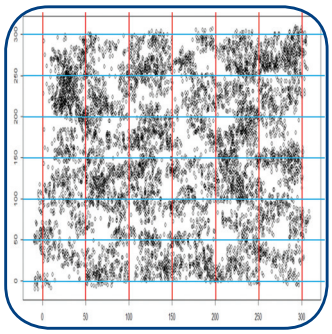
Special issue, ICICTM 2021 ISSN 1847-6996

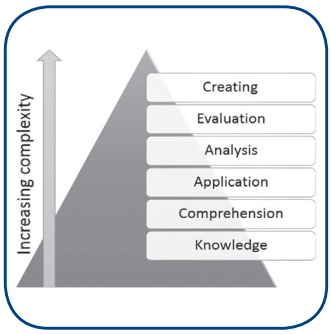
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

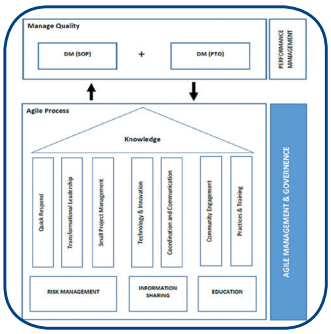
**IJECES** International Journal of Electrical and Computer Engineering Systems

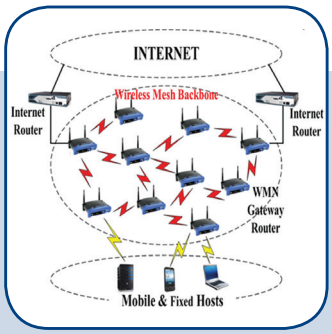
# International Journal of Electrical and Computer Engineering Systems

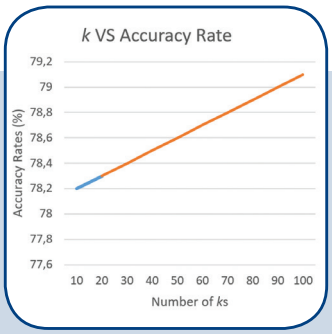
Special issue: Extended papers from the 3rd International Conference of Information & Communication Technology 2021
**ICICTM 2021**

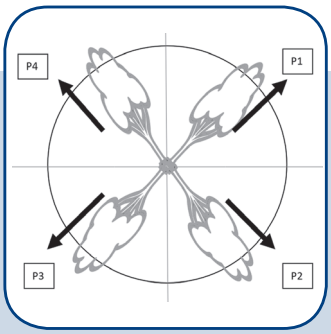




ISSN: 1847-6996

# INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Special issue, ICICTM 2021, 2021 | Pages 1-78

**CONTACT**

**International Journal of Electrical and Computer Engineering Systems (IJECES)**

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

**Subscription Information**

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777,
Croatian Postal Bank



**Journal is referred in:**

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

**Bibliographic Information**

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

**IJECES online**

https://ijeces.ferit.hr

# TABLE OF CONTENTS

# EDITORIAL PREFACE

This special issue is dedicated to selected extended papers presented at the 3rd International Conference of Information & Communication Technology 2021 (ICICTM 2021) held in a virtual environment on March 23, 2021. The conference was organized by the Malaysia Institute of Information Technology, University of Kuala Lumpur, and the Faculty of Defence Science and Technology, National Defence University of Malaysia, in collaboration with several local and overseas universities.

The special issue consists of eight selected extended papers (four original scientific papers, two review papers, one case study and one preliminary communication).

IJECES Executive Editor

Mario Vranješ

# A Comprehensive Performance Evaluation of MIPv6 and PMIPv6 Mobility Management Protocols in Wireless Mesh Network

Original Scientific Papers

**Wei Siang Hoh**
Universiti Malaysia Pahang (UMP),
Faculty of Computing,
College of Computing & Applied Sciences
Pekan, Pahang, Malaysia
weisiang@ump.edu.my

**Bi-Lynn Ong**
Universiti Malaysia Perlis (UniMAP),
Faculty of Electronic Engineering Technology,
Ulu Pauh, Perlis, Malaysia
drlynn@unimap.edu.my

