise ratio (SNR) as depicted in Fig. 8.

Figure 9 shows the values of RSSI for the two different scenarios. Location D recorded the highest RSSI of -56 dBm while both location B and C recorded the lowest RSSI value of -125 dBm. The highest RSSI achieved for the longest distance between the sensor and gateway nodes in NLOS (Location C) is -117 dBm. However, for the same distance but in LOS scenario, Location F achieves RSSI as high as -79 dBm. Comparing the results, it can be seen that all LOS scenario recorded

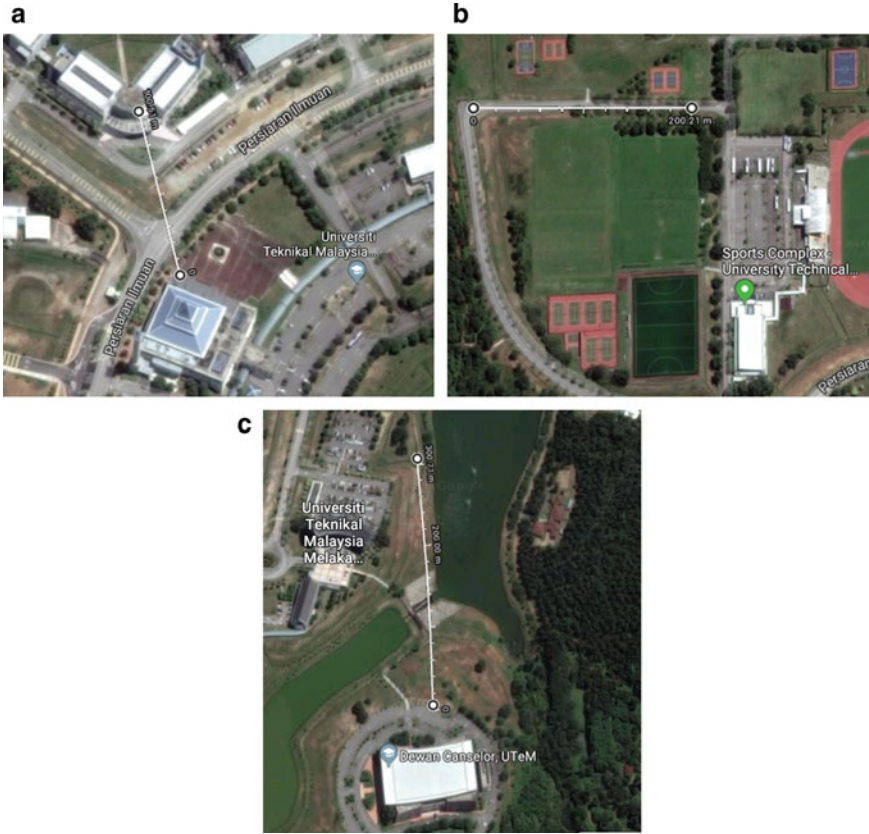
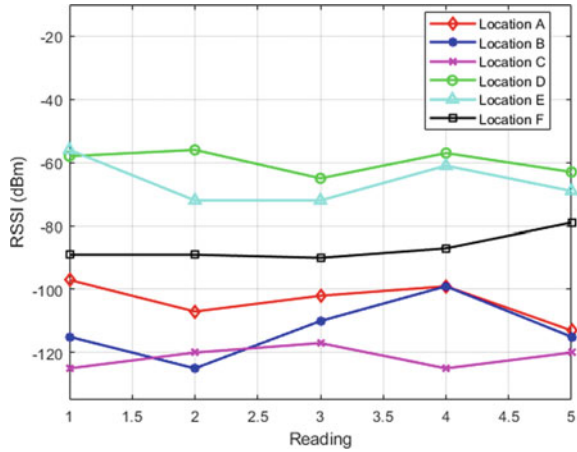


Fig. 7 Location for NLOS measurement. d 100 m, e 200 m, f 300 m

```
pi@raspberrypi:~/lora_gateway $ sudo ./lora_gateway
SX1276 detected, starting.
SX1276 LF/HF calibration
...
*****Power ON: state 0
$default sync word: 0x12
$LoRa mode 1
$Setting mode: state 0
$channel CH_5_900: state 0
$set LoRa power dbm to 14
$Power: state 0
$Get Preamble Length: state 0
$Preamble Length: 8
$LoRa addr 1: state 0
$SX1272/76 configured as LR-BS. Waiting RF input for transparent RF-serial bridge
$Low-level gw status ON
-- rxlora, dst=1 type=0x12 src=6 seq=0 len=17 SNR=2 RSSIpkt=-102 BW=125 CR=4/5 SF=12
p1,18,6,0,17,2,-102
r125,5,12
t2018-06-27T18:30:22.468
*!#3#TC/25.20
-- rxlora, dst=1 type=0x12 src=6 seq=1 len=17 SNR=2 RSSIpkt=-102 BW=125 CR=4/5 SF=12
p1,18,6,1,17,2,-102
r125,5,12
t2018-06-27T18:31:24.255
*!#3#TC/25.39
```

Fig. 8 Values of temperature, RSSI and SNR

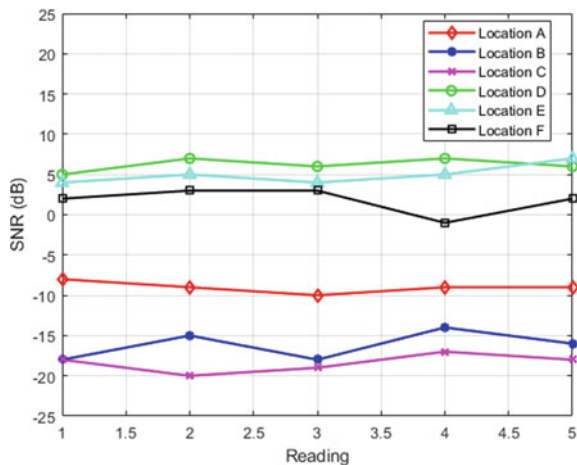
Fig. 9 Effect of distance on RSSI for NLOS and LOS



higher value of RSSI than the NLOS case. This is because higher path loss occurs in NLOS situation which result in significant degradation of the received signal strength captured by the receiver (gateway).

The effect of distance on SNR during the measurement are depicted in Fig. 10. By considering the farthest distance, Location C obtains the maximum SNR value of -17 dB while the highest SNR recorded for Location F is 3 dB. It is apparent from this figure that higher SNR can be achieved when the distance between the sensor and gateway nodes is shorter. Higher path loss which is more dominant in NLOS scenario reduces the value of SNR.

Fig. 10 Effect of distance on SNR for NLOS and LOS



4 Conclusion

In this paper, we have designed an IoT based sensor system using LoRa technology. A testbed was setup to analyze the performance of the system in terms of RSSI and SNR for non-line of sight and line of sight conditions. It was shown that higher RSSI and SNR can be achieved for lower distance between the sensor and gateway node in line of sight condition. The effect of bandwidth, spreading factor and coding rate on RSSI and SNR will be studied as future work.

Acknowledgements The authors would like to thank Centre for Research Management (CRIM), Universiti Teknikal Malaysia Melaka (UTeM) for sponsoring this paper.

References

1. Adryan B, Obermaier D, Fremantle P (2017) The technical foundations of IoT. Artech House, Boston
2. Siu C (2018) IoT and low-power wireless circuits, architectures, and techniques. CRC Press, Florida
3. Haxhibeqiri J, Poorter ED, Moerman I, Hoebeke J (2018) A survey of LoRaWAN for IoT: from technology to application. *Sensors* 18(11):3995
4. Jeong JD, Shin Y, Lee IW (2018) Long-range transmission of photovoltaic climate information through the LoRa radio. In: International conference on information and communication technology convergence (ICTC), Jeju, South Korea, pp 956–959
5. Augustin A, Yi J, Clausen T, Townsley WM (2016) A study of LoRa: long range & low power networks for the Internet of Things. *Sensors* 16(9):1466
6. Idris F, Tang J, So DKC (2018) Resource and energy efficient device to device communications in downlink cellular system. In: 2018 IEEE wireless communications and networking conference (WCNC), Barcelona, Spain, pp 1–6
7. Noreen U, Bounceur A, Clavier L (2017) A study of LoRa low power and wide area network technology. In: International conference on advanced technologies for signal and image processing (ATSIP), Fez, Morocco, pp 1–6
8. Hashim N, Abdul Razak MAH, Idris F (2015) Home security system using ZigBee. *Jurnal Teknologi (Sciences & Engineering)* 74(10):29–34
9. Iborra RS, Gomez JS, Vinas JB, Cano MD, Skarmeta AF (2018) Performance evaluation of LoRa considering scenario conditions. *Sensors* 18(3):772
10. Hashim N, Azmi NFAM, Idris F, Rahim N (2016) Smart phone activated door lock using WiFi. *ARPN J Eng Appl Sci* 11(5):3309–3312
11. Dragino LoRa Bee module. <https://www.dragino.com/products/lora/item/109-lora-bee.html>. Last accessed 22 July 2019
12. Low cost LoRa device leaflet. <https://github.com/CongducPham/tutorials/blob/master/Low-cost-LoRa-device-leaflet.pdf>. Last accessed 22 July 2019
13. A DIY low-cost LoRa gateway. <http://cpham.perso.univ-pau.fr/LORA/RPIgateway.html>. Last accessed 22 July 2019

A PSPT-MAC Mechanism for Congestion Avoidance in Wireless Body Area Network



Wan Aida Nadia Wan Abdullah, Naimah Yaakob, R. Badlishah Ahmad, Mohamed Elshaikh Elobaid, S. N. Azemi, and Siti Asilah Yah

Abstract A Remote Health Monitoring System (RHMS) is known as one of the promising applications that has been successfully developed with the help of Wireless Body Area Network (WBAN) technology nowadays. This RHMS offers a continuous monitoring of health's status by sensing and collecting the physiological signals (medical data) from bio-sensors that are attached or implanted in the body. Then, these medical data are furthered transmitted to the clinicians to diagnose the diseases. If any abnormalities are detected, a quick medical actions would be carried out. However, these collections of medical data could lead to heavy traffic which increase the risk of data congestion in the network. Congestion could severely impact the overall's network performances in terms of longer delay and packet loss. Thus, a Priority Selective Packet Timeslot (PSPT-MAC) mechanism is proposed to avoid congestion during transmitting these bulk of medical data in the network. This mechanism is initiated by classifying and prioritizing the data according to their importance through ECG Packet Classification and Prioritization (ECG-PCP) mechanism. Later, corrupted packets are earlier discarded by Prioritized Selective Packet Discarding (P-SPD) mechanism to save the limited network's resources. Finally, the remain packets (after discarding packets from P-SPD mechanism) undergo fragmentation according to slot time via Fragmentation based Slot Time MAC (FST-MAC) mechanism in the MAC IEEE 802.15.4 protocol. From the findings, this mechanism has outperformed the standard IEEE 802.15.4 protocol and FCA-MAC mechanism by yielding low delay and packet loss as well as high throughput and packet delivery ratio (PDR) under different number of nodes in the network.

Keywords WBAN · Congestion · Priority · Discarding · Fragmentation

W. A. N. Wan Abdullah (✉) · N. Yaakob · R. Badlishah Ahmad · M. E. Elobaid · S. A. Yah
ENAC, School of Computer and Communication Engineering (SCCE),
Universiti Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia
e-mail: aidanadia92@yahoo.com

S. N. Azemi
ACE, School of Computer and Communication Engineering (SCCE),
Universiti Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia

1 Introduction

A progressive advancement in bio-sensors and wireless technology has enabled the development of Remote Health Monitoring System (RHMS) nowadays. This application offers an efficient delivery of medical services to the patients since it reduce the waiting time for the patients to be treated by doctors [1]. This is because patients with RHMS technology can be wirelessly monitored anywhere and anytime. If any abnormalities are detected in the reading of vital signals, the medical practitioners would be alerted and prompt medical actions would be carried out to avoid risk of mortality among patients. In addition, this application also has been part of Internet of Things (IoT) technology [2].

RHMS is developed by using technology namely Wireless Body Area Network (WBAN) that comprises several intelligent and miniature bio-sensors are attached or implanted in the body [3]. Then, the readings of vital signals are detected and forwarded these medical data to the database of the hospital. However, transmission of these bulk medical data could lead to congestion in the network which might lead to increment in delay and packet loss.

Moreover, according to [4, 5], medical data in WBAN are significant and can be categories as normal, critical and on-demand data. Loss of normal data are still acceptable but loss of critical data may cause severe impact to the patients such as late in delivery of medical response and to be worst lead to mortality. This is due to the significant information that has been carried by critical data and thus poses strict delay and packet loss in order to avoid misdiagnosing the diseases. Thus, a mechanism of congestion avoidance should be designed to meet the requirements of low delay and packet loss in WBAN so that a reliable RHMS can be established.

The paper is organized as follows: Sect. 2 (Literature Review) provides discussions on the previous or existing mechanisms in congestion for WBAN. The design of proposed mechanism to avoid congestion is delivered in Sect. 3 (Methodology) including with the simulation parameters used in the experiments. Section 4 (Results and Discussion) elaborates the findings from the collected results of simulation experiments for the proposed mechanism in WBAN. Finally, the achievements of the proposed mechanism and some future works are concluded in Sect. 5 (Conclusion).

2 Literature Review

Congestion control mechanisms can be divided into three classes which are congestion avoidance, congestion mitigation and reliable data transport [6]. However, this paper specifically focuses on reviewing existing congestion avoidance mechanisms in the WBAN. Congestion avoidance is defined as mechanism that avoid occurrence of congestion in the first place. There are several congestion avoidance mechanisms in the network as shown in Fig. 1 as follows.

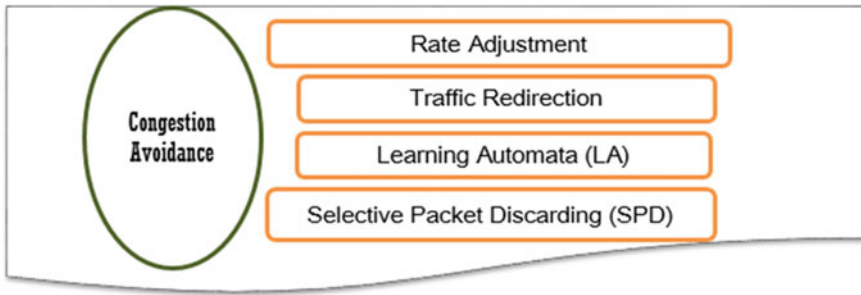


Fig. 1 Congestion avoidance mechanisms in the network

Generally, there are two techniques namely rate adjustment and traffic redirection are considered for congestion avoidance mechanism [7]. In rate adjustment, transmission rate of the nodes is regulated once the notification of congestion is received [8, 9]. Nevertheless, this technique could degrade the network's performance due to occurrence of packet drop. This packet dropping could lead to misdiagnosing the disease especially when the critical packets are dropped. This is because these type of packets carry significant information of the patient's health status. Thus, rate adjustment is not an optimum choice in handling congestion in the network. Meanwhile, the nodes are highly recommended to transmit their packets to other uncongested path in traffic redirection congestion control method. In other words, once the congestion is detected, the packets should not be routed through the affected, congested path. Thus, the packets should be redirected to the other alternative path to arrive at the destination [10]. However, there could be a possibility that the congested path might be the shortest route for the packets to reach the destination. Thus, redirection of packets may bring to time consuming that lead to increment in delay especially to the critical packets in WBAN. This is because this type of packet need to be timely delivered to the base station to avoid late delivery of medical response.

Another congestion avoidance mechanism which is termed as Learning Automata (LA) is initiated in [11]. The concept of LA is placing an automaton in every node. This automaton refers to small piece of code that is able to perform intelligent action such as controlling data flow's rate which depends on packet's drop probability when certain flow rate is kept constant as in [12, 13]. Optimal data rate is selected by learning from past performance of the flow in the network through the computation of dropped packet. In other words, the rate's flow with least number of dropped packets will be chosen as optimal rate. The implementation of Selective Packet Discarding (SPD) schemes are mainly used in voice packet and Asynchronous Transfer Mode (ATM) [14, 15]. This mechanism is selectively dropped the packets in the data transmission.

Precisely, most of these existing mechanisms are done in Wireless Sensor Network (WSN) which is not suitable to be directly implemented in the WBAN [6, 16]. This is due to strict requirements of WBAN in terms of delay and packet loss compared

to WSN. Thus, some improvements and adjustments of existing congestion control mechanisms are critically needed to be suited in WBAN environment.

3 Proposed Methodology

There are three stages of mechanisms involve in the proposed congestion avoidance for WBAN. To be exact, this proposed mechanism is only concerned on Electrocardiogram (ECG) medical data. These three stages of mechanisms are integrated and designed as Priority Selective Packet Timeslot (PSPT-MAC) mechanism as tabulated in Table 1 as follows.

At initial stage which involves ECG packet Classification and Prioritization (ECG-PCP) mechanism (Stage 1), the ECG medical data are classified and prioritized based on the value of RR-intervals in the ECG's reading. The range of interval RR-intervals plays a major role in determining the significance of the ECG medical data. For instance, if the RR-intervals is ranging from 0.6 to 1.0 s, the medical packets would be classified as Normal packets with Low Priority. Meanwhile, if the value of RR-intervals is less than 0.6 s, the medical packets are assigned with Critical packets and High Priority. The concept and details of this ECG-PCP mechanism can be referred to our previous work in [17].

After classifying and prioritizing the data through ECG-PCP mechanism, the normal and abnormal data would undergo Prioritized Selective Packet Discarding (P-SPD) mechanism (Stage 2). For the purpose of this work, the corrupted packets are referred to packets with bit errors. This type of packet is not worth to be transmitted as it might be discarded at the base station. Hence, earlier discarding process of corrupted packets could avoid waste of network resources such as bandwidth and power consumption. Also, the concept of P-SPD mechanism has been extensively discussed and elaborated in our earlier work as in [18]. It is worth to note that P-SPD mechanism is not carried out to the critical packets as these type of packet should be quickly transmitted in the network.

Later, after completing ECG-PCP (Stage 1) and P-SPD (Stage 2) mechanisms, a Fragmentation based Slot Time MAC (FST-MAC) mechanism is carried out by fragmenting the medical packets according to slot time in the IEEE 802.15.4 protocol. Generally, the typical slot time in IEEE 802.15.4 is 0.96 ms [19]. However, in this FST-MAC mechanism this value is decreased to 0.2 ms after data is being fragmented.

Table 1 Stages of PSPT-MAC mechanism for congestion avoidance in WBAN

PSPT-MAC mechanism	Stage 1	ECG Packet Classification and Prioritization (ECG-PCP) mechanism
	Stage 2	Prioritized Selective Packet Discarding (P-SPD) mechanism
	Stage 3	Fragmentation based Slot Time MAC (FST-MAC) mechanism

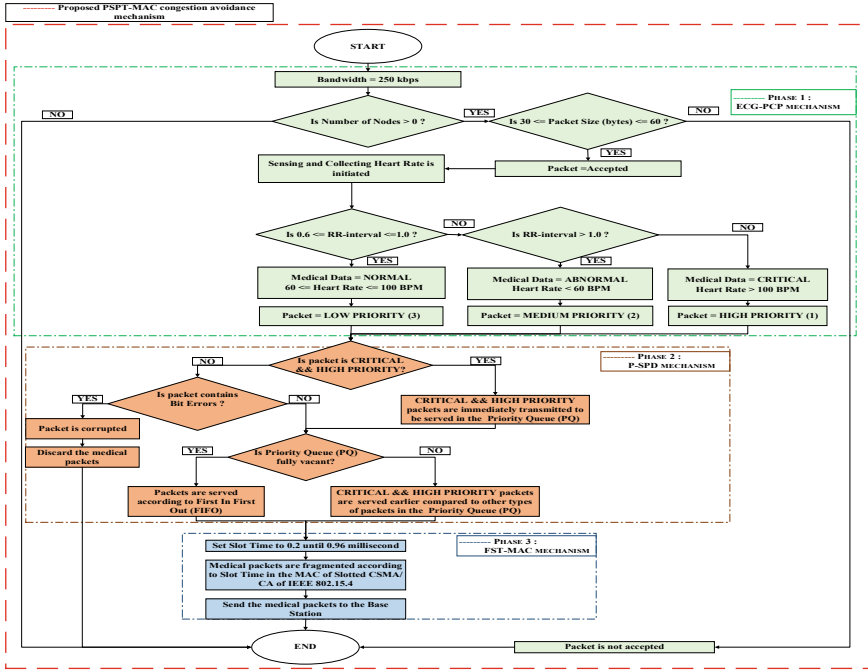


Fig. 2 Flowchart of PSPT-MAC mechanism

The details on FST-MAC mechanism are explained in our previous work in [20]. Finally, these three stages of mechanisms are integrated together to be designed as Priority Selective Packet Timeslot (PSPT-MAC) mechanism to avoid congestion in WBAN at the first place.

A flowchart of PSPT-MAC mechanism is illustrated in Fig. 2 as follows.

In addition, the proposed PSPT-MAC mechanism is carried out and analysed in OMNeT++ simulator tool with integration of INET framework which is known for its suitability to be adapted in Wireless Body Area Network (WBAN) [21]. Apart from that, most of previous studies have selected the short-range wireless technology namely IEEE 802.15.4/Zigbee for WBAN as it offers low data rate and power consumption [22, 23]. Specifically, the setting of parameters for this proposed mechanism is chosen based on [24]. In concise, the proposed mechanism is tested under different number of nodes from 10 to 60 within $50 \times 50 \text{ m}^2$ area of coverage for WBAN and compared with standard IEEE 802.15.4 protocol and FCA-MAC mechanism in [24]. The packet size is set to 30 bytes as it offers low end to end delay for WBAN [7, 25]. The network performances are measured in terms of delay, throughput, packet delivery ratio (PDR) and packet loss as tabulated in Table 2.

Table 2 Network performances in WBAN

Network performance	Definition
Delay	The duration of time required by a packet to travel through the network to its last destination and measured in time unit (second)
Throughput	The amount of data that a network system is able to transfer in a time unit and measured in bits per seconds (bps)
Packet delivery ratio (PDR)	The ability of the network to successfully delivered information as the ratio between total number of successfully received packets and total number of packets sent
Packet loss	The amount of packet loss in the delivery of the packets in the network

4 Results and Discussions

This section provides a detailed discussion and analysis from the measured performances of the proposed mechanism against standard IEEE 802.15.4 protocol and FCA-MAC mechanism.

4.1 Delay

From Fig. 3, the delay for all schemes are increasingly proportional to the number of nodes. However, PSPT-MAC mechanism has lower delay by outperforming FCA-MAC and typical IEEE 802.15.4 when number of nodes are 40 with 0.235 s compared to 0.62 s in FCA-MAC and 0.661 s in IEEE 802.15.4. In addition, it can be seen that delay for PSPT-MAC mechanism are still in range of allowable delay for WBAN for nodes between 10 and 60 compared to other two mechanisms. To be exact, the allowable delay’s threshold for transmitting one ECG packet is 500 ms [26].

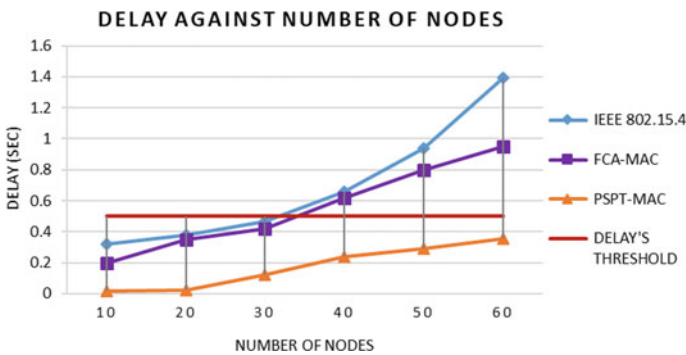


Fig. 3 Performance of delay between PSPT-MAC, FCA-MAC and IEEE 802.15.4 mechanisms

Typically, increment number of nodes directly lead to generation of high number of packets. However, PSPT-MAC mechanism has earlier discarded the corrupted packets compared to other schemes which lead to less number of traffics. After discarding process, the remaining packets would be fragmented into smaller size before furthered transmitted to the base station. Typically, no fragmentation is carried out in typical MAC IEEE 802.15.4 protocol. Thus, PSPT-MAC mechanism offered packet fragmentation which is done by using shorter slot time of 0.2 ms in MAC IEEE 802.15.4 protocol. This has resulted in many fragmented packets to easily access the superframe slots as these smaller packets do not need to wait for longer time compared to typical slot time of 0.96 ms in FCA-MAC and IEEE 802.15.4 protocol. These smaller packets lead to lower delay compared to other schemes. In contrast, FCA-MAC mechanism and IEEE 802.15.4 standard, the packets are sent without being earlier discarded and fragmented. These packets might contain corrupted packets that might be dropped and retransmitted in the network and result in longer delay of the WBAN. This could be worst if the congestion happened in the network. From the results, PSPT-MAC has shown noticeable trend in providing lower delay to alleviate the congestion in WBAN.

4.2 Throughput

Figure 4 shows the collected throughput of PSPT-MAC mechanism and IEEE 802.15.4 standard in the WBAN. High value of throughput is noticed under PSPT-MAC scheme compared to other mechanisms under increasing number of nodes. Although the throughput displayed an increment for 30 nodes which is 167.92 kbps for PSPT-MAC and 100.30 kbps for IEEE 802.15.4 standard, but the value of throughput is declined which is when the nodes are 40 for both schemes. The measured throughput for 40 nodes PSPT-MAC mechanism and IEEE 802.15.4 protocol is 155.41 kbps and 98.32 kbps respectively. In addition, lowest throughput is

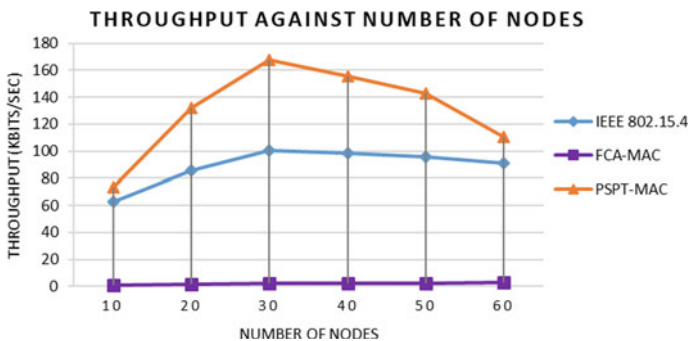


Fig. 4 Performance of throughput between PSPT-MAC, FCA-MAC and IEEE 802.15.4 mechanisms

yielded by FCA-MAC scheme for nodes between 10 and 60. This phenomenon might indicate that congestion has occurred in the network. Consequently, this has resulted to reduction in the successful number of packets received at the base station of the network. However, high throughput is still dominated by PSPT-MAC mechanism as this mechanism has discarded the corrupted packets earlier before being further transmitted and fragmented in the network. Meanwhile, IEEE 802.15.4 protocol and FCA-MAC scheme might transmit high number of packets as these mechanisms do not undergo discarding and fragmentation processes. Nevertheless, the transmitted packets in the IEEE 802.15.4 protocol and FCA-MAC scheme might be dropped or undergone retransmission if there is existence of corrupted packets. This might cause to less number of received packets in the network. Hence, earlier discarding of corrupted packets and fragmentations helps to promote high network performances in the WBAN.

4.3 Packet Delivery Ratio (PDR)

Based on Fig. 5, the resulted PDR for PSPT-MAC mechanism for 10 to 20 nodes is up to 95.48% compared to 57.24% of PDR under IEEE 802.15.4 standard. However, the PDR for both methods are declined linearly over time under increment number of nodes. The obvious changes are clearly shown when number of nodes are ranging between 40 and 60 nodes. The PDR for PSPT-MAC mechanism are half-yielded which is only 50.44% and IEEE 802.15.4 standard produced PDR of 25.05% with 40 nodes. This performance is worst for 60 nodes as PSPT-MAC provided only 27.01% of PDR and 16.38% of PDR for IEEE 802.15.4 standard. These values are clearly reflected that lower ratio of number of packets received to the number of packets sent in the network. Typically, increment number of nodes also would increase the number of packets in the network. Thus, more packets generated might lead to congestion in the network and degraded the network performance in the WBAN. However,

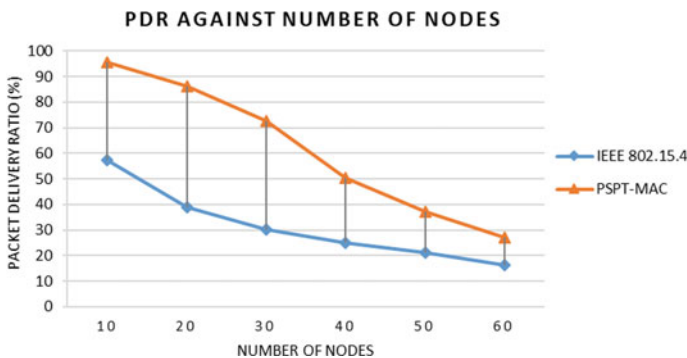


Fig. 5 Performance of PDR between PSPT-MAC and IEEE 802.15.4 mechanisms

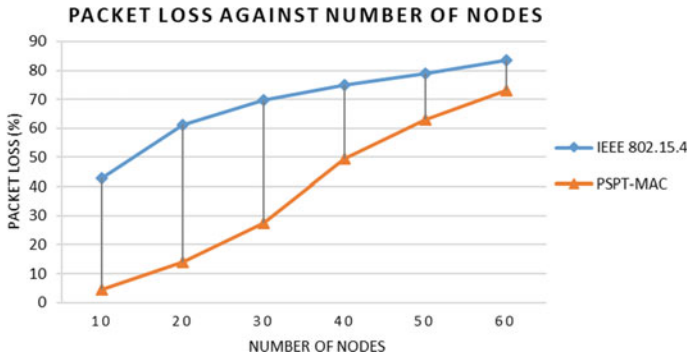


Fig. 6 Performance of packet loss between PSPT-MAC and IEEE 802.15.4 mechanisms

PSPT-MAC mechanism is still outperformed and reflected positive performances compared to IEEE 802.15.4 protocol. This has revealed the effectiveness of PSPST-MAC mechanism in alleviating the congestion of the network.

4.4 Packet Loss

Figure 6 displays the measured packet loss of PSPT-MAC mechanism and IEEE 802.15.4 standard against number of nodes. An incline trend of packet loss for both schemes can be seen as the number of nodes increased. However, PSPT-MAC mechanism still outperformed IEEE 802.15.4 standard through 30 nodes with 27.44% of packet loss compared to IEEE 802.15.4 standard with 69.62%. This resulted from less total number of packets transmitted under PSPT-MAC mechanism due to earlier discarded of corrupted packets compared to IEEE 802.15.4 standard. Later, the packets of PSPT-MAC would undergo fragmentation by using shorter slot time of 0.2 ms in the MAC IEEE 802.15.4 superframe. Hence, these packets could be served earlier in the network compared to IEEE 802.15.4 which has longer slot time of 0.96 ms before being served in the network. Thus, the possibility of congestion is lowered in PSPT-MAC mechanism than IEEE 802.15.4 standard. However, both schemes reflect high packet loss for 60 nodes which might indicate presence of congestion in the network.

5 Conclusion

Based on the findings, it is clearly shown that PSPT-MAC mechanism has yielded better performances in terms of delay, throughput, PDR and packet loss compared to standard IEEE 802.15.4 protocol and FCA-MAC mechanism. However, some

improvements need to be fulfilled for future works in order to offer high reliability in the network of WBAN such as elimination of redundant data during transmitting a bulk of medical data in the network of WBAN should be considered to avoid waste of the scarce network resources.

Acknowledgements The reported research in this paper is financially funded by Ministry of Higher Education (MOHE) Malaysia and University Malaysia Perlis (UniMAP).

Conflict of Interest The authors declare that there is no conflict of interest.

References

1. Baig MM, Gholamhosseini H (2013) Smart health monitoring systems: an overview of design and modeling. *J Med Syst* 37(2)
2. Hassanaliereagh M, Page A, Soyata T, Sharma G, Aktas M, Mateos G (2015) Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing : opportunities and challenges. In: *IEEE international conference on services computing*, pp 285–292
3. Raja KS, Kiruthika U (2015) An energy efficient method for secure and reliable data transmission in wireless body area networks using RelAODV. *Wirel Pers Commun* 83(4):2975–2997
4. Bradai N, Elhadj HBEN, Boudjit S, Chaari L, Kamoun L (2015) QoS architecture over WBANs for remote vital signs monitoring applications. In: *IEEE 12th Consumer Communications and Networking Conference (CCNC)*, pp 2–7
5. Thapa A, Shin S (2012) QoS provisioning in wireless body area networks: a review on MAC aspects. *KSII Trans Internet Inf Syst* 6(5):1267–1285
6. Kafi MA, Djenouri D, Ben-othman J, Badache N (2014) Congestion control protocols in wireless sensor networks: a survey. *IEEE Commun Surv Tutor* 16(3):1369–1390
7. Rathnayaka AJD, Potdar VM (2013) Wireless Sensor Network transport protocol: a critical review. *J Netw Comput Appl* 36(1):134–146
8. Wan C-Y, Eisenman SB, Campbell AT (2003) CODA : congestion detection and avoidance in sensor networks. In: *1st international conference on embedded networked sensor systems (SenSys'03)*
9. Tao LQ, Yu FQ (2010) ECODA: enhanced congestion detection and avoidance for multiple class of traffic in sensor networks. *IEEE Trans Consum Electron* 56(3):1387–1394
10. Kour M, Sharma B, Aseri TC (2015) A study of traffic redirection based congestion control protocols for wireless sensor networks. *Int J Comput Sci Technol* 6(2):152–158
11. Sergiou C, Antoniou P, Vassiliou V (2014) A comprehensive survey of congestion control protocols in wireless sensor networks. *IEEE Commun Surv Tutor* 16(4):1839–1859
12. Tiwari V, Misra S, Obaidat M (2009) Lacas: learning automata-based congestion avoidance scheme for healthcare wireless sensor networks. *IEEE J Sel Areas Commun* 27(4):466–479
13. Rezaee AA, Yaghmaee MH, Rahmani AM (2014) Optimized congestion management protocol for healthcare wireless sensor networks. *Wirel Pers Commun* 75(1):11–34
14. Labrador MA, Banerjee S (2002) Performance analysis of generalized selective packet discarding schemes. *Telecommun Syst* 21(1):87–101
15. Ong D, Moors T (2014) Improving video quality in congested networks through deferred discard. In: *Australasian telecommunication networks and applications conference (ATNAC)*, pp 123–128
16. Hashemzahi R, Nourmandipour R, Koroupi F (2013) Congestion in wireless sensor networks and mechanisms for controlling congestion. *Indian J Comput Sci Eng* 4(3):204–207

17. Wan Abdullah WAN, Yaakob N, Ahmad RB, Yah SA (2018) Classification of importance data for congestion control in remote health monitoring system. *Adv Sci Lett* 24(3):1767–1770(4)
18. Wan Abdullah WAN, Yaakob N, Badlishah Ahmad R, Elobaid ME, Yah SA (2019) Corrupted packets discarding mechanism to alleviate congestion in wireless body area network. *Indones J Electr Eng Comput Sci* 14(2):581–587
19. Mikhaylov K, Plevritakis N, Tervonen J (2013) Performance analysis and comparison of bluetooth low energy with IEEE 802.15.4 and SimpliciTI. *J Sens Actuator Netw* 2(3):589–613
20. Wan Abdullah WAN, Yaakob N, Badlishah R, Elshaikh Elobaid M, Yah SA, Zunaidi I (2019) Fragmentation in MAC IEEE 802.15.4 to improve delay performance in Wireless Body Area Network (WBAN). *IOP Conf Ser Mater Sci Eng* 557:012020
21. Chaari L, Kamoun L (2011) IEEE 802.15.4/Zigbee performance analysis for real time applications using OMNET++ simulator. *Adv Wireless Mob Netw Appl* 154:54–68
22. Yusuf Khan J, Yuce MR, Bulger G, Harding B (2012) Wireless Body Area Network (WBAN) design techniques and performance evaluation. *J Med Syst* 36(3):1441–1457
23. Pathak S, Kumar M, Mohan A, Kumar B (2015) Energy optimization of ZigBee based WBAN for patient monitoring. *Proc Comput Sci* 70(January):414–420
24. Nepal S, Pudasani A, Shin S (2017) A fast channel assignment scheme for emergency handling in wireless body area networks. *Sensors* 17(477):1–16
25. Yahia HS, Monnet WMA (2017) Performance of ZigBee wireless body sensor networks for electrocardiogram signal transmission under maximum payload size. *UKH J Sci Eng* 1(1):19–25

A Planar Slotted RCS Based UWB RFID Tag on a PCB and a Flexible Substrate for Packaging Application



A. K. M. Zakir Hossain , M. Z. Abidin Bin Abd Aziz, A. W. Y. Khang, and W. A. Indra

Abstract The passive chipless ultra wide band RFID tags are at the forefront of the research to replace the existing barcode technology. The attributes of the RFID system out-run various problems associated with the barcode system such as line of sight, wear and tear, tampering and etc. The chipless RFID tag in ultra wide band contains no chip that inherently makes it cheaper compared to chipped tags. However, due to their dimensions, sometime in the packaging industry, it exceeds the size of the package and also it comes with inadequate bit density. In addition, the fabricated tag is mostly designed on printed circuit boards that takes away the rapping ability on the surface of the product. The tags on flexible substrates can be a good solution for this problem. In this paper, the radar cross-section response of a new planar Dual L-Slotted ultra wide band RFID tag has been studied and investigated on a printed circuit board and a flexible substrate. The tag contains 16 bits within the compact size of $15.2 \times 7.2 \text{ mm}^2$, having a bit density of 14.62 bit/cm^2 . The tag has been realized on Taconic TLX-8 as the rigid printed circuit board and a flexible substrate, Kapton[®]HN. For both cases the code “1111111111111111” has been successfully extracted. Both are suitable for RFID applications whereby the tag realized on the flexible substrate has the potentiality to use in the packaging industry and also in sensing applications.

Keywords Passive chipless ultra wide band RFID · Barcode · Packaging industry · Flexible substrates · Radar cross-section

A. K. M. Zakir Hossain (✉) · A. W. Y. Khang · W. A. Indra
Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Teknologi Kejuruteraan Elektrik & Elektronik (FKEKE), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
e-mail: zakir@utem.edu.my

M. Z. Abidin Bin Abd Aziz
Centre for Telecommunication Research & Innovation (CeTRI), Fakulti Kejuruteraan Elektronik & Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

1 Introduction

The Radio Frequency IDentification (RFID) is a well-established technique to identify or track items/products. The RFID is a proven solution in many industries such as packaging, process, toll collection, library management, livestock and etc. [1]. The RFID provides a solid encryption and security to preserve the accurate identity and original information about the product/object to be identified. The revolution on RFID has begun back there in the World War II (WWII). Where, the first reported application of RFID was to identify the airplane that was friend or foe by the allied forces. Since then, the interests in RFID research and development have been in increasing trend day by day. One more good aspect about the RFID system is that, this does not need the line of sight (LOS) for the detection. It is less vulnerable to hacks and tempering, and also atmospheric stresses do not affect the system [2, 3]. In addition, dent, wear, tear has less effect on reading and identification whereas the conventional existing barcode system is still suffering these issues. Due to these reasons, in the last decade, many researchers have involved in developing the RFIDs that can replace the conventional barcode system [4].

The RFID system can be classified (in terms of power on board) into three categories: (i) Active, (ii) Semi active/passive and (iii) passive tags. Active tag is fully battery assisted and does not rely on the power from the reader. Similarly, the semi active/passive tags also have the battery on board but it's to power the internal circuitry (IC) only of the tag. The needed power for communication between the reader and the tag relies on the interrogation signal power from the reader. Due to this half active and half passive behavior, this type of tag sometimes called semi-active and also called as semi-passive tags. On the other hand, passive tags do not comprise any battery on board and rely merely on the radiated power from the RFID reader. Active and semi active/passive both types of tag has IC on board. However, the passive types either can have IC or can be chipless also. Since the chipless tags do not have the IC, its make them inherently cheap and put them as a potentially suitable candidate for the replacement of the barcode system [5, 6].

The passive chipless tags are mostly consists of resonators which can be read in the spectral/frequency domain (FD) or in the time domain (TD) [3]. The modulation in FD or TD follows the on off keying (OOK). Where, the presence and absence of the resonators determines the numbers of bits and the combinations for the encryption. If there are "N" numbers of resonators can be accommodated on the tag, the number of combinations can be 2^N . Due to this obvious reason, to increase the number of bits on board of the tag, the number of resonators has to be increased and consequently it makes the tag dimensions larger proportionally. In FD, the RFID chipless tags have been validated either through S-parameter (S_{21}/S_{11}) measurement [7] or the radar cross-section (RCS) of the tag which are basically designed in ultra wide band (UWB) region. The chipless tags may or may not consist of antennas. Among the chipless tags, those which do not employs the antenna(s) for communication with reader, are smaller in size due to the exclusion of the antenna on board and consequently puts them at the forefront of the research to replace the barcode. These

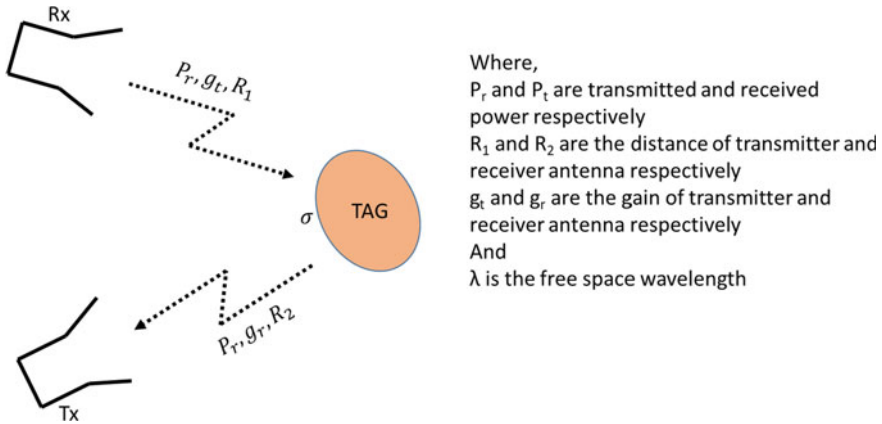


Fig. 1 Bi-static radar configuration for RCS based chipless tags

types of tags can be and usually are validated through the RCS response from the tag. Furthermore, researchers have also validated the tag through RCS response but the bit extraction has been done by different novel methods [8]. Figure 1 shows the conventional diagram for a chipless RFID system with Eq. (1) to estimate the bi-static radar based RCS (σ) [9].

In [8], so far, the highest bit density has been exhibited by the tag called slot loaded rectangular patch (SLRP). The authors have achieved bit density of 16.6 bit/cm² with dual polarized 16 bits (8-bits vertical and 8-bits horizontal) on board while the tag is partially validated with the RCS and detected by a four antenna (two cross polar at Tx and two cross polar at Rx) method. However, this combination of antennas makes this method difficult to practically implicate and makes a complex nature of detection. Also, the stated bit density decreased to half (8.3 bit/cm²) if only one polarized bits are extracted. Furthermore, although the tag is claimed to be fully printable, the authors have not

$$\sigma = \frac{(4\pi)^3 P_r}{P_t g_t g_r} \left[\frac{R_1 R_2}{\lambda^2} \right] \tag{1}$$

revealed any experimental results of the designed tag to support the claim and the tag has not been realized on any flexible substrates. As a result, the tag cannot be considered as in printable RFID family. Many other researchers have also proposed a numbers of tags with different shapes that can be validated through RCS response [10–13] and fully converted them into printable tags on flexible substrates. However, still now, through FD technique, the number of bits and bit density is not adequate as to counter the passive IC tags which has 96–124 EPC bits even though the area of the RFID tag has been shorten significantly through chipless approach. Following the same trend, in this paper, a new Dual L-slotted (DLS) rectangular shaped (LSRS) chipless RFID tag in UWB has been design that potentially can accommodate high

data bit density. The designed tag has been investigated with a rigid substrate (PCB laminate), Taconic TLX-8 and a flexible substrate, Kapton®HN. The simulation has been performed in CST microwave studio (CST MWS) 2019 to design, optimize and validate the design.

2 Tag Design

Figure 2 illustrates the design geometry of the 16-bit chipless tag. The tag comprises 16 different DLSs to make it as a 16-bit structure. One group of 8 L-slotted resonators (L_1 to L_{15}) are positioned at the right side of the tag and at the left side, 8 more inverted L-slotted resonators (L_2 to L_{16}) are placed. Since, the tag comprises two different L-slotted groups, it's called a DLS 16-bit structure. The width of the parallel lines (W) and the gap (g) between them is 0.2 mm respectively. The resonant frequency of each DLS resonator can be estimation by following Eq. 2 [8].

$$f_r = \frac{C_0}{2(L_X)} \sqrt{\frac{2}{\epsilon_r + 1}} \tag{2}$$

where, L_X is the length of the slot of each resonator and $X = 1, 2, \dots, 16$. C_0 is the speed of light at free space and ϵ_r is the dielectric constant of the substrate material. The 8-resonator group at the right side of the tag represents the odd number frequencies in the spectrum. Each resonator in their corresponding groups has a difference in length (ΔL) of 0.4 mm from there adjacent slots. The first odd resonator L_1 has a length of 14.8 mm. The following resonator (L_3) in the same group has a length of $L_3 = L_1 -$

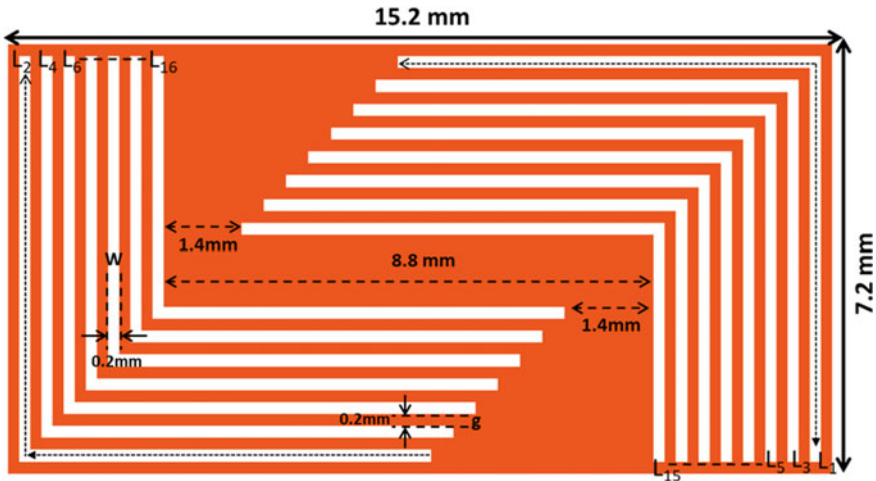


Fig. 2 Double L-Slotted (DLS) chipless tag geometry

$\Delta L = 14.4$ mm and so on as the $L_{15} = 12$ mm. Similarly, the other 8-resonator group at the left side of the tag that represents the even number frequencies in the spectrum, also has the same difference of 0.4 mm such as $L_2 = 14.6$ mm and $L_4 = L_2 - \Delta L = 14.2$ mm and so on. This combination of lengths gives a constant difference between all the resonators (L_1 to L_{16}) of 0.2 mm. Consequently, due to this difference in the slot resonators length, all the resonators represent unique resonant frequencies i.e. L_1 represents the first (lowest) resonant frequency f_1 , L_2 represents the f_2 and so on as the 16th resonator (L_{16}) represents the highest resonant frequency f_{16} in the spectrum. The tag dimension is 15.2×7.2 mm² and with 16 resonators (bits) on board, the bit density becomes 14.62 bit/cm² which is the highest so far reported among the chipless tags on printable/flexible substrates.

While exposed in plane wave in the spectral domain, the variation from the constant amplitude and/or phase of the RCS response from the tag, the resonance can be detected. If one or more resonator is absence (by shorting or removing), the spectral amplitude will not have any variation for those corresponding resonant frequency points related to the resonators. This is called the OOK method for detection of chipless RFID [9]. Whether it's with S-parameter or with RCS response, this is how the detection done conventionally. In the following section the concept has been validated.

3 Tag Validation by Simulation

To prove the concept, the tag has been simulated in CST MWS 2019. The designed RCS based chipless tag has been exposed to a circular polarized plane wave in the simulator. Figure 3a shows the simulation setup to find RCS response of the tag in CST MWS. To validate the tag structure integrity, at first, the surface current accumulation has been investigated through simulation for first two resonators (L_1 and L_2) only and the response can be seen in Fig. 3b. It is observed that while the surface current accumulates on one resonator (L_1) indicating the resonance whereby at the same time, the other resonator (L_2) has a lower surface current indicating no resonance and vice versa.

For the RCS simulation, the tag needs to be placed at the far-field region from the excitation/reader and has to have a minimum distance (R_m) away from the excitation signal (or reader antennas). Too close to the reader will make the tag fall inside the nearfield region and RCS response will not appear in the spectrum. To make sure the tag falls in the far-field region, the R_m has been calculated with (3) as follows [14],

$$R_m = \frac{2D^2}{\lambda} \quad (3)$$

where, D is the largest dimension of the object (chipless tag), in this case, 15.2 mm. And the wavelength, λ for a selected center frequency of 6.5 GHz in UWB is 46.15 mm. From Eq. (3), the calculated minimum distance $R_m = 11.01$ mm. Here,

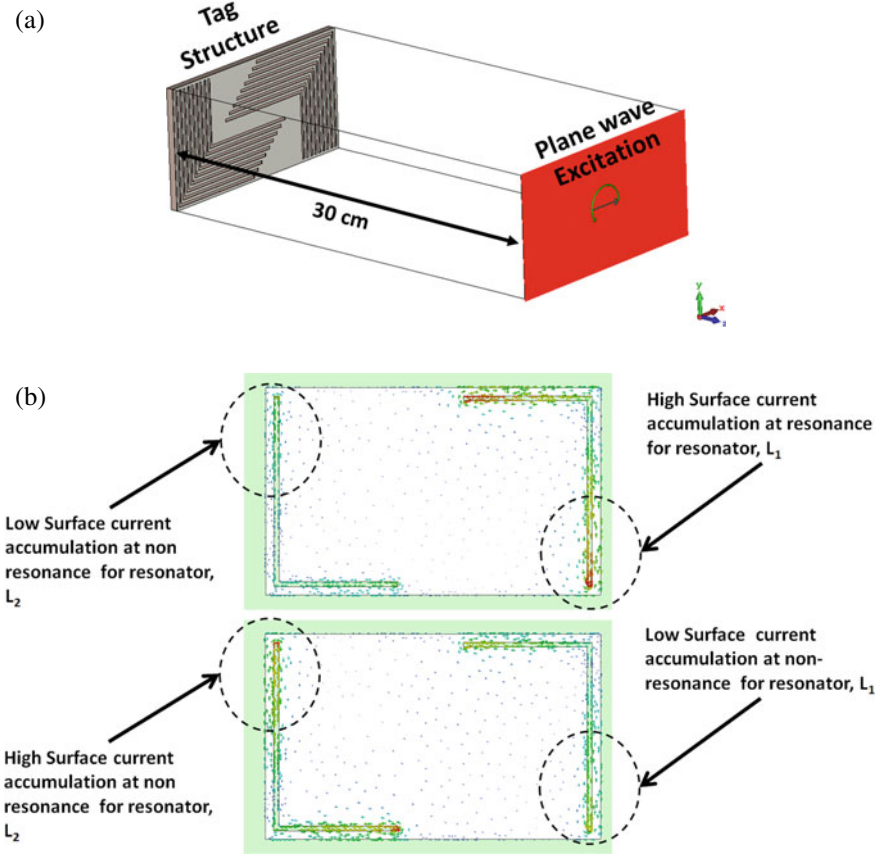
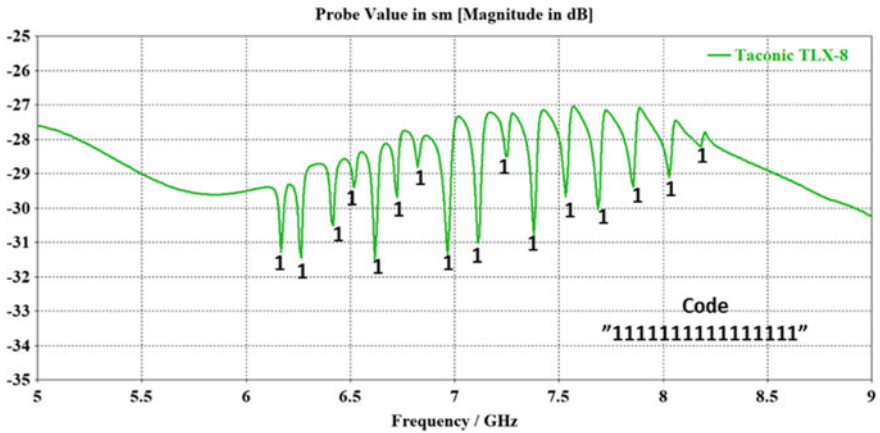


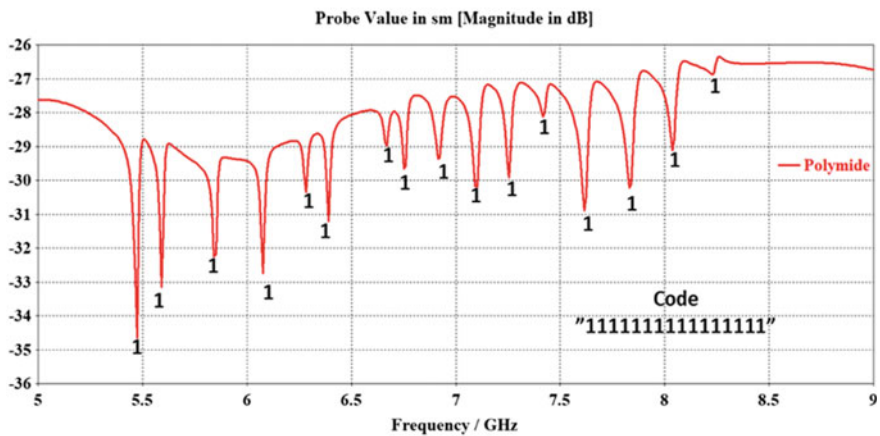
Fig. 3 Simulation setup and surface current accumulation for the chipless tag. **a** Simulation setup in CST MWS. **b** Surface current accumulation for resonant and non-resonant slots

to make sure the far-field region, the distance of the probe (RCS) has been chosen as 300 mm (30 cm) away from the tag to measure the RCS response. Figure 4a shows the simulated RCS response of the tag modeled on Taconic TLX-8 substrate. The tag structure has been designed for the band 5–9 GHz in UWB range. The reason for that is, the dimensions basically follows Eq. 2. Any inclusion of lower frequency bands (as low as 3 GHz) will ask for longer length slots. Consequently, the whole tag size will be larger and the bit density will suffer. On the other hand, it has been seen that frequency higher than 8.5 GHz gives resonance response too small to retrieve as a bit. However, the investigation has not been shown as it has been considered irrelevant for the scope.

Accordingly, the response is taken from 5–9 GHz span and can be seen that the variations in the amplitude only between 6 and 8.5 GHz. The portion that has variation in amplitude has been expanded and investigated further. It is seen that from 6 to 8.5 GHz there is a presence of 16 distinguished dips in the RCS versus



(a)



(b)

Fig. 4 RCS response on a Taconic TLX-8 and b Kapton[®]HN substrate

Frequency response as it is anticipated. The longest resonator L_1 which corresponds to the frequency f_1 has a resonance dip approximately at 6.17 GHz and the smallest resonator L_{16} that corresponds f_{16} , resonance dip at 8.2 GHz. The rest of the 14 resonators resonance in between with nominally same frequency distance apart.

The logic for one dip presence is "1" and absence is "0". Since all the 16 resonators (and their corresponding dips) are present in the spectrum, the coding bit stream can be coded as "1111111111111111". In Fig. 4b, the RCS response of the tag realized on the Kapton[®]HN film that has a dielectric constant of 3.5, the thickness of 125 μm and tangent loss of 0.0027. The response also shows exactly 16 bits same as the Taconic TLX-8 as in Fig. 4a, however, the resonance points are shifted as expected due to the change in dielectric constant as it follows the Eq. 2. Similarly, from Fig. 4b,

Table 1 Comparison between existing works

References	Number of bits	Tag area (mm ²)	Bit density (bits/cm ²)	Flexibility
[11]	24	70 × 42	0.81	Yes
[12]	24	45.2 × 43.5	1.22	Yes
[13]	24	24 × 24	4.17	Yes
[10]	24	20.6 × 19.9	5.86	Yes
[8]	16	16 × 6	16.6	No
This work	16	13.3 × 9.4	14.62	Yes

it is seen that the code “1111111111111111” has been extracted. Table 1 shows the comparison between the existing and this proposed work.

From Table 1 it can be seen that the highest bit density is proposed in [8]. However, as mentioned earlier in this paper, this is not printable/flexible on nature. Among the other works in Table 1, the highest existing proposed chipless tag with a bit density of 5.86 bits/cm² in [10] whereas the proposed work in this paper achieved 14.62 bits/cm² which is the highest so far reported and also total number of $2^{16} = 65,536$ identities can be generated with this tag.

4 Conclusion

A 16-bit high bit density passive chipless tag has been designed and validated through simulation and experimentation. Good response has been found between while modeling the tag both on Taconic TLX-8 and Kapton[®]HN. All 16 bits has been successfully retrieved from both Taconic TLX-8 and Kapton[®]HN based tags. The code of all bit presence (‘1111111111111111’) has been also extracted from the RCS versus Frequency responses. So far here, one PCB and one flexible substrate has been investigated. The proposed work has achieved the highest bit density of 14.62 bits/cm² compared to other existing works so far in terms of the printable/flexible RFID techniques. With the potentiality to be implemented in the packaging industry, more investigations will be done on this design on few more different flexible substrates for item tracking RFIDs and also for environmental parameter sensors.

Acknowledgements This work has been sponsored by the Research & Innovation Management Center (CRIM), Universiti Teknikal Malaysia Melaka (UTeM).

References

1. Hossain AKMZ, Ibrahimy MI, Motakabber SMA, Habaebi MH (2019) Planar parallel slotted circular disk passive UWB-RFID tag. *ARPN J Eng Appl Sci* 14(6):1149–1154
2. Hossain AKMZ, Ibrahimy MI, Motakabber SMA (2014) Spiral resonator for ultra wide band chipless RFID tag. In: 2014 international conference on computer and communication engineering, IEEE, pp 281–283
3. Hossain AKMZ, Ibrahimy MI, Motakabber SMA (2015) Detection of data from the UWB microstrip resonator type RFID tag. In: 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE), IEEE, pp 414–417
4. Hossain AKMZ, Motakabber SMA, Ibrahimy MI (2015) Microstrip spiral resonator for the UWB chipless RFID tag. In: *Progress in systems engineering*. Springer, Cham, pp 355–358
5. Hossain AKMZ, Ibrahimy MI, Motakabber SMA (2015) Parametric study of UWB chipless RFID tag: an impulse interrogation approach. *ARPN J Eng Appl Sci* 10(21):10138–10142
6. Hossain AKMZ, Ibrahimy MI, Motakabber SMA (2015) Tag for UWB chipless RFID: a single antenna approach. In: 2015 IEEE regional symposium on micro and nanoelectronics (RSM), IEEE, pp 1–4
7. Ibrahimy MI, Hossain AKMZ, Motakabber SMA (2017) Linear Microstrip Resonator for UWB RFID tag. *Int J GEOMATE* 13(40):167–172. <https://doi.org/10.21660/2017.40.181117>
8. Islam MA, Karmakar NC (2012) A novel compact printable dual-polarized chipless RFID system. *IEEE Trans Microwave Theor Tech* 60(7):2142–2151
9. Karmakar NC (2016) Tag, You're it radar cross section of Chipless RFID tags. *IEEE Microwave Mag* 17(7):64–74. <https://doi.org/10.1109/MMM.2016.2549160>
10. Javed N, Habib A, Amin Y, Loo J, Akram A, Tenhunen H (2016) Directly printable moisture sensor tag for intelligent packaging. *IEEE Sens J* 16(16):6147–6148. <https://doi.org/10.1109/JSEN.2016.2582847>
11. Habib A, Ansar S, Akram A, Azam MA, Amin Y, Tenhunen H (2017) Directly printable organic ASK based chipless RFID tag for IoT applications. *Radioengineering* 26(2):453
12. Sajitha VR, Nijas CM, Roshna TK, Vivek R, Vasudevan K, Mohanan P (2015) Polarization independent chipless RFID tag. *Microwave Opt Technol Lett* 57(8):1889–1894
13. Rezaiesarlak R, Manteghi M (2014) Complex-natural-resonance-based design of chipless RFID tag for high-density data. *IEEE Trans Antennas Propagation* 62(2):898–904
14. Javed N, Habib A, Akram A, Amin Y, Tenhunen H (2016) 16-bit frequency signed directly printable tag for organic electronics. *IEICE Electron Express* 13(11):20160406–20160406

Design and Analysis of Automated Inspection System for Relays Fault Detection



W. H. M. Saad , W. Y. Tai, N. M. Saad, M. S. Karis, and S. A. A. Karim

Abstract Relay is an electrical switch that consists of coil and movable iron armature and contactor that act to open and close the switch respectively. Normally relay will degrade after being operated for certain period. The conventional way to diagnose the relay condition is a time consuming and rather ticklish. The aim of this project is to provide a solution to overcome the difficulty of testing a relay using the conventional way. The embedded system is designed to be able to perform the functional test and parameter test of the relay. Functional test is done to determine the relay contacts whether it is functioning or not whereas parameter test is done to check on every pickup and drop out voltage for all contact of the relay. The embedded design developed using Arduino board with keypad and LCD for user interface and device under test (DUT) adapter and interfaced board to place the relay under testing. The relay is plugged into the suitable DUT adapter and then the DUT adapter in plug into the interfaced board. The LCD shows the menu and the result of the test. The keypad is to allow the user to select the options for relay testing. Lastly, the relay tester is successfully constructed and able to assist in detecting the faulty relay with less time taken and achieve 100% accuracy in functional test and 85% of accuracy in parameter test.

Keywords Relays testing · Automatic test equipment (ATE) · Functional test · Parameter test

W. H. M. Saad (✉)

Centre for Telecommunication Research & Innovation (CeTRI), Durian Tunggal, Malaysia
e-mail: wira_yugi@utem.edu.my

W. H. M. Saad · W. Y. Tai · N. M. Saad

Fakulti Kejuruteraan Elektronik & Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), 76100 Durian Tunggal, Melaka, Malaysia

M. S. Karis

Fakulti Kejuruteraan Elektrik (FKE), Universiti Teknikal Malaysia Melaka (UTeM), 76100 Durian Tunggal, Melaka, Malaysia

S. A. A. Karim

Department of Electrical Engineering, Politeknik Ibrahim Sultan, 81700 Pasir Gudang, Johor, Malaysia

1 Introduction

Relay is an electrical switch that consist of three component which are coil of copper wire wrapped around soft iron core, movable iron armature and contactor as open or close switch [1]. Relays are commonly having one or more poles that act as a switch when it is being activated. Each switch in the relay controls a contact between one or more circuit poles to provides electrical power between two or more circuits. Common nomenclatures of the relays are single pole-singe throw relays, single pole-double throw relays, multiple pole-single throw relays, and multiple pole-multiple throw relays [2]. Typical construction of relay includes a housing cover protection from dust and moisture for the internal circuit that consist of an electromagnetic switch coil [3].

Relays will be operated when the voltage is applied across the coil. When the current flow through the coil, it generates a magnetic field that will attract the contactor to either close or open the gap between iron armature and contactor. The relay offers a very useful function when it comes to controlling the high-power load with the low power devices such as controlling motor using microcontroller. The relay isolates the microcontroller board from the high voltage circuit to prevent it from burning out.