**Si-Kee Yoon**
Universiti Malaysia Perlis (UniMAP),
Faculty of Electronic Engineering Technology,
Ulu Pauh, Perlis, Malaysia
sikeeyoon@gmail.com

**R Badlishah Ahmad**
Universiti Malaysia Perlis (UniMAP),
Ulu Pauh, Perlis, Malaysia
badli@unimap.edu.my

***Abstract*** *– Wireless communication is becoming essential due to the dramatic increase in the usage of mobile devices. The high demand for real-time or instant services requires wireless Internet networks which can support different Quality of Service (QoS) guarantees and different traffic characteristics. All Internet network mobile device services are supported by mobility management protocols. In this paper, we compare the performance of the MIPv6 and PMIPv6 mobility management protocols in the Wireless Mesh Network (WMN) environment. We identify and analyze the MIPv6 and PMIPv6 mobility management protocols' characteristics by using performance indicators. The performance comparison of MIPv6 and PMIPv6 mobility management protocols was conducted in terms of throughput, latency, and packet loss ratio. Based on the conducted experimental results, we summarize the performances for MIPv6 and PMIPv6 mobility management protocols in the Wireless Mesh Network environment. The results obtained indicate that PMIPv6 generally outperforms MIPv6. In future work, the evaluation of HMPlv6, FMIPv6, and FHMIPv6 is proposed.*

***Keywords****: PMIPv6, MIPv6, Network-Based & Host-Based Mobility Management Protocol, Wireless Mesh Network*

## 1. INTRODUCTION

In recent years, the number of Internet users in wireless environments has grown tremendously, causing network distortion issues. A pressing issue is that mobile wireless ecosystems have proliferated rapidly in the wireless environment [1]. These wireless ecosystems play an important role in mobility management protocols. This has led to various mobility management protocols for enabling mobility services. Mobility support protocols can be divided into two main categories: host-based and network-based. Mobile Internet Protocol version 6 (MIPv6) [2], and its enhancements, such as Fast Handover Mobile Internet Protocol version 6 (FMIPv6) [3], Hierarchical Mobile Internet Protocol version 6 (HMIPv6) [4] and Fast Handover for Hierarchical Mobile Internet Protocol version 6 (FHMIPv6), are categorized as host-based. Network-based mobility management protocols have been designed and introduced to address the shortcomings of host-based mobility management protocols [5].

Network-based protocols include Fast Proxy Mobile Internet Protocol version 6 (FPMIPv6) and Proxy Mobile Internet Protocol version 6 (PMIPv6) [6].

In MIPv6, the Mobile Node (MN) plays an important role in the mobile scenario. MN allows the alteration of its network attachment points without disturbing IP packet delivery to or from the node. Access Network Procedures are introduced to maintain the current location of all the MNs in the network. PMIPv6 allows an MN to alter its point of attachment without requiring mobility signaling to be processed at MN [5]. Hence, IP packet delivery is not interrupted, and the MN remains reachable in the topology. There are two types of mobility service provisioning entities: Local Mobility Anchor (LMA) and Mobility Access Gateway (MAG). Fast Proxy Mobile IPv6 (FPMIPv6) [6] was designed and introduced to increase the handover performance by preventing the loss of packets and reducing latency during the handover. Fig. 1 depicts the design structure of network-based and host-based mobility systems.

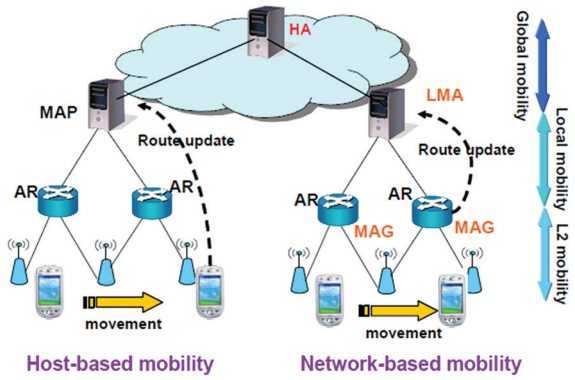



**Fig. 1.** Network-based vs. Host-based mobility

In this research paper, the characteristics and performance profiles of PMIPv6 and MIPv6 are analyzed in the WMN topology environment. These two mobility management protocols have been designed, developed, and analyzed. The performance parameters include Packet Delivery Ratio (PDR), throughput, and latency. The paper content is organized as follows: related works, terminology, simulation design, results, and finally the conclusion.

## 2. RELATED WORKS

M. Skořepa et al. [7] investigated "Analytical Method for L3 handover Latency Evaluation". The researchers conducted an analytical comparison based on MIPv6 handover schemes. They completed the comparison for the four most common handover schemes, including the cost of packet delivery of FHMIPv6, FMIPv6, HMIPv6, and MIPv6. The research access network focused on the IEEE 802.11b family. The transport of the core network was focused on Ethernet (IEEE 802.3 100BaseT). The researchers implemented the analytical methods to obtain the comparison results. Handover cost and handover latency were taken into account by the researchers in the main performance matrices.

A. Ahmad and D. Sasidharan [8] investigated "Handover efficiency improvement in Proxy Mobile IPv6 (PMIPv6) networks". The researchers aimed to reduce the handover delay for PMIPv6 through a communication state-dependent chaining scheme. Chaining based PMIPv6 (CBPMIPv6) was introduced, in which the Mobility Access Gateways (MAGs) were chained to support movements within the PMIPv6 domain. The analytical simulation was conducted using NS2. The results demonstrated that using buffering schemes was able to reduce the packet loss and that handover latency can be reduced through a triggering scheme.

Yan Zhang, et al. [9] investigated "The Simulation of Hierarchical Mobile Ipv6 with Fast Handover using NS2". The researchers performed simulation in the Network Simulator version 2 (NS2) software. Four types of mobile routing protocols were compared to determine which offered the best performance for the mobile network. These included FHMIPv6, HMIPv6, FMIPv6, and MIPv6. Ultimately, FHMIPv6 performed the best against other MIPs in terms of jitter and handover delay. However, this research did not present or discuss any potential reason for the low performances of the other MIPs.

W.K. Jia [10] investigated "A unified MIPv6 and PMIPv6 route optimization scheme for heterogeneous mobility management domains". The researchers designed a unified approach to Route Optimization (RO) scheme. This scheme is based on a simplified MIPv6 Return Routability Procedure (RRP) protocol. It is called Traffic Driven Pseudo Binding Update (TDPBU). The analytical framework of TDPBU for performance analysis included signaling cost, end-to-end latency, throughput, and other performance metrics. The proposed scheme ensured immediate route optimization without considering the residing location of MNs in the heterogeneous MIPv6/PMIPv6 environment. In conclusion, the TDPBU was able to significantly enhance the overall performance of mobility management schemes.

S. Muthut et al. [11] investigated the performance of MIPv6, HMIPv6, and FMIPv6 with WMN. The performance matrix included end-to-end delay, throughput, and packet delivery ratio. The MIPv6, FMIPv6, and HMIPv6 all perform inter-handover to measure the performances under the same network condition. The result showed that HMIPv6 performed better than FMIPv6 and MIPv6 in terms of throughput and packet delivery ratio. In terms of end-to-end delay, MIPv6 outperformed FMIPv6 and HMIPv6 when the network only performed inter-handover. The low end-to-end delay attained by MIPv6 with WMN contributed to low throughput and PDR. However, the authors focused only on performance evaluation. There were no proposed improvements or enhancement schemes to overcome the problems found in their research.

A. Yadav and A. Singh [12] performed performance analysis and optimization of FMIPv6 and HMIPv6 handover. The researchers proposed a new analytical model for the MIPv6 optimization protocol. The researchers used MATLAB 7 software to complete the simulation by collecting a sample size of 100, for both FMIPv6 and HMIPv6. According to the researchers, the new analytical model could reduce the handover by 50% of the original MIP working mechanism. In conclusion, the proposed analytical method is able to significantly reduce handover latency in HMIPv6 and FMIPv6. However, the researchers were only concerned with handover latency as the sole performance metric. The results obtained from this research may have been affected by the mobile node moving with random motion.

## 3. TERMINOLOGY

In this section, all the general terms are discussed in detail. MIPv6 and PMIPv6 protocols, WMN characteristics and behaviors, and handover management are also discussed in detail.