Pick up and drop out voltage of the relay are important characteristics to determine each type of relay function within a certain voltage range. Pick up voltage is defined as a voltage on a deenergized relay that is increasing until a certain value or lower till its starts to function. The drop out voltage is defined as a voltage on an operated relay that is decreasing until a certain value or higher which all contacts must revert to their unoperated position [4]. To design a proper circuit with the relay on it, pick up and drop out voltage must be taken into a consideration to ensure the switching mechanisms are working well with the voltage range applied to the coil [5]. It is widely use in the automatic test equipment (ATE) for testing IC in production line [6]. In the event of faulty detection operation of the electronic equipment, it is custom to locate the printed circuit board (PCB) or group of boards that might contain the defect component including diodes, transistors, and relay. Although a test equipment to identify defective components on a PCB is relatively simple for other component but for the relay, there are still no dedicated equipment's that can test the relay pick up and drop out voltage accurately with ease and time efficient as relays has many pins on it.

Relay is a very common electronic component and widely used in electronics circuit. In electromagnetic relay, electric current can flow through or be interrupted from flowing through an electric circuit by opening and closing its contact [7]. With an ever-increasing use of relays in present day devices, it is a must to test the relays to determine a relay condition [8]. An electromagnetic relay's electrical performance parameter often needed to be tested in order to assess whether it satisfies its working requirements before being put into usage [9]. Conventionally, testing electrical relays is accomplished using a device that incorporating an analog meter. The DC relay manufacturer needs to test the switching mechanism of electromagnetic relays more

than a hundred times to verify its performance [10]. The standard procedure on testing the relays is by connecting the source of the variable voltage power across the coil of the relay and gradually increasing it. The amount of current or voltage passing through or across the coils can be measured by using the analog multimeter. When enough current is supplied through the coil, the electromagnetic field created by the coil causes the armature of the relay to shift or pick up from its original position, thereby closing the relay switch. Relay's life span is short due to the inconsistent opening and closing of the mechanical switch [11]. It could also happen when oxidation occurs to the switching contact of the relay after it is being operated for a certain period [12]. Therefore, the specification and the functionality of the relay will degrade as time passes, but that could also happen to the new relay too. If a pickup or drop out voltage of a relay runs out of its order, it will either fail to close or open the switch with the desired voltage range supply to its coil. With this failure, the possible issues that could be happened are the short circuit or unwanted open circuit and eventually causing the board to faulty.

The basic instrument to test the relay pickup and drop out voltage are just the power supply unit and a multimeter, but the way to test it are very hard to be done alone since a single relay often has several pins on it. Troubleshoot the relay faulty in industrial PCB (refer to Fig. 1) and device under test (DUT) board often face a problem related to the relay itself. This type of boards typically uses relays such as pin-through-hole relays to perform switching to minimize signal distortion. These relays are expensive and have a limited life span. Faulty relays are the main cause of the test failures for the board. Due to the hectic process of testing each relay, all relays on the board are replaced at regular intervals whether they are faulty or not. The common issue encountered by the relay on the board like this is that the pickup and the drop out voltage do not meet the specification with respect to its datasheet. Usually it takes 3–5 min to test a two form C type relay, and normally there are more than 50 relays in the PCB board like this. It will take a lot of times to check the entire relay one by one to troubleshoot it faulty.



Fig. 1 Example of industrial PCB use in automated test equipment as a load board

The aim of this study is to design a device to provide a convenient way for testing the pickup and drop out voltage for different types of relay and at the same time is to shorten the testing time. The duration of troubleshooting must be as fast as possible, because the relay is widely used in the load board of automated test equipment thus manufacturer would lose a lot of money if the testing process is stopped for a long period of time.

The relay tester proposed here can help in shortening the troubleshooting time and by that the profit of the manufacturer can be secured and the stress of the technician whom perform troubleshooting can be reduced either. In addition, with the effort of repairing the faulty PCB board can reduce the e-waste. According to [13], the issue about e-waste are getting very serious and the awareness must be emphasized by everyone. With the help of the relay tester proposed in this project, the rate of load board getting scrap can help in reducing the e-waste and create a greener environment.

This manuscript consists of four sections including this section. On the methodology section, the development of the proposed relay tester is described in terms of its hardware development and software development. Next, the result of the relay tester accuracy is demonstrated and discussed based on two types of testing, which are functionality and parameter tests. Last section is the conclusion made based on the project outcome.

2 Methodology

The embedded system is designed to enable the functional test and parameter test of the relay to be done automatically. Functional test is done in order to determine the relay contacts whether it is functioning or not whereas the parameter test is done to find every pickup and drop out voltage for all contacts of the relay. The embedded system for relay tester is constructed with Arduino board as a main controller and LCD panel is used to show the menu and the result of the test. Keypad is used to enter the input setting and configuration whereas the sample of relay is plugged into a DUT adapter that connected to the interface board and then to the microcontroller for testing purposes. Figure 2a shows the hardware block diagram of the automated inspection system and Fig. 2b the complete hardware component. The PCB used in the developing of interface board is custom design into the commercialized standard to make the relay tester more reliable for long lasting against the oxidation occur. Besides, the guide of using the relay tester to perform relay testing is provided with simplified version and user friendly.

Arduino Uno incorporates with the VI board for variable voltage supply, interfaced board to centralize all connections from different components and DUT adapter to fit the relays while doing the test for pick up and drop out voltage. Arduino Uno acts as the brain to control and decide the operations of the tester. The DUT adapter is replaceable based on the dimension of relays to be tested while the VI board supplies the voltage to the DUT adapter's coil terminal gradually. This tester only covers relay with terminology of single pole double throw and below voltage rating of 12Vdc.

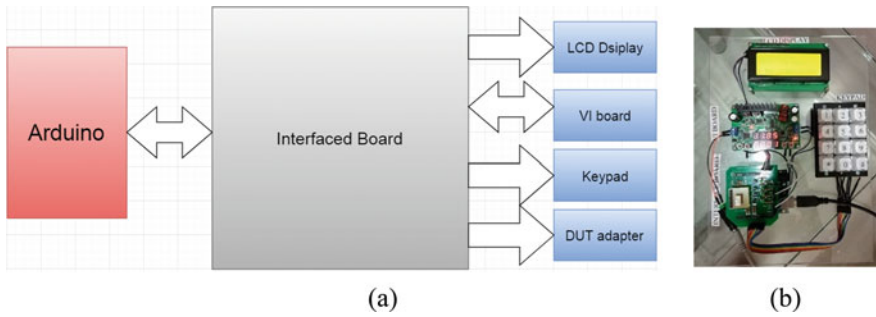


Fig. 2 Design of the relay tester system. **a** Hardware block diagram of the main controller and all available interface. **b** The hardware assembly of all component in the design

The programming code is developed under Arduino IDE. The coding is written allowing the microcontroller to execute the functional and parameter test on the relay. The purpose of functional test is to find out the relay is faulty or not whereas parameter test is function in order to find out what are the exact value of relay’s pick up and drop out voltage. The flowchart in Fig. 3 shows the sequence of the test program.

3 Result and Discussion

The test sequences were carried out to determine the relay tester prototype performance in term of its reliability. The experiments is done to compare the manual test performance with the test by using the developed relay tester in the aspect of accuracy and time consumption. The experiment includes 100 loops functional test, 100 loops parameter test, time taken to complete a whole test and the result is recorded and analyzed.

3.1 Functional Test

The functional test is repeated on a single relay for 100 times. The experiment is to ensure the test result obtain by the relay tester is reliable and trustable. In this experiment the three types of relays which are TN2, EA2 and UA2 relays were undergone the same procedure. The test results were monitored, logged and recorded in the serial monitor of Arduino to analyze it further.

Figure 4a shows the graph of each switches in a good condition UA2 relay versus the count of the result getting pass in 100 times functional loop test whereas Fig. 4b shows the graph of each switches in a bad condition UA2 relay versus the count of the result getting fail in 100 times functional loop test. Based on the result obtained

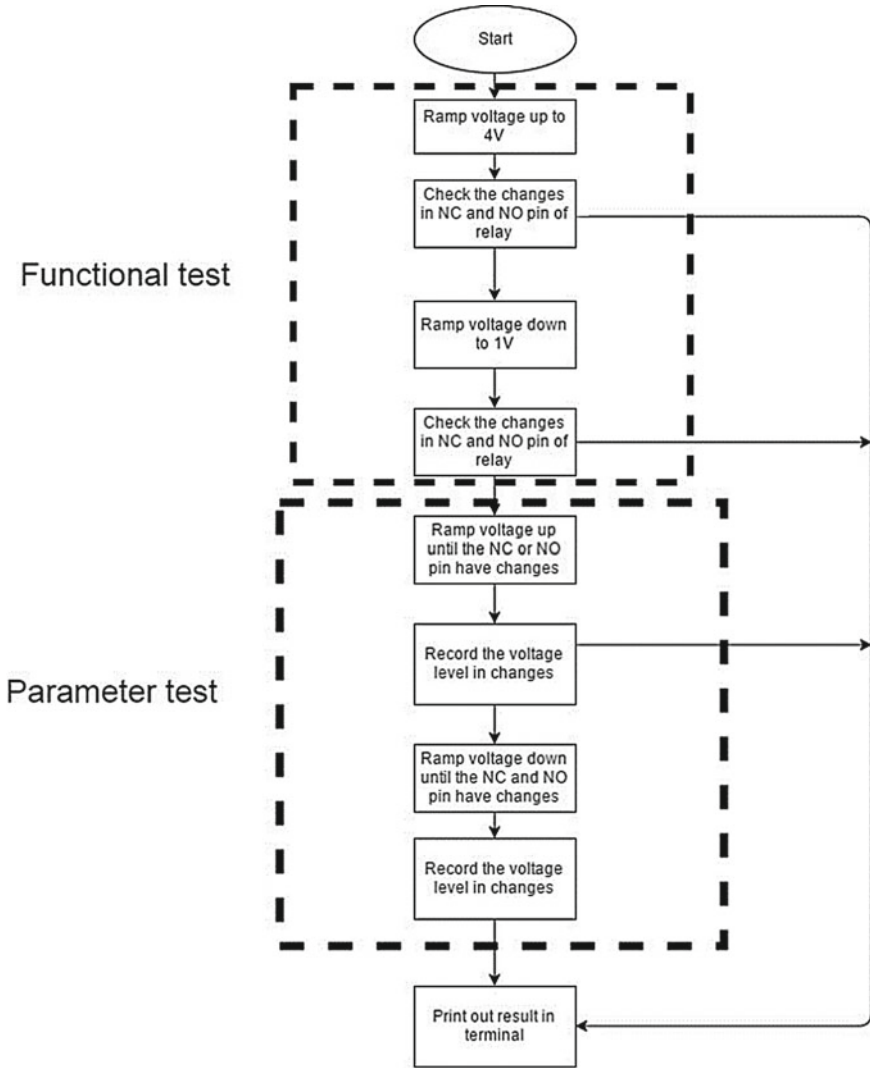


Fig. 3 Flowchart of the microcontroller setting for functional and parameter test

in both plots, it shows that the relay tester performance is 100% in accuracy and precision for UA2 relays tested. The same performances are shown for the other two types of relays which are TN2 and EA2 types.

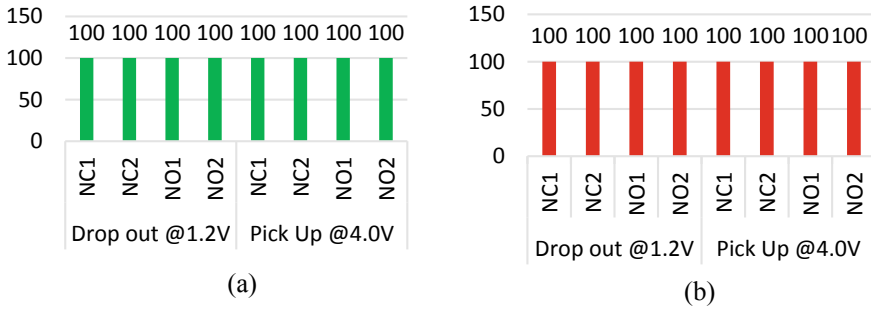
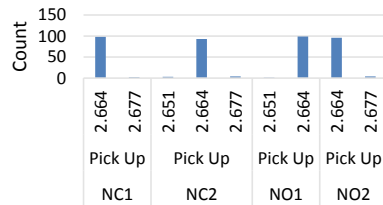


Fig. 4 Functional test demonstration in 100 loops on UA2 type of relays **a** good condition relays **b** bad condition relays

Fig. 5 Pick up voltage result in 100 loops parameter test on a good condition UA2 relay types



3.2 Parameter Test

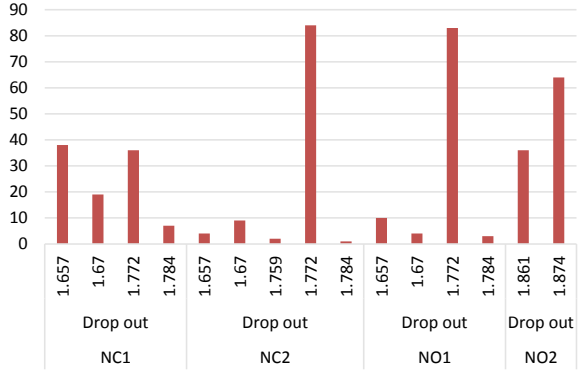
In this experiment, the pickup and drop out voltage test is carried out and repeat for 100 times as well and demonstrate only for a good condition of UA2 type of relay. The results are then recorded using serial monitor of Arduino IDE for further analysis. Figure 5 shows the graph of the count on a pickup voltage for 100 times loop test. Its show that the pickup voltage value obtain in the test were mostly consistent which is around ~2.664 V for all relay switch contact with a very minimal difference sometime while the test is repeated in a short period of time.

The result for drop out voltage shown in Fig. 6 demonstrate the same consistency in which that the value of the drop out voltage is mostly around ~1.772 V for all switch contact in relays for 100 times drop out voltage test on a single UA2 relay. The result shows the drop out voltage obtain in contact NC1, NC2, NO1, and NO2 are not far from each other and the consistency is high, therefore the parameter test is trustable too for the proposed relay tester.

3.3 Time Taken for Complete a Test by Using Relay Tester

By comparing the manual test and using the relay tester, the time taken for complete a whole test for a single relay is way shorter. The time taken for manual test is taken

Fig. 6 Drop out voltage result in 100 loops parameter test



around 6 min for a single test but using relay tester, the average needed for testing is within 2 min for diagnose any types of 5 V coil rating relay in 100 repetition.

4 Conclusion

The study shows that the relay tester is successfully constructed. The relay tester is able to assist in detecting the faulty relay with less time taken and better accuracy than manual testing. The relay tester consists of the LCD screen, keypad, interfaced board and Arduino microcontroller to work together in order to diagnose the condition of the relay. The different DUT adapter can fit different types of relays which are TN2, EA2, and UA2 relays. The accuracy of the proposed relays tester is demonstrated in 100 repetition loop tests for both functional and parameter test. Based on the result, it obtains 100% consistency within certain acceptable range for both tests.

Acknowledgements Authors would like to thank the Machine Learning & Signal Processing (MLSP) research group under Center for Telecommunication Research and Innovation (CeTRI) and Rehabilitation Engineering & Assistive Technology (REAT) research group under Center of Robotics & Industrial Automation (CeRIA) of Universiti Teknikal Malaysia Melaka (UTeM), Center of Edu-Tourism, UTeM and Faculty of Electronics and Computer Engineering (FKEKK), UTeM for sponsoring this work and providing the use of the existing facilities to complete this project. Authors would also like to thanks to Center for Research Innovation and Management (CRIM), UTeM for their sponsor on the publication fees.

References

1. Allen DE, Smith D (2012) Current zero cross switching relay module using a voltage monitor. US 12/191,641, Legrand Home Systems, Inc., 10 Apr 2012
2. Mashaney R (2016) Relay testing assembly. US 14/861,607, 24 Mar 2016
3. Wilferd RL (2016) Quick reference relay diagnostic circuit. US 13/660,558, 26 Jan 2016

4. Panasonic Electric Works Europe AG (2012) Relay technical information. Retrieved from https://www.panasonic-electric-works.com/pew/eu/downloads/ds_x61_en_relay_technical_information.pdf, 31 May 2019
5. Smith T, Childers M, Caldwell P (2014) Testing the tester, common pitfalls testing microprocessor based relays. In: 67th annual conference for protective relay engineers, College Station, TX, pp 618–626, 2014
6. Ruan MC, Shen J, Wheeler CB (2001) Latching micromagnetic relays. *J Microelectromech Syst* 10(4), Dec 2001
7. Kubono K, Yuba T, Hasegawa Y, Murakoshi T (2015) Electromagnetic relay. US 13/597,469, Fujitsu Component Ltd., Japan, 1 Apr 2015
8. Eziyi I (2014) Apparatus for testing an electromechanical relay. US 14/459,658, 4 Dec 2014
9. Po LZ, Lee Z, Wu YJ, Wu JC, Lee XJ, Yu W (2015) Electromagnetic relay testing equipment and method. CN 201310322157, Third Research Institute of the PLA. Second Artillery Equipment Academe, China, 2 Dec 2015
10. Rong TL (2013) Electromagnetic relay test device. CN 201220315006, China, 23 Jan 2013
11. Fukumoto T, Higuchi T (2016) Relay unit and method for controlling relay circuit. US 14/879,223, Omron Corp., Japan, 4 Aug 2016
12. Overton MS, Davies R, Chabreck T (2016) Relay failure detection system. US 14/242,673, Tektronix, Inc., US, 1 Dec 2016
13. Hossain MS, Al-Hamadani SMZF, Rahman MT (2015) E-waste: A challenge for sustainable development. *J Health Pollution* 5(9):3–11

Visible Light Communication-Based Indoor Notification System for Blind People



J. X. Jerry Chong, S. Saon , A. K. Mahamad , M. B. Othman, Nawir Rasidi, and M. Ikhsan Setiawan

Abstract Visible Light Communication (VLC) is a wireless communication which utilizes the visible light as the medium for data transfer. In this project, a VLC-based Indoor Notification System is proposed to help the blind people, and as a proof the concept (PoC) of VLC-based Indoor Notification System (ILS) through the developing of a prototype system. Furthermore, the performance of the developed system is evaluated in terms of distance between the transmitter and the receiver, line of sight for the VLC system, and stabilities of the system. Arduino Uno is used as the microcontroller to drive the phosphor converted LED bulb for audio transmission. The receiver circuit contain of solar cell to receive the audio signal and Divoom iFit-1 portable speaker to playback the audio instruction. Besides, the programming code to send the audio instruction via pc-LED bulb for the Arduino Uno has been developed to ensure the successfulness of the project. As the result, the transmitted audio instruction is successfully playback by the portable speaker. By increasing the distance between the transmitter and the receiver more than 50 cm or the solar cell is slightly out of the line of sight of the light beam emitted by the pc-LED bulb, the audio volume is getting weaker and furthermore causing the communication to be disconnected. The system can operate in any indoor environment because it is not getting any interruption from the radiation frequency. In conclusion, VLC has the potential to replace others wireless communication system in the future.

Keywords VLC · Indoor notification · Audio notification · Blind people · Arduino

J. X. Jerry Chong · S. Saon (✉) · A. K. Mahamad · M. B. Othman
Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
e-mail: sharifa@uthm.edu.my

S. Saon · A. K. Mahamad
Internet of Things Focus Group, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

N. Rasidi
Politeknik Negeri Malang, Soekarno Hatta No. 9, Malang, Indonesia

M. Ikhsan Setiawan
Department Civil Engineering, Narotama University, Surabaya, Indonesia

1 Introduction

The indoor positioning or navigation technique which uses the communications in between the infrastructure inside a building has become a concern nowadays [1]. Although Global Positioning System (GPS) is used extensively recently, the positioning accuracy of GPS is located in the order of several meters [2] which is unacceptably large for indoor scenarios. Since the outdoor positioning technology cannot be used in an indoor environment. Thus, further positioning techniques should be used to facilitate the Indoor Positioning System (IPS). Indoor notification system (INS) able to utilize the sound system to notify the blind people whether there is any obstacle around them and get them to the correct path.

With the combination of Visible Light Communication (VLC) technology, the INS has the unique advantages compared with the other INS which uses the others type of wireless technologies such as Frequency Modulation (FM), Cellular Networks, ZigBee, Wi-Fi, Infrared (IR), Ultrasound, Bluetooth, Radio-Frequency Identification (RFID) and Ultra-Wideband (UWB) [1, 2]. VLC does not generate RF interference, thus VLC-based INS can be used in RF sensitive areas such as mines, hospitals, gas stations and air crafts. VLC utilized the simplest component which is the light emitting diode (LED) as the transmitter that has a narrow bandwidth which enables the Angle of Arrival (AoA) information to be more precise and accurate at the receiver side [3]. The accuracy of VLC-based notification system is much higher than traditional notification technologies in an indoor environment. For example, the research shows that the number of LED luminaires is ten times more than Wi-Fi in a typical indoor building [4], which contributes to the higher precision and accuracy. The INS can work more effectively with these advantages available in the VLC technology especially in the areas where it is sensitive to Radio Frequency. This system can be helpful to blind people to move around in a building with the attachment of the sound system to acknowledge the blind people.

Blind people have difficulties to move around in unfamiliar indoor or outdoor environment by themselves [5]. There are many skills and aids considered by professionals working in the orientation and mobility to help visually impaired people go outdoor safely such as the use of canes, guide dogs, and mobility training. Although these skills and aids able to help the visually impaired people to travel outdoor, but these skills and aids is not practical and not suitable to apply inside the unfamiliar buildings. In addition, heavy obstacles such as cupboards, desks, and others may hurt the blind people if they did not realize they are going near to the obstacle and hit the obstacle. For some buildings like offices, hospitals and others may have renovation or shifting objects which may cause the blind people to get injured because they thought that they are still moving around in a familiar building. Therefore, to solve this problem, an INS for blind people should be develop so that the blind people can move and travel easily inside an unfamiliar building. Besides that, this system can be attached to the heavy obstacles to avoid the blind people from hitting the obstacles.

The objective of this project is to develop the VLC-based INS via hardware prototype with the attachment of the portable speaker. Furthermore, the performance of the developed system is evaluated in terms of distance between the transmitter and the receiver, line of sight for the VLC system, and stabilities of the system.

1.1 Visible Light Communication (VLC) Technology

VLC is a medium for data communication which utilized the visible light in between the bandwidth of 400 and 800 THz as the transmitting signal and able to be received by the solar cell, photodiode or light dependent resistor [6].

VLC technology utilizes the fluorescent lamps and LEDs to transmit signals at the rate of 10 kbits/s and 500 Mbits/s respectively [7]. The uses of LEDs had increased rapidly [1] for illumination purposed in most of the buildings such as houses, hospitals and supermarkets. This provided many opportunities for VLC technology to be widely applied in our daily lives. The LEDs nowadays can change to different light intensity levels at a very fast rate in which human eyes might also unable to perceive it. Thus, the data can be encoded in the emitting light of LEDs to execute communication. Therefore, LEDs can be used for illumination purposed and at the same time, information can be sent through the emitting light [1]. Widely used of VLC technology was previously reported. [8] apply this technology for underwater VLC and was proven beneficial in term of cost and energy consumption, yet increase the data rate. While VLC technology for indoor positioning were successfully manage to calculate the distance and estimate the actual coordinate thus reduce the positioning error [9, 10].

1.2 Notification System

Notification system is a system that utilize the simple mechanism to alert or notify the people through visual [11], vibrations [12], sounds [13], or texts [14], once emergency or events had occurred. Sound and vibration notification are very important for the visually impaired people to travel in an indoor environment. While emergency event happened such as the building is on fire, the sound notification from the fire alarm not only helps to notify the normal people but also the blind people to quickly exit the building. If the building has the real time INS for the blind people, then the blind people can utilize the system to exit the building easily and safely.

Thus the advantage of VLC technology and notification system are taken as a motivation of development of this project, in order to assist the visual impairment people to navigate in the indoor environment.

2 The Development of VLC-Based INS

In this project, there are two main parts should be concerned to ensure the successfulness of the project. Programming code development is one of the VLC-based INS development part. Furthermore, the hardware development is another part of this project. In the hardware development part, two circuits have been developed, which are the circuit for transmitter and receiver. Each circuit consists of the main components, which are Arduino Uno microcontroller and a pc-LED bulb for transmitter, whereas a 6 V solar cell and a portable speaker are requested for receiver. Once the programming code and the prototype are developed, the coding must integrate with the prototype. Before the integration of the programming code and hardware, the hardware is being tested to certify the performance of the transmitter and receiver of the VLC system. This is to ensure the range of the VLC system to be work effectively. Figure 1 shows the overall development process of the project.

2.1 Programming Code Development

This VLC-based INS utilizes the audio instructions to notify the blind people, thus generated audio signal is requested to be received by receiver. Audio generation method utilizes the encoded 16-bits pulse-code modulation (PCM) format audio for the Arduino Uno. A software called “Audacity” is used to convert the recorded audio instruction into a 16-bit PCM format with project rates of 8 kHz MP3 audio. Once the audio is down-sampled, an application of audio encoder is used to convert the down-sampled audio into numeric values so that it is able to integrate with the Arduino programming code. Besides that, the Arduino Uno microcontroller able to execute the audio in numeric values, with a library named PCM was installed in the Arduino IDE software.

2.2 Hardware Development

Transmitter and receiver circuit are built individually and the performance of this circuits were tested and analysed in term of the distance between the transmitter and receiver, and the angle of surface of the solar cell receiver with the fixed position of the LED bulb. Figure 2 shows the schematic diagram of VLC transmitter circuit used for the performance testing.

Figure 3 shows the schematic diagram of VLC receiver using solar cell and buzzer. Equations (1) is used to further analyse of the VLC power loss according to the data collected in different distance and angle of the transmitter and receiver.

$$P_{Loss}(\text{dB}) = 10 \log_{10}(P_R/P_T) \quad (1)$$

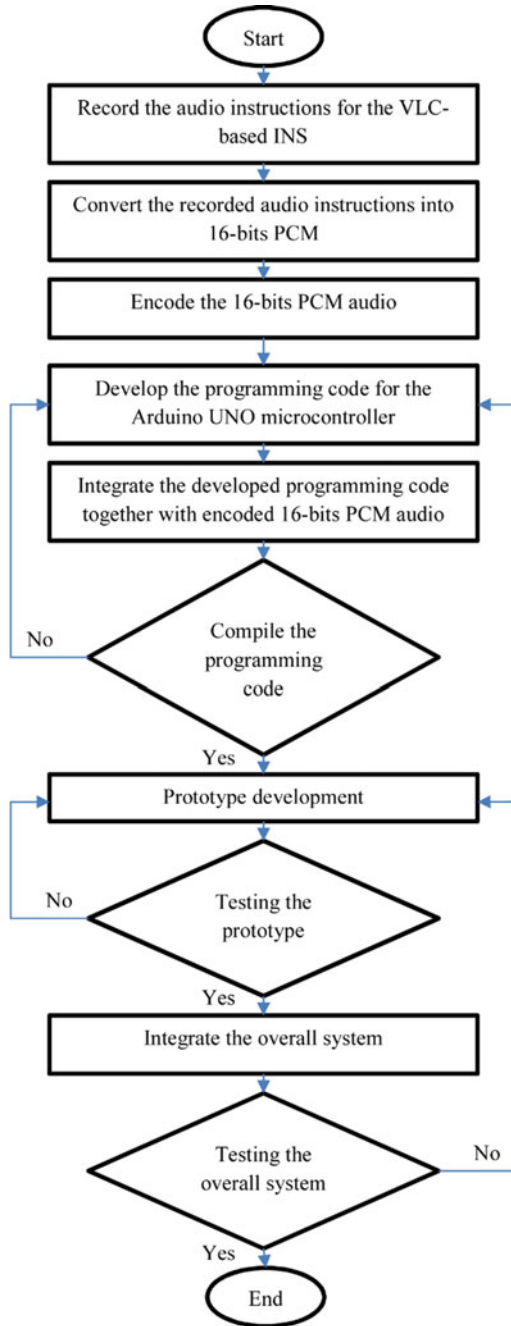


Fig. 1 Overall development process of the project

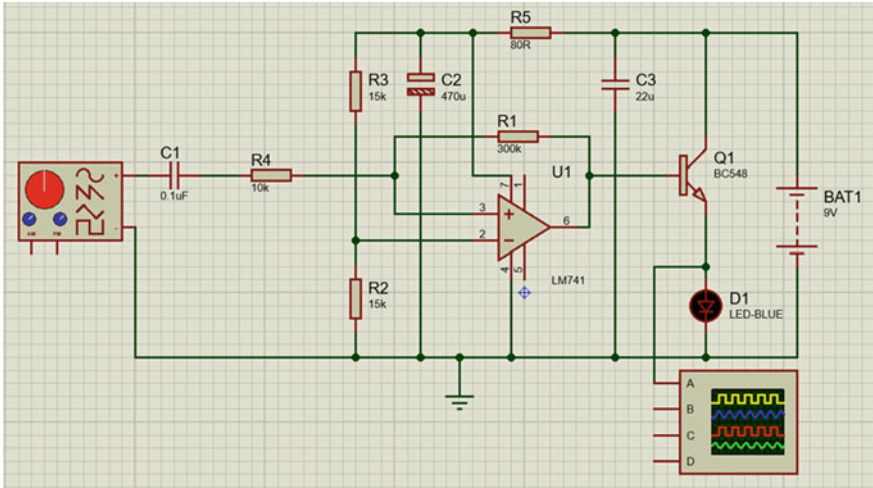
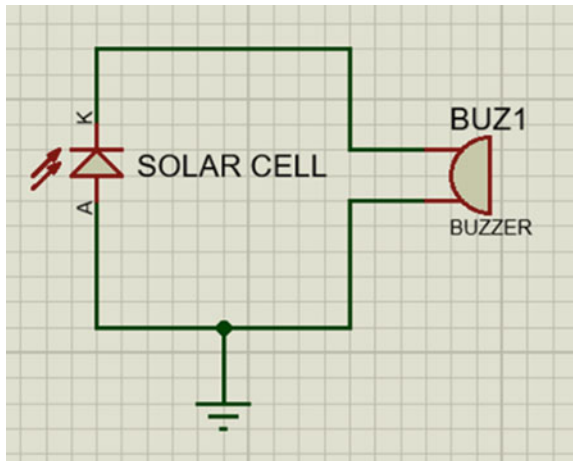


Fig. 2 The schematic diagram of the VLC transmitter circuit

Fig. 3 The schematic diagram of the VLC receiver circuit



where P_R is the power of receiver and P_T is the power of transmitter.

Once the analysis is done, the prototype of the VLC-based INS is developed. The prototype for the VLC-based INS utilize the Arduino Uno to send the audio instruction repeatedly through the LED bulb to the solar cell. Then, the audio is being playback by the portable speaker connected to the solar cell that carried by the blind people. The Arduino Uno required a 12 V, 2200 mA battery for the operational purpose. The prototype system is shown in Fig. 4.

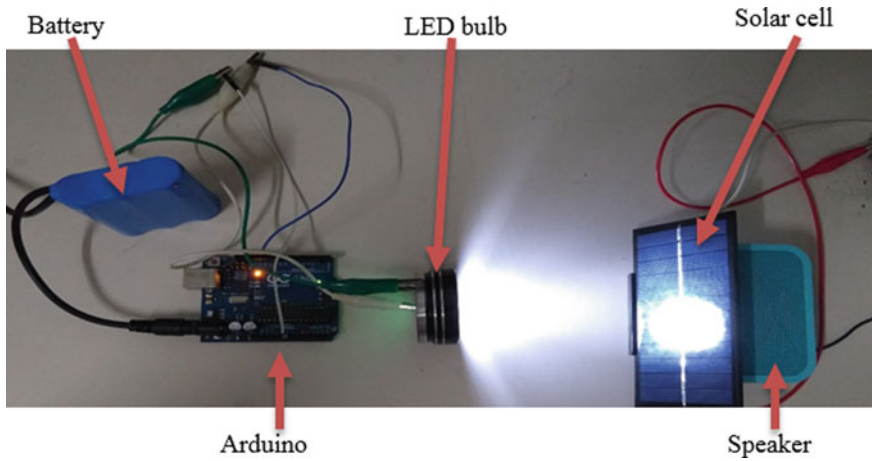


Fig. 4 Prototype of VLC-based INS

3 Performance Analysis of VLC Interfacing System

Experiment on the performance of the developed system was conducted on the various distance and angle; which are 0, 25, 50, 100 cm and 0° until 90°, respectively. The results show that the voltage and current were changing at the receiver while the distance and angle between the LED bulb and solar cell were changeable. Table 1 shows the results obtained from the transmitter and receiver that placed apart in 100 cm of range with various angles. Justification of experiment for distance range up to 100 cm is mentioned in [1].

Table 1 VLC system performance at 100 cm with various angles

Angle of solar cell (°)	Transmitter			Receiver			
	Voltage (V)	Ampere (A)	Power (W)	Voltage (V)	Ampere (mA)	Power (mW)	Power loss (dB)
0	2.82	0.26	0.73	3.21	0.08	0.26	-34.54
10	2.82	0.26	0.73	3.56	0.12	0.43	-32.33
20	2.82	0.26	0.73	3.58	0.13	0.47	-31.95
30	2.82	0.26	0.73	3.62	0.14	0.51	-31.58
40	2.82	0.26	0.73	3.65	0.15	0.55	-31.25
50	2.82	0.26	0.73	3.73	0.17	0.63	-30.61
60	2.82	0.26	0.73	3.76	0.18	0.68	-30.33
70	2.82	0.26	0.73	3.78	0.18	0.68	-30.31
80	2.82	0.26	0.73	3.80	0.19	0.72	-30.05
90	2.82	0.26	0.73	3.80	0.19	0.72	-30.05

Results summarization is shown as in Table 2 with the parameters of angle of solar cell, the distance between LED bulb and solar cell, and the power loss (dB) at the receiver. This summarization able to represent the comparisons between the power loss at the receiver based on different distance between the LED bulb and solar cell, and the angle of solar cell. From the comparisons shown in Table 2, the position of the receiver contributed the value of power loss of the system. Figure 5 shows the graph plotted based on the summarized results from Table 2 with the distance and angle between the transmitter and the receiver are represented.

Table 2 Summary of power loss at receiver against the distance and angle

Angle of solar cell (°)	Power loss at receiver (dB)			
	0 cm	25 cm	50 cm	100 cm
0	-27.73	-32.53	-33.47	-34.54
10	-27.05	-27.56	-30.93	-32.33
20	-26.56	-26.24	-29.19	-31.95
30	-26.26	-25.11	-28.14	-31.58
40	-26.24	-24.31	-27.32	-31.25
50	-26.68	-23.56	-27.03	-30.61
60	-27.76	-23.32	-26.88	-30.33
70	-29.13	-22.78	-25.91	-30.31
80	-32.16	-22.71	-25.81	-30.05
90	-35.80	-22.49	-25.60	-30.05

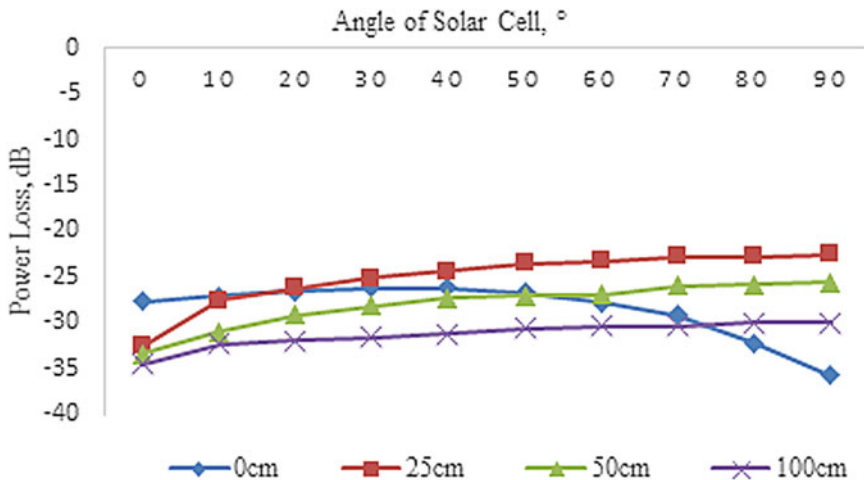


Fig. 5 Graph of the power loss at receiver against the distance and angle

Figure 5 significantly shows that the distance and angle between transmitter and receiver affect the receiver power loss. The longer the distance between the transmitter and the receiver, the lower the power of light emitted to the receiver, thus the calculated power loss at the receiver getting weaker.

4 Performance Analysis of VLC-Based INS

There are seven different audio instructions are applied to the VLC-based INS prototype, and playback by the portable speaker which connected to the receiver. This audio instructions are listed in Table 3 with the playback durations.

The `delay()` instruction in the `void loop()` function are useful for the playback operation. Therefore, the `delay()` instruction must have enough of delay time for the audio instruction to fully transmitted by the transmitter and playback by the portable speaker before next instruction is executed.

Each of the audio instruction listed in Table 3, can be transmitted by one transmitter individually. Therefore, the transmitter must be attached to a place with a suitable audio instruction programmed into the system. For example, if the transmitter is attached on the wall close to a heavy obstacle, the transmitter must deliver the audio instruction of “obstacle on the left/right ahead”, which is depends on the obstacle’s location.

The system can perform well at the distance of 50 cm between the transmitter and the receiver in which the volume of audio instruction can be heard clearly by the user. This is a great feature of the notification system for the blind people to avoid from hitting the obstacle before it is too late. Besides that, starting from the range of 100 cm, a weaker audio instruction already can be heard from the portable speaker. This can alert the blind people earlier that there is an instruction to be given, until

Table 3 The audio instructions with length of durations

Audio instruction sent by transmitter	Audio instruction received and playback by portable speaker	Durations of the audio instruction (s)
Go straight	Go straight	2
Obstacle on the left ahead	Obstacle on the left ahead	3
Obstacle on the right ahead	Obstacle on the right ahead	3
You may turn left or go straight	You may turn left or go straight	3
You may turn right or go straight	You may turn right or go straight	3
In one step ahead, turn left	In one step ahead, turn left	4
In one step ahead, turn right	In one step ahead, turn right	4

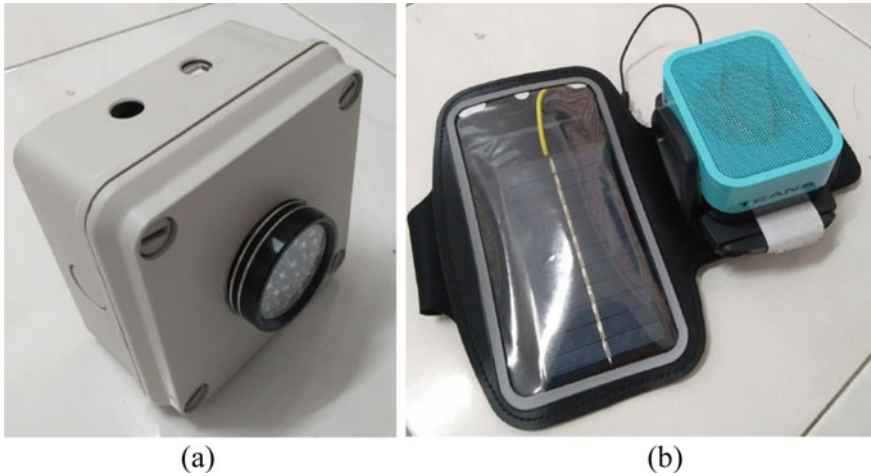


Fig. 6 The prototype system, **a** model of LED bulb transmitter and **b** receiver

they reach the certain range where they can hear the clearer instruction, then they are able to take the action as soon as possible. Figure 6a, b shows the prototype of the transmitter, and receiver, respectively.

5 Conclusion

The prototype of the VLC-based INS is successfully developed. This project has been developed with transmitter and receiver system based on the visible light communication (VLC) that utilized the phosphor converted LED bulb (pc-LED) with specific light spectrum to communicate with the receiver. The audio instruction can be sent by the Arduino Uno via the pc-LED bulb to the solar cell and playback by the Divoom iFit-1 portable speaker. Thus, this system can be used to notify the visually impaired people to travel in an indoor environment.

The performance of the developed VLC-based INS has been analysed in terms of distance between the transmitter and the receiver, and the angle of the solar cell. The distance and angle of solar cell affect the performance of the VLC system, due to the line of sight between the transmitter and receiver. Clear communication and notification information can be archive for distance up to 50 cm with 0° to 50° of solar cell angle. However, the performance of the system is decreased for the distance more than 50 cm and 50° due to lower of the power of light emitted and interference from the other light source. If the solar cell is facing the light beam emitted by the pc-LED bulb directly, the receiver totally receive the audio signal. While the solar cell is totally out of the line of sight of the light beam, thus, no communication occur between the transmitter and receiver.

Acknowledgements The financial support received from the Registrar Office of Universiti Tun Hussein Onn Malaysia is gratefully acknowledged.

References

1. Luo J, Fan L, Li H (2017) Indoor positioning systems based on visible light communication: state of the art. *IEEE Commun Surv Tutor* 19(4):2871–2893
2. Dardari D, Closas P, Djuric PM (2015) Indoor tracking: theory, methods, and technologies. *IEEE Trans Veh Technol* 64(4):1263–1278
3. Bilgi M, Sevincer A, Yuksel M, Pala N (2012) Optical wireless localization. *Wirel Netw* 18(2):215–226
4. Li L, Hu P, Peng C, Shen G, Zhao F (2014) Epsilon: a visible light based positioning system. In: 11th USENIX symposium on networked systems design and implementation, vol 1, pp 1–13
5. Riazi A, Riazi F, Yoosfi R, Bahmehi F (2016) Outdoor difficulties experienced by a group of visually impaired Iranian people. *J Curr Ophthalmol* 28(2):85–90
6. Bharadwaj V (2014) Colours: a scientific approach. *Int J Res* 1–6
7. Marwaha V (2014) Study of visible light communication. Special issue 1
8. Miramirkhani F, Uysal M (2017) Visible light communication channel modeling for underwater environments with blocking and shadowing. *IEEE Access* 6:1082–1090
9. Gligoric K, Ajmani M, Vukobratovic D, Sinanovic S (2018) Visible light communications-based indoor positioning via compressed sensing. *IEEE Commun Lett* 22(7):1410–1413
10. See YC, Noor NM, Calvin Tan YM (2017) Investigation of indoor positioning system using visible light communication. In: *IEEE region 10 annual international conference proceedings/TENCON*, pp 186–189
11. Younis O, Al-Nuaimy W, Rowe F, Alomari MH (2019) A smart context-aware hazard attention system to help people with peripheral vision loss. *Sensors (Basel)* 19(7):1–22
12. Walimbe AA, Rao SS, Sureban AK, Shah MS (2017) Survey on obstacle detection and its notification through an android app for visually impaired people. *Asian J Convergence Technol* 3(30):2–4
13. Fujihara A, Yanagizawa T (2015) Proposing an extended iBeacon system for indoor route guidance. In: *Proceedings 2015 international conference on intelligent networking and collaborative systems, IEEE INCoS*, pp 31–37
14. Ghazal M, Ali S, Al Halabi M, Ali N, Al Khalil Y (2016) Smart mobile-based emergency management and notification system. In: *Proceedings 2016 4th international conference on future internet of things and cloud workshops, W-FiCloud*, pp 282–287

The Anxiety of Parent to the Prolong Computer Games Interaction: A Survey



N. S. Razak and M. Rahmah

Abstract A video or computer game has been attracting a large number of children and young people to play games repeatedly. It has become popular since this software can be installed on personal computers and mobile devices. Even though video or computer games are friendly, but the effect is adverse consequences, especially among children. This paper aims to study the consequences of video or computer games on eye health among children. This paper review the anxious of parent to their children eye health impacts due to the prolonged usage of computer games. There are 232 survey and the respondents among the parents. As a result, this study demonstrates that 76.7% of respondents among the parent show anxiety regarding eye health. These parents also concern about their children's eye health due to the prolong computer games interaction. This survey focused on the Human-Computer Interaction (HCI) field, where the anxiety of parents to prolong usage of computer games among the children. Based on the results, the framework of the game development, which includes eye health factors, will contribute in the future.

Keywords Eye health · Computer interaction (HCI) · Computer games · Anxious

1 Introduction

Computer games now become part of children's lives since the player able to enjoy the games. However, the nature of computer game interactions has not been openly discussing in any scientific literature. Today, the exposure of computer games interaction become part of entertainment and enjoyable among the children. Therefore, these exposure computer games have become part of the Human-Computer Interaction research field. Besides computer science study, Human-Computer Interaction

N. S. Razak (✉) · M. Rahmah
Faculty of Computing, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia
e-mail: saidatul.abrazak28@gmail.com

M. Rahmah
e-mail: drmahmah@ump.edu.my

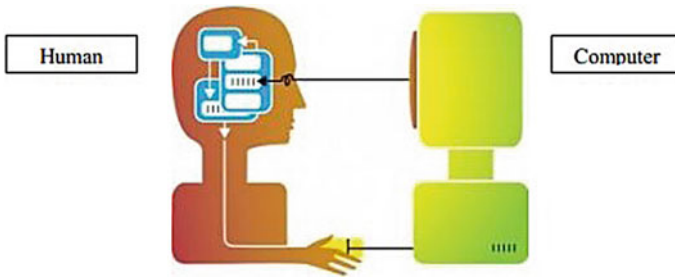


Fig. 1 Computer interaction

(HCI) also contribute to the sociology, psychology, and human factor engineering [1, 2]. Since HCI is a well-known research field in computer science since the past years, it is important in interaction study, designing, and implementing the interactive computer. For example, it able to assist our daily tasks after implementing an interactive computer system. This study focuses on eye health impact due to the prolong computer games usage. This study also focuses on anxiety among parents towards the children [3]. In consequence, the prolong usage of computer games lead to eye health impact [4]. Figure 1 shows the interaction between computer interaction and users.

Several studies have been researched the interaction and development of the child's brain. The research result demonstrates that most of the parent is concerned about their child's brain development. This is because of the prolonged usage of computer games among their children's [5]. The prolonged usage of computer games among children, especially of software applications such as games, books, social media, and history. For example, statistics show that 97% of children playing computer games spend more than one hour per day [6]. As a result, the longer spend time on computer games leads to mental fatigue. Therefore, children having a problem with executive control. This problem also called computer games transfer phenomena, which stimulate children's brains [7]. In addition to the problem, the prolong usage of computer games cause a visual effect. According to [8], most of the parents these days were concern about their children, especially their eyes' because of the prolonged usage of computer games. However, most children are more focuses on treating computer games as part of their life and utilizing the games as a medium for social communication [9, 10]. The emergence of massively multiplayer online games, encourage and stimulate children to play longer than usual. As a consequence, the children having eye problems such as redness, fatigue, and dry eye [5, 11]. Figure 2 presents the consequence of prolong usage of computer games.

Research shows that computer games used exclusively by children [12]. The increasingly widespread computer games have attracted children to the used computer during schools and at home daily. These children only used a computer for playing computer games [7]. They played computer games without knowing that it will affect their eye's health. Therefore, it has prompted the researcher to do the research o the implication of prolong usage of computer games to eye health. In

Fig. 2 Redness eye

particular, studied by Ophthalmologist show that prolongs usage of computer games affect eye health [13, 14]. Other researchers have highlighted the effect of eye health due to the poor lighting, external behavior of players, glare, and to much focus on computer games.

Therefore most of the parent's concern and feel worried about their children's eye become worst [2, 15]. Since eye health is most crucial nowadays, the games show that the increasing usage of computer games among children. Finally, this paper presents the trends by identifying substantial research efforts and emphasizing potential frameworks for game development in the future. To justify the warrant of this paper, we performed a questionnaire by distributing to 232 out of survey to parents to measure their anxiety toward their children when using prolong computer games.

This paper presents an assessment and structured analysis of a survey regarding on prolong use of computer games. This investigation is ordered as follows. Section 2 describes a literature review. Section 3 presents the methodology process, and Sect. 4 describes the results and discussion, followed by a conclusion with the significance of the result and future works.

2 Related Works

Due to the prolonged usage of computer games, a lot of children having a problem with eye visual [16]. According to the Ophthalmologist, play computer games in a long time duration could cause eye problems [17]. Table 1 lists the existing research works that have been published in the journal.

Table 1 illustrates the previous studies on the survey of anxiety among parents towards prolonging the use of computer games. Hence, it shows that the eyes might be affected by games used [5, 18]. The surveys are carried out among the parents and show that most of the parents anxious about their eye health [19]. However, as

Table 1 Analysis of research

Articles	Year	Explanation
Ghosh et al. [14]	2015	50% of children use computer games in prolong time could lead to eye fatigue
Blum and Livingstone [15]	2016	47% of children playing games for more than one hour and lead to dry eye
Randolph [5]	2017	28% of children use computer games in prolong time
Krafka et al. [18]	2016	50% similar survey on parents due to computer vision syndrome
Akinbinu and Mashalla [13]	2014	There is no survey that parents are a concern that hours spent on computer games
Rahmah and Siti Aishah [16]	2019	73.3% of preschoolers use computer games that affected emotion

mentioned earlier, there is less a survey that parents are aware of this issue. This study can be uses as a guideline to the parents in order to monitor their children.

3 Methodology

The study adopted a quantitative methodology for this research. The methodology is divided into three phases, with a detail description as follows.



3.1 Preliminary Study: Survey

The survey was carried out before the research. 233 samples of respondents were randomly collected and distributed to the parents. This study focuses on respondents in East Malaysia. These respondents are coming from the various professional fields since this study did not filter by field. To avoid any misleading or miscommunication, the respondent needs to fill up the consent form. Then they are required to answer 13 semi-structured questionnaires that were disseminated by Google form.

Table 2 Data analysis

Model		Sum of squares	df	Mean square	F	Sig.
1	Regression	299,135.916	1	299,135.916	823.056	0.000 ^a
	Residual	217,340.279	598	363.445		
	Total	516,476.195	599			

a represent Predictors: (Constant), Time. It is an independent variable

3.2 Data Analysis

After collecting all the respondents, this data will be analyzed to identify their pattern and visualize them. Data analysis is very important in research, especially for an experiment because of good data able to increase the result performance. The data collected were recorded in Microsoft Excel and then analyzed statistically using SPSS. In SPSS, we use an ANOVA statistic method to synthesis the data.

Table 2 shows the ANOVA statistics, which indicate that the regression equation of the survey. It demonstrates the value of sig. to be 0.0. This shows that the proposed model can be accepted, where it shows a significant and fit model. Hence, the new model will be developed for future work.

3.3 Experimental Results

The analysis procedure blends the research activities into a real trend of a research field as it encompasses a preliminary study. It is a process that analysis on the evaluation of the expansion of research [20]. To address this study, more respondents are supporting this research, which is 178 of respondents choose strongly agree due to the games are affected by eye health rather than 54 of respondents disagree that games are not effected on eye health. Furthermore, this survey was carried out among parents in order to investigate the anxiety of children by playing computer games in prolong time. Figure 3 illustrates the result of the survey.

4 Results and Discussion

Figure 4 illustrates the anxiety among parents related to the prolonged usage of computer games. It shows that 77% of parents strongly agree that playing computer games leads to visual problems for their children. This indicates that most of the parents acknowledge that their children having a problem with eye visuals after prolong usage of computer games. According to the previous Deputy Prime Minister of Education II, Datuk Chong Sin Woon, 6.3% of standard one experienced vision problems and suggest scheduled eye-checkup to monitor eye health. This is

Fig. 3 Experimental results

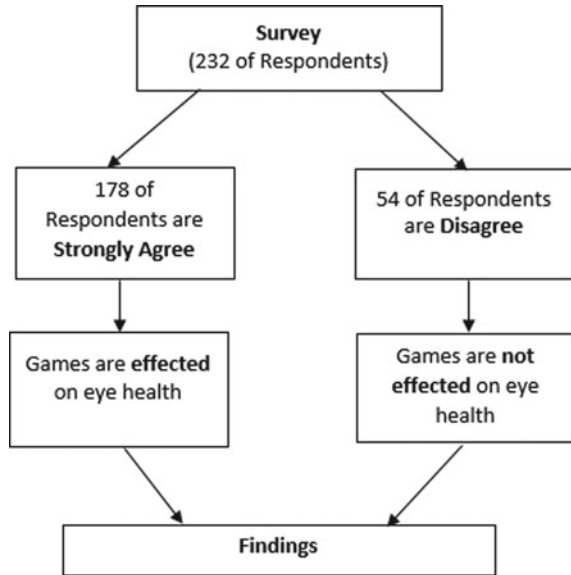
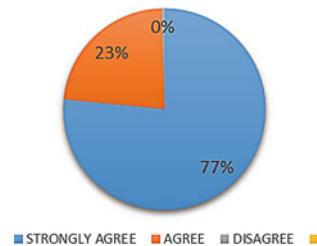


Fig. 4 Results



happening because most of the parent busy with their works and less interaction with their children [20]. Also, this problem happens because of the lack of awareness and knowledge about the effect of the computer when using too long [15]. As a conclusion, it proved that prolong usage of computer games lead to eye health problem among the children.

Figure 5 demonstrates that 76.7% of parents are very aware of these issues. These results can be a guideline for the parents in order to monitor their children in playing computer games. Besides, these cases also can increase eye blinking, reduce the attention that can cause eye problems [21]. Other than that, these problems may lead to emotional distress among children [12]. Rather than that, out 0.43% of respondents are strongly disagreed that computer games are not affected by eye health due to not frequently using computer games [12].

Based on the survey, Fig. 6 proved that playing computer games in prolong time could lead to an irritated eye or vision problem. Hence, this research will develop

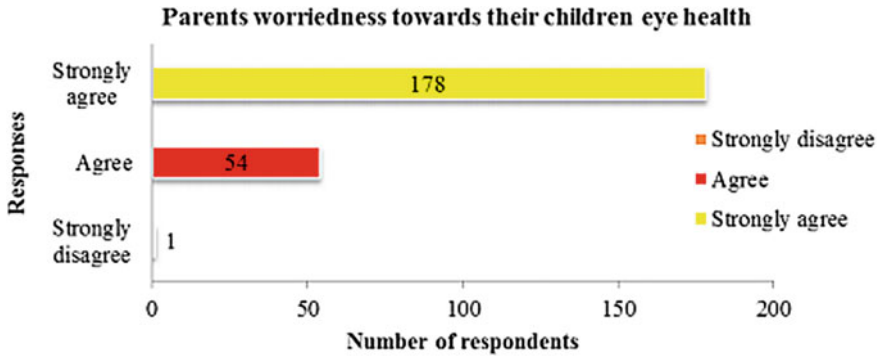


Fig. 5 Number of respondents

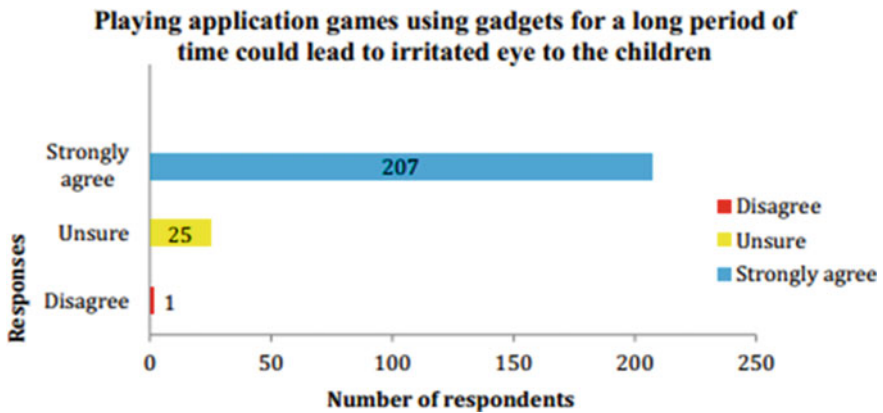


Fig. 6 Playing computer games in prolong time

a new model to protect the eyes from that problem. The result of the survey was supported for further research.

5 Conclusion

In conclusion, this study showed that the parent was anxious to prolong the usage of computer games since it affects their children’s eye health. Children who were playing games longer than one hour reported have an eye problem, such as redness, dryness, and fatigued eye. Also, this study presents children who spend more time using computers because unable to get more attention from their parents because of busywork. Therefore, this research is significantly introduced to parent’s awareness due to the effect of prolonged usage of computer games. Additionally, the great

potential is to encourage the children to play a computer game in limited time to prevent their eye health problems [18]. The new model of game development that can protect eyes will be derived for future research. This model able to help software developers to develop games with risk mitigation on prolonging computer usage to eye health.

Acknowledgements This research work is sponsored by FRGS/1/2018/ICT04/UMP/2/1 and RDU 190112 Grant. The authors would also like to thank the Faculty of Computing, University Malaysia Pahang and the respondent who involved in this research.

References

1. Carroll JM (2013) Human computer interaction—brief intro. *The encyclopedia of human-computer interaction*, 2nd edn., pp 1–69
2. Sturm C, Galal-edein GH (2015) Innovations and advances in computing, informatics, systems sciences, networking and engineering, vol 313
3. Wilkinson P (2016) Entertainment computing and serious games, vol 9970, pp 17–41
4. Rosenfield M (2016) Computer vision syndrome (a. k. a. digital eye). *Optom Pract* 17:1–10
5. Randolph SA (2017) Computer vision syndrome. *Workplace Health Saf* 65(7):328
6. Granic I, Lobel A, Engels RCME (2014) The benefits of playing video games. *Am Psychol* 69(1):66–78
7. De Gortari ABO (2016) What can game transfer phenomena tell us about the impact of highly immersive gaming technologies? In: *Proceedings 2015 international conference on interactive technologies and games, ITAG 2015*, pp 84–89
8. Allen-Scott LK, Hatfield JM, McIntyre L (2014) A scoping review of unintended harm associated with public health interventions: towards a typology and an understanding of underlying factors. *Int J Public Health* 59(1):3–14
9. Ibhari LFM, Ahmad Zaki NA, Mohamad Yatim MH (2015) Touch gesture interaction of preschool children towards games application using touch screen gadget. *Asia-Pacific J Inf Technol Multimedia* 4(1):47–58
10. Crnovrsanin T, Wang Y, Ma K (2014) Stimulating a blink: reduction of eye fatigue with visual stimulus. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp 2055–2064
11. Shapiro BMT (2016) Computers and your tired eyes. *Ophthalmology Practice*, p 48
12. González MT, Espada JP, Tejeiro R (2017) Problem video game playing is related to emotional distress in adolescents. *Adicciones* 29(3):180–186
13. Akinbinu TR, Mashalla YJ (2014) Medical practice and review impact of computer technology on health: computer vision syndrome (CVS). *Acad J* 5:20–30
14. Ghosh S, Nandy T, Manna N (2015) Advancements of medical electronics. In *Advancements of Medical Electronics Proceedings of the First International Conference (ICAME)*, vol 10, 2015, pp 193–206
15. Blum-Ross A, Livingstone S (2016) Families and screen time: current advice and emerging research. *Media Policy Brief* 17:52
16. Rahmah M, Siti Aishah Z (2019) Effectiveness of kinect-based application in gamification approach for preschooler: case study in Taska Permata Perpaduan, Kuantan, Pahang, Malaysia. In: *Proceedings of the 10th international conference on e-education, e-business, e-management and e-learning*, pp 124–129
17. Boyle EA et al (2016) An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Comput Educ* 94:178–192

18. Krafka K, Khosla A, Kellnhofer P, Kannan H (2016) Eye tracking for everyone. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2176–2184
19. Sigman A (2012) Time for a view on screen time. *Arch Dis Child* 97(11):935–942
20. Seaborn K, Fels DI (2015) Gamification in theory and action: a survey. In *International Journal of Human Computer Studies*, vol 74, pp 14–31, 2015
21. Mokhtar R, Sharif N, Zin NAM, Ihsan SN (2017) Assessing attention and meditation levels in learning process using brain computer interface. *Adv Sci Lett* 23(6):5569–5572

Design and Development of an Optical Sensor System to Measure Visibility of Air



S. Nurulain, R. M. Ramli, H. Manap, and Mohd Anwar Zawawi

Abstract Visibility is a measure of the distance at which an object or light can be clearly recognized. Visibility for a medium depends on the visible light transmission (VLT). This paper describes the development of VLT measurement system to measure the visibility of air. In recent years, there are many types of devices that can measure VLT. But the existing method and technique have their own disadvantages such as gigantic set up, costly, only can be used in small area and fail to function at night. Therefore, this research will focus on how to measure the visibility by using an open path optical method. The experimental setup for VLT measurement in this research consists of a laser pointer as a light source and spectrometer as a detector. The amount of light that passes through a few series of known-VLT thin films is measured and the percentage values of the measured thin film are set as a measurement base. This new proposed technique is believed to cater the problem mentioned above as the experimental setup is simple, small in size and can measure visibility in any light condition.

Keywords Visibility · Optical · Visible light transmission

1 Introduction

Visibility information is essential in dust and air quality monitoring which often require frequent and accurate real-time observations of visibility [1]. In the transmission process of light, an electromagnetic radiation gets impaired due to emission and absorption of different radiation wavelengths by the medium of propagation. Interaction of light packets i.e. photons with fundamental particles like electrons, atoms, ions and impurities such as fog, dust, smoke, maritime spindrift particles

S. Nurulain (✉) · R. M. Ramli · H. Manap · M. A. Zawawi
Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia
e-mail: ainnadhira93@gmail.com

H. Manap
e-mail: hadi@ump.edu.my

called aerosols etc. affects visibility. Visibility at certain medium depends on the visible light transmission (VLT). VLT is simply the measurable amount of solar visible light (daylight) that travels through a medium. A medium with a high VLT allows most of the daylight to pass through while a lower VLT restricts the majority of light from entering a medium. In a modern world, there are many devices that are used to measure VLT. Some of the devices like 3 M transmission meter is only suitable to be used to measure VLT for thin films [2]. There are also many visible light meters (VLM) distributed by a company called EDTM, Inc. However this transmission meter can measure the VLT of thin object such as glass. These devices are not suitable to be used to measure VLT of air in open space environment or at a wide area. The air VLT measurement system for a wide area and open space are already exist such as using a Laser Absorption Spectrometry (LAS) technique. However it has a gigantic set up and it is very costly [3].

There is also a device called Lux meter and it is used to measure intensity or brightness of an area [4]. However Lux meter can only be used in a small area or point measurement. It is not suitable to be used in an open wide area such to measure visibility at the airport. The other VLT measurement system such as a photography technique fails to function well at night and cannot perform real time measurement [2]. This is because of the low light condition and it faces a several errors especially in rainy, foggy and smoky conditions. Haze is one of the major factor that can interrupt the visibility. This haze phenomenon not only faced in Malaysia but also in entire world. People need to know how far they can see during haze. Therefore a new setup using a laser pointer as a light source to develop the air VLT measurement system is proposed. This new proposed technique is believed to cater the problem mentioned above as the experimental setup is simple and small in size. Besides, it can function well in a low light condition and can be used at an open wide environment. As a result, people can easily measure how far they can see.

2 Methodology

2.1 Sensor System Design

This open path optical method use laser pointer as a light source and spectrometer as a detector. At the beginning, a VLT of thin film will be measured and the percentage of this VLT thin film will be used as a measurement base in order to measure the visibility of air. Smoke will be use to increase the intensity of air. Increasing in smoke reading will be justify by using a commercialize sensor called Dustmate. A 100 cm chamber is used in order to test light source stability and to measure the VLT of the smoke. Both end part of the pipe are closed using an end caps as shown in Fig. 1. These end caps can be easily removed from the pipe. The inner side of the chamber is sprayed with black paint. This is to reduce the reflection of the light inside the chamber.

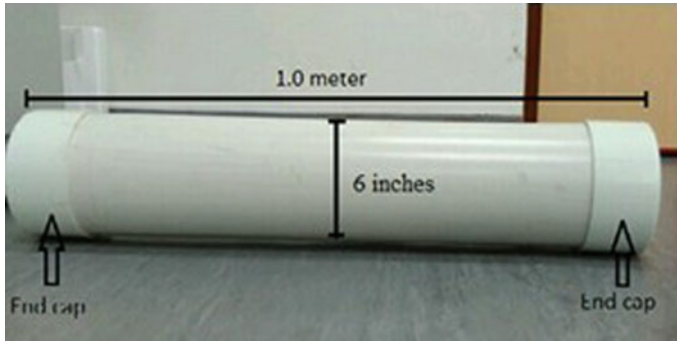


Fig. 1 Chamber

2.2 Laser Pointer

Visible light mostly comes from artificial sources such as fluorescent or tungsten devices, light-emitting diode, laser and many more [5]. This light source can be separated into two different types. The first type transmits a continuous wave as reported by Kudlinski et al. [6] and this wave requires an external modulator at its optical output. Second type of the light source transmits modulated light. According to Yuan et al. [7], a modulated type does not need any external modulator. LASER is stand for Light Amplification by Stimulated Emission of Radiation. Each laser emit light at different wavelength depends upon the material from which the laser crystal, diode, or gas composed and also the color emitted by the laser itself [5].