### 3.1. MOBILITY MANAGEMENT CLASSIFICATION

Mobility management enables a mobile device on a data network to change the attachment point. At the same time, the mobile device will have an IP address, which is also known as Mobile Host (MH). A change in IP address can be a big challenge to maintain uninterrupted data flow, ensure security, minimize loss of packets, and identify the newer location. Mobility management can be classified into four types, which are cross layer, upper layer, network layer, and lastly link layer mobility management [13]. In the network layer mobility management, it is further subclassified into 2 main groups, which are micro & macro mobility and host-based & network-based mobility. Fig. 2 below represents the classification of mobility management [14].

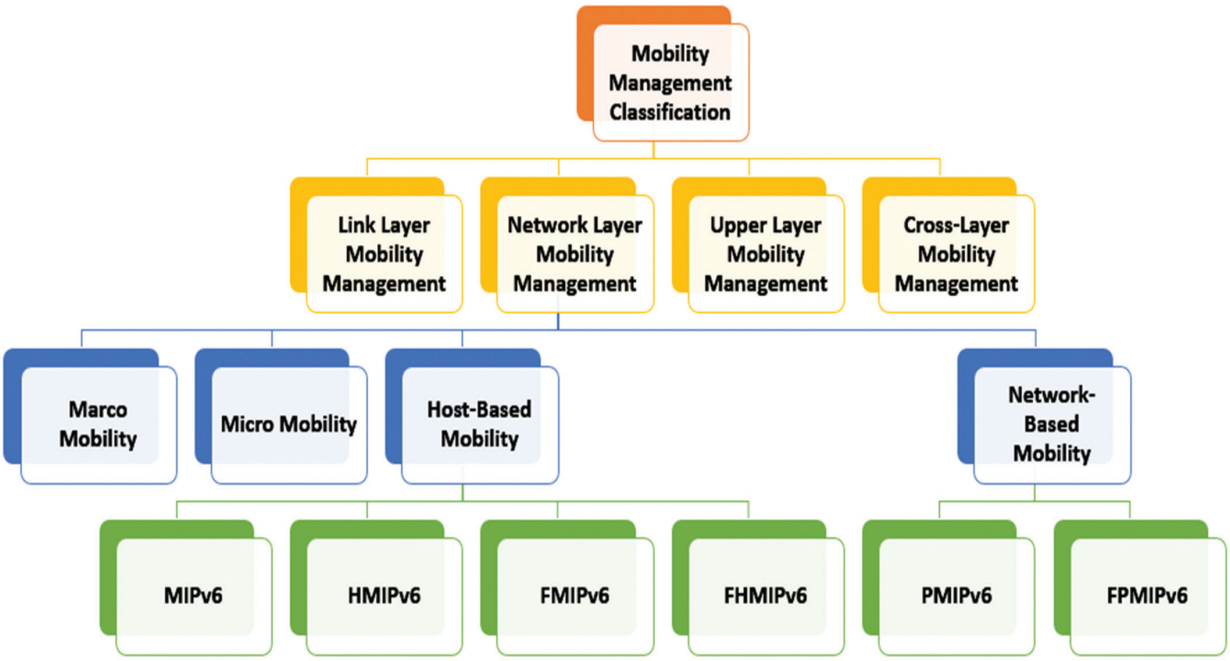



**Fig. 2.** Mobility Management [14]

### 3.2. MOBILE INTERNET PROTOCOL VERSION-6 (MIPV6)

The Internet Engineering Task Force (IEFT) established the use of Mobile Internet Protocol version 6 (MIPv6). The mechanism of MIPv6 allows MN to be reachable and to maintain an ongoing connection, even though its position keeps on changing within the topology. During the entire process, MN is still able to maintain the same allocated IP address [15]. Once the operation is started, the MN will be searching for Foreign Agent (FA) and detecting its movement. It will autoconfigure and set up itself with a New Care of Address (NCoA) through either a stateless or stateful mechanism. Binding Update (BU) is sent by MN and it will forward to its Home Agent (HA) to notify its new address which is available. The HA returns Binding Acknowledgment (BAck). With the help of HA, all packets are tunneled to MN's NCoA. Route Optimization (RO) is another mode for MIPv6. The RO will search for the shortest path and start to transfer the packets. This process requires MN to register its current Binding to Corresponding Node (CN). The CN allows the delivery of the triangulate packets to MN without getting permission from HA. In conclusion, it is able to decrease the congestion at MN's HA and Home Link [16]. Fig. 3 below represents the messages flow of MIPv6.