A battery operated laser pointer is used as a light source in this research. It can emit green color beam laser emission at range between 532–538 nm and can transmit up to 8000 m as claimed by manufacturer. Green color is chosen among all the color in visible light region because this green color is fifty times brighter than red laser. Besides, it is also have long-term reliability and stability due to the low power density and thermal effect [8]. A high and low temperature does not affect the performance of green laser itself. This laser pointer is made up from aerometal material with the dimension 150 mm × 28 mm.

2.3 Thin Film

A few series of known-VLT thin films are used in this research. At the initial stage, these thin films are measured by using a commercial transmission meter (3 M Transmission meter) in order to validate the percentage of VLT of each thin film. Then, the values of VLT thin film are measured using open path optical method in percentage value.

2.4 Detector (Spectrometer)

Spectrometer is a device that used to split light into a spectrum. It can analyse the detected substance or light in a short time and it is very easy to use. Besides, spectrometer also is used in astronomy to analyse the chemical composition of stars and planet. The disadvantage of this detector is its price and the repair cost is expensive relatively. Electronic components in this device may also generate noise that decreases the measurement accuracy and the sensitivity of device [9]. Spectrometer are use as a detector in this research.

2.5 Dustmate

A Dustmate is needed in order to validate the amount of smoke changes in the smoke chamber. DustMate from Turkey Instruments Ltd is chosen as a second measurement device in this research. Practically, the DustMate had been designed to detect sources of workplace airborne dust and fumes even at very low concentrations. It is a hand-held instrument with a very fast response where it can measure Total Suspended Particle (TSP), PM10, PM2.5 and PM1 simultaneously in real time.

3 Experimental Setup

Two sets of experimental setup was arranged in order to measure the VLT and the visibility of air by using an open-path optical method. First experimental setup consists of light source, glass plate holder, laminated thin film, spectrometer and PC installed with Spectrasuite software in order to measure VLT of thin film as shown in Fig. 2. Counts of intensity from each thin film are obtained and recorded in percentage. These data are used as a reference.

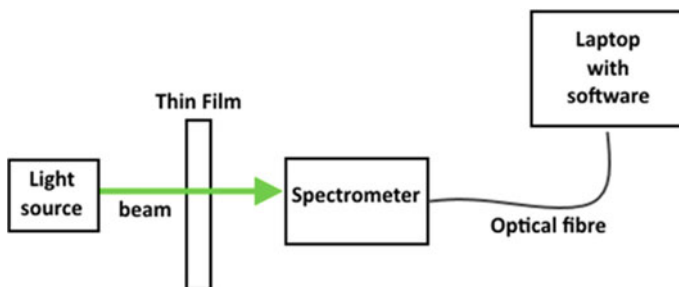


Fig. 2 Block diagram of experimental setup to measure VLT of thin film

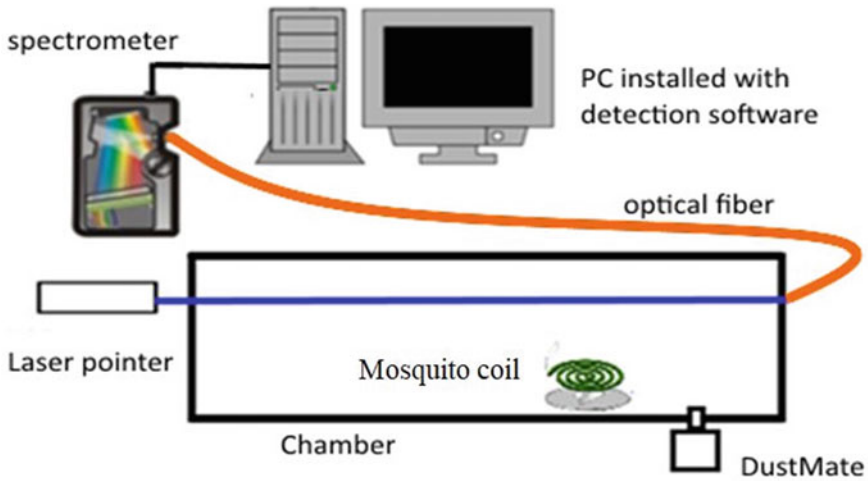


Fig. 3 Experimental setup to measure visibility of air consist smoke in chamber

Air visibility varies when a smoke is generated in a chamber using a mosquito coil. The absorption of light that passes through the air is measured and the intensity counts in the selected region was calculated and analyzed. The comparison between the references data from known-VLT thin film result and the measured data from air that consist of smoke was analyzed to measure the visibility. Dustmate is place inside the chamber to justify the increasing value of the smoke. Figure 3 shows the experimental setup to measure visibility of air consist smoke in chamber.

3.1 Beer Lambert Law

In this research study, we proposed an optical-path technique to measure the visibility of air. Visible light transmission (VLT) will affect the visibility. As the percentage of VLT increase, light intensity will also increase. The method used in this project is based on an open-path optical technique. The path length of the travelled light will influence the interaction between amount of the tested sample and the light. A principle of visibility measurement mostly depends on light propagation properties where the light is absorbed and scattered.

In optical method measurement, the Beer-Lambert law is used to relate the absorption of light to the properties of the material where light is travelling. The Transmittance, T of the travelling light can be calculated from the ratio of the transmitted Intensity, I and incident Intensity, I_0 as shown in Eq. 1. Absorbance, A as shown in Eq. 3.2 has no unit. Using Eq. 2, the absorption of light that passed through the air can be determined if the incident and the transmitted intensity are known. Light intensity that travelling through a medium will be exponentially reduced with respect

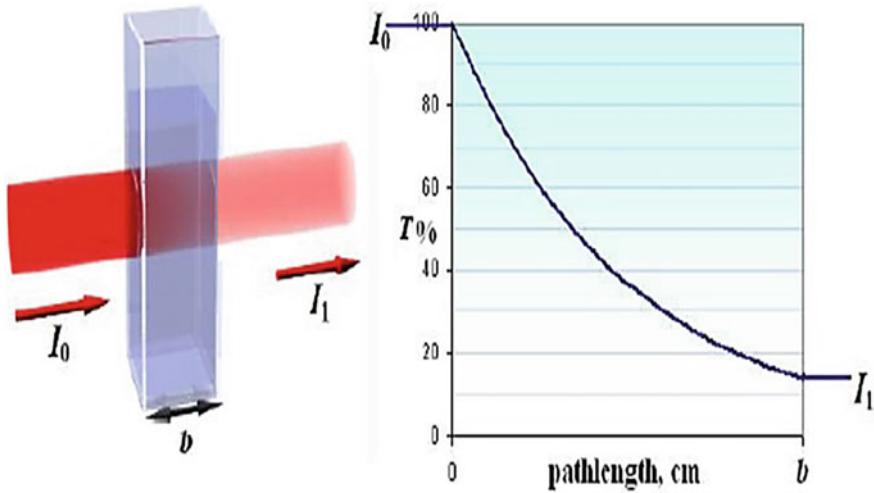


Fig. 4 Light absorption intensity over path-length

to the path-length, b as shown in Fig. 4 [10].

$$T = I/I_0 \quad (1)$$

$$A = -\ln(T) \quad (2)$$

4 Results and Discussion

VLT of thin-film is measured by using open path optical method. The wavelength is taken at 537.65 nm because the peak at that wavelength is much appropriate to be used in laboratory test. It is expected that the thin film with higher percentage of VLT will have a higher intensity of spectrometer reading. Based on Fig. 5, it is clearly shown that the higher the VLT percentage of thin film, the higher intensity of light detected by the spectrometer. The percentage of VLT-thin film labelled by manufacturer was 10, 35, 50 and 60%. But the result of thin film verification by the commercial transmission meter is are slightly lower than the percentage value labelled by the manufacturer. This is might due to the thin films are laminated on the glass plate that has different ability to absorb some of the light transmission. These measured values are acceptable as it is aligned with the transmission meter specification whereby it has $\pm 2\%$ accuracy. The percentage of VLT-thin film recorded by this transmission meter was 10%, 35%, 48% and 69% respectively.

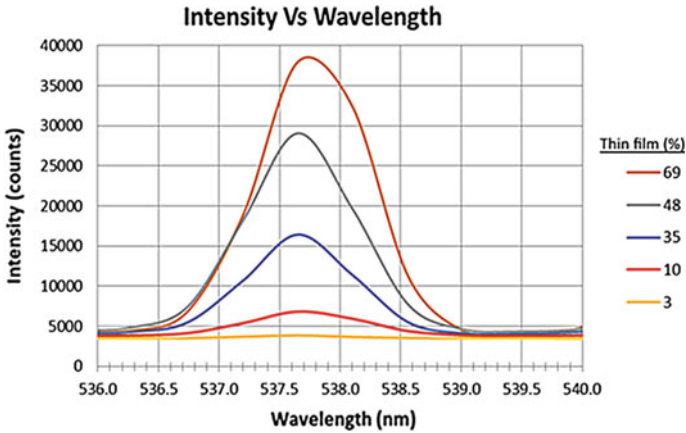


Fig. 5 Intensities of thin films at selected peak for VLT measurement

The development of this VLT measurement system is based on the data obtained in the first experiment. According to the Beer Lambert Law explained before, the transmittance, T of the travelling light can be calculated from the ratio of the transmitted, I and incident intensity, I_0 as shown in Eq. 1. To develop the VLT measurement system, a relationship between the transmittance, T and VLT percentage must be obtained. Therefore a graph of transmittance, T versus VLT percentage at selected wavelength (537.65 nm) is plotted using Microsoft Excel software as shown in Fig. 6.

In an open environment, smoke is a major factor that can affect visibility. Therefore, an experiment for VLT measurement due to present of smoke is carried out. The setup for this experiment is shown in Fig. 3. A mosquito coil is chosen because it produces smoke constantly over time when burning. It can be assumed that the smoke accumulates consistently in the chamber and the visibility will drop. The result of the VLT measurement using optical system is shown in Fig. 7. The result shows the VLT drop over time and this is aligned with the assumption that VLT will drop when more smoke accumulates in the chamber.

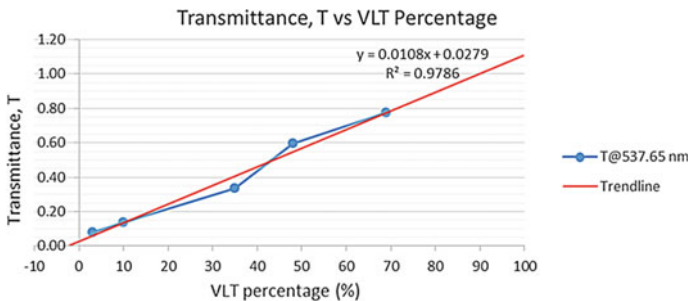


Fig. 6 Transmittance, T versus VLT percentage of thin films

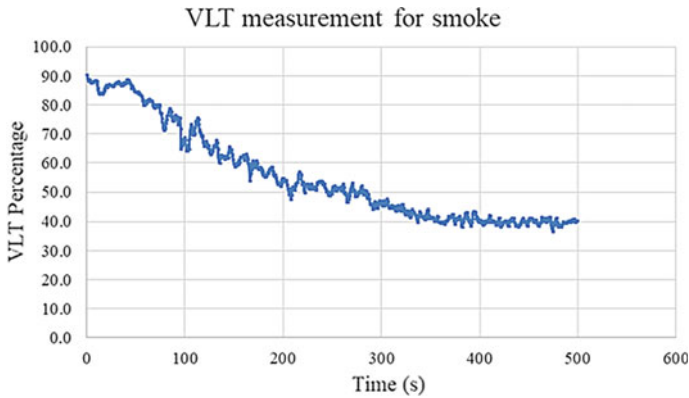


Fig. 7 VLT measurement for smoke at 537.65 nm

A commercial device called DustMate is used in this experiment to justify the VLT measurement. It is a device to measure the concentration of particle in air. The smoke particle concentration in the gas chamber is measured so that it can be relate with the opacity or VLT percentage. The DustMate sensor tip is placed in a chamber to measure the quantity of smoke. The reading of the DustMate is taken for every 20 s and the result is shown in Fig. 8.

As can be seen from Fig. 8, the amount of Total Suspended Particle in the chamber is increased over time. Thus it can be said that the smoke in the chamber is accumulated. Therefore the visibility in the chamber is reduced and the VLT percentage will drop as well. This is aligned with the result of VLT measurement done by optical sensor system as shown in Fig. 7.

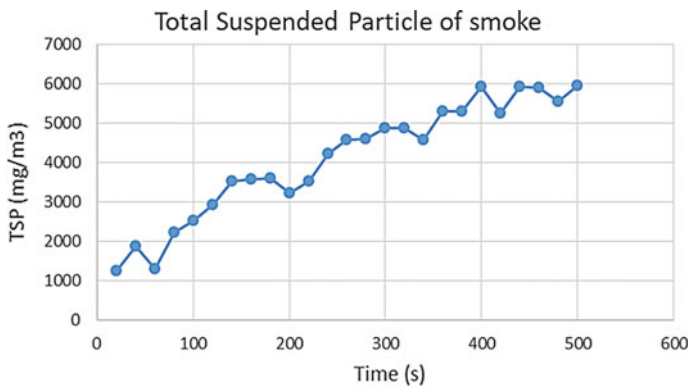


Fig. 8 Total suspended particle reading by DustMate

5 Conclusion

The open path optical VLT measurement system has shown its ability to measure VLT of air in a chamber that consist of smokes. Smoke is the major contributor to visibility measurement in an open area. It is found that the VLT measurement system is able to detect an accumulation of smoke in a chamber. However it can measure the VLT of air for a short duration only. This is due to the intensity of light source used in this experiment is not stable. Therefore a replacement of the light source could reduce the inaccuracy of this sensor reading. However the cost of replacement of a more stable light source should be compromised. For future work, many possible things can be carried out to achieve better performance and reduce the cost of the VLT measurement system. For instance, the spectrometer can be replaced by a visible range sensitive ceramic photodiodes.



Acknowledgements The author would like to thank the staff of Electrical Department, Faculty of Engineering Technology who contribute their time, energy and idea to complete this project. The author's sincere appreciation is also extended to Dr. Roshahliza and Dr. Hadi Manap for their help and guidance. Also the authors would like to acknowledge the support of the PNI, University of Malaysia, Pahang (UMP) in providing a research fund for this project.

References

1. O'L oingsigh T, McTainsh GH, Tapper NJ, Shinkfield P (2010) Lost in code: acritical analysis of using meteorological data for wind erosion monitoring. *Aeolian Res* 2:49–57
2. M Transmission Meter Alibaba.com. Retrieved from <https://www.alibaba.com/product-detail/3M-Transmission-Meter-Solar>
3. Mishra S, Ahuja B (2016) Optical parameters testing to redefine visibility for low cost transmission using channel modelling. *Optik* 127:11326–11335. <https://doi.org/10.1016/j.ijleo.2016.09.045>
4. Dias MV, Motamed A, Sergio Scarazzato P, Scartezzini J-L (2017) Toward proper evaluation of light dose in indoor office environment by frontal lux meter. *CISBAT 2017 International Conference—Future Buildings & Districts—Energy Efficiency from Nano to Urban Scale*
5. Kenneth RS Davidson MW (1998) Retrieved from <https://micro.magnet.fsu.edu/primer/lightandcolor/lightsourcesintro.html>
6. Kudlinski BB Mussot A (2010) White-light continuous-wave supercontinuum source. *Photonics society summer topical, meeting series. IEEE*. <https://doi.org/10.1109/phosst.2010.5553620>
7. Yuan ZL, Fröhlich B, Lucamarini M, Roberts GL, Dynes JF, Shields AJ (2016) Directly phase-modulated light source. *Phys Rev* 6:031044. <https://doi.org/10.1103/PhysRevX.6.031044>
8. Qi Y, Chu S, Bi Y, Zhou M, Yan W, Ying Zhang Y, Wang BY, Wang B, Tiejia W, Feng C, Liu Y, Fang T, Zheng G (2010) A compact continuous-wave green laser with line beam for laser projection. *Opt Lasers Eng* 48:737–739
9. Lee K (2017) Advantages and disadvantages of a UV-VIS spectrometer. Retrieved from <https://sciencing.com/advantages-disadvantages-uvvis-spectrometer-6466475.html>
10. Manap H (2011) An ultra violet optical fibre based sensor for ammonia detection in the agricultural sector. (PhD thesis) University of Limerick, Ireland

The Classification of Skateboarding Tricks by Means of Support Vector Machine: An Evaluation of Significant Time-Domain Features



Muhammad Amirul Abdullah , Muhammad Ar Rahim Ibrahim, Muhammad Nur Aiman Shapiee, Anwar P. P. Abdul Majeed , Mohd Azraai Mohd Razman, Rabi Muazu Musa, and Muhammad Aizzat Zakaria

Abstract This study aims to improve classification accuracy of different Support Vector Machine (SVM) models in classifying flat ground tricks namely Ollie, Kick-flip, Shove-it, Nollie and Frontside 180 through the identification of significant time-domain features. An amateur skateboarder (23 years of age ± 5.0 years' experience) executed five tricks for each type of trick repeatedly on a customized ORY skateboard (IMU sensor fused) on a cemented ground. From the IMU data a total of 36 features were extracted through statistical measures. The significant features were identified through two feature selection methods, namely Pearson and Chi-Squared. The variation of the SVM models (kernel-based) was evaluated both on all features and selected features in classifying the skateboarding tricks. It was shown from the study that all classifiers improved significantly in terms of training accuracy, prediction speed, training time and test accuracy. The Cubic-based SVM and Quadratic-based SVM demonstrated a 100% accuracy on both the test and train dataset, however, the Cubic-based SVM model provided the fastest training time and prediction speed between the two models. It could be concluded that the proposed method is able to improve the classification of the skateboarding tricks well.

Keywords Skateboarding · Machine learning · Classification · Feature selection

M. A. Abdullah · M. A. R. Ibrahim · M. N. A. Shapiee · A. P. P. Abdul Majeed (✉) ·

M. A. Mohd Razman · M. A. Zakaria

Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

e-mail: amajeed@ump.edu.my

R. M. Musa

Department of Credited Co-Curriculum, Centre for Fundamental and Continuing Education, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia

© Springer Nature Singapore Pte Ltd. 2020

M. A. Mohd Razman et al. (eds.), *Embracing Industry 4.0*,

Lecture Notes in Electrical Engineering 678,

https://doi.org/10.1007/978-981-15-6025-5_12

1 Introduction

In the 2020 Tokyo Summer Olympics, skateboarding will make its debut due to its popularity and the traction that this sport has made over the years. It is worth noting that in 2010, the net worth of the skateboarding industry is approximately \$USD 4.6 billion [1]. As the evaluation of the tricks executed are literally subjective, which is carried out by a panel of judges, that could be open for bias and imprecise assessment. Therefore, an innovative approach should be proposed in addressing the aforesaid issue and it is expected that the use of machine learning is able to provide the solution.

Machine learning has been employed for many type of sporting events [2–8] whilst in skateboarding, Groh et al. [9] in classified six different skateboarding tricks, i.e., OLLIE, NOLLIE, KICKFLIP, HEELFLIP, POPSHOVE-IT and 360-FLIP through the use of Naïve Bayes (NB), Partial Decision Tree (PART), Support Vector Machine (SVM) with a radial basis kernel and k -Nearest Neighbour (k NN). A total of 54 features were extracted combining both statistical time and frequency features including the x - y -correlation, the x - z -correlation and the y - z -correlation. The feature selection was carried out using Embedded Classification Software Toolbox (ECST) but, the number of features selected were undisclosed. The feature selection method used in the ECST was the best-first forward selection. The best overall accuracy was achieved for Naïve Bayes and SVM with a classification accuracy of 97.8%.

Groh et al. [10] further improvised the study by classifying 13 classes: the 11 trick classes, 1 class for bails and 1 rest class for all other detected events that did not contain a trick. Five (5) different classifiers, namely, Naïve Bayes (NB), Random Forest (RF), Linear Support Vector Machine (LSVM), Support Vector Machine with a radial-basis kernel (RB-SVM) and k -Nearest Neighbour (k NN) were evaluated. The best performing classifier reported for only the correctly performed tricks was the RB-SVM with a classification accuracy of 89.1%. Conversely, for the classification of all events, the Random Forest model demonstrated the best result with an accuracy of 79.8%.

Correa et al. [1] developed different Artificial Neural Networks (ANN) models by considering different axes as the input features in classifying five (5) skateboarding trick classes: NOLLIE, NSHOV, FLIP, SHOV and OLLIE. It was shown from the study that the ANN model that utilised statistical features from the Z -axis could yield a classification accuracy (CA) of 98.7%. Anlauff et al. [11] used Linear Discriminant Analysis (LDA) in classifying three (3) classes, i.e., the two (2) trick classes (OLLIE & OLLIE-180) and one (1) class for events that did not contain any trick. The 10-fold cross-validation technique was employed in the investigation. It was shown that the classifier could provide a 97% true positive classification for the OLLIE trick.

In a recent study, Abdullah et al. [12] investigated the efficacy of different machine learning models, viz. Support Vector Machine (SVM), k -Nearest Neighbour (k NN), Artificial Neural Network (ANN), Logistic Regression (LR), Random Forest (RF) and Naïve Bayes (NB) in classifying five (5) skateboarding trick classes, i.e.,

Table 1 Short summary of machine learning application on skateboarding

References	No. of target	No. of tricks	No. of features extracted	Features selection	No. of classifier	Best accuracy
[9]	6	6	54	Best-first forward, ECST	4	NB, 97.8%
[10]	13	11	–	–	5	RB-SVM, 89.1%
[1]	5	5	–	–	4	ANN Z, 98.7%
[11]	3	2	7	–	1	LDA, 97%
[12]	5	5	36	–	6	LR & NB, 95%

OLLIE, NOLLIE FRONTSIDE SHUVIT, FRONTSIDE 180, POP SHOVE-IT and KICKFLIP. A total of 36 statistical features were extracted from 6 input signals collected from the tri-axial accelerometer and gyroscope embedded in the data acquisition device. The features extraction and classification were performed by using an open source platform, Orange. The leave-one-out cross-validation technique was employed in the investigation. It was established from the study that both the LR and NB models yields the highest classification accuracy of 95% against other evaluated models (Table 1).

It could be seen from the limited literature available with regards to the employment of machine learning in classifying skateboarding tricks demonstrated commendable classification accuracy [1, 9–12]. However, it is worth noting that the investigation with regards to feature selection is rather limited [9]. Therefore, this paper aims at evaluating the significance of feature selection restricted to time-domain towards the classification accuracy of different variation of SVM models.

2 Methodology

2.1 Data Collection

The acquisition of the data from the skateboarding tricks was attained via an instrumented inertial measurement unit (IMU) device developed as shown in Fig. 1. The device is attached to the bottom front of the skateboard specifically fix behind the front truck as depicted in Fig. 2. To ensure the balance of the board after adding up the device, an equal riser pad is fixed at bottom rear of the board. Both device and riser pad are fixed with hex screw bolts and nylon lock nuts (nyloc) to ensure its stability. In addition, the device is secured with a 3D-printed casing using acrylonitrile butadiene styrene (ABS) material on Zortax M200 3D printer. The ABS material

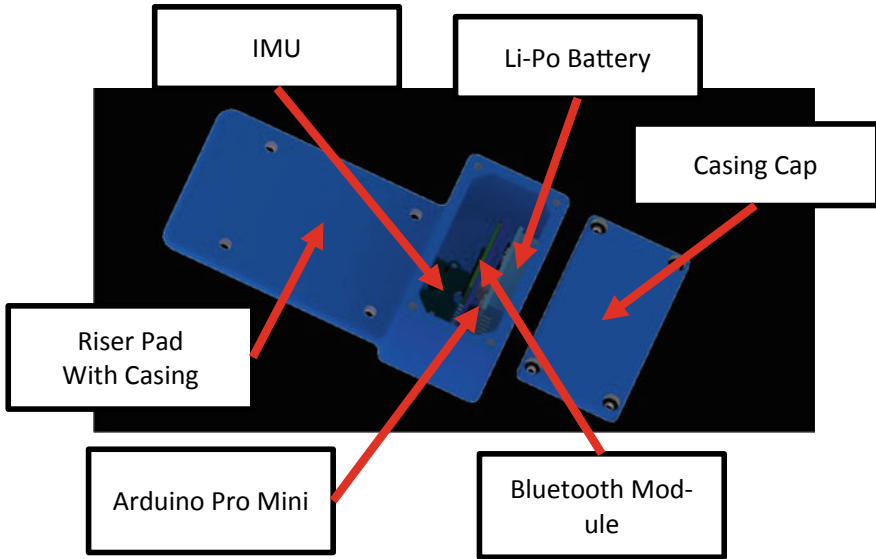


Fig. 1 The 3D model of instrumented device on CATIA



Fig. 2 Placement of the instrumented device on the skateboard

is chosen due to its desirable mechanical properties, primarily high impact strength and good absorbing as the device will be prone to impacts and shocks from the tricks. On the inside, the device consists of a MPU6050 as the IMU for raw data sensor, a Bluetooth 2.0 module (HC-05) for communication, a microcontroller (Arduino Pro Mini) for central processing unit, and a 3.7 V Lithium Polymer rechargeable battery for power supply. An amateur skateboarder (23 years of age ± 5.0 years' experience) executed five tricks for each type of trick repeatedly (Table 2).

Table 2 Description of the observed activities/tricks

Trick name	Rotation (angle and axis)	No. of samples
Ollie (O)	Board incline about the x -axis ($\approx 45^\circ +y$)	5
Nollie FS Shuvit (NFS)	Board incline about the z -axis ($\approx 180^\circ -z$)	5
Frontside 180 (FS180)	Clockwise about z -axis ($180^\circ -z$)	5
Pop Shove-it (PS)	Clockwise about z -axis ($180^\circ +z$)	5
Kickflip (KF)	Clockwise about x -axis ($360^\circ -x$)	5

2.2 Feature Selection Method

A total of 25 segmented raw signals data for five (5) different skateboarding tricks collected was used for this study. The segmented data were equally distributed for all trick which is 5 segmented raw signals per trick. The statistical features were extracted from the IMU devices, namely mean, skewness (sk), kurtosis, peak to peak, root mean square (rms) as well as standard deviation (std) for all the readings (all six degrees of freedom) which are equal to a total number of 36 features. In this study, two feature selection methods are employed, i.e., Pearson-correlation technique and chi-squared. Feature selection is non-trivial in identifying the features that significantly contributes towards the classification efficacy of the developed model, as the inclusion of irrelevant features (noises) could decrease the accuracy of the models. The features were evaluated via Spyder 3 software package in the Python 3 environment.

2.3 Classification

Different variation of SVM models are evaluated in the present investigation. The variation is based on default kernels available on the MATLAB 2016b Classification Learner toolbox, i.e., Linear, Quadratic, Cubic, and Medium-Gaussian, respectively. The five-fold cross-validation technique was employed in the study in order to mitigate the effect of overfitting on the trained models. The data was split to an 80:20 ratio for train and test, respectively.

3 Results and Discussion

By considering all features, it could be observed from Table 3, that Linear, Quadratic and Cubic SVM model yield a similar training and test accuracy of 85 and 80% with different in prediction speed and training time. Nonetheless, it is evident that the Quadratic SVM model could provide the fastest prediction speed against the other models evaluated, particularly with the best models mentioned. Through the feature selection evaluation via both Pearson and Chi-square method, the following features are identified as significant, i.e., $stdgZ$, $skgZ$, $rmsgZ$, $rmsgX$, $rmsaX$, $meangZ$, $meangX$ and $meanaX$, where a, g, X, and Z, corresponds to accelerometer readings (m/s^2), gyro readings ($^\circ/s$), X-axis, and Z-axis, respectively.

It could be seen from Table 4, that all the SVM classifier evaluated have improved in all performance indicators. It is apparent that the Cubic, Quadratic and Medium Gaussian based SVM models resulted in the best training and test accuracy of 100%. Nevertheless, it could be said that the Cubic-SVM model is the best, as it provided the fastest prediction speed as shown in Fig. 3. Moreover, from Fig. 4, it could be seen that the Linear, Quadratic and Cubic models exhibit an overfitting behavior before the considering the selected features, but upon the selection of the features, such behaviour no longer transpires, suggesting that the procedure of selecting significant features are non-trivial.

Table 3 Training and test results on SVM model on all 36 features

Performance indicators	Linear-SVM	Quadratic-SVM	Cubic-SVM	Medium gaussian-SVM
Training accuracy (%)	85	85	85	80
Prediction speed (obs/s)	~140	~130	~140	~180
Training time (s)	10.737	9.8779	9.747	11.541
Test accuracy (%)	80	80	80	80

Table 4 Training and test results on SVM models on 6 selected features

Performance indicators	Linear-SVM	Quadratic-SVM	Cubic-SVM	Medium gaussian-SVM
Training accuracy (%)	95	100	100	100
Prediction speed (obs/s)	~180	~170	~170	~310
Training time (s)	8.8811	8.6213	8.3247	11.482
Test accuracy (%)	100	100	100	100

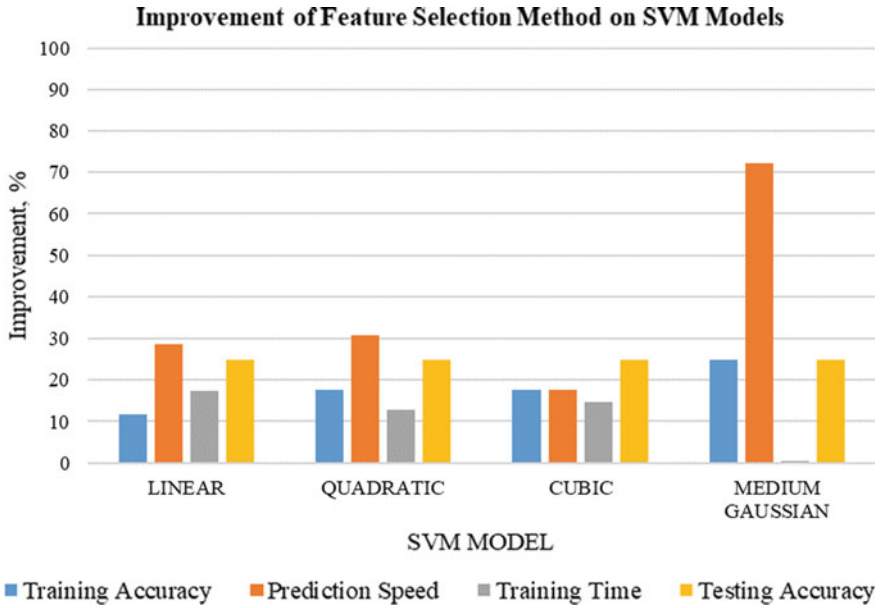


Fig. 3 Comparison of training and test accuracy before and after feature selection on SVM models

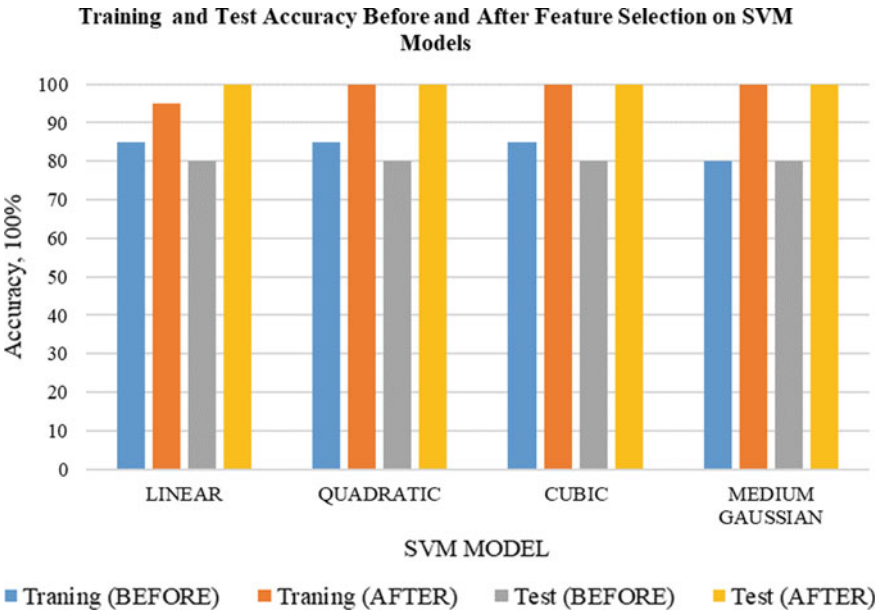


Fig. 4 Improvement of feature selection method on SMV models

4 Conclusion

It was shown in this preliminary investigation that that the features selection significantly improved the training accuracy, prediction speed, training time and test accuracy of the classification models evaluated. It was shown that the best classifier is the Cubic based SVM model as it has the fastest training time. This study suggest that the selection of the features is non-trivial in yielding a better performance of the classifiers evaluated. Future study shall evaluated other feature selection methods as well as its effect towards other classifiers that has yet been investigated in the present study.

Acknowledgements The authors would like to acknowledge the Ministry of Education, Malaysia and Universiti Malaysia Pahang for supporting and funding this research via FRGS/1/2019/TK03/UMP/02/6 (RDU1901115).

References

1. Corrêa NK, de Lima JCM, Russomano T, dos Santos MA (2017) Development of a skateboarding trick classifier using accelerometry and machine learning. *Res Biomed Eng* 33:362–369
2. Anand A, Sharma M, Srivastava R, Kaligounder L, Prakash D (2017) Wearable motion sensor based analysis of swing sports. In: 2017 16th IEEE international conference on machine learning and applications (ICMLA). IEEE (2017) pp 261–267
3. Büthe L, Blanke U, Capkevics H, Tröster G (2016) A wearable sensing system for timing analysis in tennis. In: 2016 IEEE 13th international conference on wearable and implantable body sensor networks (BSN). IEEE, pp 43–48
4. Connaghan D, Kelly P, O'Connor NE, Gaffney M, Walsh M, O'Mathuna C (2011) Multi-sensor classification of tennis strokes. In: SENSORS, 2011 IEEE. IEEE, pp 1437–1440
5. Adelsberger R, Tröster G (2013) Experts lift differently: classification of weight-lifting athletes. In: 2013 IEEE international conference on body sensor networks. IEEE, pp 1–6
6. Groh BH, Fleckenstein M, Eskofier BM (2016) Wearable trick classification in freestyle snowboarding. In: BSN 2016—13th annual body sensor networks conference pp 89–93
7. Brock H, Ohgi Y (2017) Assessing motion style errors in Ski jumping using inertial sensor devices. *IEEE Sens J* 17:3794–3804
8. Brock H, Ohgi Y, Lee J (2017) Learning to judge like a human: convolutional networks for classification of ski jumping errors. *Proceeding—International symposium on wearable computing ISWC. Part F1305*, 106–113 (2017)
9. Groh BH, Kautz T, Schuldhuis D (2015) IMU-based trick classification in skateboarding, KDD work. *Large-Scale Sport, Anal*
10. Groh BH, Fleckenstein M, Kautz T, Eskofier BM (2017) Classification and visualization of skateboard tricks using wearable sensors. *Pervasive Mob Comput* 40:42–55
11. Anlauff J, Weitnauer E, Lehnhardt A, Schirmer S, Zehe S, Tonekaboni K (2010) A method for outdoor skateboarding video games. *ACM International Conference Proceeding Series*, pp 40–43
12. Abdullah MA, Ibrahim MAR, Shapiee MNAB, Mohd Razman MA, Musa RM, Abdul Majeed APP (2020) The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning. *Lect Notes Mech Eng* 67–74

Fuzzy Coordination System for Traffic Light



Tang Ren Shyang, Adrian Loh Sheng Pang, Nureize Arbaiy,
Mohammad Haris Haikal Othman, Mohd Zaki Mohd Salikon,
and Pei Chun Lin

Abstract The traffic light is a signaling device that helps to manage the traffic. The current system uses rigid value to control the traffic timing. However, due to its unresponsiveness to the current traffic condition, the traditional traffic light system sometimes could cause a traffic jam. This is due to improper setting of timing which didn't consider the current situation i.e. heavy traffic during peak hour. Therefore, the use of human reasoning associated with the existing traffic control system is very important as it could react and change according to real-time condition. Fuzzy logic is able to do that as it is applying human logic and good at handling fuzziness in the world and produce the optimum output for the condition. Hence, this study identifies the traffic light timing setting problem and develops an intelligent traffic light system as a solution based on fuzzy logic. This web-based system is beneficial as a controller and monitoring for improving the traditional system. It is able to make changes to the green light duration according to the traffic condition at the current time of day e.g. peak, normal and off-peak as well as the density (congestion) of the road.

Keywords Traffic light system · Fuzzy logic · Fuzzy system · Coordination

1 Introduction

Traffic light plays a very important role in controlling the traffic flows and giving turn to each lane of vehicles to proceed [1]. The conventional traffic light control system is operating based on pre-set time for the lights to turn on. It is known as fixed-time traffic light control system [2]. Fixed time traffic system will be turned green after a given period of time, usually around 30 s. Fixed-time control system mostly

T. R. Shyang · A. L. S. Pang · N. Arbaiy (✉) · M. H. Haikal Othman · M. Z. Mohd Salikon
Faculty Computer Science and Information Technology, University Tun Hussein
Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
e-mail: nureize@uthm.edu.my

P. C. Lin

Department of Information Engineering and Computer Science, Feng Chia University,
No. 100, Wenhwa Rd, Taichung, Taiwan

integrated with timing plans which the traffic light is set according to the specific time of day [3]. For example, during morning and evening peak hours. Historical record or the traffic pattern have to be known in order to implement timely schedule. This type of control system is relatively stable, and one cycle length is optimal especially during two to three peak hours period. However, it is not so optimal when it comes to other periods. In fact, the traffic flow is continuous fluctuating throughout the year [4]. It makes the system of real-time traffic control less robust and scalable for today's traffic conditions that vary from time to time. As a result, traffic lights could worsen the traffic condition due to its unresponsiveness and causing the car drivers wait even longer for their turn to proceed. The negative impact of using Fixed-time system was also being proved in [5]. It is difficult to improve the performance of the traffic signal control system effectively and efficiently using conventional and traditional methods. The logic of intelligence is the need to control the red and green light periods and the decisions in traffic light settings due to fuzziness, and time fluctuations in the system [6]. The need of implementing fuzzy logic in a traffic light system to enable the traffic lights react accordingly and varyingly depend on the traffic condition is highly needed when the traffic condition is bad.

Therefore, this study proposes a smart lighting control system to efficiently and effectively manage traffic lighting periods using a fuzzy logic approach. Fuzzy logic is optimized to determine changes in green light duration that consider the congestion and road time. Using this system, the green light period can be adjusted in response to road conditions and time types such as peak, normal or off peak. Database systems are provided to store variables, rules, parameter values and log information. This is important to store that information, as it can be modified as needed. Controlling systems are developed using web-based technology to support system mobility and eligibility. In the future, it can also be further enhanced to integrate with IoT and other trusted devices.

The rest of the paper is organized as follows: Part II discusses related work. Part III presents the methodology and subsequent implementation and results in Part IV. Part V is the conclusion for the study.

2 Literature Review

Fuzzy systems are very useful systems in real life applications, especially machine control. Fuzzy systems are very flexible and can assist in handling inaccurate and incomplete data. In traffic systems, especially in big cities surrounded by an increasing number of vehicles, adaptive handling methods are needed. Monitoring and controlling traffic in the city is an important task because of its ability to control roads that also affect the quality of life. Currently, traffic signal controllers use real-time data and advanced algorithms [7]. This algorithm uses simple mathematical rules that are suitable for lower intersection load purposes. However, these methods cannot address complex road networks and increase the number of vehicles [8]. The

current algorithm needs to be able to cope with continuous and varied changes in traffic situations.

Fuzzy logic enables the implementation of rules that are very similar to the human thought process. It can make decisions even when information is incomplete. Previous work using fuzzy logic included congestion (density) and the number of inlet vehicles as fuzzy inputs [9]. It took several vehicles to wait and the number of vehicles passing between the two intersecting traffic due to fuzzy inputs shown in [10]. This method needs to consider the intersection between all roads involved in traffic. Another advanced method introduced in [11] is controlling based on receiving and analyzing traffic patterns and responding in real time. This method is a bit more efficient and advanced, but it is too complicated and expensive to implement. Giovanni Pau et al. [7] use a fuzzy-based approach to deal with the dynamic management of traffic lights in pedestrian crossing and provide some solution of possibility to change the phases of the traffic light taking into account the time of the day and the number of pedestrians about to cross the road. Patil et al. [8] use fuzzy to change the timing of the green signal with respect to the intensity of traffic and feed this to an Arduino microcontroller based fuzzy inference system which decides the duration of the green signal. This study uses the densities mentioned in [9] and time travel as a fuzzy input in implementing fuzzy light control systems.

According to [8], image processing along with fuzzy logic is found to be more efficient although it takes time to process. Hassan et al. [12] applied a type-2 fuzzy for city traffic networks where their proposed method could increase road network capacity in a timely manner and reduce congestion. Moreover, the method in [12] can optimize the queue length reduction and reduce the average wait time of cars in a queue. The techniques proposed by [13] have accelerated the development of emergency vehicles while avoiding the creation of congestion around their route by using Traffic Management System (TMS) tailored to fuzzy logic combinations. Cheng [14] used a neuro-fuzzy group junction control algorithm to subdivide vehicles in the same lane into small groups and to schedule vehicle groups via wireless communications instead of signal lights. Scheduling such vehicles can reduce wait times and increase justice, especially when traffic volume in different lanes is unbalanced. Some studies such as Suhail et al. [15] use hybrid algorithms using hybrid algorithms that combine Fuzzy Logic Controller (FLC) and Genetic Algorithm (GA) and its applications to signal light systems. Such techniques used by [15] can improve the performance of traffic light guards, reducing traffic congestion and timeout [16]. The next section describes the analysis and design of the proposed system.

3 System Design

The implementation of fuzzy coordination system for traffic light involves the development of fuzzy knowledge base that performs the action of inferencing on the inputs received. The knowledge base requires knowledge from experts of trusted sources

that are reliable or else it will affect the accuracy of decision of a fuzzy system. The following part of this section explains the system analysis and design.

3.1 *Input and Output Requirement*

Every fuzzy system must have input to be processed to produce the output. This fuzzy system requires two inputs which are density and time of day.

For the first input `density`, the system will require users to input the number of inflow vehicles, number of outgoing vehicles, and the distance between the two sensors. The formula for density [9] is calculated as in Eq. (1). For the input `time of day`, the system requires users to input any time in a day.

$$\frac{v_i - v_o}{d} \quad (1)$$

where v_i is the number of inflow vehicle, v_o is the number of outflow vehicles, and d is the distance between the two sensors.

The output of this system after the inferencing process will be a change of green light period. The fuzzy system will determine the duration of the green light based on the conclusions received.

3.2 *Fuzzy System Design*

The fuzzy system consists of 3 main components namely fuzzification, inferencing, and defuzzification. Fuzzification is the process of converting crude values into fuzzy values. Inferencing is a process of mapping formulation that produces outputs based on fuzzy logic implemented on inputs. Defuzzification is the process of converting blurred values into complex values.

Fuzzification

In this fuzzy system, three variables are used, namely `density`, `time_of_day`, and `green_light_duration` (GLC). Each fuzzy variable has been assigned the following fuzzy values:

- `Input: Density {sparse, average, dense}`.
- `Input: Time_of_day {off_peak, normal, peak}`.
- `Output: Green_light_changes {decrease, remain, increase}`.

The universe of for density are from 0 to 0.2 [5]. Meanwhile the universe discourse for `time_of_day` are from 00:00 to 23:59. Lastly the `green_light_changes` changes crisp values are from -10 to 10.

The universe of discourse for density ranges from 0 to 0.2 based on [5]. Meanwhile, the discourse of the universe for the time of day is from 00:00 to 23:59. The green_light_changes are from -10 to 10.

Membership Function Graph

Membership function graphs are produced based on the input and output variables using membership function formula. The membership function graph for density uses the linguistic values namely sparse, average, and dense. The membership function graph for time_of_day displays the off peak, normal and peak condition. Value from 00:30 to 05:30 is identified to be labelled as off peak in its fuzzy set. While the values from 06:00 to 07:00, 09:00 to 12:00, 14:00 to 16:30, 19:30 to 00:00, is labelled as normal. If the values are from 07:30 to 08:30, 12:30 to 13:00, 17:00 to 19:00, the fuzzy set is labelled as peak.

While the membership function graph for fuzzy variable green_light_changes is determined to have decrease, remain and increase condition. The decrease condition is given value by -10. If the value is 0, the fuzzy set is labelled as remain. If the value is 10, then the fuzzy set is labelled as increase.

Such fuzzy membership function graph is drawn in the developed system. The graph will react to the input value and the output.

Fuzzy inferencing

Inferencing is a process of resulting an output based on fuzzy logic formulating the mapping on the input. One of the main components in the process fuzzy inferencing is the knowledge base. Knowledge base is the combination of knowledge and a collection of rules. In this fuzzy system, Fuzzy Associative Matrix (FAM) is used for the knowledge base. Table 1 shows the FAM for the traffic fuzzy controller system. It shows the output of linguistic variable green light duration changes after combination of input linguistic variable density and time of day.

The part of rule based are extracted from the FAM and is presented with IF/THEN statement such as follows:

1. IF density is sparse and time of day is off peak, then green_light_changes remains.
2. IF density is sparse and time of day is normal, then green_light_changes decreases.
3. IF density is sparse and time of day is peak, then green_light_changes decreases.

Table 1 Fuzzy associative matrix

Density	Time		
	Off peak	Normal	Peak
Sparse	Remain	Decrease	Decrease
Average	Increase	Remain	Decrease
Dense	Increase	Increase	Remain

Fig. 1 Density table in database

ID	sparse	average	dense	default_status	session_id
28	0.05	0.1	0.2	1	1

- IF density is average and time of day is off peak, then green_light_changes increases.

The rule from FAM contains two input linguistic variable and the degree of membership of the two variables will undergo and operator by finding the minimum degree of membership or the intersection of two fuzzy sets for inferencing purpose. The formula of intersection written in Eq. (2), such that A and B are the fuzzy sets, u is an element in the universe, u :

$$\mu A \cap B(u) = \min\{\mu A(u), \mu B(u)\} \tag{2}$$

Defuzzification

Defuzzification is the process of converting fuzzy values found from inferencing phase into crisp value. In this system, the defuzzification is achieved by using the method Center of Area (COA). It is used to find the center of area under the curve of membership function. The formula of COA is as written in Eq. (3), where z^* is the fuzzified output of the system, μ_i is a membership function and x is the output variable.

$$z^* = \frac{\int \mu_i(x) \cdot x dx}{\int \mu_i(x) dx} \tag{3}$$

Database

Database is used to store the information used of the system. It stores the linguistic variables and linguistic values. The system database has tables which are `daily_session`, `density`, `green_light_changes` (`glc`), and `time_of_day`. Table `daily_session` stores the id whenever users enter a new range for each linguistic variable, table `density` stores the value for fuzzy value `sparse`, `average` and `dense`. Table `glc` stores the value for fuzzy value `decrease`, `remain` and `increase` while table `time_of_day` stores the value for fuzzy value `off peak`, `normal` and `peak`. Figure 1 shows the example of one of the tables which is table `density`.

4 Implementation

This section discusses program code writing to implement system and test results. The system was developed using HTML for mark-up design, JavaScript for injection functions, Hypertext Pre-processor (PHP) for database connection and MySQL as a database.

Fig. 2 Membership function for fuzzy variable density

```
if(x<=0.05){
    degreeDensity.degreeSparse = 1;
    degreeDensity.degreeAverage = 0;
    degreeDensity.degreeDense = 0
}
else if(0.05 <= x && x<=0.1)
{
    degreeDensity.degreeSparse = (0.1-x)/(0.1-0.05);
    degreeDensity.degreeAverage = (x-0.05)/(0.1-0.05);
    degreeDensity.degreeDense = 0;
}
else if(x==0.1)
{
```

4.1 Fuzzification

The value of degree of membership of each fuzzy value for x can be found using if else statement. Figure 2 shows the coding of finding value of degree of membership for density fuzzy value.

4.2 Fuzzy Inferencing

This process takes input from the output of fuzzy sets from the fuzzification process. Equation (2) is used to find the minimum value between the two membership levels of each fuzzy input and all the minimum values for each rule combination are then used to construct the FAM table. Figure 3 shows the encoding to obtain the minimum value and returns a three-dimensional 3×3 dimension known as the FAM table.

Fig. 3 Fuzzy inferencing and output FAM in array

```
function calcFAM(x,y)
{
    var keysDen = Object.keys(x);
    var keysTime = Object.keys(y);
    var i,j,min;
    for(i=0;i<3;i++)
    {
        for(j=0;j<3;j++)
        {
            min = Math.min(x[keysDen[i]],y[keysTime[j]]);
            FAM[i][j].value = min;
        }
    }
    return FAM;
}
```

4.3 Defuzzification

During this process, each rule in FAM is performed using its own function for each value in the green light duration changes. The de-fuzzification implementation is shown in Fig. 4. Each defuzzification function as shown in Fig. 5 is converted to crisp values as a return. Then in the final step, an optimum change of the green light in the crude value is obtained by computing Eq. (3).

```
function findCrispValueForDecreaseLabel(degree, GLC){
  // console.Log("crisp DEC: "+degree+"*("+GLC.remain+"-"+GLC.decrease+")"+GLC.remain);
  return degree*(GLC.decrease - GLC.remain) + GLC.remain;
}
function findCrispValueForRemainLabel(GLC){
  // console.Log("crisp REM: 0");
  return GLC.remain;
}
function findCrispValueForIncreaseLabel(degree, GLC){
  // console.Log("crisp INC: "+degree+"*("+GLC.increase+"-"+GLC.remain+")"+GLC.remain);
  return degree*(GLC.increase - GLC.remain) + GLC.remain;
}
```

Fig. 4 Crisp value transformation

```
function defuzzify(FAM, GLC){
  floatGLC = {};
  floatGLC.increase = parseFloat(GLC.increase);
  floatGLC.remain = parseFloat(GLC.remain);
  floatGLC.decrease = parseFloat(GLC.decrease);
  var sum = 0.0;
  var degreeSum = 0.0;
  for (i=0;i<FAM.length;i++){
    row = FAM[i];
    for (j=0;j<FAM[i].length;j++){
      col = row[j];
      if (col.ruleName == "Decrease"){
        // console.Log("sum bef: "+ sum);
        sum += col.value * findCrispValueForDecreaseLabel(col.value, floatGLC);
        // console.Log("sum after: "+ sum);
        // console.Log("this: "+col.value);
        // console.Log("col decrease: "+col.value+" * "+ findCrispValueForDecreaseLabel(col.value, GLC));
      }else if (col.ruleName == "Remain"){
        sum += col.value * findCrispValueForRemainLabel(GLC);
        // console.Log("col remain: "+col.value+" * "+ findCrispValueForRemainLabel());
      }else if (col.ruleName == "Increase"){
        sum += col.value * findCrispValueForIncreaseLabel(col.value, floatGLC);
        // console.Log("col increase: "+col.value+" * "+ findCrispValueForIncreaseLabel(col.value, GLC));
      }
      degreeSum += col.value;
      // console.Log("col sum: "+sum);
    }
    // console.Log("row sum: "+sum);
  }
  return (Math.round((sum/degreeSum) * 100)/100).toFixed(2);
}
```

Fig. 5 Defuzzification implementation

5 Result and Discussion

There are two main modules in the system that are configuration modules and simulation modules. Configuration modules are built to allow administrators to change firing values for each linguistic variable such as *Density*, *Daylight Time*, and *green_light_changes* or *GLC*. Figure 6 shows the design of the configuration module interface. System administrators can adjust values for fuzzy variables or use standards. Traffic control simulations can be performed in the simulation module.

To start simulation, user could enter number of in flow, number of outgoing vehicles, distance between *sensor1* and *sensor2* (e.g. default is set 100 m) and also the time of day in 24 hour format. Simulation is carried out based on several test cases as shown below with different user inputs to indicate the fuzzy system responses on different input.

The case study is presented to show the realization of the developed system. The output from the system is shown in the following. The input used is the number of inflow vehicles, number of outgoing vehicles, the distance between sensor 1 and sensor 2 in meters and Time of Day. The output is Green Light Duration changes in seconds as shown in Table 2. FAM table and each linguistic graph will reflect based on the input entered as well, as shown in Figs. 7, 8, 9 and 10.

Such coordination system is significant to assist traffic light formation monitoring. The database which kept the data and knowledge base efficiently supports any changes which necessary.



Fig. 6 Interface design of configuration module

Table 2 Test case

Input	Values
No of inflow vehicles	15
No of outgoing vehicles	8
Distance between s1 & s2 (m)	100
Time of day	07.55
Green light duration changes (s)	-5.2

The boldface sentence shows the output

D/T	Off-Peak	Normal	Peak
Sparse	Rem(0.00)	Dec(0.00)	Dec(0.60)
Average	Inc(0.00)	Rem(0.00)	Dec(0.40)
Dense	Inc(0.00)	Inc(0.00)	Rem(0.00)

Green Light duration changes:-5.20s

Fig. 7 FAM table and crisp output

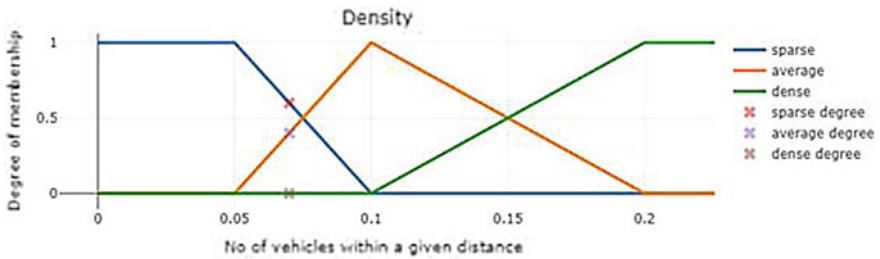


Fig. 8 Density membership function graph

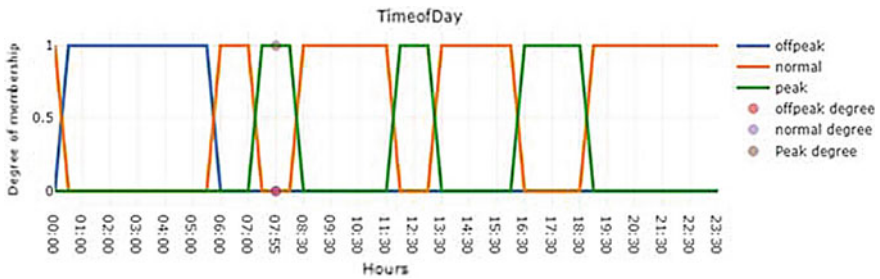


Fig. 9 Time of day membership function graph

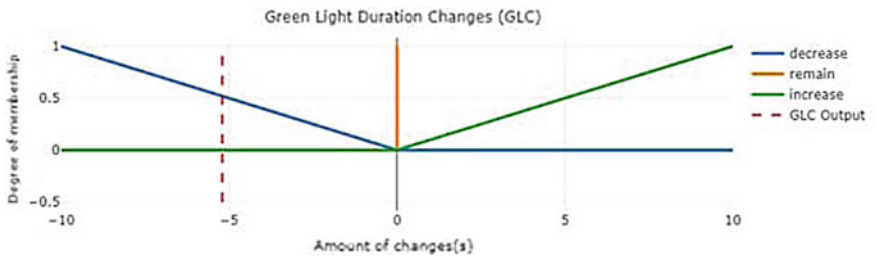


Fig. 10 Green light duration changes membership function graph

6 Conclusion

A fuzzy-based traffic light control system was developed to manage traffic with the ability to respond to daylight and traffic congestion and then determine optimal changes to green light periods. However, this system has its limits when it comes to situations with different peak sessions, regular sessions or off-peak sessions. Thus, future work can be focused on the development of features that allow for dynamic configuration of the number of sessions required and also add more vague variables such as the course of the week to integrate with the current density and time variables.

Funding This study was funded by Ministry of Education (MOE) through Fundamental Research Grant Scheme (FRGS/1/2019/ICT02/UTHM/02/7) Vote K208 and Postgraduate Research Grant (GPPS) Vote No. H332.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Li Z, Shahidehpour M, Bahramirad S, Khodaei A (2016) Optimizing traffic signal settings in smart cities. *IEEE Trans Smart Grid* 8(5):2382–2393
2. Priyadarshi S, Shekhar S, Singh TK, Gupta SK, Bansal M, Singh J Sharma K (2018) Intelligent traffic control system (ITCS). In: 2018 international conference on sustainable energy, electronics, and computing systems (SEEMS). IEEE, pp 1–3
3. Liebig T, Piatkowski N, Bockermann C, Morik K (2017) Dynamic route planning with real-time traffic predictions. *Inf Syst* 64:258–265
4. Atkinsglobal.com Adaptive Traffic Signal Control (ATSC) (2013) Retrieved from <https://www.atkinsglobal.com/~media/Files/A/Atkins-Corporate/north-america/sectors-documents/highways-and-bridges/library-docs/brochures/atkins-adaptive-traffic-signal-controls-612-2013.pdf>
5. Bandara GEMDC, Pathirana SD, Usoof TMHA (2003) Application of fuzzy logic in intelligent traffic control systems. *Small*, 5(05.0):05–0 (2003)
6. Han T, Lin C (2002) Design of an intelligence traffic light controller (ITLC) with VHDL. In: 2002 IEEE region 10 Conference on Computers, Com-munications, Control and Power Engineering. TENC0M'02. 3:1749–1752 (2002)
7. Patil R (2016) Smart traffic controller using fuzzy inference system (STCFIS), pp 335–340
8. Patil R, Srinivasaraghavan A (2016) Smart traffic controller using fuzzy inference system (STCFIS). In 2016 2nd international conference on next generation computing technologies (NGCT), IEEE, pp 335–340
9. Jhaveri CS, Perrin J, Martin P (2003) SCOOT adaptive signal control: an evaluation of its effectiveness over a range of congestion intensities. In: Transportation research board 2003 annual meeting, compendium of papers
10. Khalid M (1996) Intelligent traffic lights control by fuzzy logic. *Malaysian J Comput Sci* 9(2):29–35
11. Liu Z (2007) A survey of intelligence methods in urban traffic signal control. *IJCSNS Int J Comput Sci Network Secur* 7(7):105–112
12. Khooban MH, Vafamand N, Liaghat A, Dragicevic T (2017) An optimal general type-2 fuzzy controller for Urban traffic network. *ISA Trans* 66:335–343

13. Djahel S, Smith N, Wang S, Murphy J (2015) Reducing emergency services response time in smart cities: An advanced adaptive and fuzzy approach. In: 2015 IEEE first international smart cities conference (ISC2), pp 1–8
14. Cheng J, Wu W, Cao J, Li K (2016) Fuzzy group-based intersection control via vehicular networks for smart transportations. *IEEE Trans Ind Info* 13(2):751–758
15. Odeh SM (2015) Hybrid algorithm: fuzzy logic-genetic algorithm on traffic light intelligent system. In: 2015 IEEE international conference on fuzzy systems (FUZZ-IEEE), pp 1–7
16. Id GP, Id TC, Canale A (2018) Smart pedestrian crossing management at traffic light junctions through a fuzzy-based approach

Metaheuristic Optimization Approaches for Overcurrent Relays Discrimination: A Comparative Study



Noor Zaihah Jamal, Mohd Herwan Sulaiman, and Abdul Nasir

Abstract The discrimination for overcurrent relays has been considered as an optimization problem with complex constraints. Various meta-heuristic algorithms are presented to response with this drawback for over past few decades. This paper presented a study which compare two different meta-heuristic optimization approaches inspired by bio-nature. The approaches are tested to resolve the drawbacks of the overcurrent relays discrimination problem. With regards to this goal, two new optimization approaches called Ant Lion Optimizer and Grey Wolf Optimizer have been considered. The performances of these algorithms have been implemented on two different sizes of IEEE networks. The same initial condition and constraints' boundaries have been executing for each algorithm for fair comparison purposes. The generated results with optimal value is then identified as the best method to resolve the coordination of the overcurrent relays problem.

Keywords Pick-up current · Ant lion optimizer · Grey wolf optimizer · Overcurrent relays coordination · Time dial setting

1 Introduction

In modern distribution network, relays configuration supposedly to be arranged properly to ensure the reliability of the system's designed. Good configuration will remove the affected portion whereas keep the supply to the healthy portion during fault occurrences. Every equipment will be protected by two layer of relays which known as main relay and secondary relay. Main relays will react to the abnormal current flow within permitted time. If the main relay is fails to function well, secondary relays should take over the operation. Nevertheless, once the backup relay has taken over

N. Z. Jamal (✉) · A. Nasir
Faculty of Engineering Technology, Univesiti Malaysia Pahang, 26300 Gambang, Malaysia
e-mail: zaihah@ump.edu.my

M. H. Sulaiman
Faculty of Electrical & Electronics Engineering, Univesiti Malaysia Pahang, 26600 Pekan, Malaysia

on behalf of the primary relay, the unnecessary power outages to a bigger portion of the system will occur. This is why well-coordinated of the overcurrent relays is essential.

The optimization of the overcurrent relays (OCR) discrimination is established by two parameters that know as Plug Setting (PS) and Time Multiplier Setting (TMS). The OCR coordination is formulated as inequality constraints. In 1980s, The experimental of Trial and Error algorithm [1] has been wisely implemented to perform the OCR discrimination work. However, this approach only suitable for small distribution scale network. In early late 1980s researches has moved to topological analysis method which uses the graph theory approach to determine break points [2]. Whereas, in late 1990s, Linear Programming (LP) approach was presented in the frame of optimization method [3–6]. However, the slow convergence rate is the main drawback for the conventional methods especially for the system with large iteration no.

In early 2000, Genetic Algorithm (GA) [7–10] has pioneered the optimization method to resolve the OCR issue. GA known as the most popular method during that days. Then modification has been made to the original GA to improve the computational time of the method that called Continuous Genetic Algorithm (CGA) [11]. The evolutionary of the nature inspired approach is then continued with Particle Swarm Optimization [12–14] method which demonstrated to offer promising result rather than original GA and modified GA. The modernization of the technique is resumed by Cuckoo Search Algorithm [15]. In addition, hybrid method has been introduced in [16–19] to enhance the capabilities of the approaches. In overall, the approaches are established to explore the optimal OCR setting.

The No Lunch Free (NFL) theorem [20] highlighted that none of the method are able to solve all the problems. Since there are no limits in approaching for a new algorithm, this paper is proposing two new algorithms that known as Ant Lion Optimizer (ALO) and Grey Wolf Optimizer (GWO). Both algorithms are new for the OCR coordination field. The algorithm will be applied to search for the optimize value of the parameters TDS and Ip. The optimized parameters' value will have determined the minimize value of the objective function. From this comparative study, the best method shall be established with regards to solve the OCR coordination issue.

This research is constructed as follows; Firstly, problem formulation explained at Sect. 2. Next section explained on the background of GWO and ALO. The paper is resumed with Sect. 4 which analyzed the simulation results. The conclusion of the work is in Sect. 5.

2 Problem Formulation

In order to ensure that the total operating time of the main relays are minimized whereas sustaining the selectivity in between main and secondary relays value with regards to the position of the DG insertion. Meanwhile, the aim is to optimize the plug setting (PS) and time multiplier setting (TMS) parameters.

2.1 Objective Function

It could be expressed as below:

$$\min X = \sum_{k=1}^j T \quad (1)$$

where j is the main relays that need to be coordinated in total. T is the functioning time for the relays at the near end fault.

2.2 Characteristic Curve

Normal inverse characteristic (IDMT) will be applied where $b = 0.02$ and $k = 0.14$ as accordance to IEC standard. This is formulated as below.

$$T_i = \frac{k * TMS_a}{(I_f / PS_a)^b - 1} \quad (2)$$

PS_a is the PS, TMS_a is the setting for time multiplier, The current which may be detected by the respective relay as a fault is I_f .