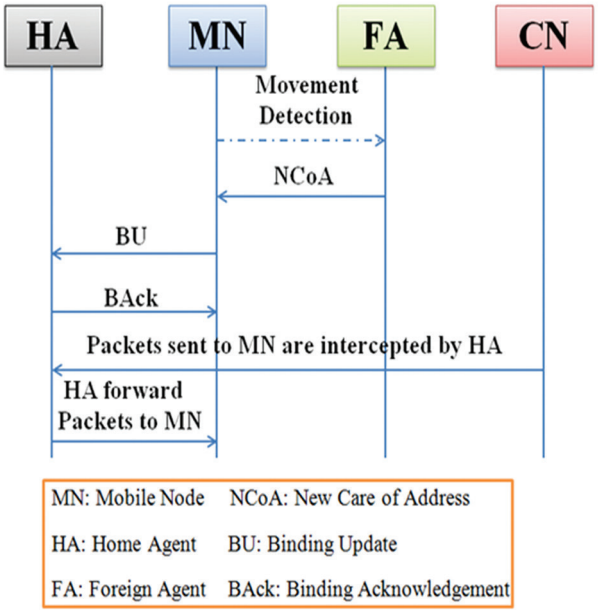



**Fig. 3.** MIPv6 Flow Diagram

### 3.3. PROXY MOBILE INTERNET PROTOCOL VERSION-6 (PMIPV6)

PMIPv6 was developed based on MIPv6's [RFC3775] design. It avoids tunneling over head over the air,

which may cause latency to increase dramatically. Such latency can be seen in MIPv6 [17]. The operation begins when Mobile Node (MN) moves and attaches to an access router which is called Mobile Access Gateway (MAG). Once the authentication is completed, MAG identifies the MN. The MAG obtains the MN's profile, which contains the Home Address and sends the Proxy Binding Update (PBU) to the Localized Mobility Agent (LMA) on behalf of MN. If MAG receives the Proxy Binding Acknowledgment (ACK) from LMA, then it sends Router Advertisements that contain MN's home network prefix. If MAG does not receive Proxy Binding Acknowledgment (ACK) from LMA, it waits and resends the Proxy binding Update to LMA. Fig. 4 below provides an overview of PMIPv6.