2.3 Constraints

TMS will determine the time delay that bound in between 0.1 and 1.1 s whereas PS will determine the current delay which bound in between 1.5 and 5 s.

$$TMS_{\min} \leq TMS \leq TMS_{\max} \quad (3)$$

$$PS_{\min} \leq PS \leq PS_{\max} \quad (4)$$

The time interval (DTI) to discriminate main and secondary relays must be fulfilled to ensure that relays are operating in sequence.

$$Ti_{se} \geq Ti_{ma} + CTI \quad (5)$$

Ti_{se} is the operating time for secondary relay, Ti_{ma} is the operating time for main relay and DTI is in between 0.2 and 0.5 s.

3 Meta-heuristic Approaches

This section explains on background of the GWO in the first subsection and in the next subsection, an overview to the ALO would be presented. Both algorithms have been introduced by Mirjalili et al. Details on the GWO and ALO can be found in [21] and [22] respectively

3.1 Grey Wolf Optimizer (GWO)

Grey Wolf Optimization method is inspired by the leadership pyramid of grey wolf initiated by Mirjalili et al. [21]. The group has average members of not more than 12 wolves. The top ranking is a leader known as alpha (α) which liable for decisions making and dominating the group. The leader is nominated due to their competency to control the members of the group.

The second ranking role as an assistance to impose instruction by the leader that called beta (β). The β is either male or female with good discipline. The last tier called Delta (δ) is once used to be α and β . The function of δ as nanny to the little members and eldest. The called not so important members are Omega (ω). The ω existence is to equalize the bio-chain of the group. There are three phases during hunting activity which are:

- Tracking: track the spot of the target.
- Encircling: surround the target in a loop.
- Attacking: repositioning to or from the target according to the terms.

The α will be placed as the best solution, secondly by β and lastly by δ . The ω will update the positioning referred to three best solutions.

Base on bio-nature capabilities, the first three best wolves are expected to have knowledge on the prey's position for mathematical modelling purposes. The formulas as below are obtained [21].

During hunting activity, the first step is encircling the prey. This could be modelled by the following equations [21]:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \quad (6)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (7)$$

C and A are vectors' coefficient, \vec{X} is the wolf positioning vector, \vec{X}_p is the prey's location and t is the iteration for present situation. The vector's coefficient of C and A are as following equation [21].

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (8)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{9}$$

\vec{a} is deducted linearly from 2 to 0 at each iterations, \vec{r}_1, \vec{r}_2 are random vectors within (0,1). According to wolf behaviors of hunting, the re-positioning of the present location is accordance to prey's location. The updated location will be depends on \vec{A} and \vec{C} with respect to the present location of the wolf. The wolf could be placed to new places by adjusting the \vec{A} and \vec{C} . Agents are allowed to shift to a random location circling the prey accordance to \vec{r}_1 and \vec{r}_2 as Eqs. (8) and (9).

In a wide search area, the optimum location of the prey is a toughest exercise to be performed. α, β and δ will assume the prey's location by the nature capabilities and could be modelled as the following [21].

$$\vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right| \tag{10}$$

$$\vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right| \tag{11}$$

$$\vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \tag{12}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{X}_\alpha) \tag{13}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{X}_\beta) \tag{14}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{X}_\delta) \tag{15}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{16}$$

The final position of the search agents is according to α, β and δ . The position would be within the search area. The ω will re-new their location in random accordance to the location of the target which have been estimated by α, β and δ .

Only when the prey stop moving, the wolves will attack and it stop the hunting process. As the wolves are forthcoming the prey, the value of \vec{a} is decreased. The decreasing of value \vec{a} will give impact to value of \vec{A} between $(-2a, 2a)$. The wolves are attacking the target if $|A| < 1$ and back-off when $|A| > 1$. For the next iteration, the process will be repeated until it is terminated by the satisfied criterion.

3.2 Ant Lion Optimization (ALO)

The method of ALO is based on the relation in between ants and antlions during hunting [22]. In nature, ants move randomly over the search area when searching for food. It could be mathematically described as below

$$X(t) = [0, \text{cumsum}(2r(t_1) - 1), \text{cumsum}(2r(t_2) - 1), \dots, \text{cumsum}(2r(t_T) - 1)] \quad (17)$$

t and T indicate the iteration (random pace) and maximum iteration numbers respectively. cumsum computes the sum of cumulative, $r(t)$ indicates the function expressed like the following

$$r(t) = \begin{cases} 1 & r > 0.5 \\ 0 & r \leq 0.5 \end{cases} \quad (18)$$

r indicates a random number within $[0, 1]$. The ants' locations are saved in matrix form as below [22]:

$$M_{Ant} = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,d} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ A_{n,1} & A_{n,2} & \dots & A_{n,d} \end{bmatrix} \quad (19)$$

where M_{Ant} shows the position matrix for each ant. Respectively, parameters n is ants' quantity and d is dimension number. $A_{i,j}$ means the j -th variable value of i -th agent. During optimization, a target function is used to evaluate the qualification of every ant and the values are arrange as below [22].

$$M_{QA} = \begin{bmatrix} f([A_{1,1}, A_{1,2}, \dots, A_{1,d}]) \\ \dots \\ \dots \\ f([A_{n,1}, A_{n,2}, \dots, A_{n,d}]) \end{bmatrix} \quad (20)$$

where M_{QA} represent the arrangement of fitness of every ant. f indicates the objective function, n indicates the ants' number whereas d indicates the variables and $A_{i,j}$ indicates j -th arrangements of i -th ant. In addition, antlions are also hiding in traps within the area which the locations and qualifications values can be saved in the matrices as following:

$$M_{Antlion} = \begin{bmatrix} A_{1,1}, A_{1,2}, \dots, A_{1,d} \\ \dots \\ \dots \\ A_{n,1}, A_{n,2}, \dots, A_{n,d} \end{bmatrix} \quad (21)$$

where $M_{Antlion}$ shows the position arrangement for each antlion. n shows the antlions and d is the arrangement. $A_{i,j}$ denotes the j -th variable of i -th antlion.

$$M_{QAL} = \begin{bmatrix} f([AL_{1,1}, AL_{1,2}, \dots, AL_{1,d}]) \\ \dots \\ f([AL_{n,1}, AL_{n,2}, \dots, AL_{n,d}]) \end{bmatrix} \quad (22)$$

M_{QAL} demonstrates the arrangement for saving the qualification of every antlion. Variables n and d indicate the number of antlions and dimension, respectively. $A_{i,j}$ signifies the j -th of i -th antlion and f represents objective function. Furthermore, ants will update their positions during each optimization. The random walk of agents within the boundary is modelled as follows [22].

$$X_i^t = \frac{(X_i^t - a_i) \times (d_i - c_i^t)}{d_i^t - a_i} + c_i \quad (23)$$

a_i represents the least pace of variable i . Parameters c_i is minimum variables and d_i is maximum variables for i -th agents. On the other hand, variables c_i^t is minimum and d_i^t is maximum variable at the iteration t .

The antlions' effect on random walk of ants is mathematically formulated as below [22].

$$c_i^t = Antlion_j^t + c^t \quad (24)$$

$$d_i^t = Antlion_j^t + d^t \quad (25)$$

$Antlion_j^t$ represents the location of the certain j -th antlion. Parameters c^t and d^t indicate the maximum and minimum of t -th iteration's variables respectively. Moreover, vectors d and c defined that ants are moving around designated antlion as hyper sphere.

In ALO, ants are captured by selected single antlion. Thus, ALO approach implemented roulette wheel operator to select the antlions upon their ability the process of optimization. This process can be mathematically described as follows where the radius of hyper sphere is decreased adaptively [22].

$$c^t = \frac{c^t}{I} \quad (26)$$

$$d^t = \frac{d^t}{I} \quad (27)$$

where I , the percentage. $I = 10^{\frac{w}{T}}$. T and t are the maximum and present iteration, correspondingly. Constant w defined the exploitation level accuracy.

Finally, the prey becomes fitter than its respective predator when the ant is being caught by antlion in the trap. The antlion will consequently re-new its location to where the hunted ant located to improve the possibility to get new prey. The formula described as follows [22].

$$Antlion_j^t = Ant_i^t \text{ iff } (Ant_i^t) > (Antlion_j^t) \quad (28)$$

Ant_i^t and Ant_j^t represent the location at t -th iteration for designated i -th and j -th ant, respectively. In addition, at t -th iteration $Antlion_j^t$ denotes the location of the designated j -th antlion.

Elitism is a vital behaviour that helps ALO algorithm to maintain the best results attained for each optimization process. The fittest antlion is known as the elite. The movements of the ants are effected by the elite. Hence, all the ants walk randomly within the designated antlion [22].

$$Ant_j^t = \frac{R_A^t + R_E^t}{2} \quad (29)$$

At t -th iteration, R_A^t and R_E^t are the walk randomly within the selected antlion by the roulette wheel and the elite, respectively. Besides, at t -th iteration, Ant_j^t is the location of the i -th ant.

The flow chart of OCR coordination using GWO and ALO as in Fig. 1

4 Simulation Results

The implementation of the GWO and ALO algorithms have been implemented on i5-6200U core intel CPU, 8 GB 2.3 GHz RAM using MATLAB software. The algorithms have been verified on different IEEE 3-bus and IEEE 8-bus systems to identify the optimal objective function. Normal inverse characteristic curve is implemented to both test cases where with $k = 0.14$ and $\alpha = 0.02$. The values of constant are accordance to standard of IEC [23].

4.1 Case I

The system is supplied by three generators and consists of three busbar (B_1 , B_2 and B_3), six overcurrent relay ($R_1, R_2, \dots R_6$) and three ring lines with 69 kV. The test case details could be retrieved in [5]. The I_p parameter varies from 1.5 to 5 [24] and TDS parameter varies from 0.1 to 1.1 s [18, 24] and. The 0.3 s of CTI value of is implemented. The results are exhibited with discrete I_p and continuous TDS models. The no. of iteration used is 1000 with 30 search agents.

The optimized results of GWO and ALO are presented as Table 1. It could be

Fig. 1 OCR coordination flow chart using GWO and ALO

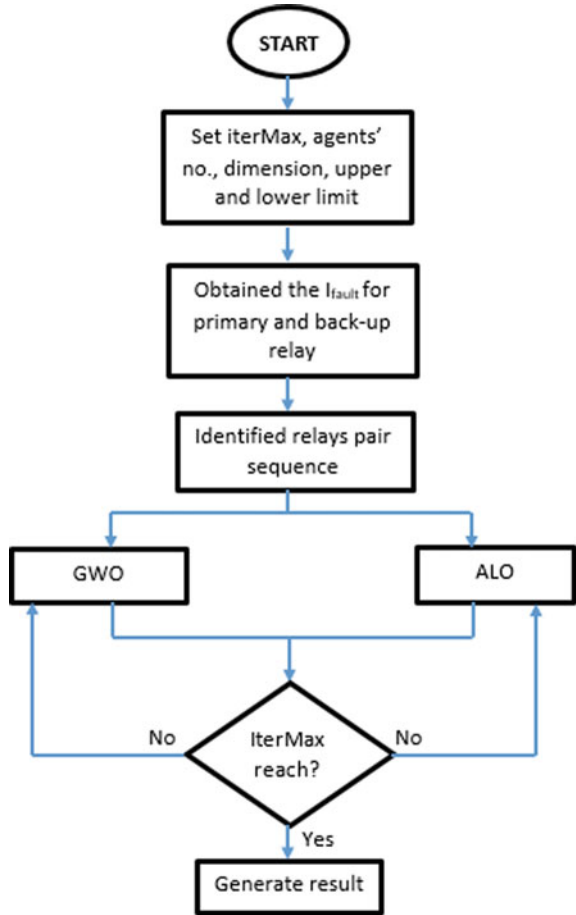


Table 1 Optimum setting for case I

Relay no.	CT	GWO		ALO	
		TMS	PS	TMS	PS
1	300	0.1000	3.0	0.1000	1.5
2	200	0.1001	1.5	0.1000	3.0
3	200	0.1000	3.0	0.1000	3.0
4	300	0.1000	3.0	0.1000	2.0
5	200	0.1001	1.5	0.1000	3.0
6	400	0.1000	1.5	0.1000	1.5
Operating time (s)		1.4779		1.4815	

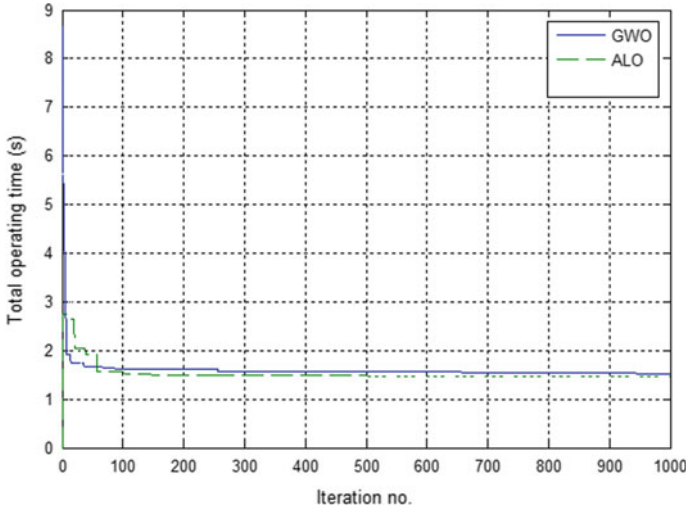


Fig. 2 ALO versus GWO results for 1000 iteration

highlighted that the GWO accomplishes better results with 0.0036 s better than ALO. This shows that GWO presents better than ALO to accomplish optimal value of TDS and Ip.

4.2 Case II

The system connect six busbars ($B_1, B_2 \dots B_6$) and 14 overcurrent relays ($R_1, R_2, \dots R_{14}$), that consists of seven ring lines as in Fig. 2. The Ip parameter bound in between 1.5 and 5 whereas TDS parameter bounds in between 0.1 and 1.1 s. The system can be retrieved from [14]. The results of GWO and ALO as in Table 2.

The simulation results obtained shows that GWO has decreases the relays' operational time for about around 0.5104 s as compared to ALO algorithm (Fig. 3).

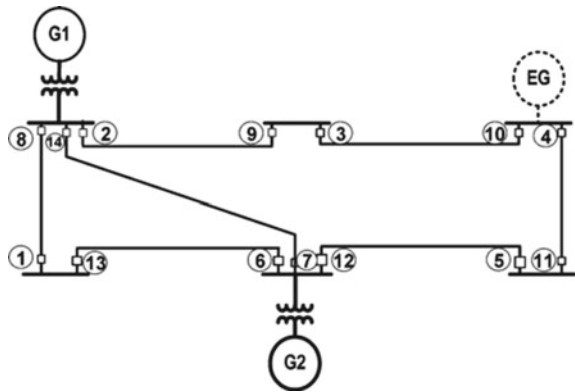
5 Conclusion

This paper simulates the performances of two metaheuristic approaches known as GWO and ALO to solve the OCR coordination problem. Both approaches have been executed with similar initial condition, characteristic curve and parameter model to two systems with different sizes. It could be highlighted that GWO approach generates the minimized relays's operating time. As the conclusion, GWO is considered the best approach to solve the OCR coordination problem.

Table 2 TMS and PS setting for case II

Relay no.	ALO		GWO	
	TMS	PS	TMS	PS
1	0.1000	1.9773	0.1018	1.5000
2	0.3359	1.5012	0.1109	4.0000
3	0.4898	2.0931	0.1720	5.0000
4	0.7147	3.6785	0.4348	5.0000
5	0.1000	1.5003	0.1009	1.5000
6	0.1000	1.5000	0.1002	4.0000
7	0.1000	1.5000	0.1003	3.0000
8	0.1001	1.6083	0.1007	4.5000
9	0.1000	1.5152	0.1002	5.0000
10	0.1000	1.5647	0.1004	4.0000
11	0.1000	1.5023	0.1059	1.5006
12	0.1000	1.5420	0.1003	4.5000
13	0.1000	1.5004	0.1006	1.5000
14	0.1001	1.9899	0.1015	2.0030
Operating time (s)	8.4536		7.9432	

Fig. 3 IEEE 8 bus system [14]



Funding This research is supported under grant RDU1703284 by Universiti Malaysia Pahang (UMP).

Conflict of Interest The authors declared, they have no conflict with the others research.

References

1. Damborg MJ et al (1984) Computer aided transmission protection system design part I: algorithms. *IEEE Trans Power Apparatus Syst* PAS-103(1):51–59
2. Mansour MM, Mekhamer SF, El-Kharbawe N (2007) A modified particle swarm optimizer for the coordination of directional overcurrent relays. *IEEE Trans Power Deliv* 22(3):1400–1410
3. Urdaneta AJ et al (1996) Coordination of directional overcurrent relay timing using linear programming. *IEEE Trans Power Deliv* 11(1):122–129
4. Abdelaziz AY et al (2002) An adaptive protection scheme for optimal coordination of overcurrent relays. *Electr Power Syst Res* 61(1):1–9
5. Urdaneta AJ, Nadira R, Perez Jimenez LG (1988) Optimal coordination of directional overcurrent relays in interconnected power systems. *IEEE Trans Power Deliv* 3(3):903–911
6. Birla D, Maheshwari RP, Gupta HO (2006) A new nonlinear directional overcurrent relay coordination technique, and banes and boons of near-end faults based approach. *IEEE Trans Power Deliv* 21(3):1176–1182
7. Cheng-Hung L, Chao-Rong C (2007) Using genetic algorithm for overcurrent relay coordination in industrial power system. In: International conference on intelligent systems applications to power systems, 2007. ISAP 2007
8. So CW et al (1997) Application of genetic algorithm for overcurrent relay coordination. In: Sixth international conference on developments in power system protection (Conf. Publ. No. 434)
9. So CW et al (1997) Application of genetic algorithm to overcurrent relay grading coordination. in *Advances in Power System Control, Operation and Management, 1997. APSCOM-97*. In: Fourth international conference on (Conf. Publ. No. 450)
10. Razavi F et al (2008) A new comprehensive genetic algorithm method for optimal overcurrent relays coordination. *Electr Power Syst Res* 78(4):713–720
11. Bedekar PP, Bhide SR (2011) Optimum coordination of overcurrent relay timing using continuous genetic algorithm. *Expert Syst Appl* 38(9):11286–11292
12. Bansal JC, Deep K (2008) Optimization of directional overcurrent relay times by particle swarm optimization. In: *Swarm intelligence symposium, 2008. SIS 2008*. IEEE
13. Moravej Z, Jazaeri M, Gholamzadeh M (2012) Optimal coordination of distance and overcurrent relays in series compensated systems based on MAPSO. *Energy Convers Manag* 56:140–151
14. Zeineldin HH, El-Saadany EF, Salama MMA (2006) Optimal coordination of overcurrent relays using a modified particle swarm optimization. *Electr Power Syst Res* 76(11):988–995
15. Ahmarinejad A et al (2016) Optimal overcurrent relays coordination in microgrid using cuckoo algorithm. *Energy Procedia* 100:280–286
16. Sueiro JA et al (2012) Coordination of directional overcurrent relay using evolutionary algorithm and linear programming. *Int J Electr Power Energy Syst* 42(1):299–305
17. Papaspiliotopoulos VA, Kurashvili TA, Korres GN (2014) Optimal coordination of directional overcurrent relays in distribution systems with distributed generation based on a hybrid PSO-LP algorithm. In: *MedPower 2014*
18. Noghabi AS, Sadeh J, Mashhadi HR (2009) Considering different network topologies in optimal overcurrent relay coordination using a hybrid GA. *IEEE Trans Power Delivery* 24(4):1857–1863
19. Chunlin X et al (2008) Optimal coordination of protection relays using new hybrid evolutionary algorithm. in *Evolutionary Computation, 2008. CEC 2008*. (IEEE World Congress on Computational Intelligence). IEEE Congress on. 2008
20. Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. *IEEE Trans Evol Comput* 1(1):67–82
21. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey Wolf optimizer. *Adv Eng Softw* 69:46–61

22. Mirjalili S (2015) The Ant Lion Optimizer. *Adv Eng Softw* 83:80–98
23. IEC (1976) Single input energizing measuring relays with dependent specified time In: IEC standard 60255–4. IEC publication
24. Amraee T (2012) Coordination of directional overcurrent relays using seeker algorithm. *IEEE Trans Power Delivery* 27(3):1415–1422

A Mutation Triggering Method for Genetic Algorithm to Solve Traveling Salesman Problem



Khandker M. Qaiduzzaman, Sabira Khatun, Maliha Afsa, Sadman Sobhan, Md. Elias Hossain, Syamimi Mardiah Shaharum, and Mostafijur Rahman

Abstract This chapter reveals the effect of combining mutation operators in Genetic Algorithm. Though mutation is a very effective genetic operator to escape the local optima, it has an adverse effect on computation time. It is quite a challenge to say how much probability is appropriate for mutation. The major contribution of this chapter is to design a mutation triggering method by combining three mutation operators (Swap, Insertion and 2-Opt) and applying adaptive probability. To decide which mutation can be activated at a given generation, a decision-making method named Mutation Triggering Method is proposed. To measure the performance of the proposed method, the classical traveling salesman problem was considered. The experimental result shows that the proposed method was able to find a better solution (travel cost) than the other approaches. For computation time, the proposed strategy did as well as other mutation strategies. In conclusion, the combination of several mutation operators can ensure the benefits of diversity as well as the benefits of faster convergence.

Keywords Triggering · Adaptive · Mutation · Genetic · Hybrid · Operator · 2-opt · Swap · Insertion · TSP

K. M. Qaiduzzaman · M. Afsa · Md. Elias Hossain · M. Rahman
Daffodil International University, Dhaka 1207, Bangladesh
e-mail: nafees.ku11@gmail.com

S. Sobhan
Dohatec New Media, Dhaka 1000, Bangladesh

S. Khatun (✉) · S. M. Shaharum
Faculty of Electrical and Electronic Engineering Technology, Universiti Malaysia Pahang, Pekan,
Pahang 26600, Malaysia
e-mail: sabirakhatun@ump.edu.my

1 Introduction

The benefit of solving traveling salesman problem (TSP) is that it has many real-world applications. The notable applications of classical TSP are vehicle routing, order picking problem in warehouses, computer wiring, drilling of the printed circuit board (PCB), X-ray crystallography, overhauling gas turbine engine, musk plotting in PCB production [1]. In this paper, the Genetic Algorithm (GA) is implemented for solving TSP. The framework of GA consists of genetic operators that could be modified easily according to the nature of the problem. The best thing is that if the operators of GA can be designed perfectly, it provides the optimal/near-optimal solution. It is inspired by the biological evaluation where each of the solutions is considered as a chromosome. Every chromosome is consisting of a set of biological elements, gene. The basic idea of GA is to generate some chromosome (i.e. solutions) randomly or based on some ideas (heuristic). These chromosomes are evaluated using a fitness function. The fitness function assigns a fitness value to each chromosome. After this phase, a selection is performed to decide which chromosomes are fittest to survive on the next generation. The literature shows that there are many strategies of selection such as Roulette Wheel Selection, Rank Selection, Tournament Selection, etc. [2]. After the selection, the fittest chromosomes are called as parents and they get the chance of biological reproduction through the crossover. Crossover is one of the most important genetic operators that helps to converge to the solution. The basic idea of a crossover is to exchange genetic elements (i.e. gene) between two parents to generate new offspring. Numerous researches have been done on crossover operators and some popular crossover strategies are Partially Mapped Crossover (PMX), Cycle Crossover (CX), Order Crossover (OX), etc. [3]. Crossover is applied with a higher probability to converge faster. However, using only the crossover might converge the algorithm to local optima. For this reason, another genetic operator named mutation is introduced to GA. The mutation is liable for bringing diversity to the solution by tweaking some genetic materials of a few chromosomes. Some well-known mutation operators are Exchange or Swap Mutation (EM), Insertion Mutation (IM), Simple Inversion or 2-opt Mutation (SIM), etc. [3]. The mutation is applied with a very low probability P_m . Higher mutation probability results in divergence of the algorithm. Still, the mutation is considered the most crucial genetic operator to find a globally optimal solution. After the mutation, the old population are replaced by the new population (generated by crossover and mutation). That is how one generation of GA is completed. GA should be run for several generations to find an optimal solution.

In this work, we solved TSP which is a well-known NP-hard problem [4]. It consists of a set of cities and a salesman. The salesman has to visit each one of the cities starting from a certain one (also known as the hometown) and returning to the same city. The challenge of this problem is that the traveling salesman wants to minimize the total length of the trip [5]. Given, n cities and the objective of TSP is to find an integer permutation $\pi = c_1, c_2, c_3, \dots, c_n$ with c_n being the city n , that minimizes the sum, $(\pi) = \sum_{i=1}^{n-1} d(c_i, c_{i+1}) + d(c_n, c_1)$, where $d(c_i, c_{i+1})$ denotes the distance between city i and city $i + 1$. When GA is implemented to solve TSP,

each city is considered as a gene and each tour is considered as a chromosome. TSP is represented as a graph and it is considered difficult to solve. For example, if the number of cities of a TSP problem is 10, there are $(n-1)! = (10-1)! = 362,880$ possibilities to try which is not wise to solve using Brute Force approach. Moreover, the complexity of the problem increases when the number of cities increases. The main challenge of this problem is to find an optimal solution within a short time. That is why we choose GA which is very effective because of its meta-heuristic nature, yet a further improvement of GA is needed. Especially, the genetic operators like selection, crossover and mutation play a vital role in the performance of GA. Good operator design can converge the algorithm quickly whereas bad operator design can take the same algorithm away from the optimal solution. We expect that the combination of genetic operators might give a better solution.

The literature shows that extensive studies have been done to solve TSP using GA in the last decades. In [6], a multi-objective solid traveling salesman problem is solved using R-MOGA. In this paper, a generation dependent mutation operator is proposed where the probability of the mutation decreases smoothly with the increase of the generation. In paper [7], an experiment is done to anticipate the effects between the 2-opt and the mutation operator for solving the Euclidean traveling salesman problem (TSP). In paper [8] a hybrid mutation operator called “IRGIBNNM” is proposed. The main idea of the paper is to combine a knowledge-based mutation and a random-based mutation. In paper [9] GA is developed by a sequential constructive crossover. Also, 2-opt search, a proposed local search and immigration were performed to find a better solution. In paper [10] an improved strategy is proposed to blend random crossover and dynamic mutation. The author claims that it increases the diversity of the population and optimizes mutation characters. Some modified and new crossover strategies for GA are found in [11, 12]. In paper [13], stochastic hill climbing algorithm was applied for mutation process. A comparative analysis is done on mutation operators in paper [14, 15]. In paper [16], two mutation operators named as Inverted Exchange Mutation and Inverted Displacement Mutation has been combined to improve the performance of GA and compared with four different mutation operators.

The contribution of this chapter is to combine several mutation operators and take the benefits of them. There are some random-based mutations such as Swap and Insertion mutation those are faster but unable to provide a high-quality solution. On the other hand, knowledge-based mutation (heuristic) such as 2-opt mutation is very effective to find an optimal solution. But, the major drawback of 2-opt is that it is computationally expensive. The novelty of this research is to develop a triggering method that can choose any mutation operator from Swap, Insertion and 2-opt Mutation. Our idea is to apply random mutation (Swap and Insertion) at the early generation of the algorithm so that it finds some feasible solution within a short time. At the later generations, we apply knowledge-based mutation (2-opt) so that the feasible solutions could be turned into an optimal solution. 2-opt is very popular for removing cycles from a TSP tour but takes a large amount of time. That is why

it is wise to invest valuable computation time only to find an optimal solution at the ending of the generation. In this work, we not only considered travel cost but also the computation time for better optimization.

The chapter is organized in a few sections. Section 1 is the introductory section that provides the theoretical background, problem definition, motivation, some related works and the novelty of this work. In Sect. 2 the methodology of this research is discussed elaborately. The experimental results are discussed in Sect. 3. It also provides a convergence analysis to better understand the performance of different algorithms. Finally Sect. 4 concludes the main remarks of this research with some future research directions.

2 Research Method

In this research, the genetic algorithm is implemented with improvement in mutation operator selection method. We proposed a Mutation Triggering Method (MTM) that can take the advantages of several prominent mutation operators such as, swap mutation, insertion mutation and 2-opt mutation.

2.1 The Structure of the Proposed Algorithm

We followed the basic structure of the genetic algorithm except for the mutation operator phase. The improved algorithm is shown in Fig. 1. The steps of the algorithms are described below:

Generate Initial Random Population. In this step, a fixed number of initial population was generated without any heuristic. As the concern of this research is to find the optimal/near-optimal solution using the mutation operators, the most commonly used, random population generation method was chosen. Some other population generation method can be found in [17–20].

Calculate Fitness of the Individuals. This step calculated the quality of each individual through a fitness function. To solve TSP, the fitness value of each chromosome is calculated by Eq. (1).

$$f = \frac{1}{D} \quad (1)$$

where, D = total travel distance of a tour = $d1 + d2 + \dots + dn$ and dn = Euclidian distance between two connected cities ci and $ci + 1$.

Stopping Criterion. This step is the stopping criterion of the algorithm. The algorithm runs for a fixed number of generations (by checking the condition, Generation < Max Generation) without considering the fact that it finds a better solution or not.

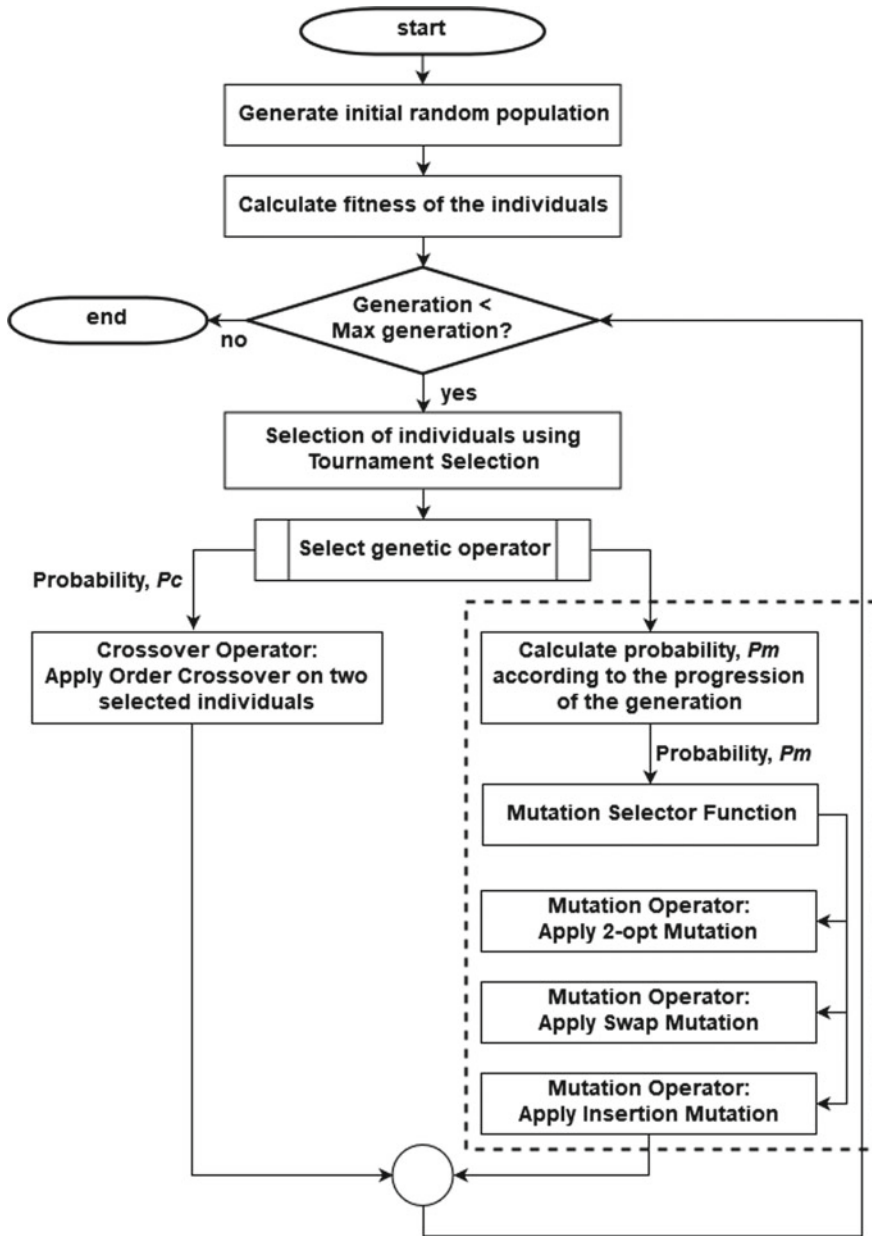


Fig. 1 The flowchart of the proposed Algorithm. The dotted line indicates the area of contribution

Selection of Individuals Using Tournament Selection. In this research, the Tournament Selection method is implemented since it has several benefits than other selection strategies [21]. Tournament Selection provides diversity as well as gives chance to all individuals to compete with each other.

Select Genetic Operator. In this step, crossover or mutation or both is selected according to their probability. Though in this work, the crossover operator has a fixed probability, the mutation probability pm is adaptable with the generation. It means that the mutation probability varies with the succession of the generation.

Crossover Operator: Apply Order Crossover on Two Selected Individuals. In this part, if the crossover probability allows the operation, two selected individuals perform recombination by following the Order Crossover strategy.

Calculate Probability, pm According to the Progression of the Generation. This step is one of the most crucial parts of the algorithm since the mutation probability pm is calculated here. The Mutation is a very sensitive operator that cannot be changed drastically. A higher probability of mutation might end up the algorithm to divergence with large computation time. On the other hand, a very smaller mutation probability might cause local optima. That is why, the mutation probability is reduced smoothly with the succession of the generation by following the Eq. (2), found in [6].

$$pm = \frac{k}{\sqrt{\text{current generation number}}} \quad (2)$$

Here, k is a constant factor chosen over a range of $[0, 1]$.

Mutation Selector Method. This step decides which mutation should be selected from 2-opt, Swap and Insertion mutation according to a selector equation. The 2-opt mutation is very effective to find cycles in the tour but the major drawback is that it is very slow. Applying 2-opt alone might cause greater computation time. On the other hand, Swap and Insertion mutations are very faster approach but the low-quality solutions. Therefore, in this work, these three mutation methods are combined to ensure diversity as well as lesser computation time. A selector variable is defined that decides when to choose 2-opt mutation according to the *factor* constant, shown in Eq. (3). For example, if the *factor* value is set as 0.2, it means that 2-opt will be triggered on the last 20% of the generations to generate a high-quality solution, sacrificing the higher computation time. On the other hand, the beginning 80% of the generation will take the benefits (faster) of Swap and Insertion mutations. Between Swap and Insertion mutation, a fair coin toss is performed.

$$\text{selector} = \text{max.generation} - (\text{factor} * \text{max.generation}) \quad (3)$$

Mutation Operator: Apply 2-opt Mutation. This mutation operator is applied according to the decision of the Eq. (3). If the generation number approaches to the end (say, last 20%) of the maximum generation, 2-opt is performed.

Mutation Operator: Apply Swap Mutation. At the beginning (say starting 80%) of the generation, Swap mutation is performed by a fair coin toss that is with 50% of probability.

Mutation Operator: Apply Insertion mutation. Insertion mutation is applied alongside the Swap mutation. It is mentioned earlier that swap mutation is applied with 50% probability. The rest of the 50% probability is applied to trigger the Insertion mutation.

2.2 The Specific Contribution of This Research: Designing the Mutation Triggering Method (MTM)

To be specific, designing MTM is the main contribution of this research. This function decides which mutation operator will be selected according to some ideas. The algorithm of MTM is shown in Algorithm 1.

Algorithm 1 MTM

```

1:  $k \leftarrow$  range [0, 1]
2:  $factor \leftarrow$  range [0, 1]
3: find the current generation number and maximum generation number
4: calculate the mutation probability,  $P_m = \frac{k}{\sqrt{\text{current generation number}}}$ 
5:  $r_1 \leftarrow$  a random number,  $r_2 \leftarrow$  a random number
6: if  $r_1 < P_m$  then
7:  $selector \leftarrow$   $max\text{ generation number} - factor * max\text{ generation number}$ 
8: if  $current\text{ generation number} > selector$  then
9:   Apply 2-opt mutation
10: else if  $r_2 < 0.5$  then
11:   Apply Swap mutation
12: else
13:   Apply Insertion mutation
14: end if
15: end if

```

First of all, a constant number k and a multiplying factor are chosen from the range of 0 to 1. Then the current generation number and the maximum generation number is found out. Next, the pm is calculated by dividing k by the square root of the current generation number. As the generation number increases, the denominator value becomes larger, resulting in the pm to be smaller. It implies that with the succession of the generation the probability of mutation, pm is decreased to ensure convergence. The constant k plays the role to distribute the value under a probabilistic range 0 to 1. Then a random number $r1$ is generated (between 0 to 1) and checked whether it is less than pm . When pm allows the mutation operation the algorithm concentrates on which mutation to be applied. Since in this work one mutation operator is applied at a time, the total generation was divided into two parts. For

example, the first part is the starting 80% of the maximum generation whereas the second part is the last 20%. If the current generation is among the last 20% of the maximum generation, the algorithm applies 2-opt mutation. This is because it is logical to spend valuable computation time on near-optimal solutions rather than random solutions. However, if the current generation lies in the first 80% of the generation, the algorithm does not concentrate on optimality rather diversity. That is why Swap mutation and Insertion mutation is applied by 50% of probability to ensure a quicker solution.

3 Results and Discussions

This work implemented the traditional genetic algorithm (T-GA), the random crossover and dynamic mutation genetic algorithm (RCDM-GA) [10] and the proposed mutation triggering method genetic algorithm (MTM-GA) to solve nine instances from TSPLIB [22] dataset. T-GA adopted swap mutation, insertion mutation and 2-opt mutation separately. RCDM-GA also adopted swap mutation and insertion mutation. Due to very high computational time, the 2-opt mutation was not used for RCDM-GA. Finally, MTM-GA was implemented with the combined mutations. Both travel cost and computation time were considered while these algorithms experimented.

3.1 Experimental Result Based on Travel Cost

The optimal solutions considering travel cost is represented in Table 1. Here, all the bold data values indicate the best/optimum solutions. From Table 1 it is seen that

Table 1 Optimal solution (travel cost) for different algorithms and mutation operators

Instances	Optimal	T-GA			RCDM-GA		MTM-GA
		Swap	Insertion	2-Opt	Swap	Insertion	Combined
eil51	426	754	871	433	551	621	432
a280	2579	20,753	21,476	2727	151,518	27,779	2756
bier127	118,282	365,293	374,119	123,294	469,925	411,091	119,416
kroa100	21,282	82,631	79,085	21,846	149,170	69,010	21,666
kroa200	29,368	199,256	190,549	30,971	234,087	242,155	30,697
lin318	42,029	384,807	388,311	43,702	1,479,572	698,061	44,563
pr226	80,369	1,025,700	976,004	81,132	1,228,103	1,089,287	80,728
st70	675	1886	1680	681	1133	31,093	678
rat195	2323	13,061	13,321	2419	19,665	87,280	2450

Table 2 Computation time for different algorithms and mutation operators

Instances	T-GA			RCDM-GA		MTM-GA
	Swap	Insertion	2-Opt	Swap	Insertion	Combined
eil51	2.3	2.48	13.23	4.09	5.94	5.3
a280	40.11	42.06	1313.95	49.09	64.69	354.79
bier127	10.67	10.63	157.77	13.68	19.65	50.54
kroa100	6.38	6.85	93.92	9.58	13.29	27.16
kroa200	21.35	22.92	657.78	27.28	29.63	185.74
lin318	51.07	52.46	2453.97	53.13	81.28	663.22
pr226	26.22	27.02	992.72	33.77	43.85	229.39
st70	4.06	3.88	31.67	6.39	9.56	11.55
rat195	21.44	21.46	570.57	25.43	36.15	155.37

T-GA adopting 2-opt mutation found best solutions 2727, 43,702 and 2419 for a280, lin318 and rat195 dataset instances respectively. But, most of the optimal solution was found by MTM-GA. It was successful to find the optimal solution for eil51 (432), bier127 (119416), kroa100 (21666), kroa200 (30697), pr226 (80728) and st70 (678) instances. However, T-GA (Swap), T-GA (Insertion), RCDM-GA (Swap) and RCDM-GA (Insertion) ended up in local optima.

3.2 Experimental Result Based on Computation Time

In this chapter, we considered computation time because it is also important to find out which algorithm finds the optimal solution in a considerable amount of time. Since Swap and Insertion mutations are the random-based approach, the algorithms adopting those mutation operators are the fastest. From Table 2 it is clear that all the minimum computation time (bold values) was taken by either the T-GA (Swap) or the T-GA (Insertion).

RCDM-GA (Swap) and RCDM-GA (Insertion) also takes very less computation time. However, the computationally most expensive approach is the T-GA (2-opt) algorithm. MTM-GA takes a considerable amount of time.

3.3 Convergence Analysis

We considered one instance rat195 to experiment how all of the algorithms converge with the progression of the generation. After 1000 generation the solution for each generation was shown in Fig. 2. It can be concluded that because of being random in nature, T-GA (Swap) and T-GA (Insertion) quickly get stuck in local optima,

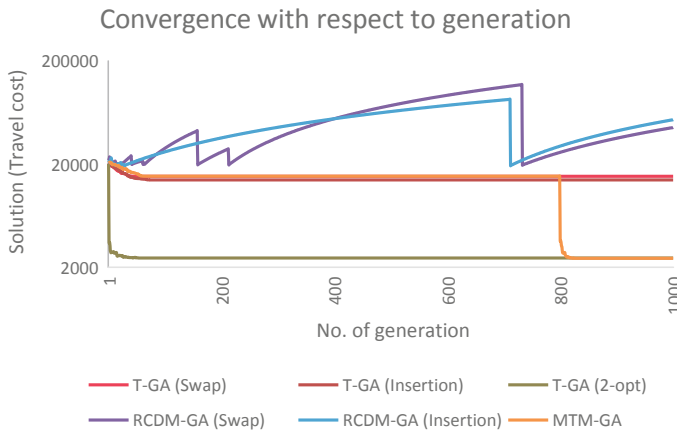


Fig. 2 The convergence of T-GA (Swap), T-GA (Insertion), T-GA (2-opt), RCDM-GA (Swap), RCDM-GA (Insertion) and MTM-GA algorithms after 1000 generation for the rat195 instance

implies that unable to bring diversity. Since T-GA (2-opt) is a heuristic approach, it finds the optimal/near-optimal solution within the first few generations. Running the algorithm for a long generation does not bring much diversity in this case. On the other hand, RCDM-GA (Swap) and RCDM-GA (Insertion) is too much random in nature. Since the mutation probability calculating function of RCDM-GA increases the probability when the solution does not vary, it fluctuates the solution for large dataset instances. It is generally effective for small dataset instances. However, MTM-GA is designed in such a way that it can use the benefits of both randomness and knowledge-based mutation operations. It quickly converges to the local optima using the random mutations. To escape from local optima, it uses heuristic mutation such as 2-opt mutation. The major improvement of TMT-GA is that it concentrates not only on optimality but also in computation time.

4 Conclusion and Future Work

Genetic algorithm is a very flexible algorithm that allows independence in operator design. Using random mutation cannot bring enough diversity when the algorithm needs it most. On the other hand, unnecessary mutation (especially when the algorithm is about to converge) might cause the wrong solution. Being too much heuristic-based again costs too much of computation time even though it is good for optimality. Our approach was to trigger the random and heuristic approach by understanding the generation of the algorithm. The experimental result proved that our proposed method can find an optimal solution. The T-GA (2-opt) is also able to find an optimal solution like the proposed method does. But, the novelty of the proposed

algorithm is that it can find the solution in a very less amount of time whereas other approaches are computationally expensive.

In future, we would like to implement more varieties of mutation operators to check the effect of the genetic algorithm. Implementing more measures to detect the rate of solution quality and finding the local optima might be a good direction to design the triggering method. It will be a great future research indication to make the triggering method recommending the appropriate mutation operator to escape from the local optima based on experience.

Acknowledgements This research work is partly supported by research grant RDU1703125 funded by Universiti Malaysia Pahang, <http://www.ump.edu.my/>. The authors would also like to thank Universiti Malaysia Pahang for publication support.

References

1. Matai R, Singh SP, Mittal ML (2010) Traveling salesman problem: an overview of applications, formulations, and solution approaches. *Traveling salesman problem, theory and applications*, InTech, Croatia
2. Shukla A, Pandey, HM, Mehrotra D (2015) Comparative review of selection techniques in genetic algorithm. In: 2015 international conference on futuristic trends on computational analysis and knowledge management (ABLAZE), pp 515–519
3. Larranaga P, Kuijpers CM, Murga RH, Inza I, Dizdarevic S (1999) Genetic algorithms for the travelling salesman problem: A review of representations and operators. *Artif Intell Rev* 13(2):129–170
4. Sathya N, Muthukumaravel A (2015) A review of the optimization algorithms on traveling salesman problem. *Indian J Sci Technol* 8(1)
5. Tollis IG (2019) Chapter 10 the traveling salesman problem. <https://www.csd.uoc.gr/~hy583/papers/ch11.pdf>. Last accessed 10 Mar 2019
6. Maity S, Roy A, Maiti M (2019) A rough multi-objective genetic algorithm for uncertain constrained multi-objective solid travelling salesman problem. *Granular Comput* 4(1):125–142
7. Chang C (2015) A 2-opt with mutation operator to the traveling salesman problem. *Int J Adv Eng Technol Comput Sci (IROSSS)* 2(1):16–21
8. Alkafaween E, Hassanat A (2018) Improving TSP solutions using GA with a new hybrid mutation based on knowledge and randomness. arXiv preprint [arXiv:1801.07233](https://arxiv.org/abs/1801.07233)
9. Ahmed ZH (2013) A hybrid genetic algorithm for the bottleneck traveling salesman problem. *ACM Trans Embed Comput Syst (TECS)* 12(1):9
10. Xu J, Pei L, Zhu RZ (2018) Application of a genetic algorithm with random crossover and dynamic mutation on the travelling salesman problem. *Procedia Computer Science* 131:937–945
11. Singh DR, Singh MK, Singh T, Prasad R (2018) Genetic algorithm for solving multiple traveling salesmen problem using a new crossover and population generation. *Computación y Sistemas* 22(2)
12. Hussain A, Muhammad YS, Nauman Sajid M, Hussain I, Mohamd Shoukry A, Gani S (2017) Genetic algorithm for traveling salesman problem with modified cycle crossover operator. *Comput Intell Neurosci*
13. Katayama K, Sakamoto H, Narihisa H (2000) The efficiency of hybrid mutation genetic algorithm for the travelling salesman problem. *Math Comput Model* 31(10–12):197–203
14. Abdoun O, Abouchabaka J, Tajani C (2012) Analyzing the performance of mutation operators to solve the travelling salesman problem. arXiv preprint [arXiv:1203.3099](https://arxiv.org/abs/1203.3099)

15. Albayrak M, Allahverdi N (2011) Development a new mutation operator to solve the traveling salesman problem by aid of genetic algorithms. *Expert Syst Appl* 38(3):1313–1320
16. Deep K, Mebrahtu H (2011) Combined mutation operators of genetic algorithm for the travelling salesman problem. *Int J Comb Optim Prob Inf* 2(3):1–23
17. Toğan V, Daloğlu AT (2008) An improved genetic algorithm with initial population strategy and self-adaptive member grouping. *Comput Struct* 86(11–12):1204–1218
18. Maaranen H, Miettinen K, Mäkelä MM (2004) Quasi-random initial population for genetic algorithms. *Comput Math Appl* 47(12):1885–1895
19. Diaz-Gomez PA, Hougen DF (2007) Initial population for genetic algorithms: a metric approach. In *Gem* 43–49
20. Deng Y, Liu Y, Zhou D (2015) An improved genetic algorithm with initial population strategy for symmetric TSP. *Math Prob Eng*
21. Saini N (2017) Review of selection methods in genetic algorithms. *Int J Eng Comput Sci* 6(12):22261–22263
22. Reinelt G (1991) TSPLIB—A traveling salesman problem library. *ORSA J Comput* 3(4):376–384

Optimized Weight Selection in Levenberg-Marquardt Back-Propagation Using Wolf Search



M. Z. Rehman, Aisha Tanveer, Nazri Mohd Naw, and Kamal Z. Zamli

Abstract Levenberg-Marquardt back-propagation (LMBP) algorithm is one of the most widely used alternative algorithms for training of artificial neural networks (ANN). In some cases, learning in LM can become a challenging machine-learning task owing to poor selection of weights and biases. In this manner, the network may become unstable and cause the search process to get stuck in local minima (i.e. resulting into network stagnancy). Recently, meta-heuristic algorithms have proven themselves to alleviate such problem and improve the convergence slow convergence in ANN. In this paper, Wolf Search (WS) that simulates the hunting behavior of wolves in their natural habitat is implemented with LMBP algorithm to optimize the weight selection in ANN. The performance of the proposed Wolf Search with Levenberg-Marquardt (WSLM) is compared with back-propagation and other stochastic LM algorithms by means of simulations on benchmark datasets. The simulation results show that the proposed WSLM performs superior to BPNN, ABC-BP, Bat-BP, and ABC-LM algorithms in terms of convergence speed and local minima avoidance.

Keywords Weight optimization · Hybrid neural networks · Wolf search · Levenberg-Marquardt · Back propagation

1 Introduction

Back-propagation is an optimization algorithm used to train the Artificial Neural Networks (ANN) and to speed-up its convergence to global minima during the learning process [1, 2]. Back-propagation Neural Network (BPNN) follows the ANN's basic principles, which imitates human's learning abilities. Unlike ANN,

M. Z. Rehman (✉) · A. Tanveer · K. Z. Zamli
Faculty of Computing, Universiti Malaysia Pahang (UMP), 26300 Gambang, Malaysia
e-mail: zrehman862060@gmail.com

N. M. Naw
Soft Computing and Data Mining Centre (SMC),
Universiti Tun Hussein Onn Malaysia (UTHM), 86300 Parit Raja, Malaysia

© Springer Nature Singapore Pte Ltd. 2020
M. A. Mohd Razman et al. (eds.), *Embracing Industry 4.0*,
Lecture Notes in Electrical Engineering 678,
https://doi.org/10.1007/978-981-15-6025-5_16

BPNN has a fully connected architecture in which a node in one layer is connected to all other node in the neighboring layer. In comparison to other ANN structures, BPNN learns to identify the errors in the hidden layers by measuring the errors from the output layer. It makes BPNN highly applicable for problems, where there is no relationship established between an output and an input. It has been successfully applied in a wide range of applications due to its high level of elasticity and learning abilities [3]. However, BPNN still needs to carefully select the initial constraints such as network size, number of layers, training weights, values for learning rate, momentum, and gain parameters, and activation functions etc. [4–7]. An improper use of these constraints can lead to slow network convergence or network failure.

In earlier days, only parametric level improvements were performed in BPNN such as; momentum, learning rate and gain parameter's enhancement etc. [8]. After that BPNN network architecture was enhanced with the second order Levenberg Marquardt (LM) algorithm BPNN derivative is utilized besides Gauss Newton (GN) [9]. Despite all these advancements, deterministic algorithms face a serious dilemma of poor approximation of weights that causes exhaustive search and leads them towards less optimal solutions or even network failure [5]. On the other hand, nature inspired stochastic algorithms perform much better due to the information sharing mechanism and the avoidance of exhaustive search [10].

Recently, hybrid combinations of different algorithms have become quite popular in the computer sciences field. With hybridization, the limitation of one algorithm can be addressed with the strengths of its counterparts. Owing to this alluring process, the hybridization process has laid some foundations for the emergence of Hybrid Metaheuristic algorithms or simply Hybrid Artificial Neural Networks (HANN). In HANN, usually the Swarm intelligent algorithms are used to approximate the initial weights with high precision instead of using the random ones [11–13]. In 2007, Karaboga laid the foundations for ABC to be used for BPNN weight training. The resultant hybrid ABC-BP algorithm performed with high accuracy on classification problems [14]. Similarly, Zhang et al. (2007) proposed a hybrid duo of PSO with BPNN to solve the slow convergence in PSO [15]. The proposed PSO-BP algorithm adopted a heuristic way to transit from particle swarm to gradient descend search. The results show that the proposed PSO-BP algorithm is better in convergence speed and accuracy than the adaptive PSO and BPNN algorithms. In the same way, Huang proposed a hybrid Genetic algorithm based BPNN (GA-BP) algorithm to improve the convergence performance of BPNN. Their research focused on the collection and re-evaluation of weight matrices in BPNN optimized by genetic algorithm. The results were quite beneficial in avoiding the over fitting drawback of BPNN [16]. In 2011, ABC was again used to train weights in LM and tested on the several benchmark problems with great success [17]. Mirjalili et al. (2012), proposed a hybrid of PSO and Gravitational Search Algorithm (GSA) to resolve the problem of slow convergence speed in the GSA [18]. The resulting PSO-GSA was used for training BPNN's and reducing its convergence to local minima and slow convergence. The simulation results illustrates the PSO-GSA's superior performance over both PSO and GSA for BPNN training and accurate avoidance of local minima.

Much later in 2013, Differential Evolution based back propagation (DE-BP) algorithm is proposed by [11] to train the weights in BPNN the feed-forward neural (FNN) network. In the proposed DE-BP algorithm, initially the DE is used to move towards global optima in the search space and then BP is used to get the optimal solution efficiently in the gradient. The performance of the proposed DE-BP was better than the GA-BP algorithm. More recently, Cuckoo search (CS) has been found to be quite beneficial in training weights in BP [19]. The proposed CS-BP is compared with the basic BPNN and General Regression Neural Network (GRNN) to study the prediction performance on Wine dataset. Lately, CS is also hybridized with LM and ERNN algorithms. The CSLM and CSRNN algorithms have performed well and successfully converged on average size datasets [20, 21]. From the above review on HANN algorithms, it can be easily realized that a proper hybrid combination of two or more search algorithms can greatly improve the weights for training the network.

Despite the availability of vast majority of hybrid algorithms available, the problem of local minima in LMBP still exists [20]. In addition, the no-free-lunch theorem (NFL) states that there is no superior optimization algorithm present for all optimization problems [22, 23]. Thus motivated by NFL, a new wolf search based algorithm is proposed for reconditioning weights in LMBP algorithm. In this paper, the proposed WSLM algorithm is used to evaluate & test on seven popular classification datasets taken from University California Irvine (UCI) repository. The results are compared with other hybrid LM algorithms from the literature and the basic back propagation algorithm.

The paper is organized as follows: A brief introduction to LMBP algorithm is given in Sect. 2. Section 3 presents the Wolf Search algorithm along with the details of the proposed WSLM algorithm. The evaluations on statistical results are made in Sect. 4 and finally the paper is concluded in Sect. 5.

2 The Levenberg-Marquardt (LM) Algorithm

Although BPNN algorithm has its benefits but it also has a slow convergence problem that can be improved with the help of Gauss-Newton (GN) method [24]. The GN utilizes appropriate steps in all directions to converge to global minima. If the error function approaches a quadratic surface during search for optimum, it simply converges. But this phenomena can only occur when the quadratic estimate of error surface is accurate, otherwise, the GN would be characteristically different [25].

Therefore, an intermediate algorithm is implemented using both gradient descent and GN. The Levenberg-Marquardt (LM) algorithm is more reliable than the GN model because it can converge in many cases even if the error surface is more complex [26, 27]. The basic principle of the LM is that it alters its state to the steepest descent until the local curvature is appropriate to make a quadratic estimation with GN algorithm to speed up the convergence [28]. LM uses Hessian matrix for approximation of error surface. The error function is given as:

$$E(t) = \frac{1}{2} \sum_{i=1}^N e_i^2(t) \quad (1)$$

Here, $e(t)$ is the error and the number of vector elements are denoted with N then the gradient descent is calculated in Eq. (2):

$$\nabla E(t) = J^T(t)e(t) \quad (2)$$

$$\nabla^2 E(t) = J^T(t)J(t) \quad (3)$$

Meanwhile, the gradient descent is denoted with $\nabla E(t)$, Hessian Matrix with $\nabla^2 E(t)$ and Jacobain Matrix with $J(t)$. The Jacobain Matrix is calculated in Eq. (4), GN method in Eq. (5) and LM in Eq. (6):

$$J(t) = \begin{bmatrix} \frac{\partial v_1(t)}{\partial t_1} & \frac{\partial v_1(t)}{\partial t_2} & \cdots & \frac{\partial v_1(t)}{\partial t_n} \\ \frac{\partial v_2(t)}{\partial t_1} & \frac{\partial v_2(t)}{\partial t_2} & \cdots & \frac{\partial v_2(t)}{\partial t_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial v_n(t)}{\partial t_1} & \frac{\partial v_n(t)}{\partial t_2} & \cdots & \frac{\partial v_n(t)}{\partial t_n} \end{bmatrix} \quad (4)$$

$$\nabla w = -[J^T(t)J(t)]^{-1}J(t)e(t) \quad (5)$$

$$w(k+1) = w(k) - [J^T(t)J(t) + \mu I]^{-1}J(t)e(t) \quad (6)$$

Here, μ is a constant and I is identity matrix. The algorithm will approach the GN, which ought to deliver rapid convergence to global minima. In addition, it should be noted that when the parameter λ is large, Eq. (6) uses gradient descent otherwise for a small λ , the algorithm acts like GN method.

3 The Wolf Search (WS) Algorithm

Proposed by Xin. She Yang, Wolf Search (WS) algorithm is a swarm intelligent metaheuristic that mimics the searching ability of the wolves in their natural habitat and the effective avoidance of enemies in case of danger [29]. The WS is used in this study to initialize the weights for LM algorithm effectively. The process is explained in the following section.

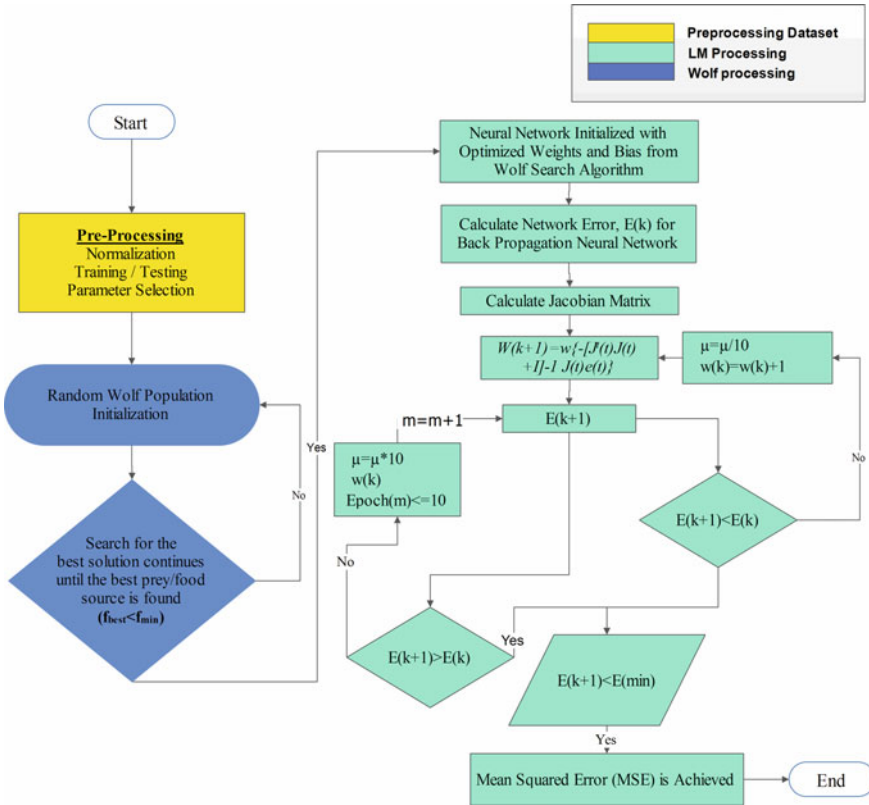


Fig. 1 The proposed WSLM algorithm

3.1 The Proposed Wolf Search Levenberg-Marquardt (WSLM)

In the proposed Wolf Search Levenberg-Marquardt (WSLM) algorithm, each wolf represents a possible solution (i.e. weights and biases for LM). The problem of optimizing weight and pack-size shows the consistency of the solution. In the first epoch, WS initializes the optimized weights and then transfers them to the LM. Then LM uses weight initialized by WS and converge to global minima. The proposed WSLM algorithm is illustrated in Fig. 1.

4 Results and Discussions

In this segment, the proposed WSLM performance is tested on benchmark classification datasets. The system used for simulation comes equipped with an Intel Core i5

processor with 8 GB RAM. The proposed WSLM algorithm is implemented using MATLAB 7.14 on Windows 10 Home Edition. The performance of the proposed WSLM is compared with the simple Back-propagation [1], ABC-BP [17], Bat-BP [30], and ABC-LM [14] algorithms. The classification performance of the all the algorithms are tested on seven classification problems taken from UCI repository. Jack-Knife resampling is used to divide datasets into training and testing samples. The dataset features and their neural network structures are given in Table 1.

The performance measures set for evaluating an algorithm's performance are Mean squared error (MSE), standard deviation (SD), Epochs, and CPU time. All the algorithms are initialized with weights from their respective optimization algorithms. For all networks, only the inputs and output neurons are varied according to the dataset given while the hidden layers are kept fixed to five neurons. The maximum iterations for all the algorithms are set to 1000 and 20 trials are run on each dataset.

During Simulation process, the proposed WSLM algorithm is compared with other algorithms in terms of MSE convergence, CPU time and maximum epochs. Table 2 shows the statistical results of the proposed WSLM, BPNN, ABC-BP, Bat-BP, and ABC-LM on the given datasets.

The proposed WSLM algorithm is outperforming BPNN, ABC-BP, Bat-BP, and ABC-LM algorithms in terms of small average MSE of 6.42×10^{-05} , 5.17×10^{-05} , 0.0201, 9.7×10^{-07} , 7.0×10^{-06} , 0.012, 5.4×10^{-06} . A smaller value of CPU time, less no of epochs, and SD by WSLM for all datasets show that the algorithm is quite stable during convergence and avoids local minima. For thyroid and Australian credit card problems, MSE for WSLM algorithm can be improved, if provided with more epoch size. The original BPNN failed to converge to global minima on most datasets, which shows that initial weights must be selected carefully to ensure global convergence. The convergence MSE, Epochs and CPU time are given in Fig. 2. The Figure shows that WSLM is quite a reliable algorithm in terms of CPU resources and consumes less number of epochs during convergence to global minima.

Overall, the proposed WSLM algorithm has integrated the enemy avoidance strategy of the wolf search algorithm and converged to global minima successfully. Due to its enhanced exploration capability coupled with Brownian motion, it is able

Table 1 Classification datasets and neural network structure

Dataset(s)	#Features	#Classes	Train samples	Test samples	ANN
Breast Cancer [31]	9	2	489	210	9-5-2
Iris [32]	4	3	105	45	4-5-3
Thyroid [33]	21	3	5040	2160	21-5-3
Diabetes [34]	8	2	537	231	8-5-2
Glass Identification [35]	10	7	149	65	10-5-7
Australian Credit Approval [36]	14	2	483	207	14-5-2
Seven Bit Parity	7	2	89	39	7-5-2
Breast Cancer [31]	9	2	489	210	9-5-2

Table 2 Average CPU time, epochs, MSE and accuracy for all datasets

Dataset/Algorithm		BPNN	ABC-BP	Bat-BP	ABC-LM	WSLM
Breast Cancer	CPU time	95.46	1482.90	345.12	1880.65	52.11
	Epochs	1000	1000	1000	1000	14
	MSE	0.271	0.271	0.0219	0.014	6.42×10^{-05}
	SD	0.01	0.01	0.0045	0.0011	2.10×10^{-05}
Iris	CPU time	28.47	156.43	475.38	171.52	19.13
	Epochs	1000	1000	1000	1000	14
	MSE	0.312	0.155	0.0194	0.058	5.17×10^{-05}
	SD	0.022	0.023	0.0036	0.005	2.70×10^{-05}
Thyroid	CPU time	1630.12	1747.23	4610.79	1382.91	36,448.77
	Epochs	1000	1000	1000	1000	1000
	MSE	0.311	0.046	0.0065	0.041	0.0201
	SD	0.033	0.0006	0.0054	0.0007	0.0033
Diabetes	CPU time	57.05	4257	532.17	2805.09	1949.87
	Epochs	1000	1000	1000	1000	538
	MSE	0.27	0.201	0.0093	0.141	9.7×10^{-07}
	SD	0.026	0.002	0.0091	0.033	5.9×10^{-06}
Glass	CPU time	32.73	1717.90	268.60	1336.19	91.79
	Epochs	1000	1000	1000	1000	14
	MSE	0.364	0.0258	4.80E-03	0.005	7.0×10^{-06}
	SD	0.048	0.009	2.30E-02	0.003	2.4×10^{-06}
Australian Credit-Card	CPU time	24.42	6894	892.31	4213.01	34,699.68
	Epochs	1000	1000	967	1000	1000
	MSE	0.271	0.173	0.0033	0.055	0.012
	SD	0.015	0.043	0	0.0058	0.0018
Seven Bit Parity	CPU Time	142.07	183.39	155.06	69.13	94.63
	Epochs	1000	1000	1000	1000	19
	MSE	0.26	0.12	0.0222	0.083	5.4×10^{-06}
	SD	0.014	0.008	0.0076	0.012	2.4×10^{-06}

to avoid local minima most of the time. In addition, the other evolutionary algorithms were able to perform well because of their intrinsic exploration capabilities but sometimes they suffered due to poor initial population selection.