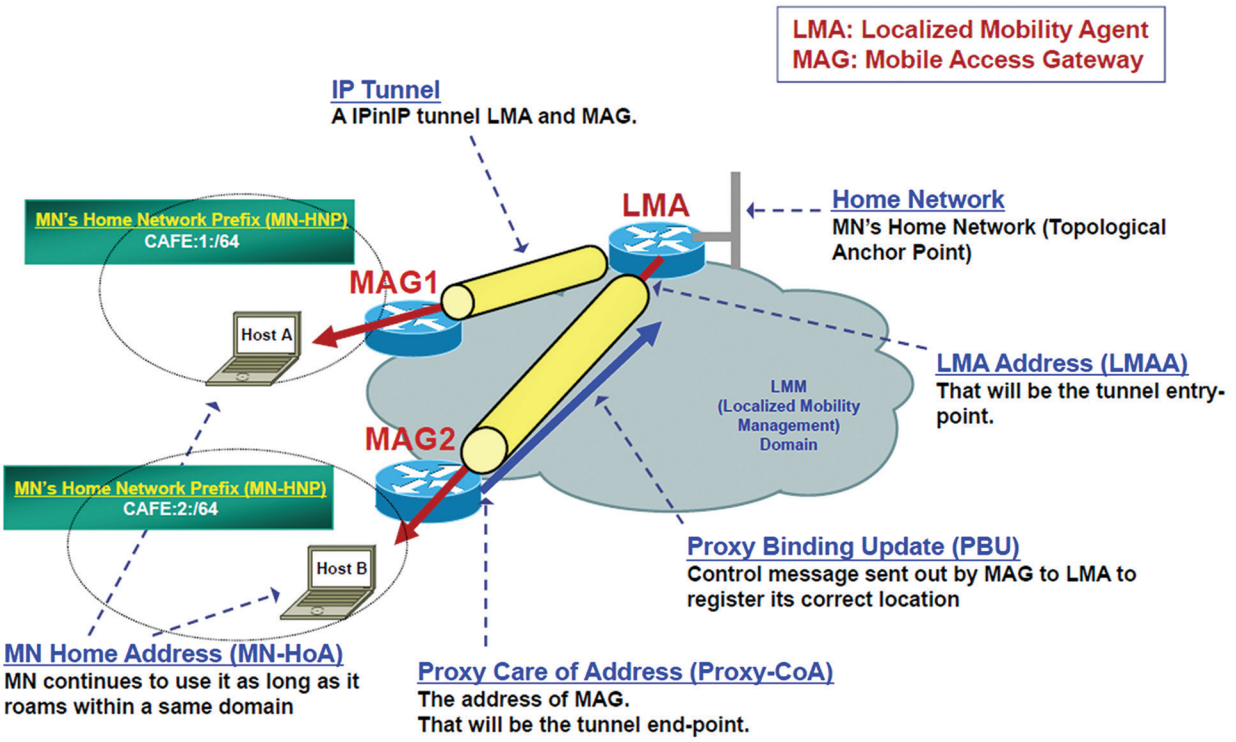



**Fig. 4.** PMIPv6 Protocols [18

This concludes the discussion of the classification of mobility management, mobility management protocols, MIPv6, and PMIPv6. Table 1 shows the summarized characteristics of MIPv6 and PMIPv6. In this paper, MIPv6, and PMIPv6 are analysed in terms of packet delivery ratio, latency, and throughput.

**Table 1.** The summarized characteristic of MIPv6 and PMIPv6 in each protocols criteria

| Protocol Criteria | MIPv6 | PMIPv6 |
|---|---|---|
| Mobility Scope | Global Mobility | Local Mobility |
| Location Management | Yes | No |
| Required Infrastructure | Home Agent (HA) | LMA, MAG |
| MN Modification | Yes | No |
| Handover Latency | Poor | Good |
| Localized Routing | Yes | No |
| Tunneling Over Wireless Link | Required | Not Required |
| Route Advertisement Type | Broadcast | Unicast |
| Return Routability | Required | Not Required |

Wireless Mesh Network (WMN) is defined as an infrastructure-based network. A WMN is a communications network which has various wireless nodes which are sorted in a mesh topology. WMN consists of gateways, mesh routers, and mesh clients. Mobile phones, laptops, tablets, intelligent machines, and other wireless devices are the examples of mesh clients. Mesh routers forward traffic to and from the gateway, but are not connected to the Internet.

The total coverage area of the radio nodes functioning as a single network is defined as a mesh cloud. The mesh cloud is dependent on the radio nodes, which operate in harmony with each other to create a radio network. WMN offers trust and provides good redundancy. When one single node cannot operate or is damaged, the rest of the nodes are still connected and the communication is maintained[19].

Fig. 5 shows how a WMN can self-organize and self-configure without any command from the network operator. WMN can be implemented with various wireless technologies, such as IEEE 802.11s WiFi Mesh and IEEE 802.16 WiMAX. A telecommunication company can expand, replace, and adapt their network based on the requests of the end users [20].

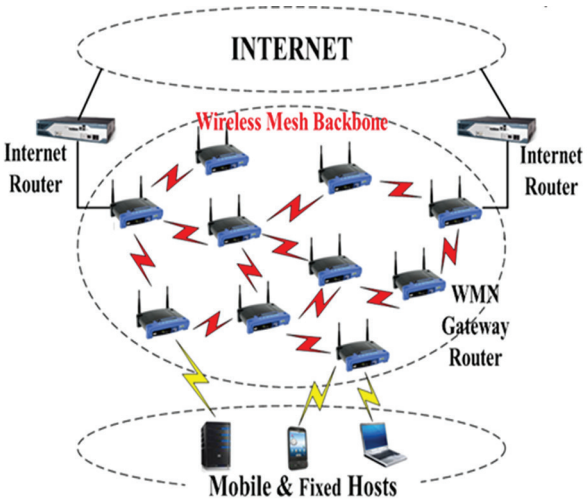



**Fig. 5.** Wireless Mesh Network Architecture [20]

A WMN can provide a more planned and systematic configuration. The deployment of WMN can provide cost-effective and dynamic connectivity over forested or mountainous areas. The network architecture consists of a mesh router that is able to support large areas at a low cost compared to single-hop routers. An alternate routing or dynamic connectivity allows traffic loads to be balanced and minimizes bottlenecks. This may also be able to significantly increase network reliability in WMN. Table 2 shows the difference between WMN and Wireless Ad Hoc Network (WANET).

**Table 2.** Difference between WMN and WANET

| Issues | WANET | WMN |
|---|---|---|
| Infrastructure requirement | Infrastructure-less | Partial or Fully Fixed Infrastructure |
| Network Topology | Highly Dynamic | Relatively Static |
| Energy Constraint | High | Low |
| Application Characteristic | Temporary | Semi-permanent or Permanent |
| Routing Performance | Fully Distributed | Fully or Partially Distributed |
| Geographical | Do Not Consider | Well Perform |

### 3.5. MOBILITY MANAGEMENT IN WMN

As mentioned above, WMN serves as an access network that implements multi-hop wireless forwarding. Hence, the nonmobile nodes relay data to and from the Internet. IETF has also announced that WMN can provide a data transmission rate up to 134.4 Mbps. Hence, it is capable of satisfying the requirements of next generation wireless networks with high speed and low latency. The commercialization of WMN is inevitable. Several working groups focus on the WMN technologies and corresponding specifications (e.g., IEEE 802.16a and 802.11) are being standardized. However, mobility management for mobile users in WMN is not specified clearly.

### 3.6. HANDOVER MANAGEMENT

A handoff or handover is one of the essential parts of mobility management. These terms refer to the same process of changing the point of the connection while a call is in progress. The objective of a handover is to provide seamless handover between mobile terminal and BS. A smooth handover can minimize the loss of data, while a fast handover can decrease the delay to and provide seamless handover service. A handover is needed to meet user preferences. Handover can be classified into two main categories, namely horizontal handover (intracell) and vertical handover (intercell). The main differences between horizontal handover and vertical handover are complexity and symmetry. Due to the different access technologies and their diverse characteristics, a vertical handover is asymmetric and more complex than a horizontal handover.

## 4. SIMULATION DESIGN

In this research paper, the MIPv6 and PMIPv6 protocols were designed, developed, and simulated in Network Simulator version 2 (NS2). Both mobility management protocols are set up in the same WMN environment for comparison and analysis.

### 4.1. THE PARAMETER OF THE NETWORK TOPOLOGY DESIGN

In order to compare the two different types of mobility management protocols, a few configurations and parameters need to be in constant value to obtain the optimum results for both mobility management protocols. The network environment for both mobility management protocols is set up to Mac 802.11. The data rate is fixed to 10 Mb. The interface queue type is drop tail mode.

After setting up the wireless environment, the number of nodes of MIPv6 and PMIPv6 needs to be built up. For MIPv6, the number of nodes is 5 nodes and these consist of one Home Agent (HA) and one Client Node (CN). For PMIPv6, it consists of one Home Agent (HA), one Client Node (CN), one LMA and two MAGs. Table 3 represents the detail of the parameter and its values for setting up the MIPv6 and PMIPv6 mobility management protocols.

**Table 3.** Type of parameters and value

| | |
|---|---|
| Link Delay | 2 ms |
| Data Rate for Mac 802.11 | 10 Mb |
| Window Size (Byte) | 32 |
| Duration | 100 s |
| Transport Protocol | TCP |

### 4.2. PERFORMANCE METRICS

The characteristics and behavior of the network topology of PMIPv6 and MIPv6 can be understood through few performance metrics. The metrics are:

1. Latency mean, which represents the delay when the packet sent from the source passes through the router and base station to the destination; and
2. Throughput, which represents the total data transmitted from one source to receiver in time duration and is normally measured in kilobits per second (Kbps).
3. Packet Delivery Ratio (PDR) represents the ratio of received packets with sent packets between the receiver and source.