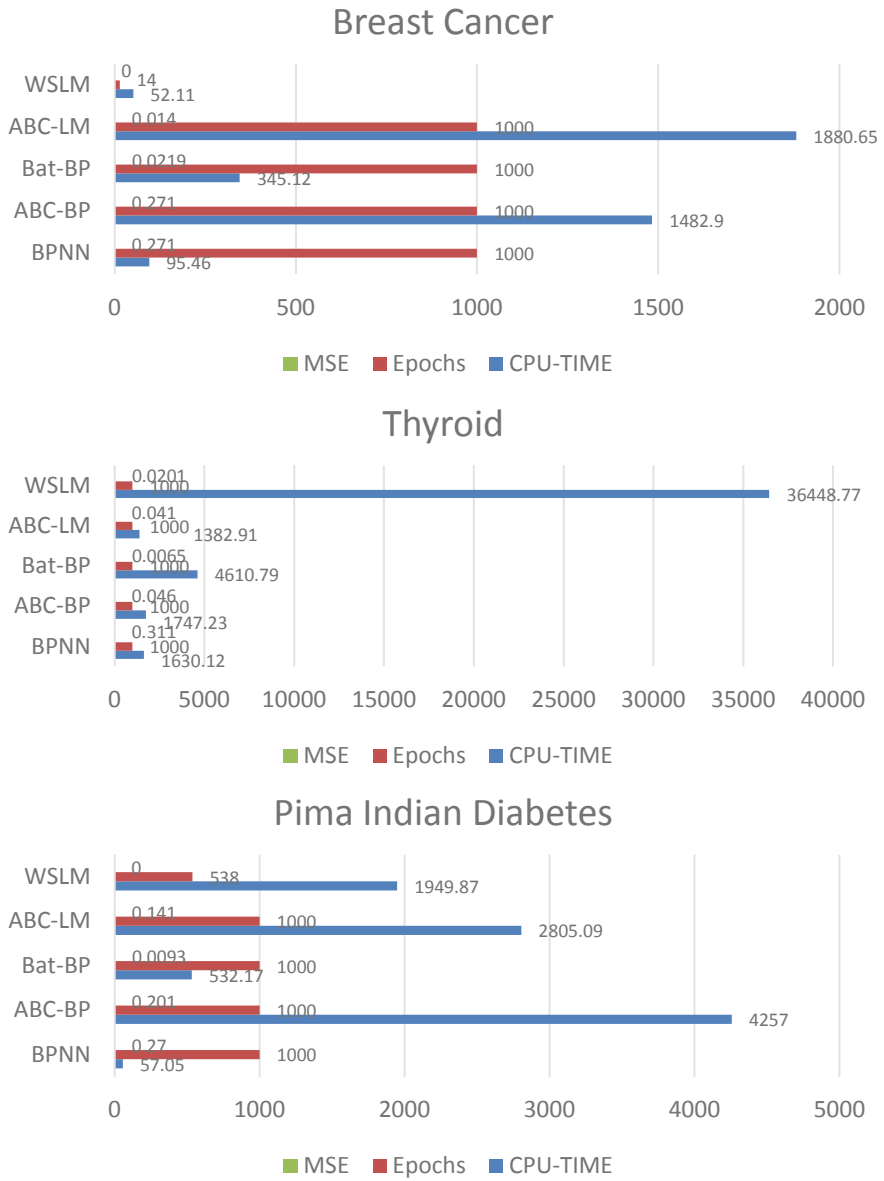


Fig. 2 MSE, CPU time, and epochs for the proposed WSLM, BPNN, ABC-BP, Bat-BP, ABC-LM

5 Conclusions and Future Works

To improve the slow convergence and network stagnancy in Levenberg-Marquardt Back Propagation (LMBP) algorithm, this paper proposed a swarm optimized improvement with Wolf Search algorithm. The resultant Wolf Search Levenberg-Marquardt (WSLM) objective was to improve Mean Squared Error (MSE) and overall convergence through the introduction of optimized weights with wolf search. Datasets of varying complexities were selected to benchmark the performance of the proposed WSLM algorithm against conventional BPNN, and other hybrid swarm intelligent first order and second order neural networks. The results proved that the proposed WSLM converged to global minima on majority of the datasets with small MSE. WSLM also proved that it is able to train and generalize well and it is highly reliable than BPNN, ABC-BP, Bat-BP, and ABC-LM. However, there is still more work required to increase the classification and prediction accuracy in the proposed WSLM algorithm. In the future, wolf search will be applied to enhance the weight vectors in recurrent neural network models for improved training time.

Acknowledgements This work is funded by FRGS Grant from the Ministry of Higher Education Malaysia titled: A Reinforcement Learning Sine Cosine based Strategy for Combinatorial Test Suite Generation (grant no: RDU170103).

References

1. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Parallel distributed processing: explorations in the microstructure of cognition. pp 318–362
2. Wang G, Guo L (2013) A novel hybrid bat algorithm with harmony search for global numerical optimization. *J Appl Math* 2013. <https://doi.org/10.1155/2013/125625>
3. Nawi NM, Khan A, Rehman MZ, Aziz MA, Abawajy JH, Herawan T (2014) An accelerated particle swarm optimization based Levenberg Marquardt back propagation algorithm
4. Nawi NM, Ransing RS, Ransing MR (2008) A new method to improve the gradient based search direction to enhance the computational efficiency of back propagation based Neural Network algorithms. In: Proceedings—2nd Asia international conference on modelling and simulation, AMS 2008, pp 546–552. <https://doi.org/10.1109/AMS.2008.70>
5. Kolen JF, Pollack JB (1990) Back propagation is sensitive to initial conditions. *Complex Syst* 4:269–280
6. Lahmiri S (2011) A comparative study of back propagation algorithms in financial prediction. *Int J Comput Sci Eng Appl* 1
7. Zhang L, Pu J (2011) An improved back propagation neural network in objects recognition. In: IEEE international conference on automation and logistics, ICAL. pp 507–511. <https://doi.org/10.1109/ICAL.2011.6024772>
8. Nawi NM, Ransing MR, Ransing RS (2006) An improved learning algorithm based on the Broyden-Fletcher-GoldfarbShanno (BFGS) method for back propagation neural networks. In: Proceedings—ISDA 2006: sixth international conference on intelligent systems design and applications, pp 152–157. <https://doi.org/10.1109/ISDA.2006.95>
9. Yu H, Bogdan M (2010) Levenberg–Marquardt training. *The Industrial Electronics Handbook* 5

10. Yang XS (2010) A new metaheuristic bat-inspired algorithm. In: *Studies in computational intelligence*. pp 65–74. https://doi.org/10.1007/978-3-642-12538-6_6
11. Sarangi P, Sahu A, Panda M (2013) A hybrid differential evolution and back-propagation algorithm for feedforward neural network training. *Int J Comput Appl* 84:1–9
12. Zamli KZ, Din F, Ahmed BS, Bures M (2018) A hybrid Q-learning sine-cosine-based strategy for addressing the combinatorial test suite minimization problem. *PLoS ONE* 13:1–29. <https://doi.org/10.1371/journal.pone.0195675>
13. Alamri HS, Zamli KZ (2019) PMT: opposition-based learning technique for enhancing metaheuristic performance. *IEEE Access* 7:97653–97672. <https://doi.org/10.1109/access.2019.2925088>
14. Karaboga D, Akay B, Ozturk C (2007) Artificial Bee Colony (ABC) optimization algorithm for training feed-forward neural networks. In: *Modeling decisions for artificial intelligence*. Springer, Berlin Heidelberg, pp 318–329
15. Zhang J, Lok T, Lyu MR (2007) A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training. *Appl Math Comput* 185:1026–1037. <https://doi.org/10.1016/j.amc.2006.07.025>
16. Huang CY, Chen LH, Chen YL, Chang FM (2009) Evaluating the process of a genetic algorithm to improve the back-propagation network: a Monte Carlo study. *Expert Syst Appl* 36:1459–1465. <https://doi.org/10.1016/j.eswa.2007.11.055>
17. Ozturk C, Karaboga D (2011) Hybrid Artificial Bee Colony algorithm for neural network training. 2011 IEEE Congr Evol Comput 84–88. <https://doi.org/10.1109/CEC.2011.5949602>
18. Mirjalili S, Mohd Hashim SZ, Moradian Sardroudi H (2012) Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. *Appl Math Comput* 218:11125–11137. <https://doi.org/10.1016/j.amc.2012.04.069>
19. Yi JH, Xu WH, Chen YT (2014) Novel back propagation optimization by cuckoo search algorithm. *Sci World J* 2014. <https://doi.org/10.1155/2014/87862>
20. Nawi NM, Khan A, Rehman MZ (2013) A new cuckoo search based Levenberg-Marquardt (CSLM) algorithm. In: Murgante B, Misra S, Carlini M, Torre CM, Nguyen H-Q, Taniar D, Apduhan BO, Gervasi O (eds) *Computational science and its applications—ICCSA 2013*. Springer, Berlin, Heidelberg, pp 438–451
21. Nawi NM, Khan A, Rehman MZ, Herawan T, Deris MM (2014) CSLMEN: a new cuckoo search Levenberg Marquardt Elman Network for data classification. https://doi.org/10.1007/978-3-319-07692-8_17
22. Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. *IEEE Trans Evol Comput*. <https://doi.org/10.1109/4235.585893>
23. Aljarah I, Faris H, Mirjalili S (2018) Optimizing connection weights in neural networks using the whale optimization algorithm. *Soft Comput* 22:1–15. <https://doi.org/10.1007/s00500-016-2442-1>
24. Nawi NM, Khan A, Rehman MZ, Herawan T, Deris MM (2014) Comparing performances of cuckoo search based neural networks. https://doi.org/10.1007/978-3-319-07692-8_16
25. Rehman MZ, Nawi NM (2011) The effect of adaptive momentum in improving the accuracy of gradient descent back propagation algorithm on classification problems. *Softw Eng Comput Syst* 179:380–390. <https://doi.org/10.1007/978-3-642-22170-5>
26. Levenberg K (1944) A method for the solution of certain problems in least squares. *Quart Appl Math* 2:164–168. <https://doi.org/citeulike-article-id:2946351>
27. Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. <https://doi.org/10.1137/0111030>
28. Yu H, Wilamowski BM (2012) Neural network training with second order algorithms. *Hum-Comput Syst Interact Backgr Appl* 2:463–476
29. Rui T, Simon F, Yang XS, Suash D (2012) Wolf search algorithm with ephemeral memory. In: *Seventh international conference on digital information management (ICDIM 2012)*. pp 165–172 (2012)
30. Nawi NM, Khan A, Rehman MZ (2013) The effect of bat population in Bat-BP algorithm. In: *The 8th international conference on robotic, vision, signal processing & power applications*. Springer Singapore, Penang, Malaysia, pp 295–302

31. Wolberg WH, Mangasarian OL (1990) Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proc Natl Acad Sci USA* 87:9193–9196. <https://doi.org/10.1073/pnas.87.23.9193>
32. Fisher R (1936) The use of multiple measurements in taxonomic problems. *Ann Eugen* 7:179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>
33. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1:81–106. <https://doi.org/10.1007/BF00116251>
34. Smith JW, Everhart JE, Dickson WC, Knowler WC, Johannes RS (1988) Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. *Proc Annu Symp Comput Appl Med Care* 261–265
35. Evett IW, Spiehler EJ (1987) Rule induction in forensic science. *KBS Government*. 107–118
36. Quinlan JR, Compton PJ, Horn KA, Lazurus L (1986) Inductive knowledge acquisition: a case study. In: *The second Australian conference on applications of expert systems*

The Classification of Skateboarding Trick Manoeuvres: A Frequency-Domain Evaluation



Muhammad Ar Rahim Ibrahim, Muhammad Nur Aiman Shapiee,
Muhammad Amirul Abdullah, Mohd Azraai Mohd Razman,
Rabiu Muazu Musa, and Anwar P. P. Abdul Majeed

Abstract The growing interest in skateboarding as a competitive sport requires new motion analysis approaches and innovative ways to portray athletes' results as previous techniques in the identification of the tricks was often inadequate in providing accurate evaluation during competition. Therefore, there is a need to introduce an unprejudiced method of evaluation in skateboarding competitions. This paper presents the classification of five different skateboarding tricks (Ollie, Kickflip, Frontside 180, Pop Shove-it, and Nollie Frontside Shove-it) through the identification of significant frequency-domain signals collected via Inertial Measurement Unit (IMU) and the use of machine learning models. One male skateboarder (age: 23 years old) performed five different tricks repeatedly for several times. The time-domain data acquired from the IMU were converted to frequency-domain by employing Fast Fourier Transform (FFT) and a number of statistical features (mean, kurtosis, skewness, standard deviation, root mean square and peak-to-peak corresponding to x-y-z-axis of the IMU) were then extracted. Significant features were then identified from the Information Gain (IG) scoring. It was shown from the study that the Naïve Bayes (NB) classifier is able to acquire the highest classification accuracy of 100% on the test data compared to the other evaluated classifiers, namely Artificial Neural Network (ANN) and Support Vector Machine (SVM), by utilising the selected features, suggesting that the proposed methodology could provide an objective-based evaluation of the tricks.

M. A. R. Ibrahim (✉) · M. N. A. Shapiee · M. A. Abdullah · M. A. M. Razman ·
A. P. P. Abdul Majeed
Innovative Manufacturing, Mechatronics and Sports Laboratory,
Faculty of Manufacturing and Mechatronic Engineering Technology,
Universiti Malaysia Pahang, 26600 Pekan, Pahang Darul Makmur, Malaysia
e-mail: rahim2g@gmail.com

M. A. M. Razman
e-mail: azraai@ump.edu.my

R. M. Musa
Department of Credited Co-Curriculum, Centre for Fundamental and Continuing Education,
Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia

Keywords IMU · Frequency-domain · Fast fourier transform · Feature selection · Machine learning · Skateboarding tricks · Classification

1 Introduction

It has been reported that the skateboarding industry is worth approximately USD 4.8 billion [1]. It is also worth noting that in 2016, skateboarding will make its debut at the 2020 Summer Olympic Games in Tokyo. Nevertheless, it is important to keep in mind that the evaluation of the skateboarding tricks has always been carried out subjectively by the judges based on their previous experiences that may lead to bias and even inaccurate evaluation.

To date, there is limited literature with regards to the classification of skateboarding. Groh et al. [2] recruited seven skateboarders (age: 25 ± 4 , stance: 4 goofy, 3 regular, gender: male) to classify six tricks (Ollie, Nollie, Kickflip, Heelflip, Pop Shove-it, 360-Flip) and employed four different machine learning models namely, Naïve Bayes (NB), Partial Decision Tree (PART), Support Vector Machine (SVM) and k -nearest neighbour (k -NN) through the use of IMU sensors and motion capture system. Several time-domain features were extracted, i.e., mean, variance, skewness, kurtosis, dominant frequency, bandwidth, and the correlation between x-y-axis, x-z-axis and y-z-axis. The feature selection was performed by using the Embedded Classification Software Toolbox (ECST) in Weka software. Only the significant features out of the 54 features extracted were employed then used for the classification of the tricks. It was shown from the study that both NB and SVM models achieved a high classification accuracy of 97.8%.

In an extended study, Groh et al. [3] attempted in classifying 11 tricks (Ollie, Nollie, Kickflip, Heelflip, Pop Shove-it BS, Pop Shove-it FS, 360-Shove-it BS, Vari-alflip, Hardflip, Double-Kickflip and 360-flip), one rest and one bail through the utilisation of different machine learning algorithms. Eleven (11) skateboarders (age: 23 ± 4 years, height: 179 ± 5 cm, stance: 5 goofy; 6 regular) participated in the study. The research demonstrated that the Radial Basis Function (RBF) based-SVM model could classify the tricks with a classification accuracy of 89.1% against other models evaluated.

Nonetheless, it is worth pointing out that the use of machine learning models with time and frequency domain features have been utilised in other fields and has been well documented [4–12]. To illustrate a few, Ashqar et al. [13] investigated the efficacy of different classifiers in the identification of five different transportation modes (driving a passenger car, bicycling, taking a bus, running and walking) through data collected via smartphone. Ten travellers participated in the study. A total of 345 features were extracted from both time and frequency domain data, namely, mean, max, min, variance, standard deviation, range, interquartile range, signChange, energy, spectralEntropy, and also derivative of the mean, max, min, variance, standard deviation, range, interquartile range and signChange, respectively. The authors utilised Random Forest (RF) as a feature selection method (often known

as Embedded Method). The significant features identified via the RF was then classified with an SVM classifier. It was shown that through the proposed approach, a classification accuracy to up to 97.02% could be achieved.

An RF classifier was used to classify six diverse activities (sitting, lying, standing, reclining, walking and running) through the integration of two accelerometers [14]. A total of 42 children (age: 7 ± 8 years, gender: male and female) and 33 adults (age 42 ± 11 years, both gender) took part in the study. Different time- and frequency-domain data features were extracted, viz. mean, standard deviation, coefficient of variation, median, 25th and 75th percentiles, minimum, maximum, skewness, kurtosis, axis correlations (between-axis and between-sensor), roll, and pitch, yaw, the magnitude of frequency and total signal power. A feature selection was performed through the use of the Gini index to identify the significant features. The Leave-One-Out cross-validation (LOOCV) technique was utilised in the study. The investigation illustrated that through the selected features, the RF model could achieve a classification accuracy (CA) of 99.1% on the adult whilst 97.3% on the child sample.

This paper intends to evaluate different machine learning models in classifying the following skateboarding tricks, i.e., Ollie, Kickflip, Pop Shove-it, Nollie Frontside Shove-it and Frontside 180 through a number of significant statistical-based frequency-domain features selected via information gain score. The remaining of the paper will be demarcated as follows; Sect. 2 shall discuss the methodology undertaken for the study. Section 3 shall deliberate on the results obtained from the study, whilst Sect. 4 shall draw the conclusion derived from the study.

2 Methodology

2.1 Instrumented IMU Device

Figure 1 depicts the rendered assembly view of the instrumented device modelled using CATIA. The casing of the device is printed using a fuse deposition modelling technique via Zortrax M200 Plus. The material chosen for the printing of the instrumented IMU device is Acrylonitrile Butadiene Styrene (ABS) due to its favorable mechanical properties, high impact resistance and good shock absorption as the device is vulnerable to impact from the tricks performed. The instrumented device is well equipped with a microcontroller (Arduino Pro Mini), an inertial measurement unit sensor (MPU6050), a Bluetooth Module (HC-06) and a 3.7 V Lithium Polymer battery. The signals obtained from the IMU are acceleration (m/s^2) as well as the angular velocity ($^\circ/\text{s}$) with a sampling time of 50 milliseconds.

Figure 2 illustrates the location of the instrumented IMU unit positioned on the skateboard. It is placed at the bottom front of the skateboard (Nose), and behind the front truck. As the device is built alongside with a riser pad, it is easy to mount the device on the deck using the existing fastener to ensure its stability. The choice of the device's location is non-trivial as it does not impede the skateboarders' movement

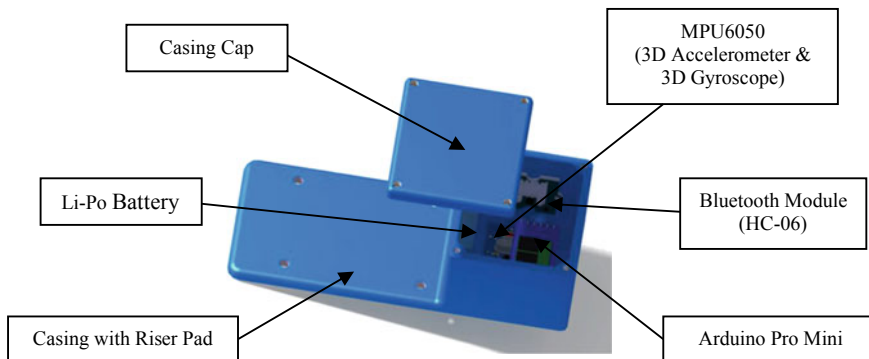


Fig. 1 Instrumented IMU device



Fig. 2 The instrumented IMU attached at bottom front of the board, **a** Attachment of the IMU to the skateboard, **b** the 3D printed IMU device

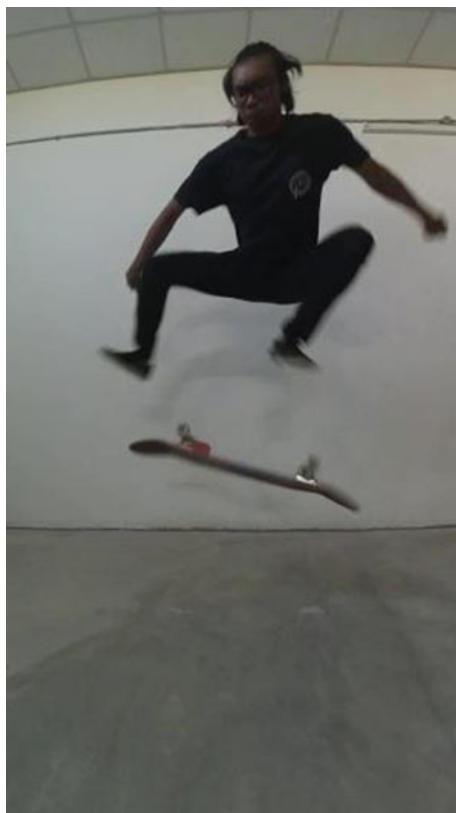
when executing a particular trick. In fact, the location of the device reduces the risk of damage to the system throughout the process of data collection.

2.2 Data Collection

A 23 years old skater from University Malaysia of Pahang with a height of 170 cm and a weight of 54 kg was recruited to perform five diverse tricks (as shown in Table 1) that is repeated five (5) times for each trick. The tricks performed by the skateboarder were selected based on their experience. The stance direction used by the skateboarder was goofy. An illustration of one of the tricks as well as its corresponding signals (acceleration and angular velocities) are shown in Figs. 3, 4 and 5, respectively.

Table 1 List of the skateboarding tricks performed

Name	Rotation (angle and axis)
Ollie (O)	Board incline about the x-axis (Approximately $45^\circ + y$)
Nollie FS Shuvit (NFS)	Spin incline about the z-axis (Approximately $180^\circ - z$)
Frontside 180° (FS180)	Clockwise about z-axis ($180^\circ - z$)
Pop Shove-it (PS)	Spin clockwise about z-axis ($180^\circ + z$)
Kickflip (K)	Rotate clockwise about y-axis ($360^\circ + y$)

**Fig. 3** A Kickflip (K) trick performed

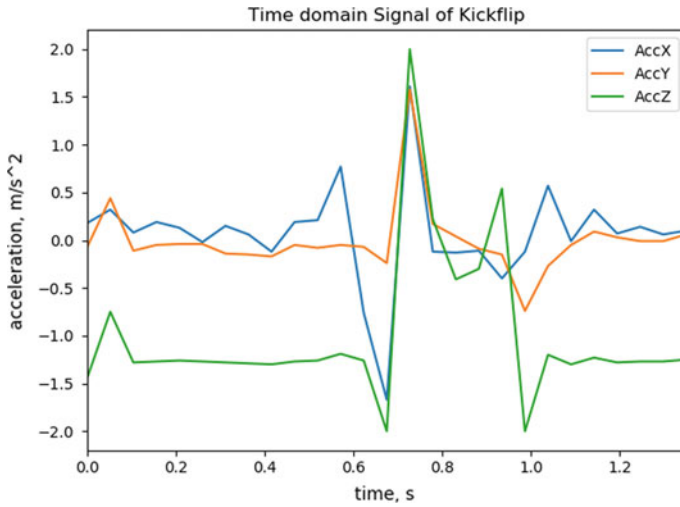


Fig. 4 Acceleration signals for Kickflip

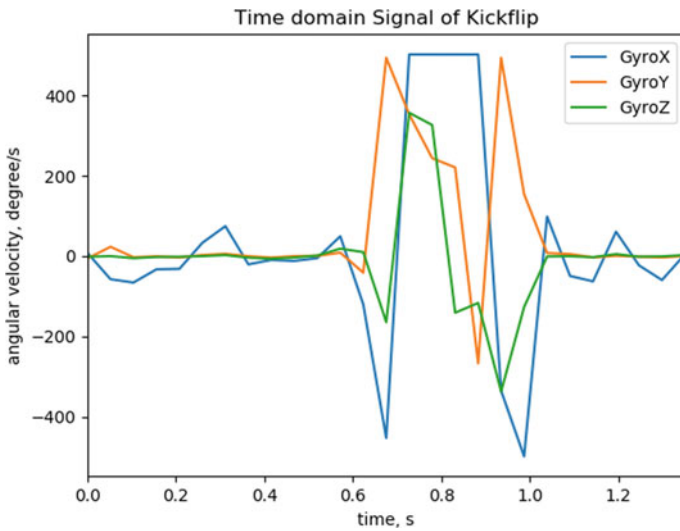


Fig. 5 Gyro signals for Kickflip

2.3 Data Transformation

The collected data from the IMU sensor are inherently in the time-domain. Fast Fourier Transform (FFT) was applied to transform time-domain signals to frequency-domain. In the frequency response, we the magnitude of the amplitudes are of interest. The Nyquist-Shannon Sampling Theorem is employed to scrutinise the frequency

Fig. 6 Converted acceleration signal for Kickflip

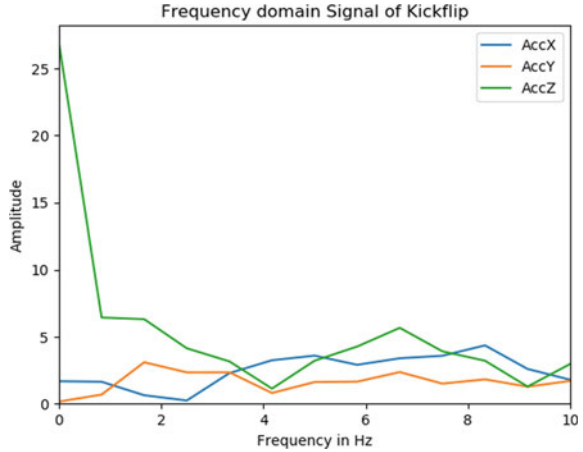
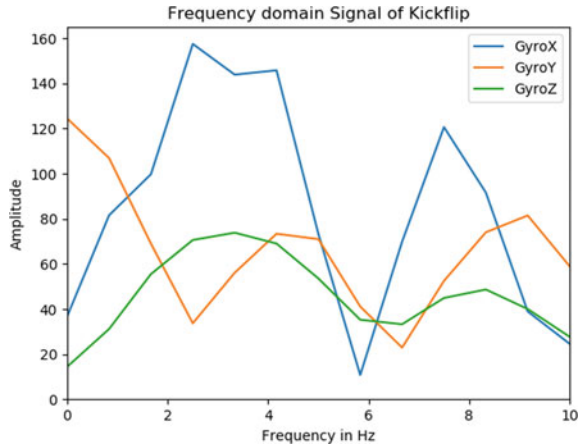


Fig. 7 Converted gyro signals for Kickflip



signals, and hence, 10 Hz is selected for each window. Figures 6 and 7 illustrate the frequency domain transformation for the corresponding acceleration and angular velocity signals acquired for Kickflip.

2.4 Feature Extraction and Selection

Statistical-based features, namely, mean, skewness, kurtosis, peak to peak, root mean square and standard deviation are then extracted from the transformed data (frequency domain) for all readings (all six degrees of freedom) through MATLAB 2016b. Subsequently, a sensitivity analysis was carried out through the use of Information Gain

(IG) to identify the significant features via an open source platform, Orange V3.11. IG is an entropy-based feature evaluation method that measures the dependence of a given feature towards its target variable [15].

2.5 Classifiers

Several of classifiers were investigated to evaluate its efficacy in classifying the skateboarding tricks, i.e., Artificial Neural Network (ANN), Naïve Bayes (NB) and k -Nearest Neighbors (k -NN). It should be noted that in this preliminary investigation, the default settings (hyperparameters) from the Orange platform is employed. The classifiers were evaluated on the basis of their classification accuracy (CA), area under the curve (AUC), precision, sensitivity, recall, accuracy as well as the F1 score derived from the confusion matrix. In the present study, leave-one-out cross-validation (LOOCV) technique was employed.

3 Results and Discussion

The skateboarder carried out a total of 40 trick events and the success rate of the landings was documented. Only 25 tricks were identified to be successful and the data obtained were used to create the machine learning models of machine learning based on the features extracted. The dataset was split randomly with a 70:30 ratio for the train and test dataset.

The efficacy of the classifiers were evaluated in two manners, namely, by considering all features and significant features that was identified via the IG feature selection method. Through the feature selection method, all except for standard deviation, mean, skewness, kurtosis and root mean square corresponding to y -axis of acceleration readings and the kurtosis of x - z -axis of the gyro readings were found to be significant. It was shown from the study that the NB and SVM models trained with all features demonstrated an overfitting trend, as the test CA was significantly lower than the train CA, whilst the ANN model does not exhibit appreciable CA for both train and test evaluations as shown in Fig. 8.

Conversely, through the selected features, the both the ANN and SVM models yielded an exact CA for both the train and test dataset, with a noteworthy zero misclassification on the test dataset. Nonetheless, the NB model exhibited a better CA against other evaluated models with regards to the train dataset, with a CA of 94.4% and in addition to a zero misclassification on the test dataset, suggesting that the NB model is the best classifier based on the identified significant features through the IG method. Table 2 lists the performance of the models evaluated based on different performance measures by considering all features on the test dataset. On the other hand, Table 3 lists the performance based on the selected features. It is

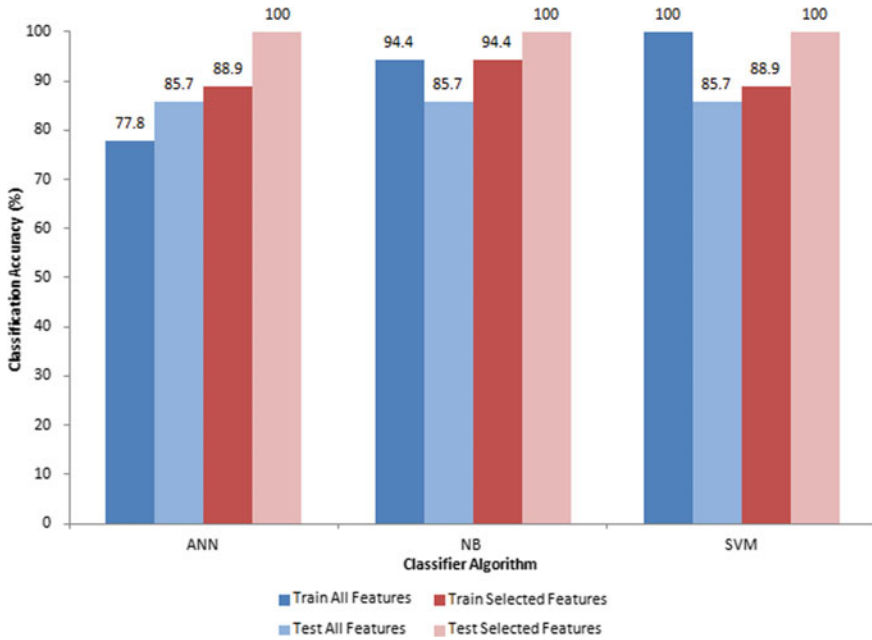


Fig. 8 Graph of comparison of performance between 3 classifiers

Table 2 Evaluation of the developed classifiers with all combined features

Method	AUC	CA	F1	Precision	Recall
ANN	0.974	0.857	0.857	0.929	0.857
NB	1.000	0.857	0.800	0.762	0.857
SVM	1.000	0.857	0.857	0.929	0.857

Table 3 Evaluation of the developed classifiers with features selected

Method	AUC	CA	F1	Precision	Recall
ANN	1.000	1.000	1.000	1.000	1.000
NB	1.000	1.000	1.000	1.000	1.000
SVM	1.000	1.000	1.000	1.000	1.000

evident that by identifying the significant features, excellent performance is exhibited across different performance measures for all evaluated models on the test dataset.

The confusion matrix depicted on Figs. 9, 10 and 11 further substantiates the findings remarked above. The misclassification is highlighted in red colour which is apparent across all trained models (see the (a) parentheses). It could be observed that the misclassification from the ANN and SVM models arises Pop Shove-it (PS) and Nollie Frontside Shove-it (NFS) tricks which was misclassified as Frontside 180

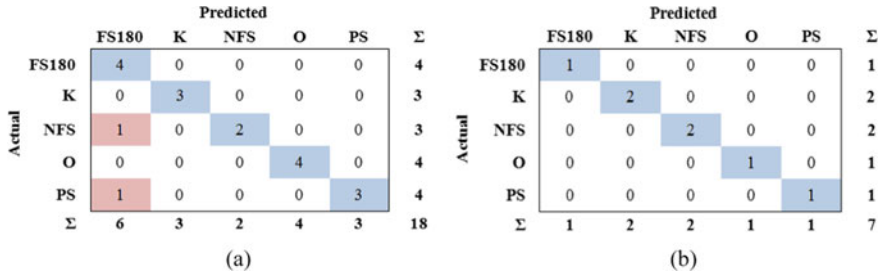


Fig. 9 Confusion matrix of the train and test of ANN model with selected features

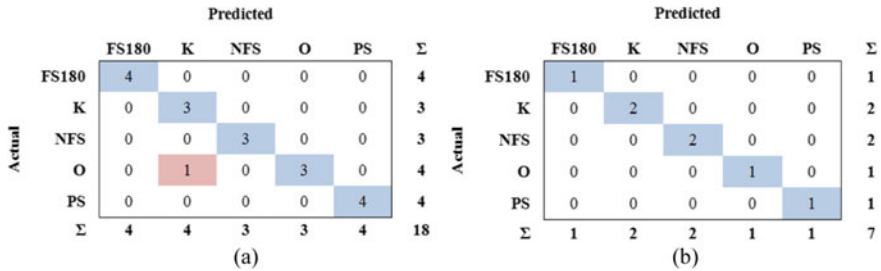


Fig. 10 Confusion matrix of the train and test of NB model with selected features

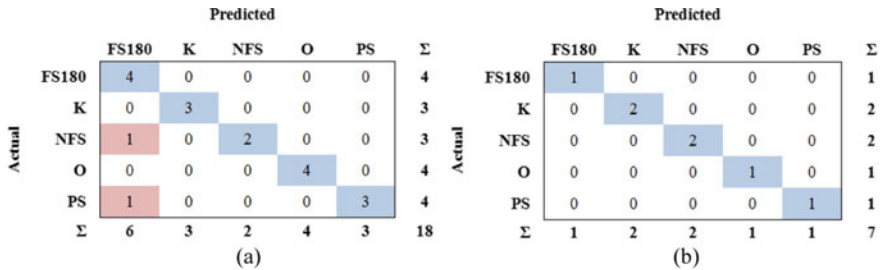


Fig. 11 Confusion matrix of the train and test of SVM model with selected features

(FS180). Conversely, the misclassification recorded by the trained NB model was the Ollie (O) which was misclassified as Kickflip (K). Moreover, from the test dataset (see Figs. 9b, 10b and 11b), all the models developed based on the identified features does not exhibit any missclassification (highlighted in blue). Therefore, from this preliminary investigation, it could be inferred that the combination of the selected features with various machine learning models can provide a fair classification accuracy of the skateboarding tricks performed.

4 Conclusion

In this preliminary investigation, an offline skateboarding tricks classification system was developed. It was shown from the investigation that the selection of the features is non-trivial in providing a reasonably accuracy classification of the evaluated skateboarding tricks. It was also demonstrated that with the selection of the features identified via the IG, could yield a CA of 100% on all the evaluated models (SVM, ANN and NB) on the test dataset. It is worth noting that the study is rather in its preliminary stage, and hence future study will be carried out by including more subjects, engineering different features, as well as performing hyperparameter optimisation on different machine learning models. The preliminary results further suggest the applicability of the proposed system in providing an objective based judgment on skateboarding tricks. This will assist the judges in providing a more accurate evaluation of trick performance as opposed to the subjective and conventional techniques currently applied in this sport.

Acknowledgements The authors would like to acknowledge the Ministry of Education, Malaysia and Universiti Malaysia Pahang for supporting and funding this research via FRGS/1/2019/TK03/UMP/02/6 (RDU1901115).

References

1. Abdullah MA, Ibrahim MAR, Shapiee MNA Bin, Mohd Razman MA, Musa RM, Abdul Majeed APP (2020) The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning. In: *Lecture Notes in Mechanical Engineering*. Springer, pp 67–74
2. Groh BH, Kautz T, Schuldhuis D (2015) IMU-based trick classification in skateboarding. *KDD Work Large-Scale Sport Anal*
3. Groh BH, Fleckenstein M, Kautz T, Eskofier BM (2017) Classification and visualization of skateboard tricks using wearable sensors. *Pervasive Mob Comput* 40:42–55
4. Anand A, Sharma M, Srivastava R, Kaligounder L, Prakash D (2018) Wearable motion sensor based analysis of swing sports. In: *Proceedings—16th IEEE international conference on machine learning and applications ICMLA 2017* 2018-January, pp 261–267. <https://doi.org/10.1109/ICMLA.2017.0-149>
5. Connaghan D, Kelly P, O'Connor NE, Gaffney M, Walsh M, O'Mathuna C (2011) Multi-sensor classification of tennis strokes. *Proceedings IEEE Sensors*, pp 1437–1440. <https://doi.org/10.1109/ICSENS.2011.6127084>
6. Corrêa NK, de Lima JCM, Russomano T, dos Santos MA, Corrêa NK, de Lima JCM, Russomano T, dos Santos MA (2017) Development of a skateboarding trick classifier using accelerometry and machine learning. *Res Biomed Eng* 33:362–369. <https://doi.org/10.1590/2446-4740.04717>
7. Friday H, Ying T, Mujtaba G, Al-garadi MA (2019) Data fusion and multiple classifier systems for human activity detection and health monitoring: review and open research directions 46:147–170. <https://doi.org/10.1016/j.inffus.2018.06.002>
8. Gani O, Fayezeen T, Povinelli RJ, Smith RO, Arif M, Kattan AJ, Iqbal S (2019) A light weight smartphone based human activity recognition system with high accuracy. *J Netw Comput Appl* 141:59–72. <https://doi.org/10.1016/j.jnca.2019.05.001>

9. Groh BH, Fleckenstein M, Eskofier BM (2016) Wearable trick classification in freestyle snowboarding. BSN 2016—13th annual body sensor network conference, pp 89–93. <https://doi.org/10.1109/BSN.2016.7516238>
10. Hollecsek T, Schoch J, Arnrich B, Tröster G (2010) Recognizing turns and other snowboarding activities with a gyroscope. In: Proceedings—international symposium wearable computer ISWC. <https://doi.org/10.1109/ISWC.2010.5665871>
11. Hölzemann A, Van Laerhoven K Using wrist-worn activity recognition for basketball game analysis. <https://doi.org/10.1145/3266157.3266217>
12. Iwase T, Nozaki Y, Yoshikawa M (2015) Detection technique for hardware Trojans using machine learning in frequency domain. 2015 IEEE 4th global conference on consumer electronics, pp 185–186. <https://doi.org/10.1109/GCCE.2015.7398569>
13. Ashqar HI, Almannaa MH, Elhenawy M, Rakha HA, House L (2019) Smartphone transportation mode recognition using a hierarchical machine learning classifier and pooled features from time and frequency domains. IEEE Trans Intell Transp Syst 20:244–252. <https://doi.org/10.1109/TITS.2018.2817658>
14. Stewart TOM, Narayanan A, Hedayatrad L, Neville J, Mackay L, Duncan S (2018) A dual-accelerometer system for classifying physical activity in children and adults, pp 2595–2602. <https://doi.org/10.1249/MSS.0000000000001717>
15. Lei S (2012) A feature selection method based on information gain and genetic algorithm. Int Conf Comput Sci Electron Eng 2:355–358. <https://doi.org/10.1109/ICCSEE.2012.97>

The Identification of RFID Signal Using k -Means for Pallet-Level Tagging



Chun Sern Choong, Ahmad Fakhri Ab. Nasir, Anwar P. P. Abdul Majeed, Muhammad Aizzat Zakaria, and Mohd Azraai Mohd Razman

Abstract Radio Frequency Identification (RFID) applications are becoming increasingly popular in a myriad of areas, and therefore, an effective RFID technology-based location would offer a much-needed additional in tracking system. This research focuses on the identification of the location of passive RFID at the pallet-level, which uses the RFID signal strength to cluster the pallet level tagging through k -means. A comparison between the actual and the predicted level attained via the k -means clustering is evaluated through a multi-class performance metrics. It was demonstrated from the investigation that the k -means model is capable of achieving a classification accuracy of 69% and 67% for the train and test data, respectively.

Keywords RFID · Pallet-level tagging · Unsupervised machine learning · K -means

1 Introduction

In recent years, there has been an increasing interest in Radio Frequency Identification (RFID) that provides significant benefits to consumers throughout the supply chain [1–3]. The RFID item-level has been adopted and provides an intelligence-rich environment that enables businesses to improve the transport of goods, predict demand, promote inventories effectively, prevent write-downs and trade-offs as well as ultimately increasing revenue.

RFID item-level tagging (ILT) is the tagging of individual objects typically used to track product pallets through the supply chain, rather than individual items [4]. Retailers can track the inventory of a given item by scanning the tag, and the location will be provided by the RFID reader. Typically, the ultra-high frequency (UHF) RFID

C. S. Choong (✉) · A. F. Ab. Nasir · A. P. P. Abdul Majeed · M. A. Zakaria · M. A. Mohd Razman

Innovative Manufacturing, Mechatronics and Sports Lab (IMAMS), Faculty of Manufacturing & Mechatronic Engineering Technology, University Malaysia Pahang, Pekan, Pahang, Malaysia
e-mail: cs_choong@hotmail.com

is applied in the inventory, readers send a request signal to the RFID tag, which is then responded with a backscatter. Thus, the reader analyses the response and records the tag information along with the Received Signal Strength Indicator (RSSI) signal. The RSSI is a power level indicator of an RFID signal (RSS) obtained by a reader from the signal transmitted by an RFID tag. In general, the RSSI values are used for UHF RFID applications to provide the tag response within the reading area [5].

Conversely, pallet-level tagging is defined as RFID tags attached as individual pallets but stacked into one, and one could argue that the tagging level facilitates mixed-pallet loads. Nevertheless, particularly in stacked pallet-level tagging, tracking individual items is challenging. This is even more so, in the event that the individual pallet is taken out of the stack or misplaced. Identifying the level using the RFID signal will, therefore, be a critical issue to be addressed in order to determine the correct signal pattern for each level.

This study aims at investigating the pattern of RFID signal for stacked pallet-level tagging using an unsupervised machine learning model, in particular, k -means clustering technique in its ability to correctly classify the pallet-level tagging. It has been reported in the literature that the aforesaid clustering technique is often associated with the ability to assign an arbitrary tag for each cluster created by it [6–10].

2 Data Processing

This section details on the collection of RSS information, feature extraction methods as well as the evaluation metrics utilised for pallet-level tagging classification. Initially, the UHF reader is mounted on the station where RSS, timestamp, and electronic product code (EPC) is recorded. During the experiment, ten (10) pallets are tagged with RFID in which they are stacked. The period during which the RFID data is collected is 60 s, and the tag is approximately scanned for 600 times by the RFID station. Upon the completion of the data collection process, different statistical features will be extracted from the readings, prior to the employment of the k -means technique.

2.1 Statistical Feature Extraction

The time-domain features are extracted through the RSSI data via statistical means to characterise the pallet-level. The statistical analysis describes the reading of the RSS with a sampling frequency of 10 Hz, which is then reduced to a sample size of 1 Hz. The following equations are used to extract features:

$$\text{Mean} = \frac{\sum_{i=1}^n x_i}{N} \quad (1)$$

$$\text{Median} = \left(\frac{N + 1}{2} \right) \quad (2)$$

$$\text{Variance} = \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3)$$

$$\text{Range} = \text{Maximum RSS value} - \text{Minimum RSS value} \quad (4)$$

$$\text{Skewness} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^3} \quad (5)$$

$$\text{Kurtosis} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} \quad (6)$$

In the algorithms as depicted in Eqs. 1–6, we have N = number of variables, \bar{x} as the mean of variables. In addition to the equations above, the maximum RSS value, minimum RSS value and the mode RSS value were also collected as features for classifying the pallet-level. Therefore, a total of nine features are extracted for this study.

2.2 Classification Through Unsupervised Learning

This study intends to use k -means to classify the location of the pallet instead of using supervised machine learning models. The number of classes is determined by an unsupervised learning algorithm that defines data points in different classes by segmentation. The disparity between the data points used in this study is calculated by the Euclidean metric which can be expressed as:

$$D(x, \bar{x}) = \sqrt{\sum_{i=1}^n (x_i - \bar{y}_i)^2} \quad (7)$$

The distance D between the data points is computed by the mean of all the data point y and the sum of the data point x in the cluster. The cluster size is checked in conjunction with the cluster reliability so that the centre of the cluster starts with the nearest data points. The average of all cluster data points is then added to the new cluster in the centroid. If the total variability is less than the previous cluster in the new cluster, the number of clusters will be increased. In classification, the categories divided by k -means are contrasted with the original tag number by means of the

experimental setup, in this case, the respective levels. Consequently, the confusion matrix shows the results of the actual and predicted class of k -means. Eventually, having 80% of the data in the training set and 20% of the data in the testing set in order to evaluate the classification performance.

2.3 Multi-class Performance Metrics

To evaluate the performance of the k -means model, different multi-class performance metrics are used, namely precision, recall, F1-score as well as classification accuracy. F1-score is calculated by combining precision and recall in a single number using a harmonic mean. In contrast to the arithmetic mean, the F1 score is between precision and recall. The performance of the tag for training and testing at each level in k -means is displayed according to precision, recall and F1-score. However, the performance with the corresponding metric is difficult to explain [11]. For instance, the macro-averaged F1-score (macro-F1) is computed as simple arithmetic means of our F1-scores per class where n is the number of the classes:

$$\text{macro F1} = \frac{F1_1 + F1_2 + \dots + F1_n}{n} \quad (8)$$

Likewise, the macro-averaged precision and the macro-averaged recall can also be calculated:

$$\text{macro Precision} = \frac{\text{precision}_1 + \text{precision}_2 + \dots + \text{precision}_n}{n} \quad (9)$$

$$\text{macro Recall} = \frac{\text{recall}_1 + \text{recall}_2 + \dots + \text{recall}_n}{n} \quad (10)$$

The macro-F1 gave the equal weight to each class, and another weighted-average F1-score (weighted-F1) is the amount of sample from that class, N is the weight of each class F1-score, thus computed as follows:

$$\text{weighted F1} = \frac{(F1_1 \times N) + (F1_2 \times N) + \dots + (F1_n \times N)}{N} \quad (11)$$

weighted Precision

$$= \frac{(\text{precision}_1 \times N) + (\text{precision}_2 \times N) + \dots + (\text{precision}_n \times N)}{N} \quad (12)$$

$$\text{weighted Recall} = \frac{(\text{recall}_1 \times N) + (\text{recall}_2 \times N) + \dots + (\text{recall}_n \times N)}{N} \quad (13)$$

3 Results and Discussion

It has been reported from previous research that RFID signal has a positive correlation towards classification efficacy [12], this work is carried out to complement the study by utilising unsupervised learning for classification. The unsupervised learning, i.e., *k*-means were employed to discern the distance between the tagged target of the reader antenna. The procedure begins with the data captured from the reader antenna and is based on the assumption that the predicted class of the target is identified by the working process. The predicted class undergoes a cross-mapping process to match the correct level of the RFID tag.

The multi-class metrics in Table 1 demonstrates the overall performance metric of *k*-means for the training set and testing set. It illustrates the performance that shows 0.69 and 0.67 in classification accuracy in the train set and test set, respectively. It can be seen clearly that the precision result implies that distinctive percentages of the number of classes (macro average) and the respective number of the sample (weighted average) are the same with 0.63 for the train and 0.62 for the test. The performance of recall reduces slightly due to the weight of the number of samples, where on macro and weighted average are the same with 0.69 for the train set, in contrast, the test results are different for macro and weighted with 0.68 and 0.67 respectively. The F1 score is located in between precision and recall, with 0.64 for macro and weighted on the train set, whilst 0.63 and 0.62 for macro and weighted, respectively on the test set.

Figure 1 illustrates the performance of precision, recall and F1 score for each pellet level tagging. It is obvious to note that level two (L2) and ten (L10) have zero performance rate, suggesting a total misclassification for both train and test. Conversely, L1 and L8 show the best classification rate. However, the performances between L3 and L9 demonstrated an erratic trend. This classification of the tags can be clearly seen through confusion matrices of train and test tabulated in Tables 2 and 3, respectively (Fig. 2).

The confusion matrix in Tables 2 and 3 indicate the classification type that is colour coded to blue, and red, respectively. The blue colour represents the correct classification of the respective classes, whereas the red colour is referred to the misclassification. For example, the level one data points in the train set have 475 data

Table 1 Multi-class Metrics in *k*-means for train and test

		Performance metric				
		Precision	Recall	F1-score	Classification accuracy	Number of samples
Train	Macro average	0.63	0.69	0.64	0.69	4800
	Weighted average	0.63	0.69	0.64		4800
Test	Macro average	0.62	0.68	0.63	0.67	1200
	Weighted average	0.62	0.67	0.62		1200

Fig. 1 The experimental setup for pallet-level tagging in stacked



Table 2 Confusion matrix of each level in comparing actual and *k*-means for train

Level		Predicted									
		1	2	3	4	5	6	7	8	9	10
Actual	1	475	0	0	0	0	0	0	0	0	0
	2	0	0	399	0	0	0	86	0	0	0
	3	0	0	462	0	0	0	13	0	0	0
	4	0	41	0	442	0	0	1	0	8	0
	5	0	0	0	0	477	0	0	0	1	1
	6	0	59	0	1	0	305	0	0	53	57
	7	0	0	12	37	0	0	448	0	0	0
	8	0	0	0	0	0	0	0	480	0	0
	9	0	26	0	0	0	98	0	0	225	125
	10	0	0	0	0	468	0	0	0	1	0

Table 3 Confusion matrix of each level in comparing actual and *k*-means for test

Level	Predicted									
	1	2	3	4	5	6	7	8	9	10
1	125	0	0	0	0	0	0	0	0	0
2	0	0	95	0	0	0	20	0	0	0
3	0	0	115	0	0	0	10	0	0	0
4	0	16	0	90	0	0	0	0	3	0
5	0	0	0	0	121	0	0	0	0	0
6	0	18	0	0	0	78	0	0	17	12
7	0	0	2	9	0	0	92	0	0	0
8	0	0	0	0	0	0	0	120	0	0
9	0	4	0	0	1	24	0	0	64	35
10	0	0	0	0	131	0	0	0	0	0

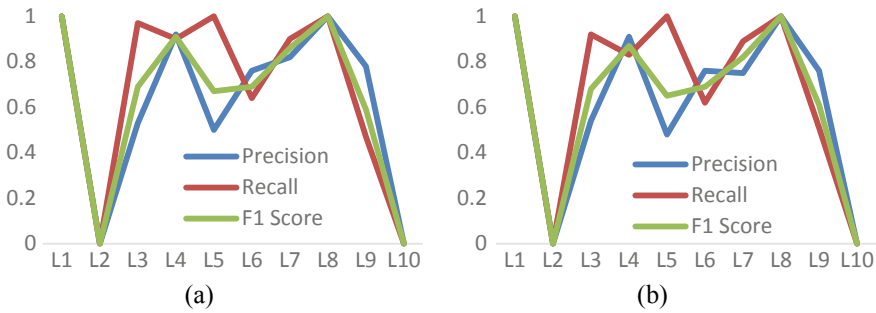


Fig. 2 The precision-recall and F1-score performances at the varying level for *k*-means in **a** training and **b** testing

and are correctly classified into level one from the *k*-means prediction. However, it could be seen that for level two, it was totally misclassified to level 3 and level 7, with a data point of 399 and 86, respectively. The same observation is made for level ten, in which it is misclassified to level five and level nine, probably due to the loss of the signal to the floor [13]. On the other hand, the rest of the levels have a good performance to correctly identify the level that most of them have a true positive with above 400 data points.

The test set had similar performance compared to the train set, which considered the issue at level two and level ten to be misclassification. Conversely, level one and level eight are correctly classified corresponding to its pallet-level. The major misclassification could be attributed by the open space experienced by the RFID tag at the top-level with not much obstacle, therefore, the misclassification happened

for identified the 2nd level into 3rd level. On the other hand, the bottom layer is also easily misclassified as the radio frequency could not measure the backscatter from the absorption losses on the floor [13].

4 Conclusion

This study attempts to employing unsupervised machine learning instead of supervised machine learning to classify the pallet-level through RSSI signals. Nine statistical-based time-domain features were extracted in this study. It has been demonstrated from the study that the *k*-means model is capable of achieving a classification accuracy of 69% and 67% for the train and test data, respectively. However, it is worth noting that this study is rather in its preliminary stage; further, an improvement could be made by using a more appropriate method of selection of features, such as Principal Component Analysis, Univariate Selection or Feature Importance techniques to evaluate its efficacy in the classification of pallet-level.

References

1. Toyoda K, Takis Mathiopoulos P, Sasase I, Ohtsuki T (2017) A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain. IEEE Access 5:17465–17477. <https://doi.org/10.1109/ACCESS.2017.2720760>
2. Gu Q, Shen B (2019) Fashion supply chain management in Asia: concepts, models, and cases. In: Fashion supply chain management in Asia: concepts, models, and cases. pp 63–70. <https://doi.org/10.1007/978-981-13-2294-5>
3. Yan B, Chen X, Yuan Q, Zhou X (2019) Sustainability in fresh agricultural product supply chain based on radio frequency identification under an emergency. Cent Eur J Oper Res. <https://doi.org/10.1007/s10100-019-00657-6>
4. Motorola (2010) Item-level RFID tagging and the intelligent apparel supply chain
5. Alfian G, Syafrudin M, Yoon B, Rhee J (2019) False positive RFID detection using classification models. Appl Sci 9. <https://doi.org/10.3390/app9061154>
6. Kotu V, Deshpande B (2019) Data science concepts and practice. [https://doi.org/10.1016/S0922-3487\(08\)70226-0](https://doi.org/10.1016/S0922-3487(08)70226-0)
7. Choi SH, Cheung HH, Yang B, Yang YX (2014) Item-level RFID for retail business improvement. In: RFID technology integration for business performance improvement. IGI Global, pp 1–26. <https://doi.org/10.4018/978-1-4666-6308-4.ch001>
8. Kukard W, Wood L (2019) Consumers' perceptions of item-level RFID use in FMCG. In: Cyber Law, Privacy, and Security. pp 1384–1407. IGI Global. <https://doi.org/10.4018/978-1-5225-8897-9.ch069>
9. Xu X, Chen X, Ji J, Chen F, Sanjay AV (2018) RETaIL: a machine learning-based item-level localization system in retail environment. In: Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST. Springer, pp 221–231. https://doi.org/10.1007/978-3-030-00916-8_21
10. Gaukler GM, Seifert RW, Hausman WH (2007) Item-level RFID in the retail supply chain. Prod Oper Manag 16:65–76. <https://doi.org/10.1111/j.1937-5956.2007.tb00166.x>

11. Yang Y, Liu X (1999) A re-examination of text categorization methods. In: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval, SIGIR 1999. pp 42–49. <https://doi.org/10.1145/312624.312647>
12. Choong CS, Fakhri A, Nasir A, Majeed APPA, Zakaria MA Automatic identification and Categorize Zone of RFID reading in Warehouse Management System
13. Rath HK, Timmadasari S, Panigrahi B, Simha A (2017) Realistic indoor path loss modeling for regular WiFi operations in India. In: 2017 23rd national conference on communications, NCC 2017. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/NCC.2017.8077107>

Flat Price Prediction Using Linear and Random Forest Regression Based on Machine Learning Techniques



Julakha Jahan Jui, M. M. Imran Molla, Bifta Sama Bari, Mamunur Rashid, and Md Jahid Hasan

Abstract Flat price prediction is an important topic of real estate. Flat price in a city depends on different criteria such as, the crime rate of that location, total populations on that area, number of bedrooms, bathrooms, the total size of the flat, location of the flat, etc. People feel confused and face different harassments with unreliable information during purchasing a flat in a city. By taking consideration of this scenario, we have proposed here flat price prediction framework. In this study, we have used our own data set that we have collected from Dhaka, Bangladesh. Two regression algorithms namely the linear regression and the regression tree/random forest regression have been used for building the prediction model. We have also checked the validity of the model using boxplot analysis, residual analysis, error checking and cross-validation. Finally, the performance of two methods has been compared which shows that the random forest regression model gives the best prediction result than the linear regression model.

Keywords Flat price prediction · Machine learning · Linear regression · Random forest regression

J. J. Jui (✉) · B. S. Bari · M. Rashid
Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang, 26600
Pekan, Pahang, Malaysia
e-mail: julakha.ump@gmail.com

B. S. Bari
e-mail: biftasama_120@yahoo.com

M. Rashid
e-mail: mamun110218@gmail.com

M. M. Imran Molla
Faculty of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajgonj,
Bangladesh
e-mail: mmimranmolla@gmail.com

M. J. Hasan
Faculty of Mechanical and Manufacturing Engineering Technology, Universiti Malaysia Pahang,
26600 Pekan, Pahang, Malaysia
e-mail: sawikot@gmail.com

1 Introduction

The demand for flats is gradually increasing with the race of civilization. A wide variety of factors affects price of a flat. Real estate business has a great impact on national economy. Therefore, forecasting flat price is beneficial for buyers, real estate agents and economic professionals. Actual prediction of the flat price is important for expected flat owners, developers, investors and tax assessors. Therefore, the accurate flat price prediction models make easy to decide for the buyer as well as a seller by removing the information gap. In recent years, machine learning techniques are widely used for flat price prediction due to its remarkable advantages as compared to traditional methods [1–5]. Machine learning algorithms build models using the train data set and test data set is employed to these models for the prediction. Machine learning algorithms can be sorted in different types and among these types, regression, classification and neural network are used in flat price prediction.

In order to predict the house/flat price, different machine learning models have been implemented in [6–8]. Wang proposed linear regression model and random forest regression for predicting the house price in [9]. Authors utilized online dataset to build the proposed models, and they achieved the R-squared value of 0.6391 for random forest and 0.3442 using linear regression whereas the RMSE of 367.05 for random forest and 407.94 for linear regression.

In [10] house price was predicted using different machine learning regression algorithms (linear, random forest, neural network and so on). The implemented data in their research taken from Melbourne Housing Market dataset and downloaded from Kaggle website [11]. In their study, among the algorithms regression tree gives the best result than the other algorithm for the dataset and the MSE rate was 0.0985.

Tang et al. [12] have been proposed ensemble machine learning algorithms for house price prediction. In their experiment, the MRE rate for the random forest was 0.87. Lu et al. [13] proposed the similar algorithm for the house price prediction. They used hybrid Ridge and Lasso regressions for their prediction. They found the best regression value with the test dataset, that is 0.11260 from the combination of 65% Lasso and 35% gradient boosting.

In the last few decades, researchers were also trying to predict the daily stuff price by utilizing different machine learning algorithms. In [14] various machine learning algorithms was compared for the prediction of house price in Virginia City and their results illustrated that the RIPPER algorithm gives best results than the other algorithms.

Van-gao and Guangwen [15] proposed an electricity price prediction model using the support vector machine algorithm which is trained by genetic algorithm and their experimental results show that their algorithm gives better prediction accuracy than the radial basis function neural network with 2.22% error rate. Varma et al. [16] used machine learning algorithm and neural networks algorithm for predicting the house price. But they does not provide any statistical value for the accuracy of their proposed algorithm. Additionally, in [17] an ensemble machine learning technique is proposed for predicting the price of oil. They constructed an ensemble model by considering

IBL, KStar and SMOReg models for oil price prediction. They demonstrate from their findings that the ensemble method established has done better than SMOReg and IBL.

Parasich and Kaftannikov [18] predict the sales price of a house by machine learning regression methods and they used classical machine learning algorithm. For their prediction they use Ames Housing Price data set [19]. In [20] house prices in China, a modified genetic algorithm was predicted based on long short-term memory (LSTM). For their prediction, they collect data from Shenzhen, China and consider eight features.

In this paper, linear regression and random forest regression-based machine learning algorithms have been implemented for the prediction of flat price. We have collected the complete dataset from the different areas of Dhaka, Bangladesh and these dataset have been utilized to build the prediction models. This paper is organized as follows; research methodology is presented in Sect. 2 with data collection, processing and model building. Section 3 compares the test result of the used algorithms, followed by a conclusion in Sect. 4.

2 Methodology

The proposed flat price prediction technique has been shown in Fig. 1. The first step of this study is data collection and the full data set has been collected from Dhaka, Bangladesh. After data collection, data have been preprocessed and normalized. Moreover, the validity of the data set has been checked using ANOVA. Then the dataset has been divide into two dataset a training set and a testing set. Now, the proposed model has been built up using training data set and the built model has been tested using testing data set. Finally, the accuracy of the proposed model has been calculated. In the rest of the section, the data collection, preprocessing, model building parts have been broadly discussed.

2.1 Dataset Collection

The dataset for this study has been collected from four different locations namely Mirpur, Dhanmondi, Bashundhara and Mohammadpur of Dhaka city which is the capital of Bangladesh. The sources of this dataset are different real state websites [21, 22] and bikroy.com [23]. In Bangladesh, Bikroy.com is one of the most popular websites for online buyer and seller. The features set of this dataset have been tabulated in Table 1.

Fig. 1 Flow chart of flat price prediction process



Table 1 Attributes descriptions for the experiment

Attributes	Description	Attributes	Description
Crime	Crime rate of the location	Pop	Total number of populations
Bed	Number of bedrooms	Area	Total area of the location
Bath	Number of bathrooms	Price	The asking price of the flat
Size	Size of the flat	–	–

Web scrapper software has been employed for collecting necessary features from the source of dataset. This dataset contains total 1199 observations and each observation contain seven features namely crime rate, total bedrooms, total bathrooms, size of the flat, population of the location, total area of the location and asking price of the flat. Table 2 present the sample dataset of 10 observations.

2.2 Data Pre-processing

Data in raw form is not ideal for analysis [24]. To be useful for predictive modeling the data must be cleaned at first. Then, errors and inconsistencies of the dataset have been removed to improve the quality of data. Then the missing values in the

Table 2 Sample dataset of 10 observations

Observations	Name of the features						
	Crime	Bed	Bath	Size	Pop	Area	Price
1	-0.34679	1	1	360	1.699294	6.45	20.55
2	-0.58503	2	2	600	1.551258	6.6	42.33
3	-0.284	2	2	660	1.512192	5.8	30.5
4	-0.284	4	5	3100	1.512192	5.8	72
5	-0.40894	5	6	4800	1.169213	6.9	125
6	-0.58503	2	2	700	1.551258	6.6	30
7	-0.34679	2	2	650	1.699294	6.45	29.23
8	-0.40894	4	4	1790	1.169213	6.9	60
9	-0.284	3	4	1765	1.512192	5.8	75
10	-0.34679	2	2	650	1.699294	6.45	32.3

dataset has been checked. After checking the missing value of the dataset, stepwise regression process has been employed. Basically, stepwise regression is employed for the multiple independent variables and this ability is accomplished by measuring statistical values such as R-squared, p, t-status and AIC metric in order to determine relevant variables. From the summary table, it has been seen that the highest adjusted R-squared of 0.3646 and the p -value of $2.20e^{-16}$ were achieved. When the R-squared value is higher and the p -value is lower, this indicates a good data set, but from the obtained R-squared value from the dataset is quite poor for building any model. Hence, it is necessary to observe the histogram of the features. The histograms of seven features are shown in Fig. 2.

In the histogram, it has been shown that there is no normal distribution in the attribute of population and crime rate. Hence, in order to make the histogram as normally distributed, \log_{10} operation has been performed on the attribute of population and crime rate. After performing \log_{10} operation, the histogram become normal distributed and the adjusted R-squared improved to 0.5988.

Another way to check the data quality is ANOVA test. Analysis of variance (ANOVA) has been employed on preprocessed data set. The F-value and p -value for the attributes are shown in Table 3. From Table 3 it has been shown that the p -value for all attributes is below 0.005 which indicates that the feature sets are perfectly preprocessed and capable to make the model.

2.3 Model Building

In this study, two machine learning model namely linear regression (LR) model and random forest (RF) regression model has been employed for the prediction of flat price.

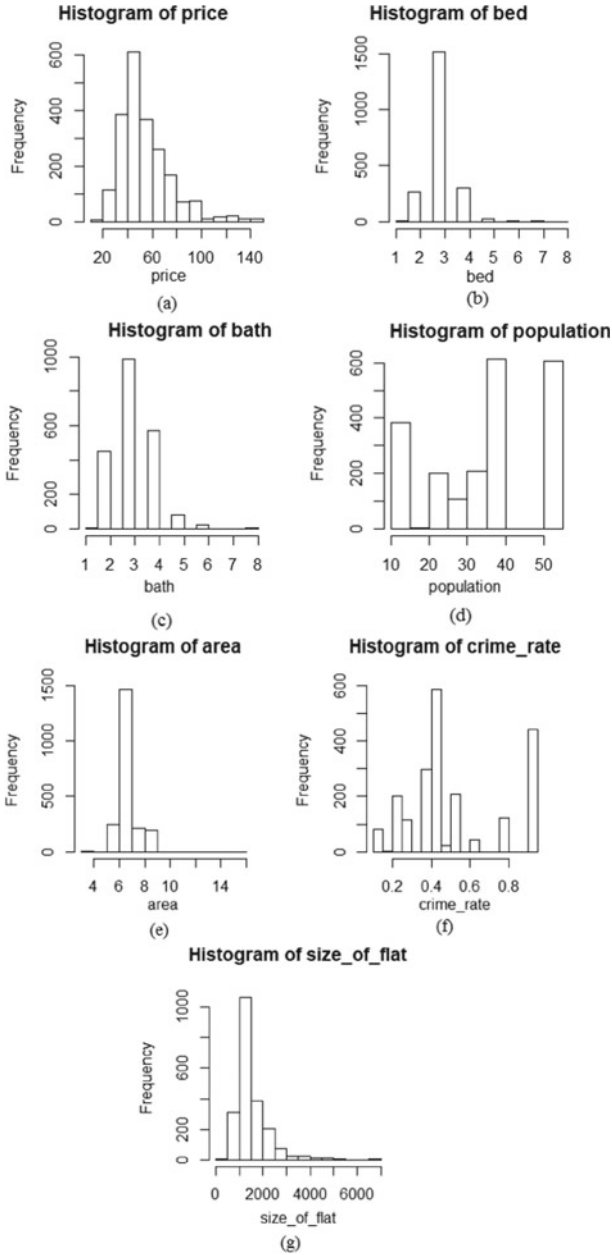


Fig. 2 Histogram of the attributes before \log_{10} operation; **a** price, **b** bed, **c** bath, **d** population, **e** area, **f** crime rate, **g** size

Table 3 ANOVA test table

Attributes	F-value	p-value
Crime	F (1,1192) = 8.589	0.0034
Bed	F (1,1192) = 403	$2e^{-16}$
Bath	F (1,1192) = 497	$2e^{-16}$
Size	F (1,1192) = 529	$2e^{-16}$
Population	F (1,1192) = 331	$2e^{-16}$
Area	F (1,1192) = 10	0.0010

Linear Regression (LR). LR is the commonly employed framework to build prediction models. Basically, it is a linear approach where a dependent variable is going to be related with one or more independent variables. Equation (1) indicates the regression relationship with one dependent variable and one independent variable

$$y = a + b * x \tag{1}$$

where, y = estimated value for dependent variable, a = constant, b = coefficient, and x = independent variables value. The performance of a linear regression model should be assessed in term of mean standard error. Moreover, the dataset should be normalized and split randomly into train and test set where training dataset is utilized to build the model and testing dataset is used to evaluate the model. K fold cross validation technique has also been performed to split dataset.

There are different ways to check a model whether it shows acceptable performance or not for a specific data set. Among these techniques, residual plots significantly able to evaluate the regression model. Basically, it's the difference between the dependent variable's observed value and its expected value. There are many residual plots including, the residual versus fitted plot is widely used. A good residual versus fitted plot must be shapeless with simple data patterns without apparent outliers and usually distributed symmetrically around the 0 line against big residuals [25]. Positive residual values (on the y-axis) mean that the prediction is too low and negative values mean that the prediction was too high; 0 means that the prediction was accurate. Figure 3a shows residual versus fitted plot and it seems excellent due to fulfill of all criterion of a good residual versus fitted plot. Figure 3b shows the normal Q-Q plot. This plot used to check whether the residuals are normally distributed or not. A Q-Q plot is a scatter plot that is measured against one another by two sets of quartiles. This plot can be considered as effective when both sets of quartiles come from the same range and the points form an approximately straight line. Figure 3c is the scale-location plot showing whether or not the residuals are distributed equally along with the predictor ranges. A good scale location plot generates a horizontal line of similarly (randomly) distributed points [26]. Figure 3d is the residuals versus leverage plot that allows the linear regression model to classify influential data points [26]. Outliers can be influential, although they do not always have to do so and some points in the model within a normal range could be very influential. The points can be in the top right or bottom right corners, which are outside the red dashed Cook's

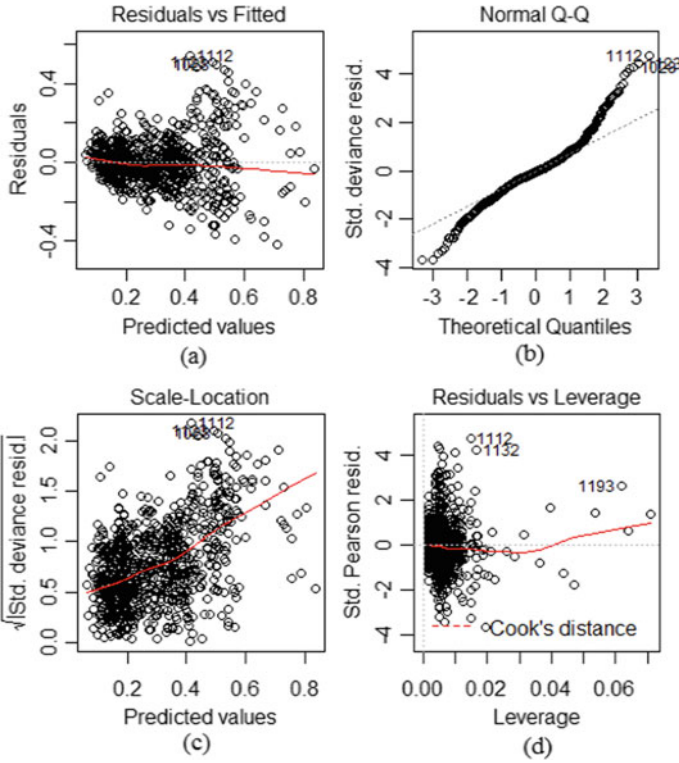


Fig. 3 Diagnostic plots for linear regression; **a** residuals versus fitted plot, **b** normal Q-Q plot, **c** scale-location plot, **d** residuals versus leverage plot

distance line. These points are influence the model and also by removing them the regression results would be noticeably altered. The diagnostic plot for the linear model has been shown in Fig. 3.

For analyzing the regression efficiency, it is important to perform statistical analysis test such as the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean relative error (MRE), to evaluate how far the predictions are distant from the actual data. In this study, 10-fold cross validation technique has been performed for calculating the MSE.

Random Forest Regression. Random Forest (RF) method is an ensemble machine learning method that uses multiple decision trees to perform both regression and classification tasks, and bagging is such type of statistical technique to performed regression and classification simultaneously [27]. Bagging also with boosting is the most common strategy of ensemble designed to address high variance and high bias.

However, RF model constructs multiple decision trees focusing on the data and merges together with their predictions to achieve a more accurate and reliable prediction rather than focusing on single decision trees [27].

For this prediction the complete dataset has been split into two groups including training dataset and testing dataset. The training data is composed of 80% of the total observations and the remaining observations have utilized for testing purpose. The response variable “price” indicates the asking price of flat. The model fitted with 300 trees. MSE and out of Bag error are estimated to evaluate the performance of the model. The 2/3rd training data is used here to train the trees and the remaining 1/3rd is used to test the trees. The number of variables selected randomly on each split is 2. The errors versus the number trees plot is shown in Fig. 4.

Figure 4 shows that the error is dropping by adding more and more trees. This paper also compared the “out of bag sample errors” and “error on test set” which has been shown in Fig. 5. The red curve represents the “out of bag errors” whereas the blue curve represents the test error. When the number of predictor at each split is 3–4, a low MSE value is observed.

Before log operation, the accuracy of the RF model was around 41% and after log₁₀ operation the accuracy was improved to 64%.

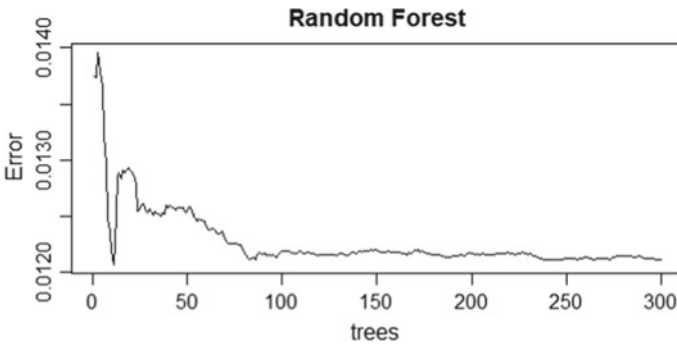
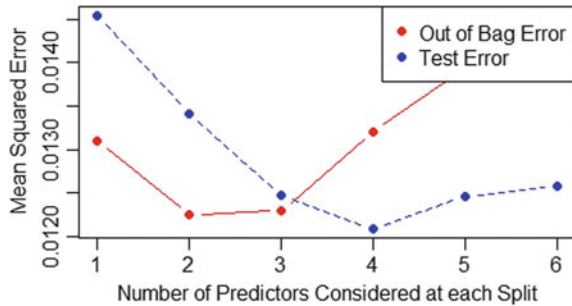


Fig. 4 Error versus N tree plot

Fig. 5 Out of bag error versus test error



3 Result and Discussion

Evaluating the R-Squared and p -Value: The R-squared and p -value for the two-model shown in Table 4. The R-squared value of linear regression is 0.6134, which shows that the linear model explains 61% of the variation in the data.

The R-squared value of random forest regression model is 0.6498, which shows that the model explains 65% of the variation in the data. Here, the p -value for linear regression is low (≤ 0.05) and this also supports the alternative hypothesis that each covariate is significant.

Real Versus Predicted Plot Analysis: In order to visual analysis, the real versus predicted plot has been shown in Fig. 6. It is obvious from the plot that the random forest regression plot (see Fig. 6a) provides the prediction line which is generally more straight around the real line than the linear regression prediction lines (Fig. 6b).

Boxplot Analysis: The performance of the employed models has also been explained with help of Box plot, which is useful in visualizing skewness in data. Figure 7a, b show the MSE for LM and RF respectively, where the distribution is positively skewed. This means, their mean is considerably higher than their median. The MSE for random forest is more positively skewed than the linear model. In case of MAE (shown in Fig. 7c, d) for both LM and RF are negatively skewed whereas the MRE (pictured in Fig. 7e, f) both for LM and RF are positively skewed.

The MSE, RMSE, MAE and MRE of the two algorithms has been tabulated in Table 5. The MSE of linear and random forest regression is 0.0055 and 0.0036

Table 4 Evaluating the R-squared and p -value

Algorithms	R-squared	p -value
Linear regression	0.6134	$2.2e^{-16}$
Random forest regression	0.6498	–

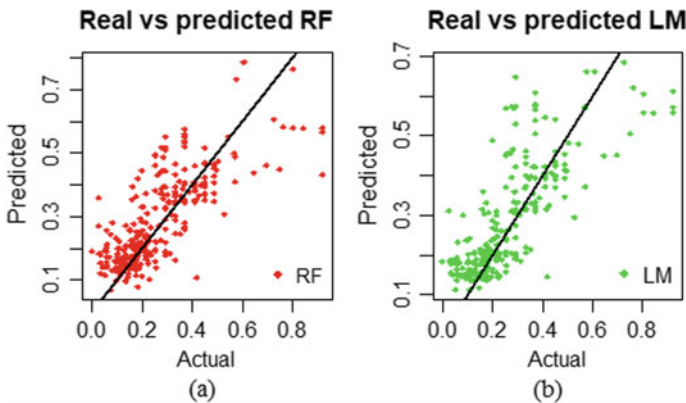


Fig. 6 Real versus predicted plot; **a** random forest, **b** linear model

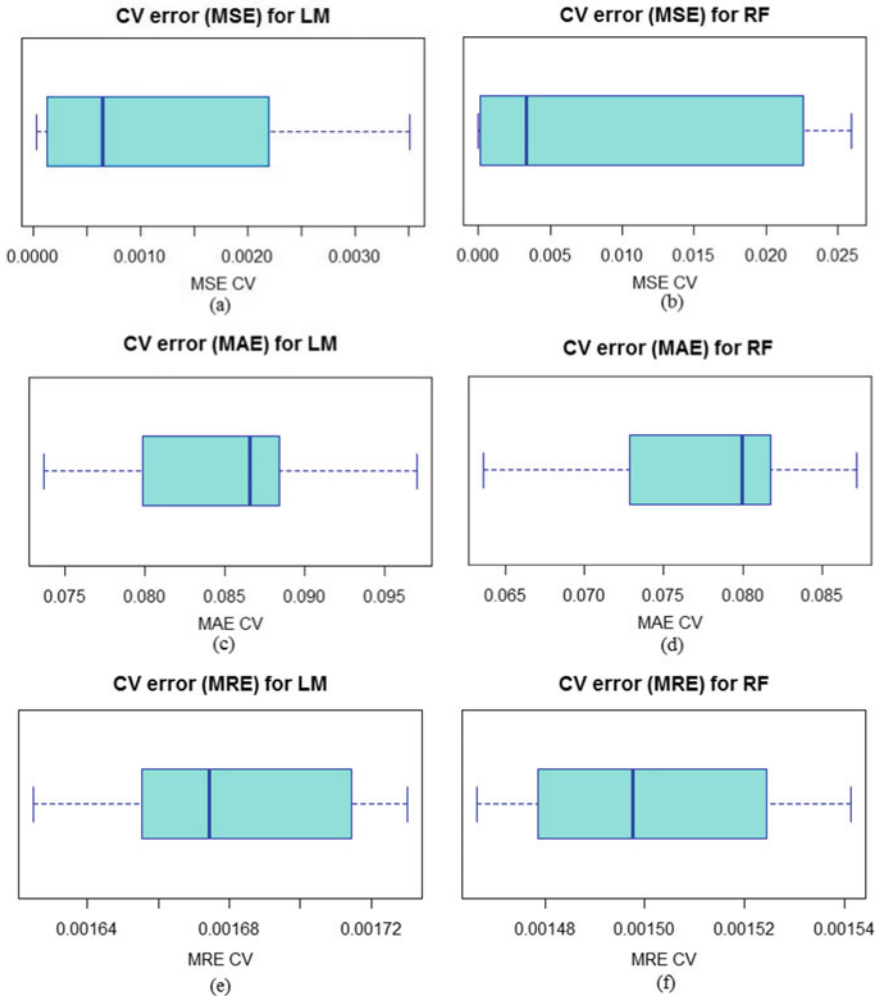


Fig. 7 Boxplot of CV error (MSE, MAE and MRE) for LM and RF; **a** MSE for LM, **b** MSE for RF, **c** MAE for LM, **d** MAE for RF, **e** MRE for LM, **f** MRE for RF

Table 5 Comparison of the linear model, random forest regression model

Algorithms	MSE	RMSE	MAE	MRE
Linear regression	0.0055	0.1168	0.0826	0.0016
Random forest regression	0.0036	0.1097	0.0750	0.0014

respectively. The MSE is achieved through the difference between the actual value and the expected value which means the model gives better prediction when the difference becomes less.

4 Conclusion

In this study, two machine-learning regression-based techniques were proposed to predict the flat price, namely linear regression and random forest regression. Here, the data set that has been collected from four different cities in Dhaka, Bangladesh. The data has been extracted from several real-state websites using the web scrapper (Data Toolbar) software. For building the model, seven features have been considered that may affect the flat price. Here, the quality of data has been analyzed by checking histogram, residual plotting, and ANOVA. After preprocessing the dataset, the linear and random forest model has been built. The performance of both techniques has been evaluated in terms of MSE, RMSE, MAE, and MRE of both techniques have been computed. From the obtained error rate, it has been concluded that the performance of random forest regression model is much better than the linear regression model. Hence, the outcome of this study indicates that random forest regression-based machine learning technique can predict the flat price efficiently. However, the selection of most appropriate features and optimization of this algorithm should be studied.

Acknowledgements The author would like to acknowledge the great supports by the Faculty of Electrical and Electronics Engineering and Universiti Malaysia Pahang, Malaysia.


References

1. Crawford GW, Fratantoni AMC (2003) Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. *Real Estate Econ* 31:223–243
2. Al-Janabi S (2017) Pragmatic miner to risk analysis for intrusion detection (PMRA-ID). In: *International conference on soft computing in data science*. Springer, Singapore, pp 263–277
3. Al-Janabi S (2018) Smart system to create an optimal higher education environment using IDA and IOTs. *Int J Comput Appl* 1–16
4. Al-Janabi S, Al-Shourbaji I (2017) Assessing the suitability of soft computing approaches for forest fires prediction. *Appl Comput Inf* 14:214–224
5. Al-Janabi S, Rawat S, Patel A, Al-Shourbaji I (2015) Design and evaluation of a hybrid system for detection and prediction of faults in electrical transformers. *Int J Electr Power Energy Syst* 67:324–335
6. ParkJae B, Kwon Bae K (2015) Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data. *Expert Syst Appl* 42:2928–2934
7. Jirong G, Mingcang Z, Liuguangyan J (2011) Housing price forecasting based on genetic algorithm and support vector machine. *Expert Syst Appl* 38(4):3383–3386
8. Wang X, Wen J, Zhang Y, Wang Y (2014) Real estate price forecasting based on SVM optimized by PSO. *Optik* 125(3):1439–1443

9. Wang C, Wu H (2018) A new machine learning approach to house price estimation. *New Trends Math Sci* 6(4):165–171
10. Danh P (2018) Housing price prediction using machine learning algorithms: the case of Melbourne City, Australia. In: 2018 international conference on machine learning and data engineering (iCMLDE), Dec 3–7
11. Pino A (2018) Melbourne housing market data. Kaggle. <https://www.kaggle.com/anthonypino/melbourne-housing-market>
12. Tang Y et al (2018) Predicting housing price based on ensemble learning algorithm. In: International conference on artificial intelligence and data processing (IDAP), Sept 28–30
13. Lu S et al (2017) A hybrid regression technique for house prices prediction. In: IEEE international conference on industrial engineering and engineering management, Dec 10–13
14. Park B, Bae JK (2015) Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data. Pergamon Press Inc, Oxford
15. Van-gao C, Guangwen M (2009) Electricity price forecasting based on support vector machine trained by genetic algorithm. In: 2009 Third international symposium on intelligent information technology application, vol 2, IEEE
16. Varma A, Sarma A, Doshi S, Nair R (2018) House price prediction using machine learning and neural networks, pp 1936–1939
17. Gabralla LA, Jammazi R, Abraham A (2013) Oil price prediction using ensemble machine learning, pp 674–679
18. Parasich AV, Kaftannikov IL (2018) Predicting sales prices of the houses using regression methods of machine learning. In: 2018 3rd Russian-pacific conference on computer technology and applications (RPC), IEEE
19. De Cock D, Ames I (2011) Alternative to the Boston housing data as an end of semester regression project. *J Stat Educ* 19(3)
20. Liu R, Liu L (2019) Predicting housing price in China based on long short-term memory incorporating modified genetic algorithm. *Soft Comput* 1–10
21. Bdhousing homepage, <https://www.bdhousing.com/>
22. pbazar.com, <https://pbazaar.com/en/>
23. Bikroy.com, <https://bikroy.com/en/ads/dhaka-division/property>
24. Lavrakas PJ (2011) Raw data. In: *Encyclopedia of survey research methods*
25. Hagerman I (2017) Residual plots part 1—residuals vs. fitted plot. In: *Data distilled*
26. Kim B (2015) Understanding diagnostic plots for linear regression analysis. University of Virginia Library
27. Drakos G (2019) Random forest regression model explained in depth. In: *Data science*

The Classification of Skateboarding Tricks by Means of the Integration of Transfer Learning and Machine Learning Models



Muhammad Nur Aiman Shapiee , Muhammad Ar Rahim Ibrahim, Mohd Azraai Mohd Razman, Muhammad Amirul Abdullah, Rabiuh Muazu Musa, and Anwar P. P. Abdul Majeed 

Abstract Generally, the assessment of skateboarding tricks executions is completed abstractly dependent on the judges' understanding and experience. Hence, an objective and means for assessing skateboarding tricks, especially in the big competition are important. This research aims at classifying skateboarding flat ground tricks, namely Ollie, Kickflip, Shove-it, Nollie Front Shove and Frontside 180 through camera vision and pre-trained convolution neural network for feature extraction coupled with a conventional machine learning model. An amateur skateboarder (23 years of age \pm 5.0 years' experience) executed five tricks for each type of trick repeatedly on a skateboard from a camera with a distance of 1.26 m. From the images captured, the features were engineered and extracted through Transfer Learning, particularly VGG-16 and then classified by means of Logistic Regression (LR) and k -Nearest Neighbour (k -NN) models. The observation from the preliminary investigation demonstrated that through the proposed methodology, the LR and k -NN models attained a classification accuracy of 99.1% and 97.7%, on the test dataset, respectively. It could be shown that the proposed strategy can classify the skateboard tricks well and would, in the long run, support the judges in providing an increasingly objective-based judgment.

Keywords Image processing · Transfer learning · Machine learning · Skateboarding tricks · Classification

M. N. A. Shapiee (✉) · M. A. R. Ibrahim · M. A. M. Razman · M. A. Abdullah · A. P. P. Abdul Majeed

Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan, Pahang Darul Makmur, Malaysia
e-mail: nuraimanshapiee@gmail.com

A. P. P. Abdul Majeed
e-mail: amajeed@ump.edu.my

R. M. Musa
Department of Credited Co-Curriculum, Centre for Fundamental and Continuing Education, Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia

1 Introduction

The net worth of the skateboarding industry is approximately \$USD 4.8 billion in 2010 [1]. It is worth noting that skateboarding will make its debut in the 2020 Tokyo Summer Olympics, suggesting the popularity and the traction that this game has made over the years. This growth somewhat demands new innovative approaches in evaluating the game in such competitive competitions, especially with the evaluation of the tricks executed are conventionally carried out by judges through subjective means that is often liable to bias and imprecise assessment.

Convolution Neural Networks (CNN) has been employed for human activity recognition through the use of single tri-axial accelerometer in Android phones [2]. One hundred healthy subjects (68 males and 32 females) were recruited in the study and they were asked to carry out eight (8) typical activities which includes falling, running, jumping, walking, walking quickly, step walking, walking upstairs and walking downstairs. The authors compared the developed CNN model with a Support Vector Machine (SVM) and Deep Belief Network (DBN) where different feature extraction techniques namely Fast Fourier Transform and Discrete Cosine Transform apart from time-domain features. It was shown that the CNN model could achieve an average classification accuracy (CA) of up to 93.8% and triumphed against the other models evaluated.

Park et al. [3] also exploited the use of CNN for human activity recognition. The dataset used in the study was obtained from the Microsoft Research Cambridge-12 (MSRC-12) database, which discriminates twelve (12) actions, i.e., lift arms, duck, push right, goggles, wind it up, shoot, bow, throw, had enough, change weapon, beat both, and kick, respectively. The features extracted are the joint angles from different parts of the body. DBN, and Hidden Markov Model (HMM) was compared to the CNN developed by the authors. It was demonstrated from the study that the CNN model could yield a CA of up to 98.59% followed by DBN and HMM with a CA of 97.54% and 92.49%, respectively, suggesting the remarkable capability of CNN. Nonetheless, it is worth noting that CNN often requires extensive hyperparameter tuning to achieve such results and hence, researchers have exploited the use of pre-trained CNN models through transfer learning.

The utilization of transfer learning has been investigated for the classification of both human activities as well as of sports-related motion by Sargano et al. [4]. The authors utilised AlexNet as a feature extraction method and then classified the different types of activities through Support Vector Machine (SVM) and k -Nearest Neighbour (k -NN) as well as a hybrid of both. The dataset for sports and human activities was solicited from University of Central Florida (UCF) and Kungliga Tekniska Hogskolan (KTH) datasets, respectively. The Leave-One-Out (LOO) cross-validation (CV) scheme has been proposed in the investigation. It was shown from the study that the hybrid scheme outperformed the conventional SVM and k -NN models with a classification accuracy of 98.15%.

There is limited literature with regards to the classification of skateboarding tricks through the employment of machine learning and to a certain extent a representation

of deep learning. Amongst the earlier notable works were carried out by Groh et al. [5, 6], Shapiee et al. [7] and Abdullah et al. [8] dealt the classification of skateboarding tricks through the integration machine learning with image processing technique and Inertial Measurement Unit (IMU) sensors, respectively. It is worth noting that the notion of transfer learning particularly the use of Inception V3 as a feature extraction method was initially explored in [7] with regards to skateboarding. The features extracted was then used to evaluate a number of classifiers and it was shown from the study that the classification accuracy of the Logistic Regression (LR) and k -NN yield a classification accuracy of 98.6% and 95.8%, respectively.

This paper aims at evaluating two machine learning models, namely LR and k -NN on their ability for classifying several skateboarding tricks, i.e. Ollie, Kickflip, Shove-it, Nollie and Frontside 180 based on featured engineered via the a pre-trained CNN model, particularly VGG-16 as a form of transfer learning (TL). This outcome of this investigation may serve useful to a more objective-based evaluation by the judges as well as providing a means for skateboarders to improve their performance further.

2 Methodology

2.1 Experimental Setup

The YI Action Camera with Full HD 1080p 60 frames per second was used to record the skateboarding tricks. The specifications of the camera; (model: YI action camera, lens: F2.8 aperture/155° wide-angle lens, resolution: full HD videos 1920 × 1080 (60 FPS), dimension: 60 × 42 × 21 mm, weight: 72 g) whilst the experimental setup, with regards to the camera orientation, is depicted in Fig. 1 and the view of the trick’s execution can be viewed in Fig. 2.

Fig. 1 Experimental setup

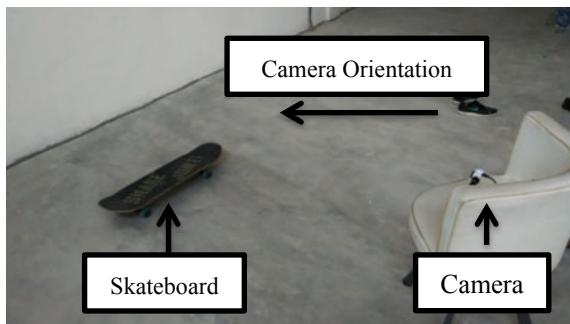


Fig. 2 Execution of the tricks



2.2 Data Collection

One skateboarder (23 years old, 170 cm and 54 kg) was recruited from the Skilled Extreme Sports Hub Club (SESH), University Malaysia Pahang. The skateboarder is required to perform five different tricks (as shown in Table 1) and to be repeated 5

Table 1 Description of the tricks class

Tricks	Description
Ollie (O)	The skater “pops” (striking to the ground) the tail of the board down whilst jumping and slid the nose of the skateboard to achieve liftoff to make the board “pop” into the air
Nollie Frontside Shove (NFS)	The trick of mirror image of a Pop Shove-it, the difference being that you will be snapping from nollie stance. The snap of the tail rotating the board 180 or 360 beneath the rider and catching it in mid-air
Front-side 180 (FS180)	An Ollie is where the skater and his board rotate 180° (front-side) after liftoff from the ground with the skateboard sticking to the skateboarder’s feet
Kickflip (K)	The skater Ollie the board into the air and then flips it fully 360° along the axis. Then landing the board approximate as its original position
Pop Shove-it (PS)	An Ollie variation with the snap of the tail rotating the board 180 or 360 beneath the rider and catching it in mid-air

times per trick. The tricks were chosen based on the experience of the skateboarder. The axis of the rotation is referring to the goofy stance direction.

2.3 Image Processing

The video of the tricks was recorded from the YI Action Cam. The extraction of the video was carried out using VLC media player 2.2.6. The video that is subsequently used for the image extraction process was taken for a period between 2 and 3 s (the execution of the trick only). It is worth noting that unrelated video scenes were removed. The videos were then processed by using the video to jpg converter v. 5.0.101 to obtain the frame per frame images with dimension 1080×1920 pixels then been resized to 300×300 pixels using Caesium software. The extraction of the images was set to 30 frames per video. A total of approximately 750 images were extracted from the 25 videos captured for the tested tricks.

2.4 Feature Extraction Through Transfer Learning

In the present investigation, the VGG-16 architecture shall be used for the purpose of feature extraction. VGG16 (also known as OxfordNet) is a CNN-based architecture pre-trained on ImageNet developed by the Visual Geometry Group. The default input size for this model is 224×224 as the input shape has to be (224, 224, 3) as it should be exactly three inputs channels, and width and height should be no smaller than 32. It has been reported that the employment of such strategy is useful in training models with the limited dataset and could avoid the notion of overfitting if a model is to be built from scratch apart from being computationally inexpensive [9].

2.5 Classifiers

In this study, the features extracted via transfer learning are then fed into different two classifiers, namely Logistic Regression (LR) as well as k -Nearest Neighbour (k -NN), respectively. As for the LR model, the L-2 regularisation or Ridge Regularisation is employed to regulate the overfitting nature of the model [10]. Conversely, the number of k used for the k -NN model developed is five (5), whilst the distance metric selected for the study is the Euclidean distance. The train to test ratio employed is 70–30, respectively. The five-fold cross-validation technique is used to train the models [11, 12]. Different performance metrics were used to evaluate the efficacy of the method employed, namely, Area-Under the Curve (AUC), classification accuracy (CA), F1-score, as well as Precision and Recall.

3 Results and Discussion

A total of 40 trick events were carried out by the skateboarder, and the success rate of the landings was recorded. From the 40 tricks, only 25 tricks were found to be successful, and the data (herein the images, i.e., 46 for FS180, 48 for K, 39 for NFS, 37 for O and 46 for PS) obtained were then fed into the VGG-16 to extract its features. Table 2 lists the performance the models developed on the train dataset. It could be observed that the LR and *k*-NN models achieved a CA of 98.88% and 93.33%, respectively. Further evaluation could be made on the confusion matrix of both the LR and *k*-NN models (Figs. 3 and 4, respectively). It could be seen from Fig. 3 that the misclassification arises from the PS images trick that is misclassified as K and O, whilst for K, NFS and O are misclassified as O, O and PS, respectively. Meanwhile, it is evident from Fig. 4 that the misclassification transpired across all classes.

Conversely, Table 3, lists the performance of the models developed on the test dataset, this phase is carried out to provide an unbiased evaluation of a final model that was developed based on the training dataset. It could be seen that both the LR and *k*-NN models yielded a higher CA of 99.1% and 97.7%, respectively as compared to the train evaluation. Moreover, from the Confusion Matrix of LR and *k*-NN models (Figs. 5 and 6), it could be observed that the O images were misclassified as NFS for the LR model, whilst for the *k*-NN model, FS180 was

Table 2 Evaluation of the developed classifiers on the train dataset

Method	AUC	CA	F1	Precision	Recall
LR	1.000	0.988	0.988	0.988	0.988
<i>k</i> -NN	0.994	0.933	0.933	0.934	0.933

Fig. 3 Confusion matrix of the developed LR model on train dataset

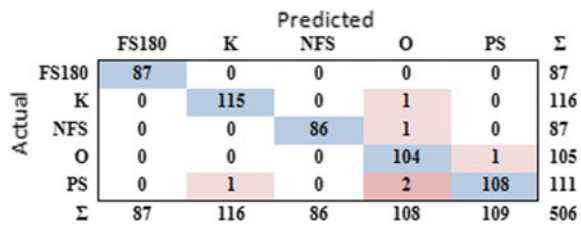


Fig. 4 Confusion matrix of the developed *k*-NN model on train dataset

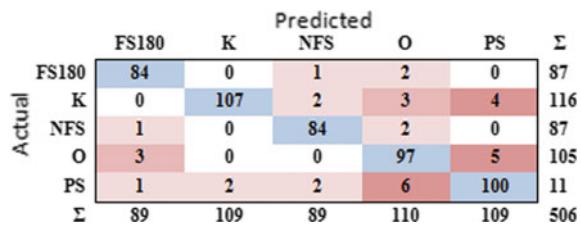


Table 3 Evaluation of the developed classifiers on the test dataset

Method	AUC	CA	F1	Precision	Recall
LR	1.000	0.991	0.991	0.991	0.991
<i>k</i> -NN	1.000	0.977	0.977	0.977	0.977

Fig. 5 Confusion matrix of the developed LR model on train dataset

		Predicted					Σ
		FS180	K	NFS	O	PS	
Actual	FS180	62	0	0	0	0	62
	K	0	34	0	0	0	34
	NFS	0	0	40	0	0	40
	O	0	0	2	39	0	41
	PS	0	0	0	0	39	39
Σ		62	34	42	39	39	216

Fig. 6 Confusion matrix of the developed *k*-NN model on train dataset

		Predicted					Σ
		FS180	K	NFS	O	PS	
Actual	FS180	61	0	1	0	0	62
	K	0	33	0	1	0	34
	NFS	0	0	40	0	0	40
	O	0	1	0	38	2	41
	PS	0	0	0	0	39	39
Σ		61	34	41	39	41	216

shown to be misclassified as NFS, K as O and O as K and PS, respectively. It is apparent that the proposed methodology could provide appreciable classification on skateboarding tricks, particularly by fusing VGG-16 with LR classifier. In addition, it could also be concluded from this preliminary investigation that the VGG-16 transfer learning model is much better at extracting the significant features, as compared to Inception-V3 model on the same classifiers evaluated in [7].

4 Conclusion

In this preliminary investigation, an offline skateboarding tricks classification system was developed via image processing. It was shown from the investigation that the features extracted by the pre-trained VGG16 deep CNN model are non-trivial in providing a reasonable CA of the evaluated skateboarding tricks. It was also demonstrated that the LR and *k*-NN model developed yielded a high CA of 99.1% and 97.7%, respectively. The present investigation will be further investigated by evaluating different TL models, namely AlexNet, ResNet and VGG-19 amongst others apart from optimising the hyperparameters of different ML models. The preliminary results further suggest the appropriateness of the proposed system in providing an

objective-based judgment on skateboarding tricks as opposed to the conventional subjective techniques that are currently being applied in this sport.

Acknowledgements The authors would like to acknowledge the Ministry of Education, Malaysia and Universiti Malaysia Pahang for supporting and funding this research via FRGS/1/2019/TK03/UMP/02/6 (RDU1901115).

References

1. Corrêa NK, de Lima JCM, Russomano T, dos Santos MA (2017) Development of a skateboarding trick classifier using accelerometry and machine learning. *Res Biomed Eng* 33:362–369. <https://doi.org/10.1590/2446-4740.04717>
2. Chen Y, Xue Y (2015) A deep learning approach to human activity recognition based on single accelerometer. In: 2015 IEEE international conference on systems, man, and cybernetics, pp 1488–1492. <https://doi.org/10.1109/SMC.2015.263>
3. Park JH, Park SU, Uddin MZ, Al-Antari MA, Al-Masni MA, Kim T-S (2017) A single depth sensor based human activity recognition via convolutional neural network. In: International conference on the development of biomedical engineering in Vietnam. Springer, Berlin, pp 541–545
4. Sargano AB, Wang X, Angelov P, Habib Z (2017) Human action recognition using transfer learning with deep representations. In: 2017 international joint conference on neural networks (IJCNN), pp 463–469. <https://doi.org/10.1109/IJCNN.2017.7965890>
5. Groh BH, Fleckenstein M, Kautz T, Eskofier BM (2017) Classification and visualization of skateboard tricks using wearable sensors. *Pervasive Mob Comput* 40:42–55. <https://doi.org/10.1016/j.pmcj.2017.05.007>
6. Groh BH, Kautz T, Schuldhaus D (2015) IMU-based trick classification in skateboarding. In: KDD work large-scale sports analytics
7. Shapiee MNA, Ibrahim MAR, Mohd Razman MA, Abdullah MA, Musa RM, Hassan MHA, Majeed APA (2020) The classification of skateboarding trick manoeuvres through the integration of image processing techniques and machine learning. In: Nasir ANK, Ahmad MA, Najib MS, Wahab YA, Othman NA, Ghani NA, Irawan A, Khatun S, Ismail RMTR, Saari MM, Daud MR, Faudzi AAM (eds) InECCE2019 proceedings of the 5th international conference on electrical, control & computer engineering, Kuantan, Pahang, Malaysia, 29th July 2019, 1st edn. Springer, Singapore
8. Abdullah MA, Ibrahim MAR, Shapiee MNA Bin, Razman MAM, Musa RM, Majeed APPA (2019) The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning. In: Symposium on intelligent manufacturing and mechatronics. Springer, Berlin, pp 67–74
9. Ng H, Nguyen VD, Vonikakis V, Winkler S (2015) Deep learning for emotion recognition on small datasets using transfer learning. pp 443–449
10. Musa RM, Majeed APPA, Taha Z, Chang SW, Nasir AFA, Abdullah MR (2019) A machine learning approach of predicting high potential archers by means of physical fitness indicators. *PLoS ONE* 14:e0209638
11. Razman MAM, Susto GA, Cenedese A, Majeed APPA, Musa RM, Ghani ASA, Adnan FA, Ismail KM, Taha Z, Mukai Y (2019) Hunger classification of Lates calcarifer by means of an automated feeder and image processing. *Comput Electron Agric* 163:104883
12. Musa RM, Taha Z, Maje Anwar PPA, Abdullah MR (2019) Machine learning in sports: identifying potential archers. Springer, Singapore

π 1a and π 1b Impedance Matching for Capacitive Power Transfer System



Yusmarnita Yusop, Shakir Saat, Khairul Kamarudin, Huzaimah Husin, and Sing Kiong Nguang

Abstract This paper presents a capacitive power transfer (CPT) system for low power applications. This CPT system utilized π impedance matching to provide reactive compensation and voltage gain, and is driven by a Class-E resonant inverter. In this work, two basic π matching network topologies that can be integrated into the Class-E resonant inverter was analysed. By utilising the proposed method, the zero voltage switching (ZVS) condition can still be achieved even if the load is larger than that optimum and hence yielding high efficiency. Simulation comparison between the two matching networks was performed to determine their performances and later, to select the best matching network that can be integrated into the proposed CPT system prior to practical implementation. The validity of the proposed concept was verified experimentally. A 10 W CPT system lab scale prototype was designed and implemented to demonstrate efficient wireless power transfer across 0.25 mm gap. The experimental work showed that 10 W output power was transferred successfully to the 50 Ω load with only 913 pF capacitive coupling plates at 91.9% efficiency.

Keywords Capacitive power transfer · Impedance matching · Compensation network · Class-E inverter · Zero voltage switching

Y. Yusop (✉) · S. Saat · H. Husin

Faculty of Electronics and Computer Engineering, Centre for Telecommunication Research and Innovation (CeTRI), Universiti Teknikal Malaysia Melaka (UTeM), Alor Gajah, Melaka, Malaysia
e-mail: yusmarnita@utem.edu.my

K. Kamarudin

Faculty of Electrical Engineering, Universiti Teknologi MARA, Johor Branch, Pasir Gudang Campus, Masai, Johor, Malaysia

S. K. Nguang

Department of Electrical and Computer Engineering, The University of Auckland, Auckland, New Zealand

1 Introduction

The most pervasive methods of the wireless power transfer (WPT) technology are Inductive Power Transfer (IPT) and capacitive power transfer (CPT) [1]. The IPT technique has achieved great success in both theoretical development and industrial applications for power delivery in certain areas, besides being applicable to many power levels and gap distances. Since the 1990s, IPT has been widely employed with successful applications in biomedical [2–4], consumer electronics [5–7], power supply to electric vehicles [8], and others [9, 10]. Nonetheless, the IPT is unable to transfer power across metal barriers, and power losses are of concern when metal objects are present close to the magnetic field. For example, the presence of metal debris in the air gap is dangerous in practice due to the generated eddy current losses that may cause substantial increment in temperature [11]. This drawback, however, can be addressed by applying the CPT technology. In comparison to the IPT system, the CPT system offers two advantages. First, the CPT system is not sensitive to the presence of any metal nearby as it does not generate extra heat in the metal due to eddy current losses. The second advantage is that the CPT system utilizes metal plates to transfer power, instead of coils made of Litz wire, hence a system that is cost-effective. With that, the CPT technology is an attractive alternative solution to the conventional IPT system, particularly in the areas where IPT technology is not convenient [12, 13].

The CPT technology research has become popular since recent years. Despite all the successful attempts and investigations mentioned in the previous research [4, 14–16], CPT is still a very new technology that involves a number of challenges, including theoretical development, practical implementation, as well as concerns in both social and economic impacts. With that, several essential issues need to be addressed before it can be widely used in real life applications. All the present CPT systems; from low power range to high power range, share a similar fundamental challenge that derives from the trade-off between transfer distance, transfer power, and its efficiency. It is very difficult to achieve high power over long distance and high efficiency, concurrently. The main objective of the WPT technology is to transfer as much power as possible over a longer distance. In the CPT technology, power transfer across the capacitive coupling depends on the operating frequency, the plate capacitance, and the plate voltage. The coupling capacitance may be correlated inversely proportional to the coupling plate distance. With the prescribed coupling plate area, when the power transfer distance increases, the coupling capacitance becomes small. As a consequence, it requires increasing either the switching frequency or the voltage to achieve high power transfer, which is difficult to realise in practice. Therefore, the selection of compensation circuit and designing of an efficient power converter to operate at high frequency and high voltage appear to be challenging tasks.

This work explores efficient power converter, namely Class-E resonant inverter to drive a CPT system. In relation to the outcomes revealed in [17] the CPT system based on Class-E resonant inverter appeared to be highly sensitive towards the variation of its circuit parameters. In fact, variation in capacitive coupler and loads usually exists

in the real CPT situations that would surely result in lower system efficiency. Therefore, the utilisation of compensation network to serve as an impedance converter that enables efficient power transfer between two stages with non-matching impedances is investigated in this paper. To be specific, two basic π matching network topologies were selected to be integrated into the Class-E resonant inverter. Simulation comparison between the two matching networks was performed to determine their performances and later, to select the best matching network that can be integrated into the proposed CPT system prior to practical implementation. Performance comparisons in terms of ZVS, system output power and efficiency with respect to load response were observed. Finally, the validity of the proposed concept was verified experimentally.

The paper is presented in the following order. Section 2 provides theoretical analysis of the Class-E resonant inverter with impedance matching. Section 3 presents the impedance matching load characteristics. The simulation results and performance comparisons are given in Sect. 4 while the experimental results are discussed in Sect. 5. The conclusions that summarize the findings are presented in the last section.

2 Theoretical Analysis of Class-E With Impedance Matching

Figure 1a illustrates a diagram of the Class-E inverter with an impedance matching circuit. Impedance matching can be accomplished by tapping either resonant capacitor or resonant inductor in series or in parallel with the load resistor. Nevertheless, this work only investigated $\pi 1a$ and $\pi 1b$ matching networks, as presented in Fig. 1b. These particular impedance matching types were selected due to the following reasons: first, the capacitor that was connected in series with the load was modified easily into the capacitor coupling plate to fit the actual CPT system. Second, the additional capacitor that was connected in parallel before or after the capacitive coupler had managed to increase the mutual coupling capacitance that can increase the level of power transfer in the CPT system, as mentioned in [18, 19]. Hence, the details of mathematical analysis for Class-E with matching components equations were gathered.

2.1 Class-E with $\pi 1a$ Impedance Matching

This section presents the analysis of Class-E resonant inverter that was carried out with tapped capacitor matching circuit, namely, $\pi 1a$ impedance matching. When the circuit displayed in Fig. 2a is compared with the traditional Class-E circuit, one can observe an additional capacitor denoted as C_3 connected in parallel with the load resistance, R_L . In this case, the load had been assumed to be known. Since the

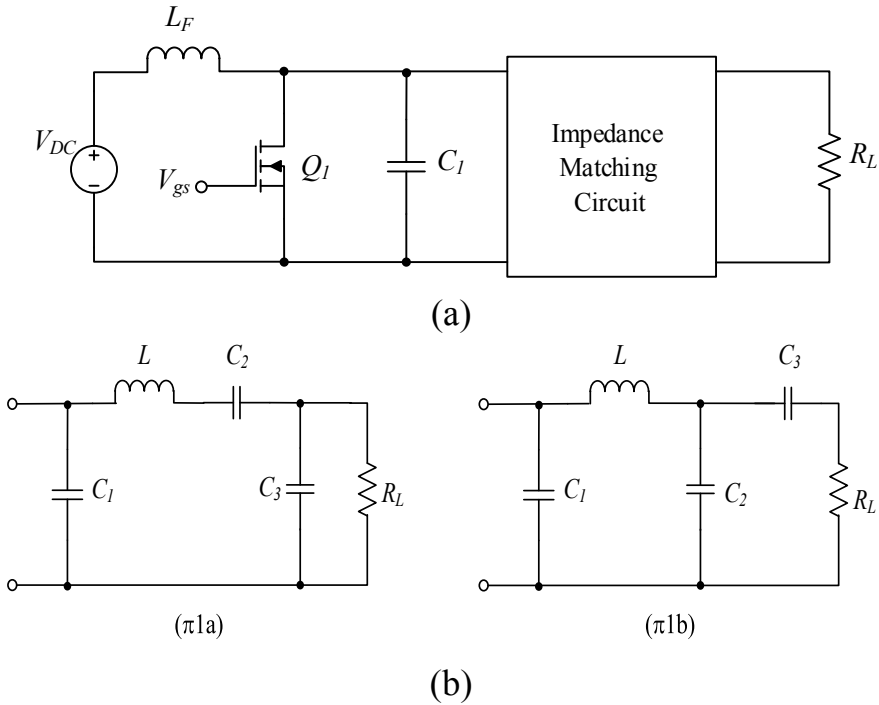


Fig. 1 a Class-E with impedance matching block diagram, b π matching network

matching type was for downward impedance matching transformation, the given load resistance, R_L , must be greater than the source or the series resistance, R_S . The basic Class-E components such as L_F , C_1 , and L for optimum operation were calculated using equations and assumptions in [20].

Next, as for the π 1a matching network components values, the key calculation aid for this analysis had been based on the series to parallel transformation theorem. The first step was to determine the reactance factor of the series circuit, R_L - C_3 , denoted as q .

$$q = \frac{R_L}{X_{C3}} \tag{1}$$

Based on Fig. 2b, the resistances, R_S and R_L , as well as the reactances, X_{C_S} and X_{C_3} , were associated with the following equations:

$$R_S = \frac{R_L}{1 + q^2} = \frac{R_L}{1 + \left(\frac{R_L}{X_{C3}}\right)^2} \tag{2}$$

and

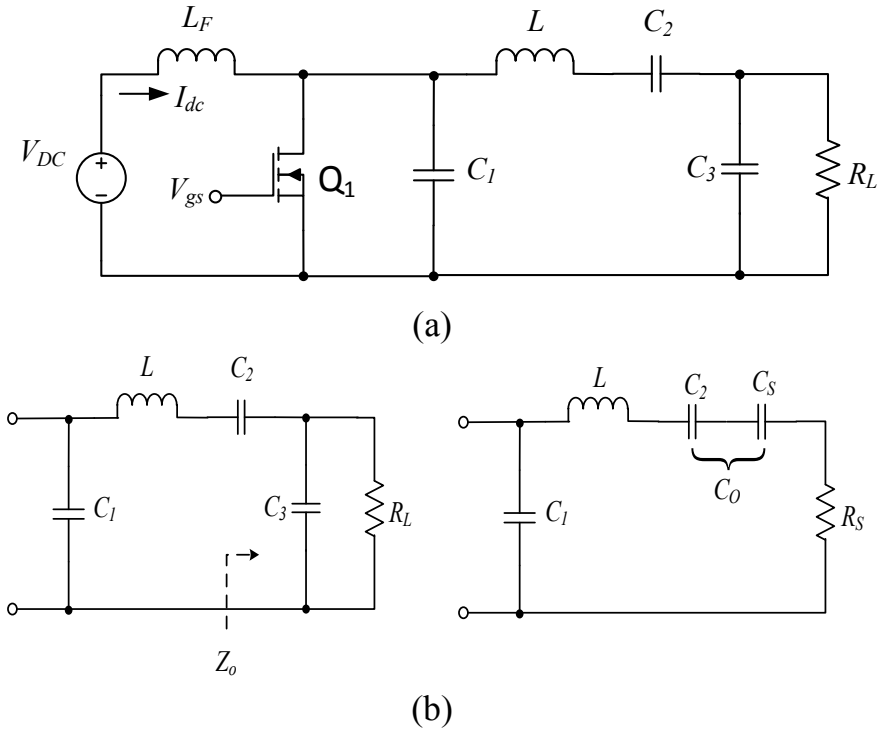


Fig. 2 a Class-E with $\pi 1a$ matching network, b equivalent circuit of the matching circuit $\pi 1a$

$$X_{CS} = \frac{X_{C3}}{1 + \frac{1}{q^2}} = \frac{X_{C3}}{1 + \left(\frac{X_{C3}}{R_L}\right)^2} \tag{3}$$

Hence,

$$C_S = C_3 \left(1 + \frac{1}{q^2}\right) = C_3 \left[1 + \left(\frac{X_{C3}}{R_L}\right)^2\right] \tag{4}$$

When Eq. (2) was rearranged, the following was obtained:

$$q = \sqrt{\frac{R_L}{R_S} - 1} \tag{5}$$

From Eqs. (1) and (5), X_{C3} can be expressed as,

$$X_{C3} = \frac{1}{\omega C_3} = \frac{R_L}{q} = \frac{R_L}{\sqrt{\frac{R_L}{R_S} - 1}} \tag{6}$$

Thus,

$$C_3 = \frac{R_L}{\omega R_S \sqrt{\frac{R_L}{R_S} - 1}} \quad (7)$$

When Eq. (5) was substituted into (1) to seek X_{C3} , the following was retrieved:

$$X_{C3} = q R_S = R_S \sqrt{\frac{R_L}{R_S} - 1} \quad (8)$$

Based on Fig. 2b and using Eq. (8), the capacitance, C_2 , was obtained,

$$X_{C2} = \frac{1}{\omega C_2} = \left[Q_L - \frac{\pi(\pi^2 - 4)}{16} \right] R_S - q R_S \quad (9)$$

$$= R_S \left[Q_L - \frac{\pi(\pi^2 - 4)}{16} - \sqrt{\frac{R_L}{R_S} - 1} \right] \quad (10)$$

Hence,

$$C_2 = \frac{1}{\omega R_S \left[Q_L - \frac{\pi(\pi^2 - 4)}{16} - \sqrt{\frac{R_L}{R_S} - 1} \right]} \quad (11)$$

Based on Eq. (6), the circuit displayed in Fig. 2a seemed to match the resistance that satisfied the inequality of $R_S < R_L$. As such, the voltage and the current for C_2 and C_3 could be determined. The amplitude of the voltage across capacitor C_3 appeared similar with the amplitude of the voltage across the load resistance, R_L . Therefore, V_{C3} and I_{C3} were calculated as follows:

$$V_{C3} = V_{RL} = \sqrt{2R_L P_O} \quad (12)$$

$$I_{C3} = \frac{V_{C3}}{X_{C3}} \quad (13)$$

The amplitudes of both the voltage and the current across capacitor C_2 are given as follows:

$$V_{C2} = X_{C2} I_m \quad (14)$$

$$I_{C2} = I_{C3} + I_{RL} \quad (15)$$

2.2 Class-E with $\pi 1b$ Impedance Matching

As portrayed in Fig. 3, a tapped capacitor matching circuit had been embedded to provide downward impedance transformation by integrating an additional capacitor C_3 in series with the load. The mathematical analysis for matching components equations of Class-E $\pi 1b$ was omitted here since it can be obtained in previous work done by [17]. However, certain important equations were included here for better understanding of this particular circuit topology. The additional matching capacitor C_2 and C_3 are determined as follows:

$$C_3 = \frac{1}{\omega R_L \sqrt{\frac{R_S \left[\left(Q_L - \frac{\pi(\pi^2-4)}{16} \right)^2 + 1 \right]}{R_L} - 1}} \tag{16}$$

$$C_2 = \frac{Q_L - \frac{\pi(\pi^2-4)}{16} - \sqrt{\frac{R_S \left[\left(Q_L - \frac{\pi(\pi^2-4)}{16} \right)^2 + 1 \right]}{R_L} - 1}}{\omega R_L \left[\left(Q_L - \frac{\pi(\pi^2-4)}{16} \right)^2 + 1 \right]} \tag{17}$$

Next, the analysis was continued to seek the amplitude of the current through the load resistance, R_L , and the capacitance, C_3 .

$$I_{R_L(peak)} = I_{C_3(peak)} = \sqrt{\frac{2P_o}{R_L}} \tag{18}$$

Hence,

$$V_{C_3(peak)} = X_{C_3} I_{R_L(peak)} \tag{19}$$

The amplitudes of both the voltage and the current through capacitor C_2 had been:

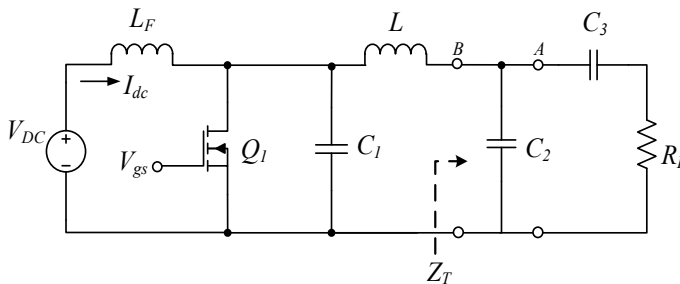


Fig. 3 Class-E with $\pi 1b$ matching circuit

$$V_{C2(peak)} = \sqrt{(V_{RL(peak)})^2 + (V_{C3(peak)})^2} \quad (20)$$

$$I_{C2(peak)} = \frac{V_{C2(peak)}}{X_{C2}} \quad (21)$$

3 Load Variation Characteristics

It is essential to arrive at the correlation of the Class-E output power with its matching network components. This section provides the details of mathematical analysis pertaining to the output power, as well as the effects of load variations on the system output power with $\pi 1a$ impedance matching that is represented as dP_o/dR_L . The dP_o/dR_L expression demonstrates the rate of changes or sensitivity level of the output power with respect to load variation. Based on this expression, the maximum output power point can also be determined by equating dP_o/dR_L to zero. The target of dP_o/dR_L must be as small as possible to show that the output power is less sensitive to the load.

3.1 dP_o/dR_L Analysis for $\pi 1a$ Impedance Matching

In order to determine the correlation between Class-E output power and impedance matching network, the equivalent impedance, Z_o , of capacitor C_3 parallel with load resistor, R_L , was identified first and it can be expressed as below,

$$Z_o = \frac{R_L - j\omega C_3 R_L}{1 + (\omega C_3 R_L)^2} \quad (22)$$

The expression of output power for basic Class-E is written as,

$$P_o = \frac{8V_{dc}^2}{(\pi^2 + 4)R_S} \quad (23)$$

where

$$R_S = \frac{R_L}{1 + (\omega C_3 R_L)^2} \quad (24)$$

When Eq. (24) was substituted into (23), the output power is expressed in terms of C_3 and R_L ,

$$P_o = \frac{8V_{dc}^2}{(\pi^2 + 4)} \left(\frac{1 + (\omega C_3 R_L)^2}{R_L} \right) \quad (25)$$

The maximum value of P_o in respect to load resistance was obtained by taking the derivative of (25) with respect to R_L , hence

$$\frac{dP_o}{dR_L} = \frac{8V_{dc}^2}{(\pi^2 + 4)} \left[\frac{d}{dR_L} \left(\frac{1}{R_L} \right) + \frac{d}{dR_L} (\omega C_3^2 R_L) \right] \quad (26)$$

$$= \frac{8V_{dc}^2}{(\pi^2 + 4)} \left(\omega^2 C_3^2 - \frac{1}{R_L^2} \right) \quad (27)$$

In order to determine if P_o has a local maximum, the critical points were identified by setting $dP_o/dR_L = 0$ and by solving R_L , which yielded the following:

$$R_L = \frac{1}{\omega C_3}$$

Now that dP_o/dR_L seemed to have only one critical point at $R_L = 1/\omega C_3$, P_o exhibited a maximum value at $R_L^* = 1/\omega C_3 = X_{C3}$. Therefore, if the actual load, $R_L^* < R_L$, dP_o/dR_L , is positive, thus P_o would increase. Vice versa, if $R_L^* > R_L$, dP_o/dR_L is negative, P_o would decrease.

3.2 dP_o/dR_L Analysis for $\pi 1b$ Impedance Matching

The equivalent impedance at point B was determined, which is represented as Z_T , in which C_2 appeared parallel with the summation of capacitor C_3 and load resistor R_L .

$$Z_T = \frac{1 + j\omega C_3 R_L}{j\omega(C_2 + C_3) - \omega^2 C_2 C_3 R_L} \quad (28)$$

The expression of output power for Class-E with $\pi 1b$ matching network appeared similar with Eq. (23), where

$$R_s = \frac{C_3^2 R_L}{\omega^2 C_2^2 C_3^2 R_L^2 + (C_2 + C_3)^2} \quad (29)$$

Therefore, Eq. (29) was substituted into (23) to gain the following:

$$P_o = \frac{8V_{dc}^2}{(\pi^2 + 4)} \left(\frac{\omega^2 C_2^2 C_3^2 R_L^2 + (C_2 + C_3)^2}{C_3^2 R_L} \right) \quad (30)$$

$$= \frac{8V_{dc}^2}{(\pi^2 + 4)} \left(\frac{\omega^2 C_2^2 C_3^2 R_L^2}{C_3^2 R_L} + \frac{(C_2 + C_3)^2}{C_3^2 R_L} \right) \quad (31)$$

$$= \frac{8V_{dc}^2}{(\pi^2 + 4)} \left(\omega^2 C_2^2 R_L^2 + \frac{(C_2 + C_3)^2}{C_3^2 R_L} \right) \quad (32)$$

The maximum value of P_o in respect to the load resistance was obtained by taking the derivative of (32) with respect to R_L . Let,

$$U = \omega^2 C_2^2 R_L^2 \quad \text{and} \quad V = \frac{(C_2 + C_3)^2}{C_3^2 R_L}$$

$$\frac{dU}{dR_L} = \omega^2 C_2^2 \quad (33)$$

$$\frac{dV}{dR_L} = -\frac{(C_2 + C_3)^2}{C_3^2 R_L^2} \quad (34)$$

Equation (32) was re-written in terms of U and V ,

$$P_o = \frac{8V_{dc}^2}{(\pi^2 + 4)} (U + V) \quad (35)$$

Therefore,

$$\frac{dP_o}{dR_L} = \frac{8V_{dc}^2}{(\pi^2 + 4)} \left[\frac{dU}{dR_L} + \frac{dV}{dR_L} \right] \quad (36)$$

Equations (33) and (34) were substituted into (36), thus yielding the following:

$$\frac{dP_o}{dR_L} = \frac{8V_{dc}^2}{(\pi^2 + 4)R_L^2} \left(\omega^2 C_2^2 R_L^2 - \frac{(C_2 + C_3)^2}{C_3^2} \right) \quad (37)$$

Similar to the $\pi 1a$ approach, it was essential to determine if P_o had a local maximum. Hence, following the step taken in the $\pi 1a$ matching analysis, the dP_o/dR_L was fixed at zero to solve R_L , hence

$$R_L = \frac{(C_2 + C_3)}{\omega C_2 C_3} \quad (38)$$

As the dP_o/dR_L had only one critical point at (38), hence, P_o appeared to have a maximum value at $R_L^* = R_L = (C_2 + C_3)/(\omega C_2 C_3)$. Therefore, if the actual load

is $R_L^* < R_L$, dP_o/dR_L becomes positive and P_o would increase. Vice versa, if $R_L^* > R_L$, dP_o/dR_L turns negative and P_o would decrease.

4 Simulation

The modelling of the CPT system operation based on Class-E with impedance matching network was carried out by applying the professional simulation software program in order to clarify the theoretical analysis performed in Sect. 3. All simulation modelling analyses for the CPT system with $\pi 1a$ and $\pi 1b$ matching networks were designed by using the same design specifications of $V_{dc} = 12$ V, $D = 0.5$, $f = 1$ MHz, and $Q = 10$, which was targeted to transfer 10 W across 50 Ω load that is bigger than the optimum load; 8.3 Ω . Figure 4 displays the simulation circuits with values of components given in Table 1a. It is important to highlight here that C_2 of $\pi 1a$ and C_3 of $\pi 1b$ matching networks are represented as the capacitive coupler, respectively. All voltages and currents were measured and tabulated in Table 1b, as well as in Figs. 5 and 6 in order to analyse the performances of the proposed matching circuits in terms of ZVS condition, input power, output power, and efficiency. The details of the comparison are depicted in forthcoming section.

4.1 Performance Comparison

The discussion of the results comparison is based on Figs. 5 and 6. In terms of both switching condition waveforms displayed in Figs. 5i and 6i it seemed apparent that the ZVS condition for both matching networks was perfectly achieved. As a result, this allowed 9.89 W RMS output power to be achieved from 10 W DC input power, which yielded 98.9% efficiency for $\pi 1a$ matching network. Meanwhile, as for the $\pi 1b$ matching network, 9.77 W RMS output power was successfully obtained from 10 W DC input power with 97.66% efficiency. To summarise, all input and output power results retrieved from the simulations, as depicted in Table 1b, were close to the theoretical values and verified that both the π matching networks did have the ability to produce the desired output power across $R_L = 50$ Ω perfectly well. Nevertheless, the best network had to be selected for practical implementation in the system. Therefore, further comparison analysis that revealed the matching type performance in terms of dP_o/dR_L , frequency response, and capacitive coupler behaviour was carried out.

4.1.1 Load Variation Analysis

Figure 7 illustrates the curves of output power with respect to the load variation. Both of the π matching displayed the positive sign of dP_o/dR_L due to the increasing

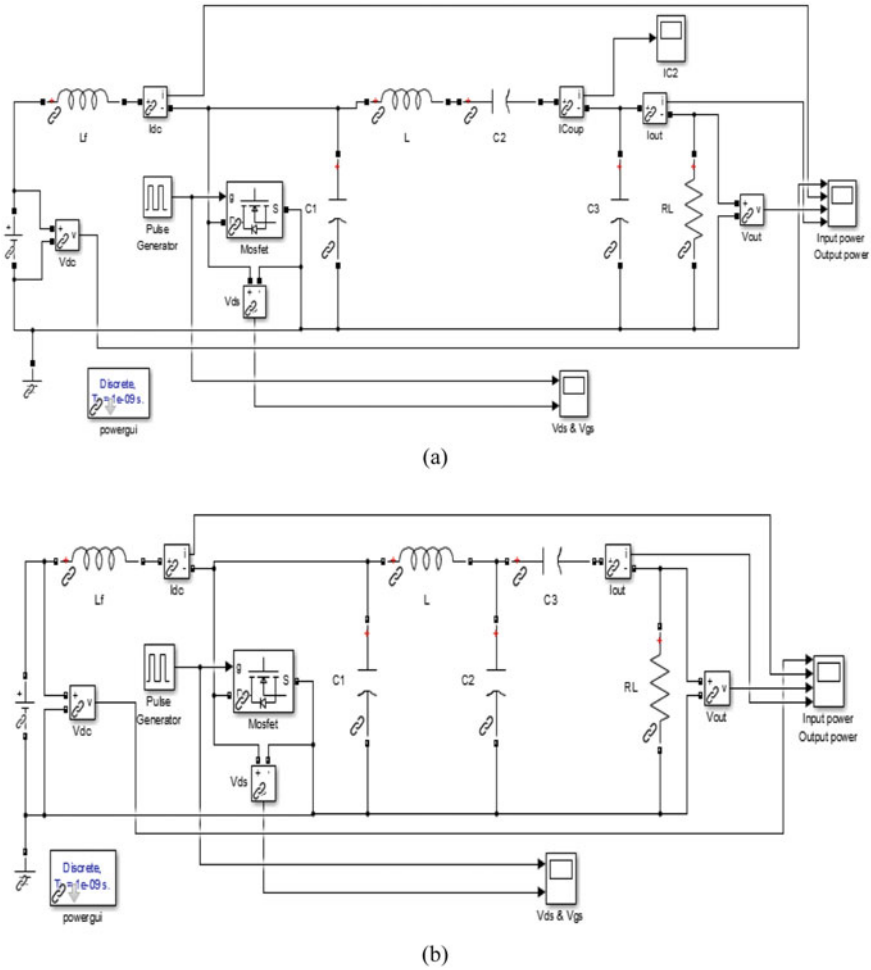


Fig. 4 CPT simulation circuit a $\pi 1a$, and b $\pi 1b$

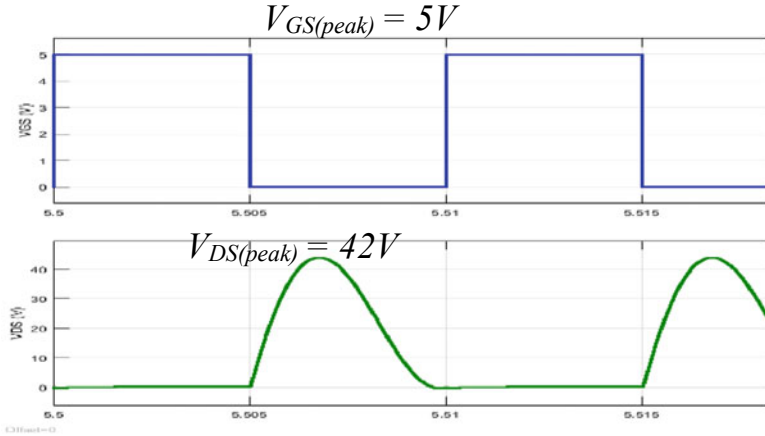
output power as the load was increased from optimum (8.3Ω) to the desired load (50Ω). It is clear that the output power of $\pi 1a$ was less sensitive in terms of load variation since the slope of dP_o/dR_L appeared flatter than $\pi 1b$. This is in line with the equation derived in Sect. 3, wherein the changes of output power were predicted by calculating the dP_o/dR_L of $\pi 1a$ and $\pi 1b$ using Eqs. (26) and (37), respectively. After substituting components parameter values presented in Table 1 into Eqs. (26) and (37), at $R_L = 50 \Omega$, the maximum dP_o/dR_L for $\pi 1a$ and $\pi 1b$ had been 0.113 and 0.189, respectively. These values confirmed that $\pi 1a$ was better than $\pi 1b$. In precise, the smaller the dP_o/dR_L , the better the system performance.

Table 1 Design specifications and performance comparison at $R_L = 50 \Omega$

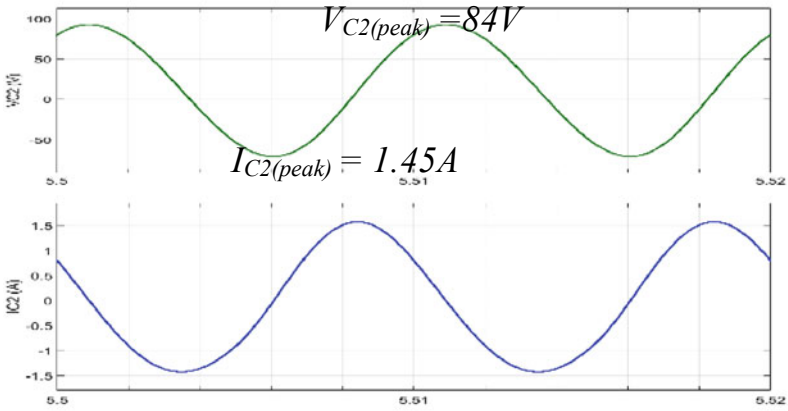
(a) Design specification				(b) Performance comparison					
Parameter		π 1a	π 1b	Parameter		π 1a		π 1b	
						Calculation	Measured	Calculation	Measured
f	MHz	1	1	$I_{RL(peak)}$	A	0.63	0.63	0.63	0.60
V_{dc}	V	12	12	$V_{RL(peak)}$	V	31.62	31.45	31.62	30.00
P_o	W	10	10	$V_{C3(peak)}$	V	31.62	31.50	110.32	110.00
D		0.5	0.5	$I_{C3(peak)}$	A	1.42	1.41	0.63	0.63
Q		10	10	$V_{C2(peak)}$	V	85.00	84.00	114.76	114.00
L_f	μ H	100	100	$I_{C2(peak)}$	A	1.55	1.45	0.93	0.93
C_1	pF	3520	3520	V_{dc}	V	12.00	12.00	12.00	12.00
L	μ H	13.22	13.22	I_{dc}	A	0.83	0.83	0.83	0.83
C_2	pF	2900	1295	P_{in}	W	10.00	10.00	10.00	10.00
C_3	pF	7130	913	P_{out}	W	9.96	9.89	9.99	9.77
R_s	Ω	8.30	8.30	η	%	99.60	98.91	99.92	97.66
R_L	Ω	50	50						

4.1.2 Capacitive Coupling Size, Voltage, and Current

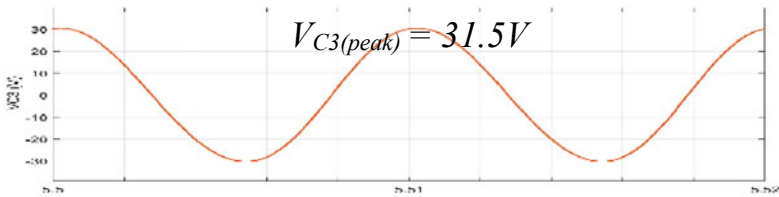
Table 1b shows a significant variance between the two matching types in terms of capacitive coupling size, voltage, and current. It was assumed that the simulation modelling circuit could be transformed to form a complete CPT system, in which one of the series capacitors would be replaced with a capacitive coupler. In this case, C_2 of π 1a and C_3 π 1b would be selected to be replaced with the capacitive coupler plates. By comparing in terms of coupler size, $C_2 = 2900$ pF of π 1a appeared larger than the coupling plate capacitance of π 1b, $C_3 = 913$ pF. Therefore, if the design size turns into a major constraint, CPT modelling with π 1a would be unsuitable for practical implementation due to its larger capacitance, which can result in a huge coupling plate area. Typically, the size of consumer electronics is relatively small, and increment of coupling capacitance by enlarging the area of the coupling plates is one that is far from desirable. Moreover, the additional capacitor C_2 in parallel with the load impedance serves to boost the transmitting voltage before the coupling plates, thus decreasing both the voltage and the current through the capacitive coupler (see Figs. 5ii, iii, and 6ii, iii). Therefore, due to the advantages mentioned above, CPT modelling with π 1b matching network had been selected in this study for practical implementation in further analysis.



(a) Switching waveforms

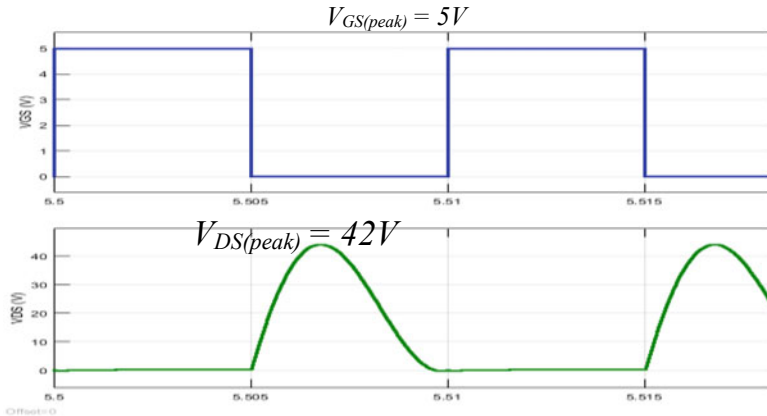


(b) Capacitive coupler waveforms: V_{C2} and I_{C2}

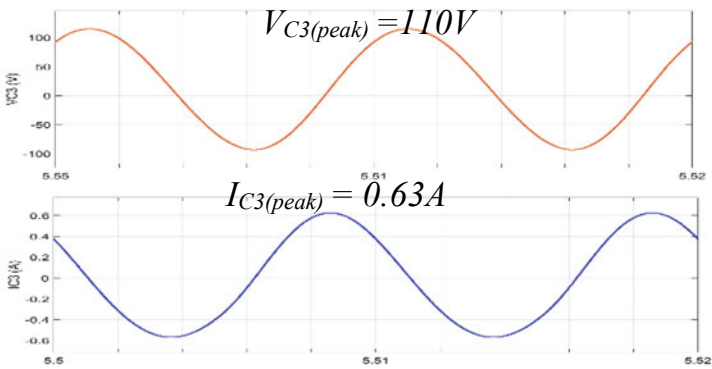


(c) Matching capacitor, C_3 voltage waveform

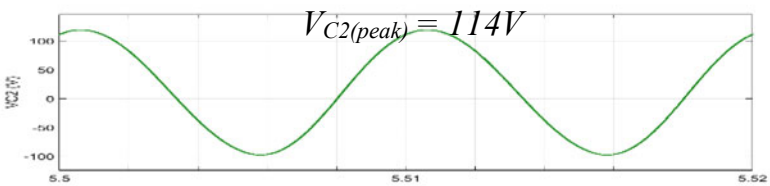
Fig. 5 π 1a matching network simulation results



(a) Switching waveforms: V_{GS} and V_{DS}



(b) Capacitive coupler waveforms: V_{C3} and I_{C3}



(c) Matching capacitor, C_2 voltage waveform

Fig. 6 π 1b matching network simulation results

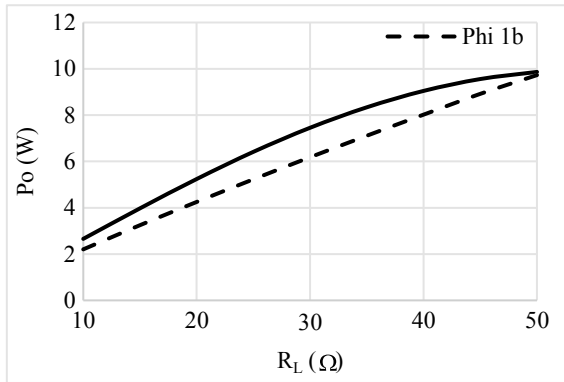


Fig. 7 Load variations

5 Experiment of Modified Class-E $\pi 1b$ for CPT System

This section presents the design implementation for a CPT system with $\pi 1b$ impedance matching circuit based on Class-E topology. The previous simulation analysis was performed under the assumption that the load connected to the receiver plate was a pure resistive AC load. This section presents the experimental work extended for the CPT system with DC load since almost all low power CPT systems used as battery charger typically require DC signal. The block diagram of the proposed CPT system is illustrated in Fig. 8. The Class-E inverter converted the input DC voltage to high frequency AC voltage. This was followed by the $\pi 1b$ matching network that boosted the voltage at the transmitter side. In order to transfer high power through the capacitive coupler, generation of high voltage was required between the capacitive plates to build up more electric fields. The capacitive interface consisted of two sets of parallel plates that were separated by air as the dielectric

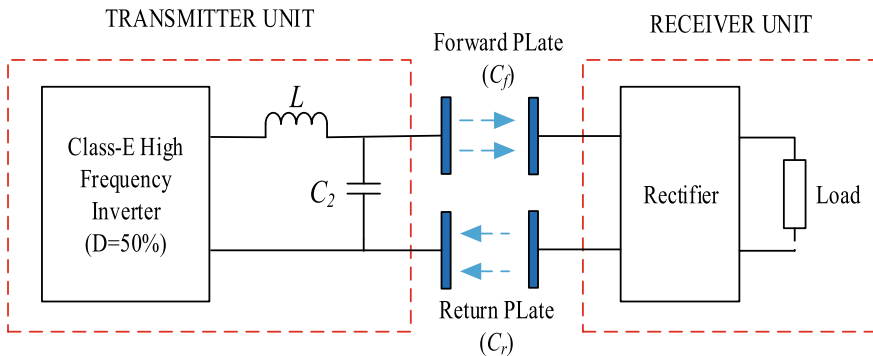


Fig. 8 Block diagram of the proposed CPT system

medium. The receiver side of the capacitive interface was connected to a full-bridge rectifier that supplied DC voltage to the load. This type of single phase full-bridge rectifier applied four individual rectifying diodes connected in a closed loop “bridge” configuration to produce the desired output. The main advantage of this bridge circuit is that it does not need a special centre tapped transformer, thus reducing its size and cost. As for the CPT operation, the load was actually connected in series with the diode rectifier, as well as both of forward and return capacitive plates.

As for the capacitive coupling plates design, the matching capacitor C_3 was split into two pairs of coupling plates to form two pairs of coupling capacitors; C_f and C_r , in the $\pi 1b$ output matching network. The capacitive interface was implemented with rectangular PCB plate capacitors and was separated by a piece of thick paper (card stock). In addition, the distance of the initial coupling plate was assigned equivalent to the thickness of the card stock, which was 0.25 mm, along with dielectric constant at 2.5. Therefore, the size of plate shown in Fig. 8 was 0.021 m² with combined interface capacitance at 913 pF. Both the coupling plates at the receiver side were connected to a rectifier circuit with 50 Ω load resistor rated at 13 W. The circuit of the bridge rectifier consisted of four high frequency diodes. Based on prior outcomes, $I_{out(peak)} = 0.63$ A and $V_{out(peak)} = 31.62$ V. Hence, in practical hardware design, Schottky diode with 1N5819 code was deemed as safe to be employed in this application since it has an average rectified forward current at 1 A and 40 V maximum working peak reverse voltage. The computer simulation of the complete CPT system using simulation software program and the prototype are displayed in Figs. 9 and 10, respectively. Several essential points of observation emerged from the simulation and prototype testing were related to ZVS conditions and DC-DC efficiency.

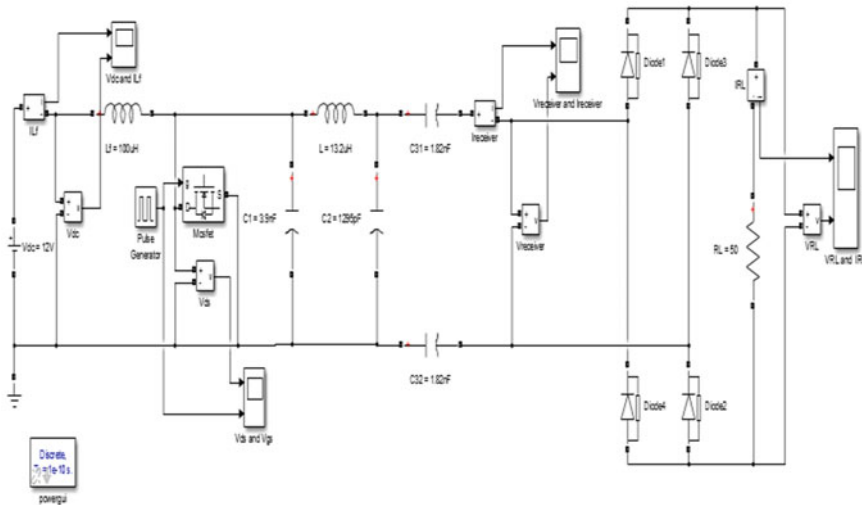


Fig. 9 Circuit simulation

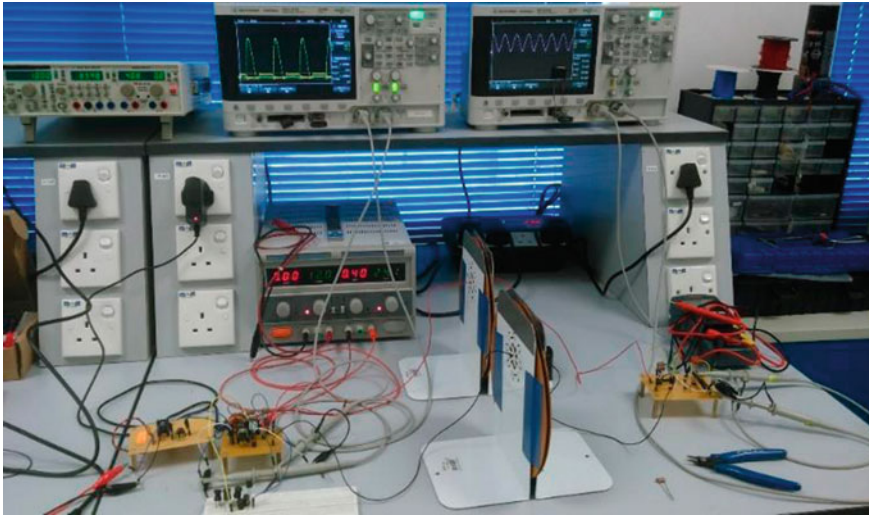


Fig. 10 Experimental setup

5.1 Performance Analysis

The analysis of the results obtained from both simulation and experiment is discussed in this section. The system performance parameters tabulated in Table 2 were analysed and compared between theoretical, simulation, and experimental works.

Table 2 System performance parameters

Parameter		Theoretical	Simulation	Experiment
V_{dc}	V	12.00	12.00	12.00
I_{dc}	A	0.83	0.83	0.80
$V_{RL(peak)}$	V	31.60	31.60	30.00
$I_{RL(peak)}$	A	0.64	0.64	0.60
$V_{RL(avg)}$	V	22.36	21.72	21.00
$I_{RL(avg)}$	A	0.45	0.43	0.42
P_{in}	W	10.00	10.00	9.60
P_o	W	10.00	9.34	8.82
η	%	100.0	93.40	91.90

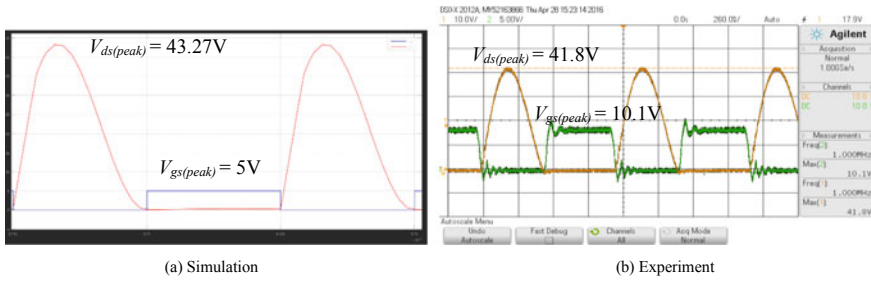


Fig. 11 Zero voltage switching waveforms

5.1.1 ZVS

Based on the results presented in Fig. 11, the maximum voltage values across the MOSFET and the shunt capacitor during turn off were $V_{ds(peak)} = 43.27$ V for simulation and $V_{ds(peak)} = 41.8$ V for experiment. Meanwhile, during turn on, $V_{ds(peak)} = 0$ V for both simulation and experiment. Hence, both waveforms proved that the ZVS condition was satisfied as no overlapping was noted in time between MOSFET drain voltage and gate voltage.

5.1.2 DC-DC Efficiency

Figure 12a illustrates the measurement results for both input and output power in absence of a smoothing capacitor. First, for the simulation results, the DC input power, P_{in} , was calculated using $P_{in} = V_{cc} \times I_{dc} = 12 \times 0.83 = 10$ W. Meanwhile, as for the DC output power, as given in Fig. 12a(i), P_o was calculated using $P_o = V_{(avg)} \times I_{(avg)} = 0.636 \times 31.6 \times 0.4 = 8.08$ W. This gave 80.8% simulation efficiency. Second, as for the experimental results displayed in Fig. 13a(ii), P_i and P_o were 9.60 W and 7.28 W, respectively. Therefore, in terms of efficiency, the CPT system circuit without a smoothing capacitor could only generate 81% efficiency.

Theoretically, the maximum efficiency of full-wave rectifier is 81.2%. This is due to the non-filtered DC output voltage. As noted in Fig. 12a, the full-bridge rectifier produced an output wave at every cycle, thus providing an average DC value ($0.637 \times$

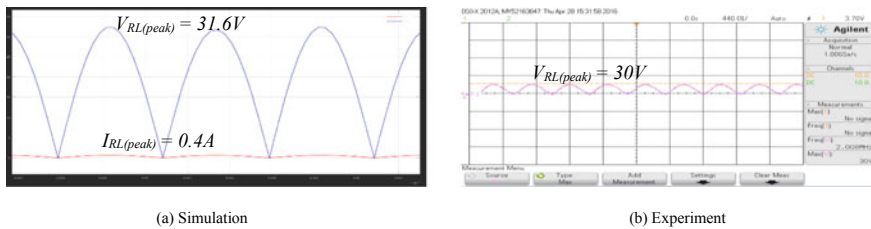


Fig. 12 Waveforms of output current and voltage without smoothing filter

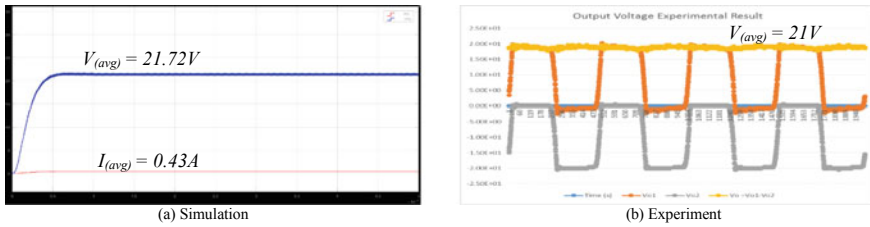


Fig. 13 Waveforms of output current and voltage with smoothing capacitor

$V_{RL(peak)}$) with less superimposed ripple, while the output waveform was twice that of the frequency of the input supply frequency. Hence, the increment in average DC output level appeared to be even higher after connecting a suitable smoothing capacitor across the output of the bridge. In this case, a 0.1 μ F smoothing capacitor was selected. Figure 12b shows the simulation output power that hiked from 8.08 to 9.34 W. On the other hand, the experiment output power increased from 8.10 to 8.82 W. Therefore, the practical work indicates that the proposed system has the ability to transfer 8.82 W to the 50 Ω load with 91.9% efficiency. As signified by the results tabulated in Table 2, all the simulation and experimental results are consistent with those found theoretically.

5.1.3 Load Characteristics

Figure 14a illustrates the comparison of the CPT system with and without π 1b matching network. A clear trend of decreasing output power curve was observed as the load increased for the CPT system without any matching network. By taking

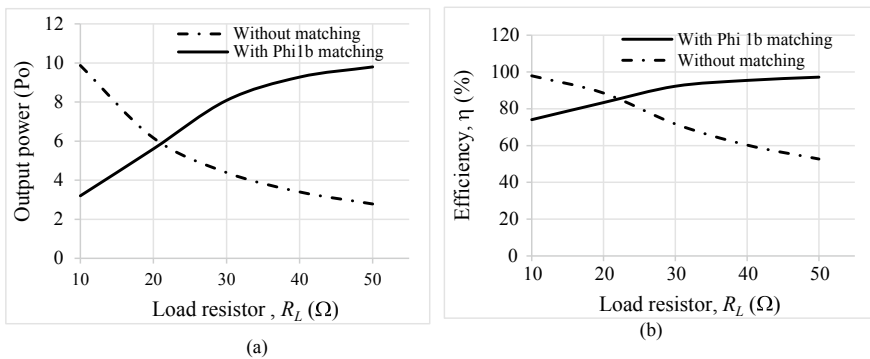


Fig. 14 a P_o versus R_L , b η versus R_L

derivative to equation of basic Class-E circuit without matching circuit, the dP_o/dR_L appeared negative. This negative sign signified the decrease in output power as the load exceeded the optimum load. On the contrary, as for the CPT with matching network, the derived dP_o/dR_L as given in Eq. (37) appeared to be positive, thus the increasing output power trend when the load was hiked up to 50Ω , with efficiency above 70% for all the tested load resistors, as presented in Fig. 14b. These findings serve as evidences that the performance of CPT system with $\pi 1b$ matching network is enhanced, in comparison to that without a matching network.

5.1.4 $R_L = X_{C3}$ Case

In relation to Eq. (37), the size of the coupling plates, C_3 , must be as small as possible in order to gain smaller dP_o/dR_L . The smallest C_3 was discovered when $R_L = X_{C3}$ and this consequently increased the additional capacitor, C_2 , to a possible larger value. Hence, the application of $R_L = X_{C3}$ generated the following:

$$R_L = X_{C3(\max)} = \frac{R_S(Q_L - 1.1525)^2 + 1}{2} = 329$$

The resistance that can be matched by the abovementioned case is $R_S < R_L < R_{L\max}$. Upon $R_{L(\max)} = 329 \Omega$, the values of the new components for $\pi 1b$ matching network were $C_2 = 1897.0$ pF and $C_3 = 483.4$ pF. As a result, the new $dP_o/dR_L = 0.0068$. This means; the changes of output power with respect to load variation had reduced. In fact, this is consistent with the plots presented in Fig. 15a, b. The maximum 10 W output power was measured across 329Ω at 1 MHz operating frequency. Due to smaller dP_o/dR_L , the current flowed through the coupling plates almost constant, while the load or the coupling capacitance seemed to be varied, as illustrated in Fig. 15c. This situation is beneficial for certain charging applications, wherein voltage-to-current conversion and load independent current output are desirable. For instance, a constant-current output is preferred for driving a light-emitting diode for stable luminance [21].

6 Conclusion

The analyses of the CPT system based on Class-E topology with $\pi 1a$ and $\pi 1b$ matching networks are presented in this work. Based on the performance comparison, $\pi 1b$ offered more advantages in terms of practical implementation due to its smaller coupling capacitance with a wider coverage of load resistor, in comparison to the coupling size explored in previous work required 2.17 nF to transfer the same amount of output power. The function of Class-E with $\pi 1b$ matching network is mainly to help the CPT system in maximising power transfer by providing downward impedance transformation. By utilising the proposed method, the ZVS condition can still be

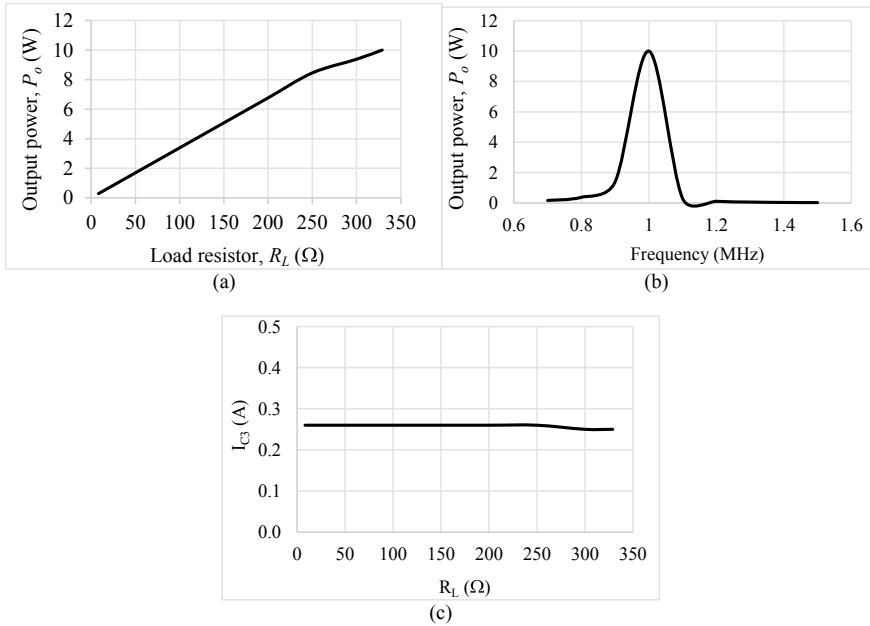


Fig. 15 Performance analysis when $R_L = X_{C3}$

achieved even if the load is larger than that optimum. The experimental work showed that 10 W output power was transferred successfully to the 50 Ω load with only 913 pF capacitive coupling plates at 91.9% efficiency. Apart from that, the analysis of dP_o/dR_L is also introduced in this research as it serves as a useful guideline in predicting the performance of the system. In the near future, the proposed work will be implemented for wireless rotary applications.

References

1. Marques EG, Mendes AMS (2015) Comparison of magnetic coupling structures for IPT systems. *COMPEL* 34(2):514–530
2. Sanni A, Vilches A, Toumazou C (2012) Inductive and ultrasonic multi-tier interface for low-power, deeply implantable medical devices. *IEEE Trans Biomed Circuits Syst* 6(4):297–308
3. Nair V, Choi J (2015) An integrated chip high-voltage power receiver for wireless biomedical implants. *Energies* 8(6):5467–5487
4. Frey JE, Kreitz A, Scheibe G (2006) Unplugged but connected. *EngineerIT* 44–45
5. Choi BH, Nguyen DT, Yoo SJ, Kim JH, Rim CT (2014) A novel source-side monitored capacitive power transfer system for contactless mobile charger using class-E converter. In: 2014 IEEE 79th vehicular technology conference, vol 2015, pp 8–12
6. Hui SY (2013) Technology for portable electronic products and Qi. *Proc IEEE* 101(6):1290–1301

7. Goto K, Nakagawa T, Nakamura O, Kawata S (2001) An implantable power supply with an optically rechargeable lithium battery. *IEEE Trans Biomed Eng* 48(7):830–833
8. Deng J, Lu F, Li S, Nguyen TD, Mi C (2014) Development of a high efficiency primary side controlled 7 kW wireless power charger. In: 2014 IEEE international electric vehicle conference (IEVC)
9. Covic JTBGA (2013) Inductive power transfer. *Proc IEEE* 101(6):1276–1289
10. Mi C (2015) Charging applications, no 619, pp 1–35
11. Huang L, Hu AP, Swain A, Kim S, Ren Y (2013) An overview of capacitively coupled power transfer; A new contactless power transfer solution. In: 2013 IEEE 8th conference on industrial electronics and applications, pp 461–465
12. Lu F, Zhang H, Hofmann H (2015) A high efficiency 3.3 kW loosely-coupled wireless power transfer system without magnetic material. In: 2015 IEEE energy conversion congress and exposition (ECCE), pp 2282–2286
13. Li S, Mi CC (2015) Wireless power transfer for electric vehicle applications. *IEEE J Emerg Sel Top Power Electron* 3(1):4–17
14. Dai J, Ludois DC (2015) A survey of wireless power transfer and a critical comparison of inductive and capacitive coupling for small gap applications. *IEEE Trans Power Electron* 30(11):6017–6029
15. Theodoridis MP (2012) Effective capacitive power transfer. *IEEE Trans Power Electron* 27(12):4906–4913
16. Kazmierkowski MP, Moradewicz AJ (2012) Contactless energy transfer (CET) systems—a review. In: 2012 15th international power electronics and motion control conference, pp Session 3–1–Session 3–6
17. Yusop Y, Saat MSM, Husin SH, Nguang SK, Hindustan I (2017) Performance assessment of class-E inverter for capacitive power transfer system. *COMPEL-Int J Comput Math Electr Electron Eng* 36(4) ISSN: 0332–1649
18. Liu C, Hu AP, Budhia M (2010) A generalized coupling model for capacitive power transfer systems. In: IECON 2010—36th annual conference on IEEE Industrial Electronics Society, pp 274–279
19. Lu F, Zhang H, Hofmann H, Mi C (2016) A CLLC-compensated high power and large air-gap capacitive power transfer system for electric vehicle charging applications. In: 2016 IEEE applied power electronics conference and exposition (APEC), vol 2016, pp 1721–1725
20. Yusmarnita Y, Saat S, Hamidon AH, Husin H, Jamal N, Kh K, Hindustan I (2015) Design and analysis of 1 MHz Class-E power amplifier 2 circuit description. *WSEAS Trans Circuits Syst* 14:373–379
21. Qu X, Zhang W, Wong SC, Tse CK (2015) Design of a current-source-output inductive power transfer LED lighting system. *IEEE J Emerg Sel Top Power Electron* 3(1):306–314

Application of Smart Rotatable Solar Powered Electrocoagulation System for Pollutants Removal from Palm Oil Mill Effluent



B. K. Zaied, Mohd Nasrullah, L. Singh, and A. W. Zularisam

Abstract Palm oil mill effluent has become a serious concern to society, which needs intensive care and treatment for the environmental pollution removal and preservation of the inadequate water source. The electrocoagulation treatment process has earned special attention in treating palm oil mill effluent (POME) because it has great versatility and ecological affinity. This study was regulated to explore the application of the electrocoagulation process continuously in treating POME by using a smart rotatable solar powered electrocoagulation system. This system is initiated to be suitable as a single unit process for the treatment of POME since utilizing solar energy reduces operational costs and shows enhanced treatment efficiency. The effects of various operational conditions for instant current density (1, 2, 3, 4, 5 A/dm²), reaction time (5, 10, 15, 20, 25, 30 min), initial pH (3, 5, 7, 9, 11), inter-electrode spacing (0.01, 0.02, 0.03 m) and electrode configuration (Fe/Fe, Al/Fe, Fe/Al, Al/Al) on chemical oxygen demand (COD), biochemical oxygen demand (BOD), and total solids (TS) removal were examined. The maximum COD, BOD and TS removal efficiency of 89, 96, and 92% respectively were obtained with the electric current density of 4 A/dm², the reaction time of 30 min, effluent pH of 7, the inter-electrode spacing of 0.01 m, and electrode arrangement of Fe/Fe. The experimental outcomes presented that the smart rotatable solar powered electrocoagulation system could be utilized as an effective and ecologically friendly technology for successful pollutants removal.

Keywords Electrocoagulation · Palm oil mill effluent · Pollutants removal · Solar energy · Rotatable solar panel

1 Introduction

The palm oil mill industry is the largest global agricultural-based industry. This industry has been developed abruptly, making Malaysia the biggest crude palm

B. K. Zaied · M. Nasrullah (✉) · L. Singh · A. W. Zularisam
Faculty of Civil Engineering Technology, Universiti Malaysia Pahang (UMP), 26300 Gambang,
Kuantan, Pahang, Malaysia
e-mail: nasrul.ump@gmail.com

oil maker in the earth and lead fundamental financial importance in the past few centuries. Several operational procedures are involved in extracting palm oil, for example, cleansing, stripping, absorption, sorting, decontamination, and void aeration [1]. So, all these growing processing and production of palm oil, as well as its by-products, have generated a large volume of wastewater, named as palm oil mill effluent (POME) which unfavorably affects animal and aquatic life expectancy [2]. Malaysia alone produces POME about 43.29 million cubic meters yearly [3]. This extreme amount of wastewater is essential to be treated for maintaining the discharge standard to keep safe the environment.

Numerous wastewater treatment techniques have been stated concerning on the treatment of POME like as anaerobic digestion [4], aerobic oxidation [5], aerobic digestions [6], chemical flocculation and coagulation [7], membrane anaerobic system [8], up-flow anaerobic sludge fixed film (UASFF) reactor [9], immobilized up-flow anaerobic sludge blanket (UASB) reactor [10], adsorption [11] etc. Though the above treatment process is employed as an effective method, they needed a long treatment duration or a considerable amount of sample. All these are less efficient for the treatment of POME when it expects a short duration for the treatment process and ecological affinity with fewer chemical effects [2]. It is essential to find another treatment method as a substitute for the traditional wastewater treatment system to comply with the necessities and rules of safe discharge of wastewater [12]. This alternate treatment method must comprise several operating factors for more suitable to operate, quick settlement time, easy installation and maintenance, energy and cost-efficient, broadly applicable for the numerous pollutant removal and required less labor [13].

The electrocoagulation process has the potential for the treatment of POME since it is becoming a prominent and convenient treatment process having the capability for the removal of various pollutants [13]. This process has several advantages like a short settlement period, quick sedimentation of flocs, fewer production of sludge, less or absence of added chemicals, simple apparatus and easy operational procedure [14]. According to the literature survey, limited research has been studied on the electrocoagulation process for the treatment of POME. Nasution et al. studies reported that by electrocoagulation the optimal operating time for the treatment of POME was 6 h for getting 42.94% of removal efficiency for COD [15]. Agustin et al. testified that the removal efficiency of COD was about 30.44% in 6 h [16]. The highest COD removal efficiency was found 74% by applying electrode configuration of using steel wool material, vertical direction, and monopolar-series arrangement [17]. Even though the operational period is relatively quicker than the conventional treatment of POME, still it is required a long duration due to the inadequacy of several operational conditions that should be taken care of. After all, for the treatment of POME comprehending high-strength pollutants concentration, the necessity of functional application is needed to boost the electrocoagulation treatment process.

Over the past decades, increased environmental pollution and social concern have resulted in the expansion of new forms of sustainable energy that can be utilized as an electricity supply source in electrocoagulation reactors. Photovoltaic (PV) solar energy is the sustainable energy sources that are the widely popular and most studied

system. As PV is fully independent and ecologically friendly, this is suitable for small operations, particularly in rural areas where there is no electrical grid. Recently, many kinds of literature have been reported that electrochemical systems can be powered by PV energy, one of them is employed on electrocoagulation [18, 19]. In PV systems, the use of solar trackers can help in minimizing the angle of incidence (the position that a solar light ray can make by a line vertical to the PV panel surface) in between the received sunlight and PV panel which will enhance the amount of collecting solar energy. Solar PV systems and solar thermal have optics which receive solar light directly, and so usage of trackers need to be adjusted properly to collect solar energy. Throughout the day, trackers make changes orientation of solar PV panels to the direction of the direct solar light and follow the path of the sun for maximizing the capture of solar energy.

The main objective of this research is to design and develop the smart rotatable solar powered electrocoagulation system for the pollutants removal from POME affecting the various operating parameters. During this study, conventional electricity will be replaced by solar PV powered electricity. The solar PV panel will be mounted on a platform that can rotate with the help of a servo motor. It will increase the receive of solar light to generate optimum power for the electrocoagulation process. The application of solar light for providing electricity is a sustainable option made by a widespread technology that makes usage of the physical proposition of sustainable energy conversion because of the prevalence of the easy solar systems operation concerning other sustainable energy sources.

2 Materials and Methods

2.1 Feedstock Collection and Preparation

A 100 L sample of raw POME was accumulated in sample collection container from the anaerobic pond of the LKPP Corporation Sdn. Bhd., No. 45/4, Jalan Teluk Sisek, 25000 Kuantan, Pahang, Malaysia. POME sample was subjected to the simple screening to eliminate coarse materials. The POME was placed in compact frost containers and transferred to the preservation cold room. The temperature was maintained at 4 °C during preservation. The characteristics of POME are obtained with the mean values \pm standard deviation in Table 1.

Table 1 Characteristics of palm oil mill effluent (POME)

Parameter	Value
pH	4.6 \pm 0.07
COD (g/L)	28.34 \pm 0.3
BOD (g/L)	15.28 \pm 0.2
TS (g/L)	39.75 \pm 0.3

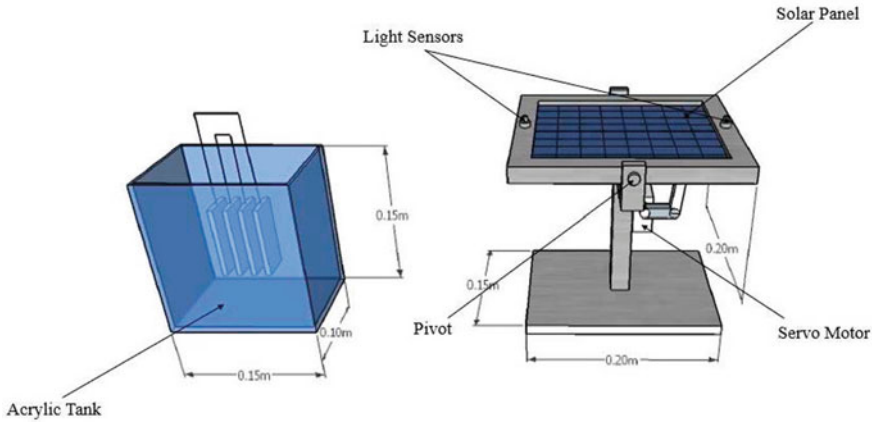


Fig. 1 Schematic diagram of smart rotatable solar powered electrocoagulation system

2.2 *Electrocoagulation Reactor Design and Development*

The project was proposed to attain the demand in the treatment of POME more efficiently. There are two main components in this project to make the system function efficiently, which is the electrocoagulation process and the photovoltaic solar panel. In the electrocoagulation part, as POME passes through the electrocoagulation cell, multiple reactions take place simultaneously. First, two metal electrodes are driven into the reactor containing wastewater. On the surface of the cathode, wastewater is hydrolyzed into hydrogen gas and hydroxyl groups. In the meantime, metal ions generated from anode flow spontaneously to coagulate pollutants of wastewater into solid particles. As the reaction continues, large flocs formed which contained solid particles, heavy metals, and other contaminants. For the photovoltaic, the light sensors act to detect the ambient solar light level then transmit the reading of light detected to the receiver part to activate the movement of the rotatable solar panel accordingly to the maximum light radiation. Thus, the solar PV panel will absorb the sunlight as a source of energy to generate electricity for the electrocoagulation process. Figure 1 is a schematic diagram of this electrocoagulation system which is used in this research study.

2.3 *Electrocoagulation Setup*

The experimental setup will consist of a $0.15\text{ m} \times 0.10\text{ m} \times 0.15\text{ m}$ acrylic tank as a reactor for holding a POME sample of 1.0 L, an electric power source from the solar panel, four electrodes, and POME sample. The dimension of electrodes will be $0.03\text{ m} \times 0.07\text{ m}$ each. The electrode thickness was 0.005 m and the inter-electrode spacing will be 0.01, 0.02, 0.03 m. The electrode arrangements will be

monopolar series (MP-S) connection. The electrode area which dipped in the sample solution will be $0.03 \text{ m} \times 0.06 \text{ m}$. Before each run, electrodes will be dipped in 15% hydrochloric acid and washed with water in order to remove dust from the electrode plates and then the electrodes will be weighed after drying. Samples will be taken at 1.0 h intervals of 8.0 h of experiment time. The removal of COD, BOD, and TS from POME will be investigated by operating the electrocoagulation process. DR 5000 spectrophotometer (HACH) was used to measure COD using 8000 and 10,212 methods, US EPA. Sension6 DO meter (HACH) was used to measure dissolved oxygen (DO) for a 5-day BOD test using the 5210 B method, APHA. DR 5000 spectrophotometer (HACH) was used to measure SS using 8006 photometric methods. Mettler Toledo Seven Compact pH meter S220 was used to measure pH.

2.4 Analytical Methods

The COD, BOD, TS, and pH measurements were followed by the standard methods of water and wastewaters (APHA, 2016). Pollutants removal efficiency, electrode consumption, and space time yield were determined with the subsequent equations:

The performance of this treatment process was assessed based on the COD, BOD and TS removal efficiencies, that will be calculated according to;

$$\text{Removal Efficiency}(\%) = \frac{C_0 - C_t}{C_t} \times 100 \quad (1)$$

where C_0 and C_t are the initial and final values of COD, BOD and TS respectively. To investigate the electrode consumption, subsequent equations are followed;

$$\text{Electrode Consumption} \left(\frac{\text{g}}{\text{m}^3} \right) = \frac{I \times t \times M}{n \times F \times V} \quad (2)$$

where I is the measurement of current usage (A), M is the atomic mass of electrode in g/mol, t is the operational time (s), F is Faraday's constant (96,487 C/mol), V is volume (m^3), and n is the number of the transferred electron ($n = 2$).

The space time yield is illustrated as the product mass (m_p) found per reactor volume (m^3) and time (h). Product mass is calculated by multiplying its relative atomic mass by the molar mass constant (1 g/mol). Y_{STY} is determined with the following Eq. 3.

$$\text{Space Time Yield, } Y_{STY} \left(\frac{\text{kg}}{\text{m}^3 \text{h}} \right) = \frac{m_p}{V \times t} \quad (3)$$

2.5 Statistical Analysis

The data were analyzed for three replicates by using Microsoft Excel 2016. All necessary statistical data were derived in this software. The mean, standard deviation, and standard error results were calculated from replicates by OriginPro 9.1 and applied to each figure and table values.

3 Results and Discussions

3.1 Effects of Current Density

Electric current density regulates the coagulant dosage, bubble formation size, rate and flocs growth and thus removal effectiveness of the pollutants can be easily influenced in the electrocoagulation treatment process [20]. The experimental results shown in Fig. 2 stated that the COD, BOD, and TS removal were improved from 64 to 89%, 76 to 96% and 70 to 92% respectively at a current density of 1 to 5 A/dm². Along with Faraday's Law, the electron charge transferred to the effluent was linearly proportional to the dissolution amount of the electrode (Fe or Al). So, it suggests that the removal of COD, BOD, and TS by electrocoagulation process may led by the production of hydroxyl metal.

Consequently, when current densities were increased, the high coagulation and high removal efficiency of the pollutants were achieved. All tests were executed

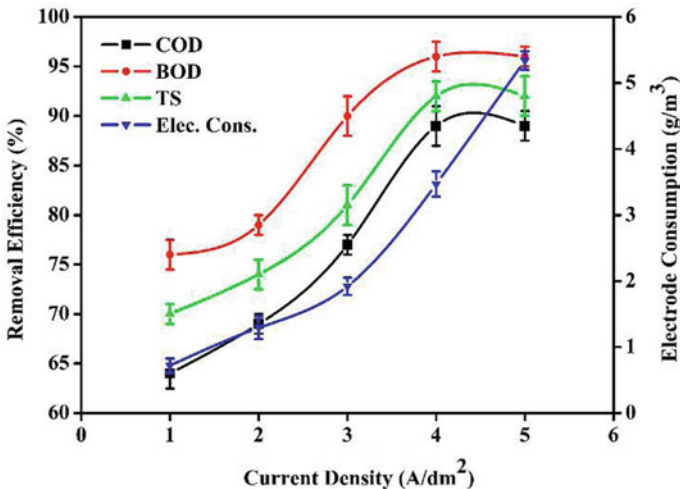


Fig. 2 Effects of current density (A/dm²) on COD, BOD and TS removal efficiency (%) and electrode consumption (g/m³)

at a 4 A/dm² of current density. Moreover, the consequence of current density in consuming electrode materials was also investigated, which are shown in Fig. 2. It can be determined that the current density was increasing from 1 to 5 A/dm², electrode consumption also increased from 0.72 to 5.34 g of electrode materials. The electrode consumption increases depending on the different operating parameters and characteristics of wastewater. According to Kobya et al., the experimental amount of electrode consumption is more than the hypothetical amount of electrode consumption [21]. The actual consumption of the electrode was 3.18 g in during operation at 4 A/dm² of current density where the theoretical consumption of electrode was measured with Eq. 2 and the values were 3.47 g of electrode materials.

3.2 Effects of Reaction Time

From Fig. 3, it can be easily said that, if the reaction time of the electrocoagulation process increases, the removal efficiency of COD, BOD, and TS are also changed. It affects the process efficiency on electrocoagulation as the more duration takes, the more formation of metal and hydroxyl ions happen on both electrodes. The impacts of reaction time were investigated at a fixed electric current density of 4 A/dm². In these trials, two steps happened in the operational process. In the first step, it is generally quick, while for the second step, it is consumed fairly long duration. As reported to the effect shown in Fig. 3, the reaction time of 20 min is satisfactory for almost full process efficiency of COD, BOD, and TS. Electrocoagulation process efficiency is almost similar and has inconsequential upgrading after 20 min of reaction time.

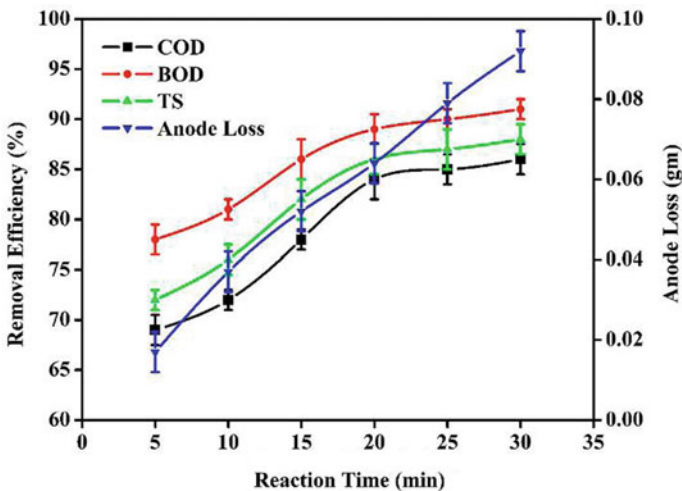


Fig. 3 Effects of reaction time (min) on COD, BOD and TS removal efficiency (%) and anode loss (g)

According to Nasrullah et al., treatment efficiency remains almost constant and has insignificant improvement above 20 min [2]. So, the optimal reaction time is 20 min.

The consumptions of the electrode are also linearly proportional to reaction time. Figure 3 displayed that the electrode consumption elevated almost gradually in every single minute. The time increment from 5 to 30 min caused a weight decreasing from 88.430 to 88.355 g anode electrode, which is equivalent to anode consumption of 15 mg/L of POME every 5 min. According to Nasrullah et al., increasing in time from 5 to 30 min resulted in a decreasing weight of cathode from 88.432 to 88.356 g stainless steel anode which is increment average of anode consumption is about 15 mg/L sewage water every 5 min [2]. Hence, this consequence specifies that reaction time remains a very significant operational parameter as its effects on the financial employability of the electrocoagulation system for the POME treatment.

3.3 Effects of Effluent pH

Effluent pH is considered as another vital operational condition which can directly impact the electrocoagulation operation. To investigate the effects of solution pH for the efficient treatment process, the pH of the POME was accustomed in between 3 and 11. The efficiency of COD, BOD, and TS removal with changing the initial pH is shown in Fig. 4. Once the initial pH was from 3 to 7, the removal efficiency of COD, BOD, and TS increased from 78 to 89%, 83 to 96% and 81 to 92% respectively. But, if the initial pH was from 7 to 11, the removal efficiency of COD, BOD and TS decreased from 89 to 72%, 96 to 77% and from 92 to 75% respectively. The experimental outcomes specified that at high initial pH of the effluent, some hydroxide ions are oxidized at the anode electrode which reduces the iron ions formation and thus the removal efficiency of COD, BOD and TS decreased. Additionally, $\text{Fe}(\text{OH})_6^3$ or $\text{Fe}(\text{OH})_4^-$ ions can be existing at a higher value of effluent pH, which reduces the removal efficiency of COD, BOD, and TS. At a lower value of the initial pH of the effluent, the solution ions reduced H_2 production at the cathode where the equal amount of hydroxide ions cannot be formed. At a lower value of initial pH of the effluent, $\text{Fe}(\text{OH})^{2+}$ and $\text{Fe}(\text{OH})_2^+$ were produced which was detrimental for the precipitation of the pollutants. In a word, the high removal efficiency of COD, BOD, and TS was found in a pH of 7, which is also stated by various researchers [22].

The impact of effluent pH for determining the space time yield (Y_{STY}) was measured, and the found results are also presented in Fig. 4. Space time yield within the molecule implies directly to the rates of reaction. This concept examines the light of rate equations in flow systems. That's why, this concept is important for determining pH effects in electrocoagulation as it is predicting reaction order and finally helps to predict the removal of COD, BOD, and TS. It was observed from the figure that the Y_{STY} increased from 25 to 58 $\text{kg}/\text{m}^3\text{h}$ if the pH value is increased from 3 to 7. But, the more pH value increase from 7 to 11 happened reduction of Y_{STY} from 58 to 45 $\text{kg}/\text{m}^3\text{h}$ respectively. So, the optimal Y_{STY} was found at the initial effluent pH value of 7.

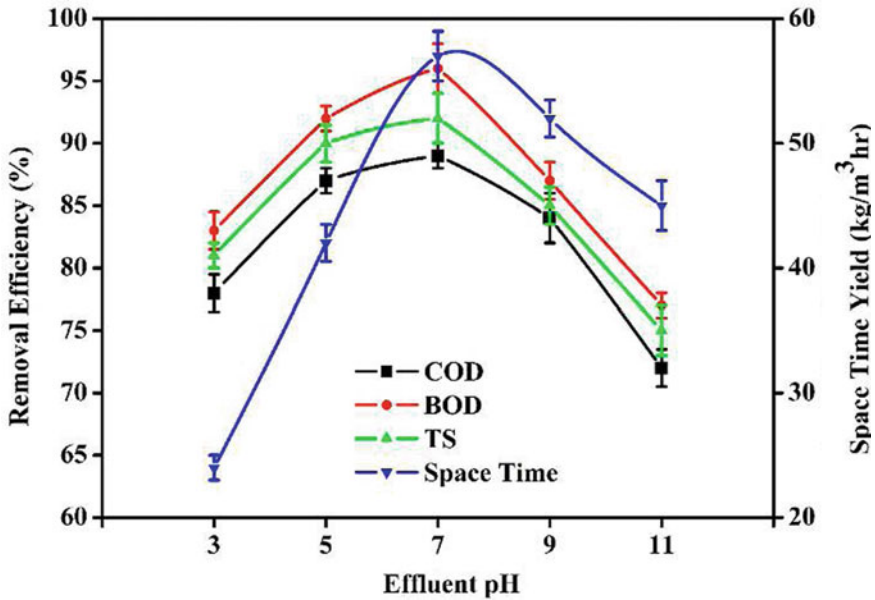


Fig. 4 Effects of effluent pH on COD, BOD and TS removal efficiency (%) and space time yield ($\text{kg}/\text{m}^3\text{hr}$)

3.4 Effects of Inter-electrode Spacing

The experimental conditions were led with the variation of the inter-electrode spacing between the electrodes maintaining other operational conditions fixed. The outcomes of this trails are exposed in Fig. 5 which displays that if the inter-electrode spacing between the electrodes is increased as of 0.01 to 0.03 m, the COD, BOD, and TS removal efficiency were decreased from 89 to 79%, 96 to 84% and 92 to 82% respectively. If the spacing between the electrodes increases, the electric current density decreases, and thus, the electric voltage is also increased. Then again, the IR-drop rises because the spacing between the two electrodes increases. With increasing for the inter-electrode spacing between the electrodes, less interaction happened for ions and electrostatic magnetism [23]. To achieve the expected removal efficiency of pollutants from POME, the inter-electrode spacing between the electrodes needs to be minimized.

According to Ohm's law, the measurement for electricity through an electrode is linearly proportional to the voltage applied to it. This electric current relationship is stated as:

$$V = IR \quad (4)$$

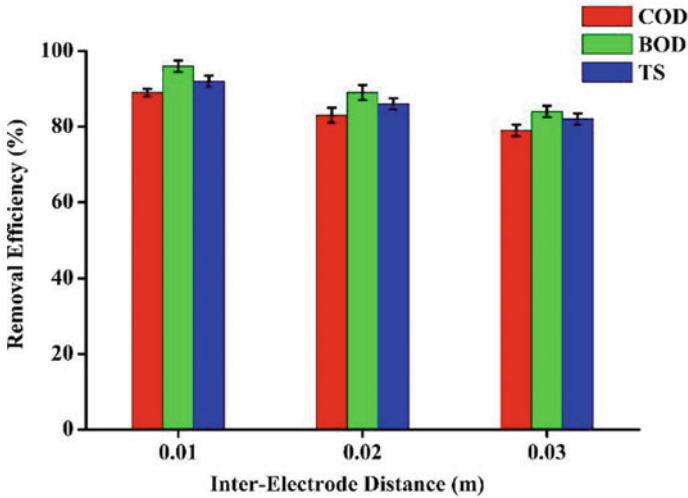


Fig. 5 Effects of inter-electrode distance (cm) on COD, BOD and TS removal efficiency (%)

IR drop or ohmic potential drop may have a noteworthy impact on electrochemical capacities. IR drop is a potential drop because of the resistance of the solution. The potential difference is essential to transfer metal ions through the solution. The difference in IR drop is measured by the following equation:

$$\eta IR = I \frac{d}{A \cdot k} \quad (5)$$

where, I = current (A), d = spacing between the anode and cathode (m), A = active area of anode (m^2), k = specific conductivity (10^3 mS/m).

From the above equation, it concludes that the IR drop is increased if the spacing of the electrodes is increased. During the investigation, the electric current all at once drop after few minutes of operation, therefore increased voltage has been applied to maintain the constant electricity. This situation happens because of the rising of IR drop, which governs the rate of anodic oxidation inhibition. So, the rise of IR drops by increasing the spacing between the electrodes is not suggested in the electrocoagulation procedure.

3.5 Effects of Electrode Arrangements

Four different arrangements of the electrode (Fe and Al electrodes) were examined, and the results found are displayed in Fig. 6. From the figure, the electrode materials had a distinct impact on the removal of COD, BOD, and TS. The removals efficiency for the COD, BOD, and TS for the anode/cathode arrangements of electrode for

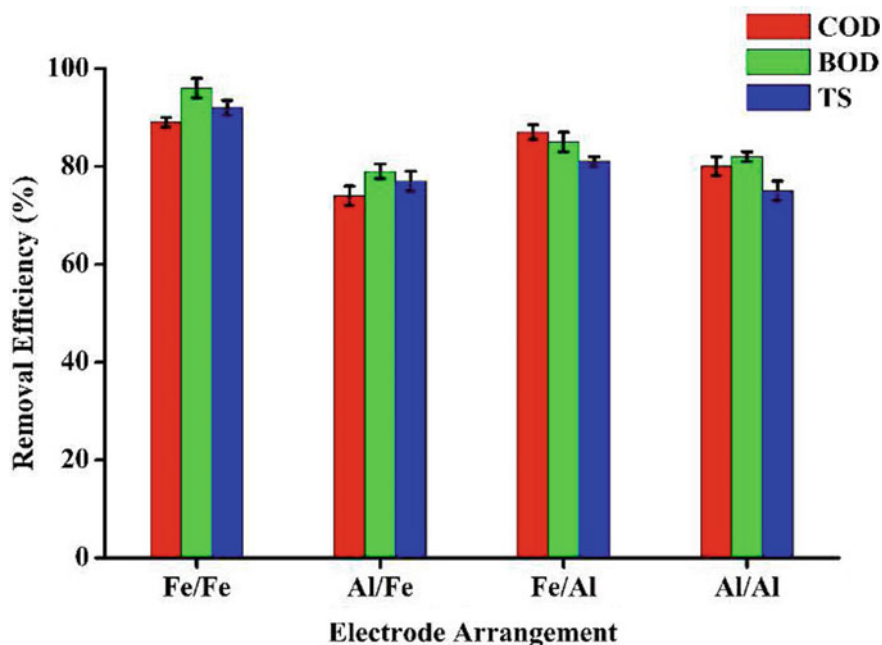


Fig. 6 Effects of electrode arrangement on COD, BOD and TS removal (%)

Fe/Fe were about 89, 96 and 92% respectively, for Al/Fe were about 74, 79 and 77% respectively, for Fe/Al was about 87, 85 and 81% respectively and for Al/Al were about 80, 82 and 75% respectively. So, the experimental outcomes are clearly indicating that the Fe/Fe electrode arrangements generated higher removal of pollutant concerning other arrangements.

The promising Fe and Al ions are effective coagulants to precipitate pollutants. As per the chemical reaction in the electrocoagulation process, Fe has a greater potential for oxidation concerning Al, and thus the rate of Fe oxidation is more than Al electrodes [24]. The formation of Fe coagulant is about three times more than Al that contributes to greater removal efficiency of COD, BOD, and TS at the Fe electrode concerning the anode electrode while the current was passed with the same density.

3.6 Electrocoagulation Kinetics

The COD removal efficiency in the smart rotatable solar powered electrocoagulation system exposed a first-order kinetic process, which was linearly proportional to the concentration of COD in POME. So, the electrocoagulation kinetics for COD removal is as follows.

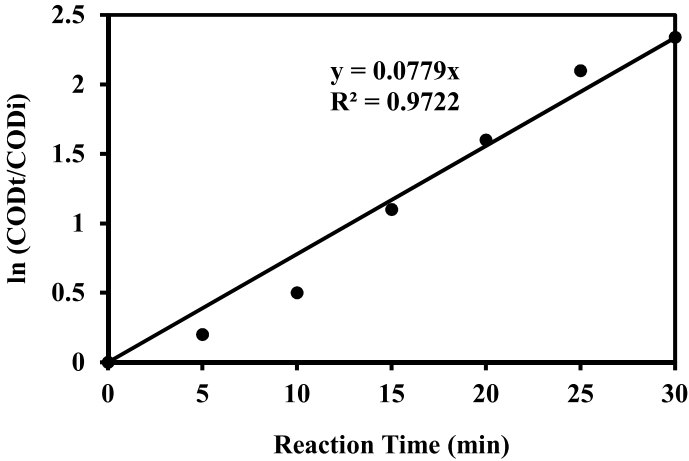


Fig. 7 Effects of reaction time (min) on kinetics COD removal (%) under optimal operational condition

$$-\frac{d}{dt}[COD] = k[COD] \quad (6)$$

After reorganization and integration, the above kinetics equation becomes;

$$\ln \frac{COD_t}{COD_i} = -kt \quad (7)$$

where COD_i is the initial concentration of COD of the POME in mg/L and COD_t is the final concentration of the COD of the POME after a reaction time of “t” in mg/L. After plotting ln(COD_t/COD_i) on the y-axis with plotting reaction time on the x-axis consequences a straight line passing through the origin with the slope value of k. from the Fig. 7, the value for rate constant k and R² for the removal of COD were 0.1076 min⁻¹ and 0.9722 respectively.

4 Concluding Remarks

By this research study, the subsequent conclusions are figured out;

- The smart rotatable solar powered electrocoagulation system in treating POME was effectively executed, and the outcomes were stated in terms of COD, BOD, and TS removal. The results disclosed that this system was successfully effective in pollutant removal.
- The impacts of several operational conditions such as current density, reaction time, effluent pH, inter-electrode spacing and electrode combinations on the

COD, BOD and TS removal efficiencies by the smart rotatable solar powered electrocoagulation process were investigated.

- This system was enhanced to make maximum removal of the pollutants (COD, BOD, and TS) with electrode consumption. The maximum removal of COD, BOD, and TS of 89, 96 and 92% were found having the electric current density of 4 A/dm², 30 min of reaction time, maintaining pH of 7 for the solution, the inter-electrode spacing of 0.01 m, and Fe/Fe electrode combination.
- This smart rotatable solar powered electrocoagulation system experimented is appropriate to be employed for the treatment of POME because this system has greater treatment effectiveness concerning the typical wastewater treatment systems.

Acknowledgements The authors would like to thank the Faculty of Civil Engineering Technology for using their laboratory and for the financial support from RDU1803143 of Universiti Malaysia Pahang (UMP).

References

1. Zhang Y, Yan L, Qiao X (2008) Integration of biological method and membrane technology in treating palm oil mill effluent. *J Environ Sci* 20(5s):558–564. [https://doi.org/10.1016/S1001-0742\(08\)62094-X](https://doi.org/10.1016/S1001-0742(08)62094-X)
2. Nasrullah M, Singh L, Mohamad Z (2017) Treatment of palm oil mill effluent by electrocoagulation with presence of hydrogen peroxide as oxidizing agent and polialuminum chloride as coagulant-aid. *Water Resour Ind* 17:7–10. <https://doi.org/10.1016/j.wri.2016.11.001>
3. Zainal NH, Jalani NF, Mamat R (2017) A review on the development of palm oil mill effluent (POME) final discharge polishing treatments. *J Oil Palm Res* 29(4s):528–540. <https://doi.org/10.21894/jopr.2017.00012>
4. Singh L, Siddiqui MF, Ahmad A (2013) Application of polyethylene glycol immobilized *Clostridium* sp. LS2 for continuous hydrogen production from palm oil mill effluent in upflow anaerobic sludge blanket reactor. *Biochem Eng J* 70:158–165. <https://doi.org/10.1016/j.bej.2012.10.010>
5. Vijayaraghavan K, Ahmad D, Ezani Bin Abdul Aziz M (2007) Aerobic treatment of palm oil mill effluent. *J Environ Manage* 82(1s):24–31. <https://doi.org/10.1016/j.jenvman.2005.11.016>
6. Del Pozo R, Diez V (2003) Organic matter removal in combined anaerobic–aerobic fixed-film bioreactors. *Water Res* 37(15s):3561–3568. [https://doi.org/10.1016/S0043-1354\(03\)00273-2](https://doi.org/10.1016/S0043-1354(03)00273-2)
7. Zinatizadeh AAL, Mohamed AR, Mashitah MD (2007) Optimization of pre-treated palm oil mill effluent digestion in an up-flow anaerobic sludge fixed film bioreactor: a comparative study. *Biochem Eng J* 35(2s):226–237. <https://doi.org/10.1016/j.bej.2007.01.018>
8. Fakhru'l-Razi A, Noor MJMM (1999) Treatment of palm oil mill effluent (POME) with the membrane anaerobic system (MAS). *Water Sci Technol* 39(10s):159–163. [https://doi.org/10.1016/S0273-1223\(99\)00268-1](https://doi.org/10.1016/S0273-1223(99)00268-1)
9. Zinatizadeh AAL, Salamatinia B, Zinatizadeh SL (2007) Palm oil mill effluent digestion in an up-flow anaerobic sludge fixed film bioreactor. *Int J Environ Res* 1(3s):264–271
10. Singh L, Wahid ZA, Siddiqui MF (2013) Application of immobilized upflow anaerobic sludge blanket reactor using *Clostridium* LS2 for enhanced biohydrogen production and treatment efficiency of palm oil mill effluent. *Int J Hydrogen Energy* 38(5s):2221–2229. <https://doi.org/10.1016/j.ijhydene.2012.12.004>

11. Ahmad AL, Sumathi S, Hameed BH (2005) Residual oil and suspended solid removal using natural adsorbents chitosan, bentonite and activated carbon: a comparative study. *Chem Eng J* 108(1s):179–185. <https://doi.org/10.1016/j.cej.2005.01.016>
12. Bajracharya S, Sharma M, Mohanakrishna G (2016) An overview on emerging bioelectrochemical systems (BESs): technology for sustainable electricity, waste remediation, resource recovery, chemical production and beyond. *Renew Energy* 98:153–170. <https://doi.org/10.1016/j.renene.2016.03.002>
13. Nasrullah M, Zularisam AW, Krishnan S (2019) High performance electrocoagulation process in treating palm oil mill effluent using high current intensity application. *Chin J Chem Eng* 27(1s):208–217. <https://doi.org/10.1016/j.cjche.2018.07.021>
14. Elazouz M, Haboubi K, Elyoubi MS (2017) Electrocoagulation flocculation as a low-cost process for pollutants removal from urban wastewater. *Chem Eng Res Design* 117:614–626. <https://doi.org/10.1016/j.cherd.2016.11.011>
15. Nasution A, Ng BL, Ali E (2014) Electrocoagulation of palm oil mill effluent for treatment and hydrogen production using response surface methodology. *Pol J Environ Stud* 23(5s):1669–1677
16. Agustin MB, Sengpracha WP, Phutdhawong W (2008) Electrocoagulation of palm oil mill effluent. *Int J Enviro Res Pub Heal* 5(3s):177–180
17. Nasrullah M, Singh L, Krishnan S (2018) Electrode design for electrochemical cell to treat palm oil mill effluent by electrocoagulation process. *Environ Technol Innov* 9:323–341. <https://doi.org/10.1016/j.eti.2017.10.001>
18. Valero D, Ortiz JM, Expósito E (2008) Electrocoagulation of a synthetic textile effluent powered by photovoltaic energy without batteries: direct connection behaviour. *Sol Energy Mater Sol Cells* 92(3s):291–297. <https://doi.org/10.1016/j.solmat.2007.09.006>
19. García-García A, Martínez-Miranda V, Martínez-Cienfuegos IG (2015) Industrial wastewater treatment by electrocoagulation–electrooxidation processes powered by solar cells. *Fuel* 149:46–54. <https://doi.org/10.1016/j.fuel.2014.09.080>
20. Secula MS, Crețescu I, Petrescu S (2011) An experimental study of indigo carmine removal from aqueous solution by electrocoagulation. *Desalination* 277(1–3s):227–235
21. Kobya M, Hiz H, Senturk E (2006) Treatment of potato chips manufacturing wastewater by electrocoagulation. *Desalination* 190(1–3s):201–211
22. Saravanan M, Sambhamurthy NP, Sivarajan M (2010) Treatment of acid blue 113 dye solution using iron electrocoagulation. *CLEAN–Soil, Air, Water* 38(5–6s):565–571
23. Dalvand A, Gholami M, Joneidi A (2011) Dye removal, energy consumption and operating cost of electrocoagulation of textile wastewater as a clean process. *Clean–Soil, Air, Water* 39(7s):665–672. <https://doi.org/10.1002/clen.201000233>
24. Chou W-L, Wang C-T, Huang K-Y (2009) Effect of operating parameters on indium (III) ion removal by iron electrocoagulation and evaluation of specific energy consumption. *J Hazard Mater* 167(1–3s):467–474

Variable Frequency Phase-Disposition Pulse Width Modulation Technique for Multilevel Inverter



Auni Syahirah Abu Bakar, Syamimi Mardiah Shaharum,
and Nur Huda Ramlan

Abstract The Multilevel Inverters (MLI) are used to obtain the desired output AC voltage from the DC voltage source. This paper implement a variable frequency pulse width modulation (PWM) technique to improve the output voltage of the MLI which is the cascaded H-bridge multilevel inverter (CHMI) by using SIMULINK/MATLAB. Some of the conventional PWM technique still causes extra losses of power due to high frequency switching. In order to overcome this problem, a variable frequency of Phase Disposition (PD) PWM is implemented in minimizing the distortion by injecting variable frequency of the carrier wave into the CHMI. It is confirmed that the variable frequency method offers lower switching losses. Furthermore, Variable Frequency Phase Disposition (VF-PD) PWM gives better total harmonic distortion (THD) for the multilevel output voltage compared to its conventional PD PWM. The results showed VF-PD PWM gives 33.45% of THD better than PD PWM. The THD performance comparison for each PWM technique is presented for five, seven and nine levels of CHMI.

Keywords Cascaded H-bridge multilevel inverter · Pulse width modulation · Variable frequency · Total harmonic distortion

1 Introduction

One of the powerful human inventions that has ever made is ‘Electricity’, due to its very important role in industrial development, socio-economic and technological evolution. Electricity can be transmitted in two different ways either in Direct Current (DC) or Alternating Current (AC). Inverter is a power electronic device that act in converting DC to AC which means to produce a sine wave AC voltage [1].

A. S. A. Bakar (✉) · S. M. Shaharum · N. H. Ramlan
Faculty of Electrical & Electronics Engineering Technology, Universiti Malaysia Pahang,
26600 Pekan, Pahang, Malaysia
e-mail: syahirahauni96@gmail.com

Nowadays, multilevel inverter is one of the circuit system that has been getting more attention among the industries for a high power and medium voltage applications. In contrast, the conventional inverter has power rating limitation and can also causes switching loss. The multilevel inverter's first efficacy is to decrease the harmonic distortion of the waveform of the multi-step AC output voltage. Therefore, a great power quality could be perform to get an ideal sine wave [2]. In addition, due to that the switching loss can be decreased thus, the effectiveness can also be increased in comparison to the standard inverters. [3]. Conventional multilevel inverters have three distinct topologies which are diode clamped (DCMI), flying capacitor (FCMI) and cascaded H-bridge (CHMI) [4]. In this paper, cascaded H-bridge multilevel inverter topology is used.

Cascaded H-bridge Multilevel Inverter (CHMI) has a simple circuit structure with fewer number of components, less production cost and simpler modularity compared to the DCMI and FCMI. Even though, CHMI needs isolated DC source for each cell but it can still become an ideal candidate for high power application in terms of reducing the harmonic distortion [5]. Moreover, to achieve a good quality of output voltage waveform with less harmonic distortion, a multilevel inverter needs different pulse width modulation techniques (PWM).

The general structure of multilevel inverter is to synthesize the sinusoidal voltage waveform. In order to regulate the frequency and amplitude parameter of the AC voltage, PWM is used by comparing the reference signal with the carrier signal [6]. The main reason for this PWM implementation is to help in minimizing the distortion. In different multilevel inverter topologies there are various type of modulation techniques that can be applied to analyze the performance based on the modulation index and frequency given. Moreover, each one of the modulation has its own features and drawbacks [7].

There are two ways in getting a good output voltage waveform of the multilevel inverter. First is by modifying the structure of the inverter circuit itself, changes such as reducing the number of the switches or by changing the position angle of the switches can be made. Secondly, improvement on the PWM technique can be made by reducing the harmonic of the output voltage with setting the switching method into the desired PWM characteristic [8].

Several implementation of standard PWM techniques for multilevel inverters have been developed: Phase Disposition (PD PWM), Phase Opposition Disposition (POD PWM) and Alternative Phase Opposition Disposition (APOD PWM) that based on the phase shifting [9]. All of the PWM techniques have some carrier wave that will be compare with sine reference wave that help in generating the sinusoidal output voltage waveform of the multilevel inverter. By considering the frequency of the PWM carrier wave as variable parameter, it is possible to reduce the output voltage harmonic. This method is call Variable Frequency carrier wave. In this paper the Variable Frequency of Phase Disposition (VF-PD PWM) is implemented and the analysis of the THD performance is compared with the conventional PD PWM with five, seven and nine level of CHMI.

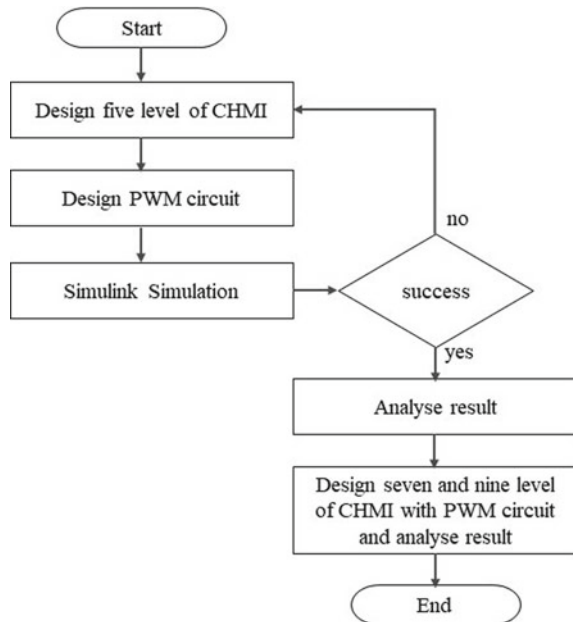
Variable frequency in switching method was one of the most proposed and popular PWM technique used in getting a better harmonic performances. Work by

Mohammad Sadegh and Govindaraju [8] applied the PWM technique with variable frequency range. The combination of this low and high switching frequency are validated through the detail analysis showed by both of this paper and prove a better THD result from VF PWM method. Meanwhile, the paper from V. Sudha analysed that VF PWM was not suitable to use for reduced switch of multilevel inverter because it produced a higher THD compared to other conventional PWM. In summary, VF PWM can help in giving a better THD for CHMI compared to the modified CHMI.

2 Methodology

The flowchart of the work can be observed in Fig. 1 starting from the designing process of the five level CHM, continued with implementation proses of the PWM method. The PWM circuit need to correspond with the switching devices in order to trigger the CHMI to produce the desired output for each type of the CHMI level used. Once the task in designing the five level of CHMI and PWM circuit is done, the simulation of the circuit is tested in order to detect problem. Finally, the obtained result from the five level of CHMI with 0.9 of modulation indices is analysed and the THD of the output voltage from each PWM method is also compared. Then, the progress is continued by alternating the level of the CHMI with seven level and nine level.

Fig. 1 Flowchart of the work progress



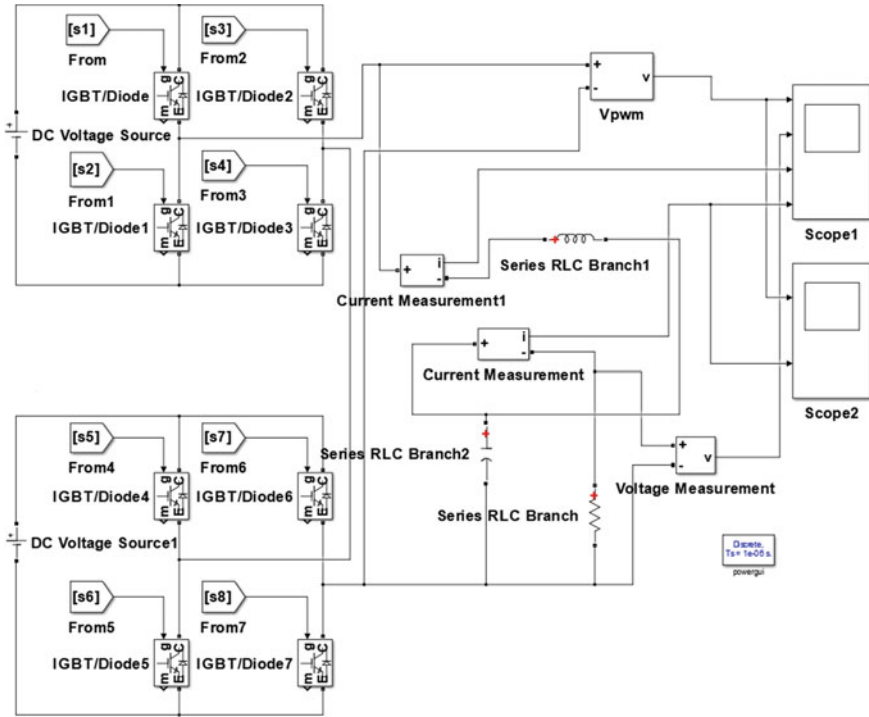


Fig. 2 Circuit design for five level of CHMI

The Simulink circuit design basically consists of CHMI Circuit, PWM circuit, load and Scope to display the Output voltage waveform. Figure 2 shows Simulink block circuit diagram for five level of CHMI with desired number of component needed.

2.1 Cascaded H-Bridge Multilevel Inverter

CHMI generates a staircase waveform of output voltage by using a series strings of single phase H-bridge cells, each providing three different levels of DC voltages (0, +Vdc and -Vdc). The output voltage is produced from the total generated voltages from each H-Bridge cell [10]. The number of output voltage levels for a CHMI are defined in term of $2n + 1$, where n is the number of cells. While, the N level of CHMI will use $2(N - 1)$ number of switching devices and N is represented as the level of the CHMI and the level for the output voltage. Figure 3 shows that five level of CHMI needed two H-bridge cell, eight switching devices and five level of stepped output voltage waveform (0, +Vdc, +2Vdc, -Vdc, -2Vdc) [11]. The term of multilevel begin with three level inverter and produced three level of output voltage. The quality

Fig. 3 Five level of CHMI [11]

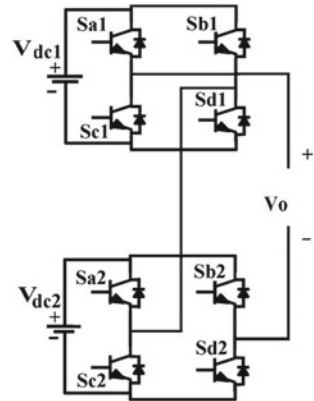


Table 1 Comparison of component for five, seven and nine level of CHMI

N-level of CHMI	5	7	9
Switching device	8	12	16
H-bridge cell	2	3	4
DC source	2	3	4

of the inverter is validated by the quality of the output voltage waveform obtained with less THD. Hence, the minimization of the THD percentage can be achieved by increasing the N level of the multilevel inverter. This paper analyzed five, seven and nine levels of CHMI in order to prove on the THD performance based on increasing the N level of the multilevel inverter. Hence, the number of component used for each level of the CHMI circuit can be refer to Table 1.

2.2 The PWM Switching Method

Selecting a suitable switching method is necessary to reduce the distortion of the output voltage. PWM with carriers signal and a reference signal will help the multilevel inverter in generating the staircase sinusoidal output voltage waveform. In this method, the amplitude of reference wave, A_r is the magnitude of the sine waveform. While, the amplitude of carrier waveform, A_c is equal to the total magnitude of carrier in each level. Modulation index, ma is obtain by dividing both amplitude of reference and carrier as following:

$$ma = \frac{A_r}{A_c} \tag{1}$$

Another important index use in multilevel inverter is the frequency modulation index, mf that is achieved form the division of the frequency magnitude of sine

waveform, fr on the frequency magnitude of carrier wave, fc as following:

$$mf = \frac{fr}{fc} \quad (2)$$

In many common switching methods for multi carrier PWM, the frequency modulation index is defined by fixing the value for frequency magnitude of carrier wave [8]. In this paper the frequency magnitude of carrier wave is considered as a control variable. Thus, it is possible to reduce the harmonic distortion of the output voltage by varying the carrier frequency.

In this paper, Phase Disposition PWM is investigated. Also a new implementation of variable frequency of PD PWM is introduced and injected to the cascaded H-bridge multilevel inverter (CHMI) with the carrier signal still has the same amplitude and in phase but with different frequency of the original PD PWM. The number of carrier signal needs for each N-level of CHMI is different, the term use in defining the carrier wave is N-1. The PWM techniques are as follows:

Phase Disposition (PD PWM)

For phase disposition method all the carrier wave will have same amplitude, fixed frequency value that are in phase as in Fig. 4. In order to generate N level inverter output voltage, PD PWM will be used (N – 1) of carrier wave. Five level of CHMI will used four carrier signal and needed to be compared with one reference signal.

Variable Frequency Phase Disposition (VF-PD PWM)

VF-PD PWM is an implemented PWM technique which implement from combination of conventional PD PWM with variable frequency range of triangular waveform of the carrier. The carrier signal still has the same amplitude and phase but with different frequency that consisted of high and low switching frequency as shown in Fig. 5.

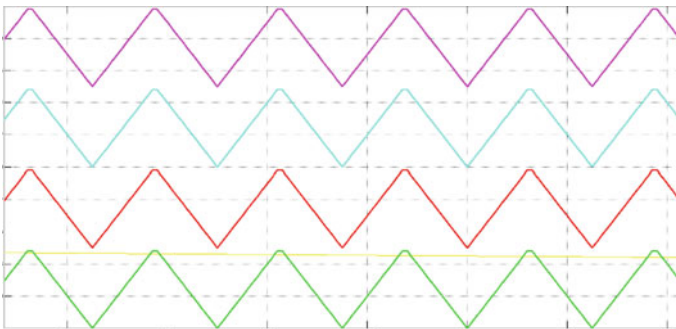


Fig. 4 PD PWM carrier wave

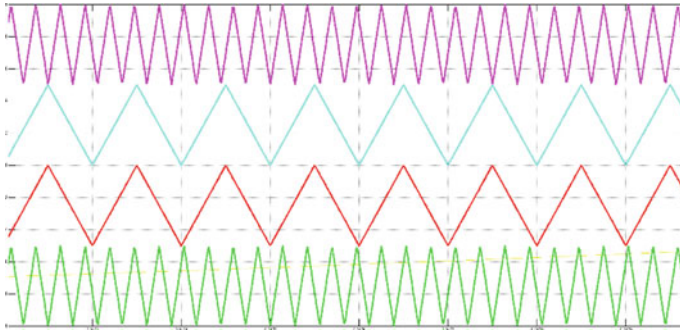


Fig. 5 VF-PD PWM carrier wave

3 Analysis and Discussion

The analysis on relationship between modulation index (ma) and N-level of the CHMI with the THD performance are discussed in this paper.

3.1 Modulation Index Analysis

In this work, the value of modulation index for the reference wave that compared with carrier wave is varied from 0.7 until 0.9. This is because the overlapping of the sinusoidal wave with the triangular shape of carrier wave will produced pulses. Thus, the ma is varied to determine which modulation indices will produces a higher rate of total harmonic distortion.

Based on Figs. 6 and 7 we can observed that the relationship of modulation index with the THD performance. It is showed that the lowest ma which is at 0.7 gave the highest value of THD for each type of PWM technique. Meanwhile ma at 1.0 gave the lowest V THD performance. This is because the sinusoidal reference wave with the highest value of ma is overlapped regularly with the triangular carrier amplitude. However, the ma with 0.9 is selected. This is because the value of modulation index at 0.9 is optimum to be choose other than it approaches the value of the amplitude of the triangular carrier wave of 1.0. Although the ma at 1.0 was the highest but it is not suitable to stimulate the CHMI because the amplitude of the triangular carrier of the PWM is crooked and can cause no overlapping between the reference signal with the carrier to produce the output voltage waveform.

Fig. 6 Variation ma with THD for PD PWM

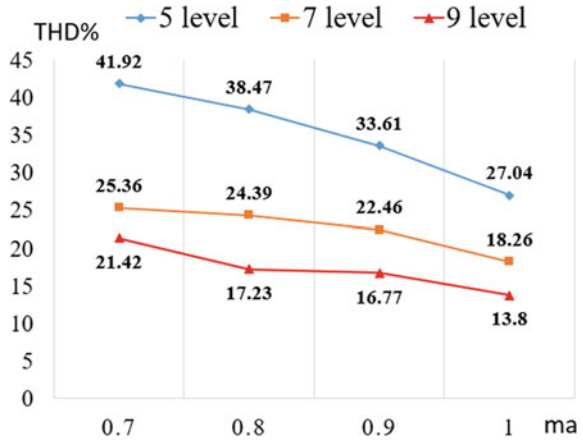
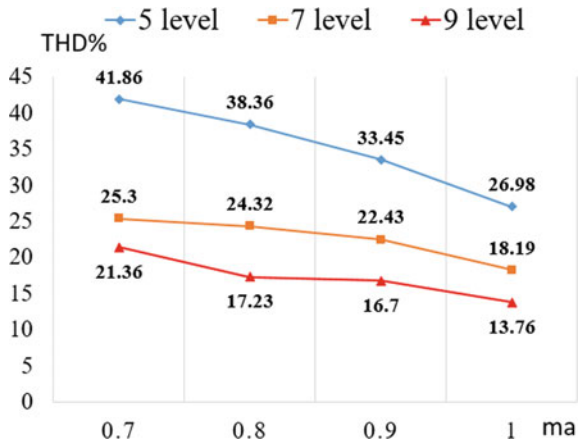


Fig. 7 Variation ma with THD for VF-PD PWM



3.2 N-Level of CHMI Analysis

Figure 8 shows the relationship patterned between the N level of the CHMI with the THD performance of the output voltage and it clearly showed that nine level of CHMI give the lowest THD compare to five and seven level of CHMI. Therefore, this can be concluded that, the increased number of the N-level of the CHMI can helped to reduce the THD voltage more. This is because the output voltage waveform with more stepped level is approaching to an ideal sine AC waveform which has less distortion. Hence, nine level of CHMI output voltage waveform representing a better sinusoidal AC rather than five and seven level of CHMI.

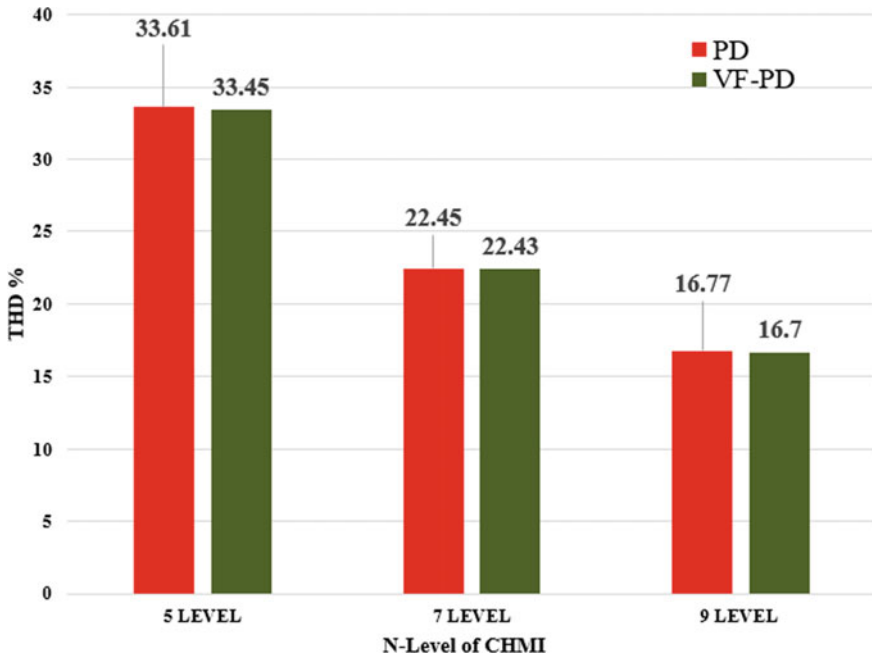


Fig. 8 THD comparison for five, seven and nine level of CHMI

4 Simulation Results

The simulation for the five, seven and nine level of CHMI configuration is done with the help of SIMULINK. The strategy is to focus on the Variable Frequency PD PWM. The constant frequency used for PD PWM is 20 kHz which is optimum for IGBTs switching frequency with 500 V of DC source and 0.9 of modulation index is chosen. While, the frequency of the carrier wave used for VF-PD PWM technique for five, seven and nine level of CHMI are selected within the range from 5 kHz up to 20 kHz as shown in Table 2 500 V DC source and 0.9 of modulation index.

Figures 9, 10, 11, 12, 13 and 14 present the results of the output voltage performance between PD PWM and VF-PD PWM that is generated from the Simulink.

Table 2 Variable frequency values for VF-PD PWM

N-level	Frequency of carrier wave
5	18,000 Hz (1st and 4th), 5000 Hz (2nd and 3rd)
7	18,000 Hz (1st and 6th), 5000 Hz (2nd and 5th), 12,000 Hz (3rd and 4th)
9	18,000 Hz (1st and 8th), 12,000 Hz (2nd and 7th), 8000 Hz (5th and 6th), 5000 Hz (4th and 5th)

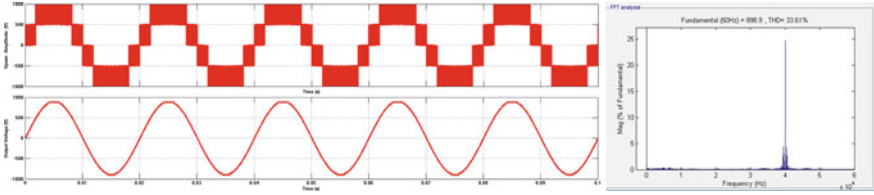


Fig. 9 THD value for of five level of CHMI with PD PWM

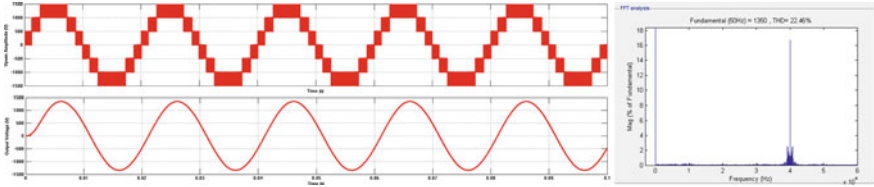


Fig. 10 THD value for seven level of CHMI with PD PWM

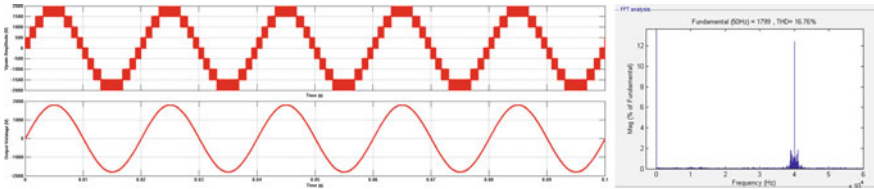


Fig. 11 THD value nine level of CHMI with PD PWM

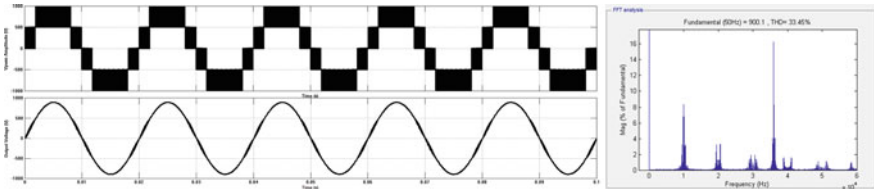


Fig. 12 THD value for five level of CHMI with VF-PD PWM

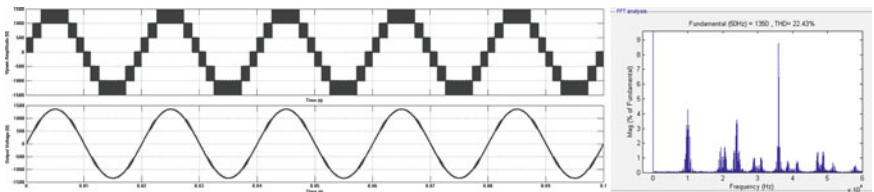


Fig. 13 THD value for seven level of CHMI with VF-PD PWM

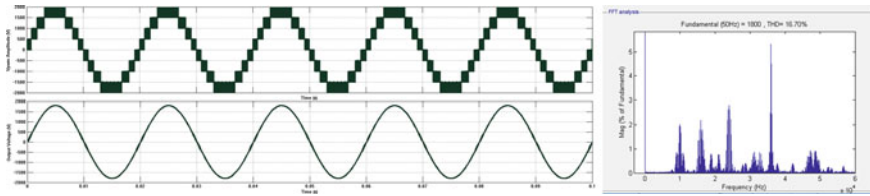


Fig. 14 THD value for nine level of CHMI with VF-PD PWM

The stepped waveform of the output voltage showed base on the desired level of the CHMI.

It is clearly showed that THD value of the implemented VF-PD PWM technique is better than the conventional which can reduce the harmonic from 33.61% into 33.45% for five level of CHMI. Meanwhile, for seven level of CHMI VF-PD PWM reduced THD from 22.45 to 22.43% and for nine level of CHMI that can reduced the THD from 16.77% into 16.7%.

5 Conclusion

In this work, all of the result and analysis that are obtained from the Simulink simulation which were the total harmonic distortion for five, seven and nine level of CHMI has been discussed, analyzed and also compared by using VF-PD PWM and PD PWM. Based on the comparison that has been made. VF-PD PWM showed a better performance with less percentage of voltage THD in compared to the PD PWM for each level of CHMI. The VF-PD PWM used was the combination between the conventional PD PWM with variable switching frequency of the carrier waveform, thus making the VF-PD PWM as a good option to reduce the THD and to produce a better output voltage waveform compared with the standard PWM.

Acknowledgements The work was partly supported by the Research and Innovation Department, Universiti Malaysia Pahang (grant number RDU170376). The authors would like to be grateful to the Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang for providing financial support as well as research facilities.

References

1. Eggleston DP, Doucet JM, Shaw JS (2007) PWM pure sine wave power inverter
2. Rodriguez J, Lai JS, Peng FZ (2002) Multilevel inverters: a survey of topologies, controls, and applications. *IEEE Trans Ind Electron* 49(4):724–738
3. Lopez O, Teodorescu R, Doval-Gandoy J (2006) Multilevel transformerless topologies for single-phase grid-connected converters. In: *IECON 2006-32nd annual conference on IEEE industrial electronics*, pp 5191–5196

4. Babaei E, Haque MT, Hosseini SH (2005) A novel structure for multilevel converters. In: 2005 international conference on electrical machines and systems, vol 2, pp 1278–1283
5. Rathore S, Kirar MK, Bhardwaj SK (2015) Simulation of cascaded H-bridge multilevel inverter using PD, POD, APOD techniques. *Electr Comput Eng: Int J (ECIJ)* 4(3):27–41
6. Peddapelli SK (2014) Recent advances in pulse width modulation techniques and multilevel inverters. *Int J Electr Comput Eng Electron Commun Eng* 8(3):593–601
7. García Franquelo L, Rodríguez J, Leon JI, Kouro S, Martín Prats MDLÁ, Portillo Guisado RC (2008) The age of multilevel converters arrives. *IEEE Ind Electron Mag* 2(2):28–39
8. Yegane MSO, Sarvi M (2017) A new improved variable frequency triangular carrier-PWM with MOPSO algorithm for carrier based PWM techniques in inverters. *TEM J* 6(1):32
9. McGrath BP, Holmes DG (2002) Multicarrier PWM strategies for multilevel inverters. *IEEE Trans Ind Electron* 49(4):858–867
10. Krishna RA, Suresh LP (2016) A brief review on multi level inverter topologies. In: 2016 international conference on circuit, power and computing technologies (ICCPCT), Mar 2016, pp 1–6
11. Balamurugan CR, Natarajan SP, Arumugam M (2015) A review on various multilevel inverter topologies

Discrepancy Resolution: A Review of Missing Tags Detection in RFID Systems for Inventory Shrinkage



Yusuf Musa, Hazalila Kamaludin, and Hairulnizam Mahdin

Abstract Radio Frequency Identification (RFID) is a technology that is widely used for inventory management to tackle inventory shrinkage. Detecting all missing tag events in this application is no doubt a paramount important. However, due to inefficient of the current missing tag detection protocols and wide number of missing tag detection approach, inventory shrinkage however exists. To address this issue, this paper makes a review from most current and relevant literatures on missing tag detection approaches and gives clear and fair understanding of the current protocols in missing tag detection. In addition, this paper compared some of the current protocols on localization, cardinality estimation and clone tag detection protocols, state efficiency and inefficiency of the current protocols. Finally, this paper, contributed to the state of art on missing tag detection that will significantly benefit the entire supply chain.

Keywords Radio frequency identification (RFID) · Missing tag · Replaced missing tag · Inventory shrinkage

1 Introduction

More than two decades RFID is promising excellent changes which make it unpredictable to displayed barcode [1]. RFID is a technology that uses radio waves and automatically, track and identifies an objects that attached to it [2]. Because of the tremendous advantages of RFID which includes automation, cost efficiency and ease

Y. Musa · H. Kamaludin (✉) · H. Mahdin
Faculty of Computer Science and Information Technology, Universiti Tun Hussein
Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
e-mail: hazalila@uthm.edu.my

Y. Musa
e-mail: yammadugu@gmail.com

H. Mahdin
e-mail: hairuln@uthm.edu.my

of deployment [3, 4]; make it more acceptable in many applications, such as supply chain management, inventory management, healthcare system, intelligence transportation, agriculture, water supply management, anti-tag cloning attacks and internet of things [5, 6]. As a result of aforementioned, RFID technology can produce a real time response [7]; which gives it the capacity of improving 40–50% management processes as compared to the traditional processes, specifically, detecting missing tags in supply chain management. Moreover, RFID also capable of reducing the operational cost and as well improve the quality of service [8]; and enhance managerial decision within and outside the organization. However, despite its enormous advantages RFID suffer a setback that hindered its efficiency [9]. One of the major setbacks of RFID technology that attract the attention of researchers is missing tags detection [10].

Actually, there are many causes of missing tags in RFID systems. However, the concern here is shrinkage which ranges from: internal fraud (employees' and supplies' fraud), shoplifting, customers' fraud and administrative errors (misplacement or wrong tagging) [11, 12]. The first scenario, internal fraud is the worst case in most of the organization. However, it is must negligible in the organization. Many attempts have been made to tackle the problems of missing tags in RFID systems using difference techniques in different perspectives [13]. However, due to wide assortment of the current solutions and underestimating of the negative impacts of shrinkage, it is very difficult to have clear understanding of the diverse solutions. Therefore, the current approaches need an improvement in order to give a clear and fair understanding of the diverse solution of missing tags techniques and to completely address the issues of missing tags detection. Moreover, classification and comparisons of divers' solution will also be considered.

RFID technology is an essential feature in supply chain for process efficiency and decision making [14]. The processes of supply chain as shown in Fig. 1 is a sequential processes that need to be checked out at any point, so that the actual point of missing tag as well as the actual tag that is missing can be detected.

The rest of this paper is organized as follows: Sect. 2 provides the background knowledge of RFID missing tags detection and inventory shrinkage, Sect. 3 systematically review the related current literatures on missing tag detection solutions, Sect. 4 discussed limitation of the prior art which highlight open research issues, Sect. 5 discussed on the critical analysis and result of the review, and finally, Sect. 6 concludes this paper.

2 Background

In this section, we discussed RFID system, missing tags, common confusions from the current literatures on missing tag detection and inventory shrinkage due to missing tag.

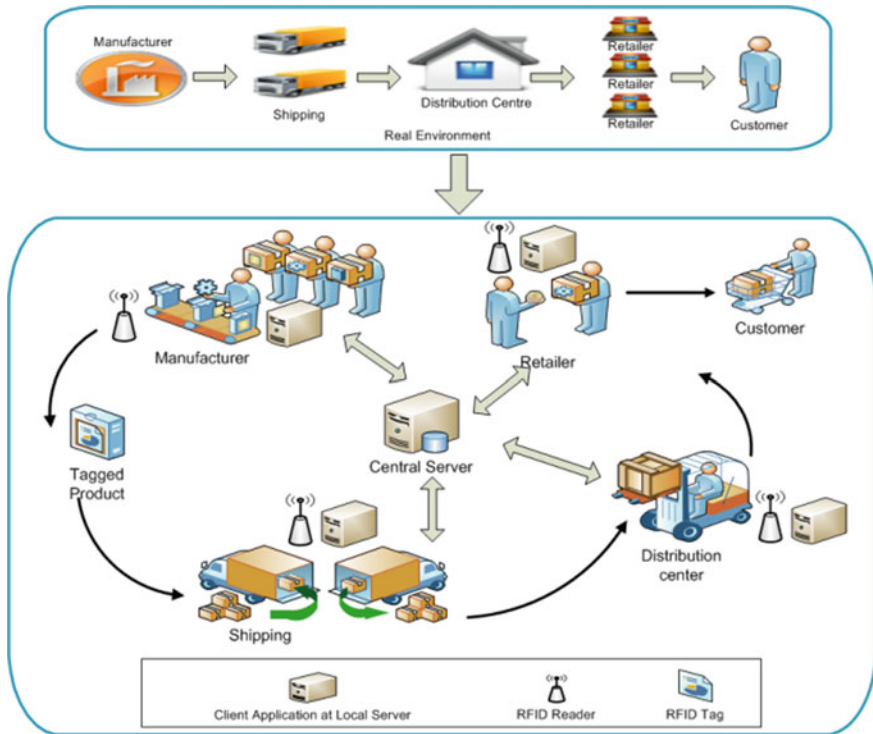


Fig. 1 RFID supply chain system [1]

2.1 RFID

RFID referred to radio frequency identification is a system that uses radio waves and automatically identifies an object without the line of sight, as compared to barcode [1]. RFID system architecture as in Fig. 2 consists of three entities: back-end server, RFID reader and RFID tag [15, 16]. An RFID tag contains an antenna to receive and send signal and short memory for storing data i.e. electronic product code (EPC).

Reader is a mediator between tag and data processing system (DPS), it communicate directly with the tags [17]. The server is the system monitoring protocol [18]. It monitors the authentication of objects.

2.2 Missing Tag

According to [19], missing tag is any tag that left unread or did not respond to reads' query. However, according to [20], missing tag might not be necessary that tag is missing completely from the system. Some tags may be missed, but counting them

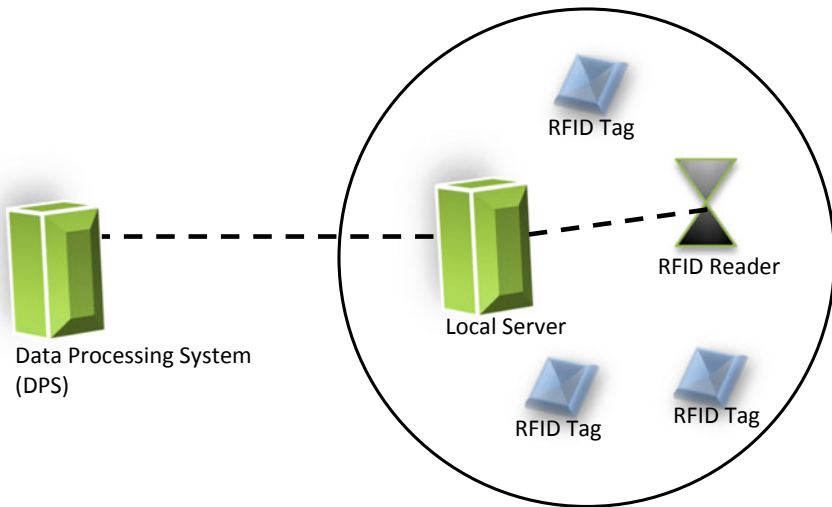


Fig. 2 Schema of RFID system architecture

as presence tag, because they have been replaced with another tags. Misplacement is another type of missing tag which may be classify as genuine while it is not [21].

Many protocols have been proposed (probabilistic and deterministic) in recent years more than the previous years [22]. Unfortunately, despite the increase in the number of solutions in missing tag detection, RFID system suffered from shrinkage due to insufficient of the current protocols. Therefore, it is challenging to systematically review the current literatures so as to clear the confusions in order to address the issue of shrinkage. To do so, this paper takes review of the current and relevant literatures on missing tag detection and clear the common confusion that emanate from the current protocols. The details explanation, classifications and comparisons, will be discussed latter.

2.3 Common Confusions

The major confusions that usually disturb the interest of the readers from the divers' solutions of missing tag detection were identified from current literatures such as [10, 21, 23, 24], and stated as follows: (1) Considering replaced missing tags as present tags (legitimate tags). (2) Regarding clone tags as replaced missing tags.

2.3.1 Present Tag Versus Replaced Tag

In missing tag detection, present tag usually refers to expected tag that its present is been confirmed [21]. While in contrast, replaced tag is the unexpected tag that replaced the legitimate tag and act as genuine tag [14]. Replaced missing tag thus can be recorded as presence tag when detecting missing tag. But in actual science, it is not the genuine tag. To this sense, replaced missing tag is different from actual missing tag. Hence many RFID missing tag detection related articles, consider replaced missing tag as present tag. Some of the articles consider responses from the tag to identify whether tag is presence or not. Since replaced tag can act as genuine tag, it can respond as well. However, using only responses from the tags cannot differentiate genuine tag from replaced tag.

2.3.2 Clone Tag Versus Replaced Tag

Clone tag is tag with all the features of genuine tag [1, 24]. Therefore, clone tag can act as genuine tag, but it causes collision and increase tag cardinality. While replaced missing tag, is a tag that replaced another tag (genuine tag), act like it, but do not causes collision or increase cardinality. Thus replaced tag, should not be treated as clone tag. Therefore, cardinality contradiction and clone tag detection are not the optimum solution to the problem of missing tag detection. To that end, a new approach for missing tag detection is therefore necessary.

2.4 Inventory Shrinkage

As stated earlier, there are numbers of reasons for missing tags in supply chain management, but the concern here is inventory shrinkage. Shrinkage is the loss of an inventory which is in the main list, but absent in shelf. Inventory shrinkage, it cause as a result of the following factors.

2.4.1 Shoplifting

One of the most popular causes of permanent missing tag in supply chain management is shoplifting which is cause by the customers' thefts [25]. In this type of missing tag, the current protocols can detect the present of missing tag event, because, the tag is appeared to be absent completely. However, in the case of replaced missing tag, the current missing tag detection approach will not be appropriate.

2.4.2 Administrative Errors

Administrative errors occur as a result of misplacement or missed reading [20, 23]. Tag may be left unread due to misplacement or misreading of tag by the reader. In this type of missing tag, the problems can be correct latter without losing the tags, because, the actual tags are not missing just the problem from the reader.

2.4.3 Internal Fraud

This type of missing tags is cause by either suppliers' or employees' fraud [10, 20]. Internal fraud is the worst among the three, and in most cases, it is been under estimated or misunderstand by some researchers. One of the examples of missing tag by internal fraud is replaced missing tag.

There are approaches related to missing tag detection, such as [1, 26, 27] for cardinality estimation, localization and clone tag detection. However, there is some common confusion that needs to be clarified. Details explanation and comparisons of the current approaches will come in the latter discussion. We reviewed the most current and relevant literatures on missing tag detection approaches to enhance missing tags detection in supply chain management systems.

3 Related Works

There are many solutions for missing tag detection, ranging from cardinality estimation, clone tag detection, and localization [28–, 30]. However, most of the solutions, do not give attention to replaced missing tags, which is one of the major cause of missing tags by shrinkage [31].

There are approaches that been presented in recent literatures to address missing tags, such as [32, 33]. Different protocols such (i.e. comprehensive localization and tracking scheme, and single slot-base estimator) were proposed using localization, cardinality estimation and clone tag detection. They used both identifiable and anonymity methods. Xiao et al. [33], proposed the use of double tags detection in indoor localization. The research is based on phase measurement report to remove phase ambiguity. The method is said to improve accuracy in the tag positioning systems. They try to optimize the localization accuracy and antenna calibration. However, using double tags localization, increase the overhead which will increase complexity and slow down the system performance, thus the required accurate localization will not be achieved. Moreover, the approach is more of localization not detection.

Also, in [4], Xie et al. proposed system for tag localization known as fast continuous scanning (FCS) protocol. The protocol is designed based on multi-categories, so that empty and collision slots will be deactivated in order to optimize the execution time, so as to improving efficiency. Although the method is better than the prior

method in terms of execution time, but the privacy of data is not protected and may not report all the missing tags. When an empty and collision slots are skip, there is tendency that some tags may not be reported.

Zhang et al. [34], proposed a method for missing tag detection by observing collision slots. The method tends to reduce the filter size using compressed filter to increase the efficiency of the system and reduce the detection time for more accuracy. The presence of unknown tags is considered in the literature. However, using collision slot to detect the presence of missing RFID tag will not be the optimum solution to the missing tag detection, because, replaced missing tag will not cause collision, yet it will respond to the query. Therefore, using collision will not be the appropriate solution.

Recently, the attentions have been shift to missing tag detection using anonymous in cardinality estimation. There are many solutions for cardinality estimation with different performances. The study of cardinality estimation was first done by Kodialam et al. [35], proposed protocols known as Unified Simple Estimator (USE) and Unified Probabilistic Estimation (UPE). An Aloha protocol was utilized for communication between tags and reader. The reader first broadcast a query by sending the frame size in the time frame. After receiving queries, all tags randomly select a slot and respond to the query. The estimated number of singleton, empty and collision slots where used to determine the cardinality of tags. Although, identifiable protocols can be used to obtain tag IDs and use it to estimate tag cardinality. However, the tag privacy may not be protected. Moreover, the concern here is not the tag IDs, but the cardinality. In such a situation, the best protocol to use is anonymous and not identifiable.

Liu et al. [32], observed cardinality estimation from privacy view. They observed un-key tags so that they reduce the execution time and optimize the system. They considered both changes in expected key tags and non-key tags to predict frame status of the expected key tags. To monitor the whole frame event is inefficient, especially, when the size of the frame is long. In an another development, Chen et al. [36], proposed a method known as sampling based key tag tracking (S-KT) protocol. They used sampling method to sample out the observed changes in expected key tags, instead of visiting all time frames. This makes protocol in [36] more efficient. However, tags that have been replaced by another tags, can't be detected.

In another development, Kamaludin et al. [1] proposed a new method to address the problem of tag missing through counterfeit detection in RFID system. They used dual hash collision and count-min sketch vector to detect the presence of collision slots. The protocol makes significant contribution in terms of detection accuracy and execution time. However, the research, only consider the consistency of dual hash collision, which means only clone tags that cause collision will be considered, those that missed and been replaced, will not be considered since it won't cause any collision to the system.

Table 1 gives the illustration of missing tag events of the current approaches. It also point out the efficiency and inefficiency of each approach. For the first row, all the approaches can detect the presence of clone tag during process. Since clone tag causes collision and increases tag cardinality when injected into the system, hence

Table 1 Missing tag detection events and approaches

Type missing tag	Cardinality estimation approach	Collision slots approach	Localization approach	Proposed approach
Clone tag	✓	✓	✓	✓
Unknown tag	✓	✓	✓	✓
Replaced tag	×	×	×	✓

all the approaches can detect its presence. For the second row, also all the approaches will detect the present of unknown tags. This is because, when unknown tag is added into the system, it will cause cardinality contradiction, and collision will be detected as well, thus it can be detected using cardinality estimation, clone tag detection, and localization. But for the last row, all the approaches can't detect its present. Replaced tag has the ability to act as the genuine tag, and will neither contradict cardinality nor cause collision. So, none of the approach can detect its presence. Cardinality estimation, consider the frame status to estimate whether tag is missing or not. Also the collision slot, observe collision in a slot and determine the missing tag event while replaced missing tag will not contradict cardinality, neither it will cause collision. Therefore, the current approaches need to be improved.

From the above scenarios, it observed that none of the approaches can detect the presence of missing tag event in all the cases, as such, the problem of shrinkage, is yet to be addressed. An efficient protocol is therefore needed so as to address the problem of missing tag detection in inventory shrinkage.

Table 2, gives the summary of the current protocols in missing tag detection. In the first and fourth rows, the authors used cardinality estimation which cannot detect all missing tag, as such, the system accuracy is decreased. In the second row, even though the authors considered privacy of tag, yet, it cannot detect presence of replaced tag. Again, the authors in the third row try to improve the detection accuracy by considering anonymity of missing key tags but forgetting the non-key tag which can also affect the detection accuracy of the system. In the fifth row, compressed filter was used by the authors so as to improve time efficiency, but still cannot detect replaced missing tag because it considered only the effect of an unknown tag. In the sixth row, clone tag was observed, but the privacy is not protected. Finally, in the last row, the authors used Mrun1 AND Mrun2 to address detection error by addressing the impact of an unknown tag, but the approach is an identifiable, as such it cannot protect data privacy and do not outperformed the existing approaches.

Table 2 Comparisons of the current approaches

Study/authors	Problems	Methods	Advantages	Disadvantages
Mike et al. [37]	Irreconcilable collisions	Randomize tag responses	Provide scanning proof for smart supply chain	Used only cardinality estimation
Mike and Jorge (2016) [38]	Readers trust	Anonymous grouping proof of integrity	Anonymous group scanning	Cannot detect replaced missing tags
Honglong et al. [23]	Ordinary tags interference (non-key tags)	Vector-based missing key tag identification	Addressed the anonymous missing key tag identification	Do not consider effects of non-key tags
Woltmann et al. [27]	Tag cardinality estimation	Expectation maximization method	Time execution is relatively fast	Used cardinality estimation
Zhang et al. [39]	1. Time efficiency 2. Unknown tag interference	Compressed filter	1. Reduce time efficiency 2. Unknown tag interference	1. The process is relatively complex 2. Can't detect replaced missing tags
Kamaludin et al. [1]	Tags cloning	Consistence of dual hash collision	Used consistency of dual hash collision which significantly enhance the detection accuracy	1. Considered collision slot which can't detect replaced missing tag 2. Do not protect privacy of data, because the approach is identifiable
Nguyen et al. [40]	1. Detection error 2. Impact of unexpected tag	Mrun1 AND Mrun2	Optimized the detection error	1. The privacy of data is not protected 2. The approach cannot outperformed the existing ones

4 Open Research Issues

There are currently two main protocols (cardinality contradiction and clone tag detection) for missing tag detection and identification which are either probabilistic or deterministic protocols [1, 10, 12, 25, 26, 41]. Cardinality contradiction is to estimate the number of participating tags in the reader's vicinity. If an expected slot is tune to be an empty slot, missing tag will be detected [10, 26]. It does not considering unexpected tag that may replace genuine tag without changing the cardinality.

While clone tag detection, report missing tag events only when there is existence of collision [1]. It does not reflect clone tag that may not cause collision.

There are two main drawbacks in the existing protocols. The first drawback is that the protocols required perfect process that considers only empty slots or existence of collision to detect missing tags. In practical, in accordance with [1], clone tag is a tag that possessed all the features of genuine tag and can act as legitimate tag. That mean, clone tag can replaced genuine tag without contradicting the cardinality or causing collision [42]. For instance, a large shopping mall, some items are very expensive (key items) while others are non-expensive (non-key items), during internal fraud which cause by employee's or supplies' fraud, the expensive items can be exchange with chip ones without contradicting the cardinality or causing collision.

The second limitation of the existing protocols is that it cannot handle a dynamic situation where changes are taking place frequently. In addition, a practical RFID system (i.e. supply chain), new items are introducing to the system while some are taking out of the system regularly [26, 27, 32]. However, using tag estimation, the existing protocol cannot be the optimum solution for missing tag detection, because the number of tags that is taking out of the system might not be the same as the number of tags that were introduced to the system, thus resulting to cardinality contradiction. We give two examples. For the first example in a large warehouse which rented to more than one tenants, their RFID readers will be over lapping each other, resulting to false positive. In that situation, a tag can be replaced by another tag and act as the genuine tag. For the second example, in poultry, chickens are regularly taking in and out of the poultry. The existing protocols cannot estimate the actual number of the dynamic system.

5 Critical Analysis and Discussion

From the earlier discussion, missing RFID tag by inventory shrinkage, is been underestimated, as such, the problem of inventory shrinkage is yet to be addressed. Inventory shrinkage can result in permanent loss of items or inaccuracy in the inventory which may result to wrong decision that may lead to destruction, privacy loss and even loss of life (i.e. in hospital under RFID monitoring) [1, 21]. In order to address the problem of inventory shrinkage, there is need to first know the actual cause of the missing tag, either it was the internal fraud, administrative error or is it shoplifting. If the cause of missing tag is an administrative error, the missing tag can be detected later because it is not a permanent loss. Also in the case of shoplifting, even though it is a permanent loss, but the missing event can be detected. However, in the case of internal fraud, if the tag is been replaced, the current protocols cannot detect the missing tag event. That is why the current protocols, need to be improved so that missing tag event can be detected as early as possible. Considering the problems discussed, the solution can be resolved as follows.

One of the ways to address the problem of inventory shrinkage is by differentiating between present tag and replaced tag. Present tag usually refers to expected tag that

its present is been confirmed [21]. While in contrast, replaced tag is the unexpected tag that replaced the legitimate tag and act as genuine tag [14]. When present tag can be differentiating from replaced tag, this issue can be addressed. Secondly, clone tag should be differentiating from replaced tag. Usually, when clone tag is been introduced into a system, tag cardinality will be contradicted and collision slot will be observed. However, in the case of replaced tag, neither the cardinality nor the collision will be observed. Because replaced tag can act as present tag haven have the features of present tag. For that, there is need of way to address the issue of clone tag and replaced tag differently.

The problem of replaced tag can be address using the combination of Track Sequence Number (TSN) and Time of Arrival (TOA). TSN is the sequence number of all the activities that take place during the process. When clone tag is injected, it said to happen at the middle of the process; as such its sequence number is less than that of genuine tag. In this situation the presence of clone tag can be detected. But in the case of inventory shrinkage, tag can be replaced at the begging of the process, as such, TSN is not enough to detect replaced missing tag, because its sequence number will not change. Thus additional parameter is therefore needed. TOA is the average time spends by tag from one slot to another. During the time slot, each tag is given an average time to arrive at the next slot. If it exceed the average time, there is a case of missing. In conclusion, the problem of inventory shrinkage can be addressed using the combination of TSN and TOA. If sequence number and arriving time are not intact, we concluded that, missing tag event has occurred.

6 Conclusion

RFID is a daily essential feature for inventory management to cater inventory shrinkage. Missing tag by shrinkage, negatively affect the performance efficiency of RFID system, thus inventory shrinkage should be given more consideration. To address this issue, this paper makes review from current and relevant literatures on missing tag detection. The result shows that internal fraud still suffered a setback from missing tag. However, if replaced tag can be differentiate from present tag and also clone tag is separated from replaced tag, this issue can be solved. In addition, combining two parameters (TSN) and (TOA), can detect the presence of missing tag in the internal fraud for inventory shrinkage. Addressing this problem will enhance missing tag detection and that will improve the performance efficiency of RFID systems. Therefore, the current approach should be adjusted so as to enhance current approaches for reliable and better result.

Acknowledgements This research is supported by Ministry of Education Malaysia (MOE) under the Fundamental Research Grant Scheme (FRGS) Vot 1611 and Universiti Tun Hussein Onn Malaysia (UTHM).

References

1. Kamaludin H, Mahdin H, Abawajy JH (2018) Clone tag detection in distributed RFID systems. *PLoS ONE* 13(3):1–22
2. Kamaludin H, Mahdin H, Abawajy JH (2016) Filtering redundant data from RFID data streams. *J Sens*
3. Musa A, Dabo AA (2016) AA review of RFID in supply chain management: 2000–2015. *Global J Flex Syst Manage* 17(2):189–228
4. Xie X, Liu X, Li K, Min G, Xue W (2017) Fast temporal continuous scanning in RFID systems. *Comput Commun* 106:46–56
5. Sutar SH, Koul R, Suryavanshi R (2016) Integration of Smart Phone and IOT for development of smart public transportation system. In: 2016 international conference on internet of things and applications (IOTA), Jan 2016. IEEE, pp 73–78
6. Mahdin H, Kamaludin H, Saedudin RR, Omar AH, Kasim S, Jailani J (2017) The application of RFID system in water level monitoring. *Int J Adv Sci Eng Inf Technol* 7(4–2):1522–1527
7. Zhou W, Piramuthu S, Chu F, Chu C (2017) RFID-enabled flexible warehousing. *Decis Support Syst* 98:99–112
8. Manzoor A (2016) RFID in health care-building smart hospitals for quality healthcare. *Int J User-Driven Healthcare (IJUDH)* 6(2):21–45
9. Li B, He Y, Liu W, Wang L (2019) Towards time-efficient localized polling for large-scale RFID systems. *Comput Netw* 150:250–262
10. Yu J, Chen L, Zhang R, Wang K (2017) Finding needles in a haystack: missing tag detection in large RFID systems. *IEEE Trans Commun* 65(5):2036–2047
11. Yu J, Chen L, Zhang R, Wang K (2016) On missing tag detection in multiple-group multiple-region RFID systems. *IEEE Trans Mob Comput* 16(5):1371–1381
12. Chen H, Ma G, Wang Z, Xia F, Yu J (2017) Probabilistic detection of missing tags for anonymous multicategory RFID systems. *IEEE Trans Veh Technol* 66(12):11295–11305
13. Gupta G, Chhabra I (2017) Optimized template detection and extraction algorithm for web scraping of dynamic web pages. *Glob J Pure Appl Math* 13(2):719–732
14. Pagoropoulos A, Pigosso DC, McAlloone TC (2017) The emergent role of digital technologies in the Circular Economy: a review. *Procedia CIRP* 64:19–24
15. Fang Y, Cho YK, Zhang S, Perez E (2016) Case study of BIM and cloud-enabled real-time RFID indoor localization for construction management applications. *J Constr Eng Manage* 142(7):05016003
16. Chai J, Wu C, Zhao C, Chi HL, Wang X, Ling BWK, Teo KL (2017) Reference tag supported RFID tracking using robust support vector regression and Kalman filter. *Adv Eng Inform* 32:1–10
17. Zhao C, Wu C, Chai J, Wang X, Yang X, Lee JM, Kim MJ (2017) Decomposition-based multi-objective firefly algorithm for RFID network planning with uncertainty. *Appl Soft Comput* 55:549–564
18. Srivastava K, Awasthi AK, Mittal RC (2018) Biometric based RFID tag mutual authentication protocol defending against illegitimate access. *Malaya J Matematik (MJM)* 1:102–106
19. Hansen KS ((2016)) U.S. Patent No. 9,504,896. U.S. Patent and Trademark Office, Washington, DC
20. Shahzad M, Liu AX (2016) Fast and reliable detection and identification of missing RFID tags in the wild. *IEEE/ACM Trans Network* 24(6):3770–3784
21. Alsinglawi B, Nguyen QV, Gunawardana U, Maeder A, Simoff SJ (2017) RFID systems in healthcare settings and activity of daily living in smart homes: a review. *E-Health Telecommun Syst Netw* 6:1–17
22. Tan H, Choi D, Kim P, Pan S, Chung I (2018) An efficient hash-based RFID grouping authentication protocol providing missing tags detection. *J Internet Technol* 19(2):481–488
23. Chen H, Wang Z, Xia F, Li Y, Shi L (2017) Efficiently and completely identifying missing key tags for anonymous RFID systems. *IEEE Internet Things J* 5(4):2915–2926

24. Belstner J, Blair A (2018) U.S. Patent Application No. 10/095,996
25. Chen H, Xue G, Wang Z (2017) Efficient and reliable missing tag identification for large-scale RFID systems with unknown tags. *IEEE Internet Things J* 4(3):736–748
26. Gu C (2017) Fast discrepancy identification for RFID-enabled IoT networks. *IEEE Access* 6:6194–6204
27. Woltmann L, Hartmann C, Thiele M, Habich D, Lehner W (2019) Cardinality estimation with local deep learning models. In: *Proceedings of the second international workshop on exploiting artificial intelligence techniques for data management*, July 2019. ACM, p 5
28. Wang S, Mao G (2018) Missing data estimation for traffic volume by searching an optimum closed cut in urban networks. *IEEE Trans Intell Transp Syst* 20(1):75–86
29. Ruan W, Sheng QZ, Yao L, Li X, Falkner NJ, Yang L (2018) Device-free human localization and tracking with UHF passive RFID tags: a data-driven approach. *J Netw Comput Appl* 104:78–96
30. Fan K, Gong Y, Liang C, Li H, Yang Y (2016) Lightweight and ultralightweight RFID mutual authentication protocol with cache in the reader for IoT in 5G. *Secur Commun Netw* 9(16):3095–3104
31. Tao F, Fan T, Lai KK, Li L (2017) Impact of RFID technology on inventory control policy. *J Oper Res Soc* 68(2):207–220
32. Liu X, Xie X, Li K, Xiao B, Wu J, Qi H, Lu D (2017) Fast tracking the population of key tags in large-scale anonymous RFID systems. *IEEE/ACM Trans Network (ToN)* 25(1):278–291
33. Xiao F, Wang Z, Ye N, Wang R, Li XY (2018) One more tag enables fine-grained RFID localization and tracking. *IEEE/ACM Trans Network (TON)* 26(1):161–174
34. Zhang Y, Chen S, Zhou Y, Odegbile O (2018) Missing-tag detection with presence of unknown tags. In: *2018 15th annual IEEE international conference on sensing, communication, and networking (SECON)*, June 2018. IEEE, pp 1–9
35. Liu X, Cao J, Yang Y, Jiang S (2018) CPS-based smart warehouse for industry 4.0: a survey of the underlying technologies. *Computers* 7(1):13
36. Xie X, Liu X, Zhao X, Xue W, Xiao B, Qi H, Li K, Wu J (2019) Implementation of differential tag sampling for cots RFID systems. *IEEE Trans Mob Comput*
37. Burmester M, Munilla J, Ortiz A, Caballero-Gil P (2017) An RFID-based smart structure for the supply chain: Resilient scanning proofs and ownership transfer with positive secrecy capacity channels. *Sensors* 17(7):1562
38. Burmester M, Munilla J (2016) An anonymous RFID grouping-proof with missing tag identification. In: *2016 IEEE international conference on RFID (RFID)*. IEEE, pp 1–7
39. Zhang Y, Chen S, Zhou Y, Fang Y (2018) Anonymous temporal-spatial joint estimation at category level over multiple tag sets. In: *IEEE INFOCOM 2018—IEEE conference on computer communications*, pp 846–854
40. Nguyen CT, Hoang TT, Hoang LT, Phan VX (2019) Efficient missing-tag event detection protocols to cope with unexpected tags and detection error in RFID systems. *Wireless Commun Mob Comput* 2019
41. Liu X, Yang Q, Luo J, Ding B, Zhang S (2018) An energy-aware offloading framework for edge-augmented mobile RFID systems. *IEEE Internet Things J*
42. Bagheri N, Alenaby P, Safkhani M (2017) A new anti-collision protocol based on information of collided tags in RFID systems. *Int J Commun Syst* 30(3):e2975
43. Tu YJ, Zhou W, Piramuthu S (2018) A novel means to address RFID tag/item separation in supply chains. *Decis Support Syst* 115:13–23