## 5. RESULTS & DISCUSSION

Simulation results are presented with detailed discussion. Based on Table 4, each performance metrics result for MIPv6 and PMIPv6 are presented in detail. The packet size starts from 256 bytes, and increases to 512 bytes, 1024 bytes, 2048 bytes, and ends at 4096 bytes.

**Table 4.** Results of MIPv6 and PMIPv6 for different performance metrics

| **Mobile Internet Protocols version 6 (MIPv6)** | | | | | |
|---|---|---|---|---|---|
| Packet Size (Bytes) | 256 | 512 | 1024 | 2048 | 4096 |
| Latency Mean (ms) | 90 | 92 | 93 | 103 | 104 |
| Throughput (Kbps) | 1.080 | 2.212 | 4.444 | 10.609 | 21.381 |
| Packet Delivery Ratio (PDR) | 75.90 | 76.87 | 76.68 | 82.48 | 82.86 |
| **Proxy Mobile Internet Protocols version 6 (PMIPv6)** | | | | | |
| Packet Size (Bytes) | 256 | 512 | 1024 | 2048 | 4096 |
| Latency Mean (ms) | 2 | 2 | 2 | 2 | 3 |
| Throughput (Kbps) | 129.823 | 234.322 | 390.124 | 592.323 | 799.130 |
| Packet Delivery Ratio (PDR) | 99.83 | 99.82 | 99.86 | 99.75 | 99.75 |

Based on the observation of Table 4, when packet size increases from 256 bytes to 4096 bytes, the throughput increases gradually. The throughput for 256 bytes packet size is 1.080 kbps and reaches 21.381 kbps when the packet size increases to 4096 bytes. For PMIPv6, the throughput for packet size 256 bytes is 129.823 kbps, increasing gradually to 799.13 kbps when the packet size reaches 4096 bytes.

Based on Fig. 6, the throughput between MIPv6 and PMIPv6 has a huge difference. The reason is in the basic MIPv6 protocol. In this protocol, during switching between different subnets, the MN needs to go through mobile testing, setting, getting the new address configuration, duplicating address detection, and finally binding the registration process. These processes cause a lot of switching delay, high rate of packet loss, and overload signaling. Throughput is very low due to these reasons. For PMIPv6, since it is completely transparent to mobile nodes, the MAG becomes the proxy and communicates with the Mobile Node. This can lower the probability of signaling overload and decrease the packet loss rate. The throughput of PMIPv6 is very high and it is completely utilized the bandwidth.

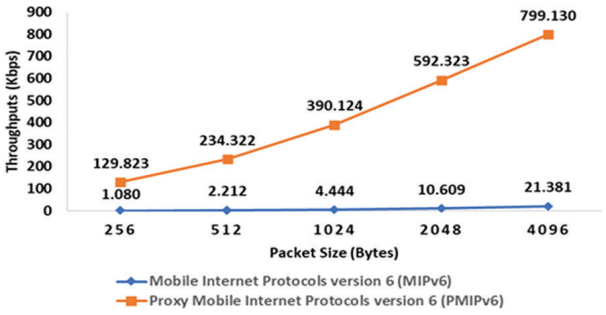



**Fig. 6.** A comparison of MIPv6 and PMIPv6 in terms of throughput (kbps)

Fig. 7 showed the packet delivery ratio of MIPv6 and PMIPv6 for various packet sizes. For MIPv6, the packet delivery ratio increases from 75.90% to 76.87% when the packet size is increased from 256 bytes to 512 bytes. When the packet size reaches 1024 bytes, the packet delivery ratio is slightly decreased to 76.68%. This may be caused by the signaling overload. When packet size increases to 2048 bytes and 4096 bytes, the packet delivery ratio is 82.48% and 82.86% respectively. As compared to PMIPv6, the packet delivery ratio for ev-

ery packet size is higher than MIPv6. When packet size is 256 bytes, the packet delivery ratio is 99.85% and it decreases to 99.82% when the packet size is 512 bytes. The packet delivery ratio slightly increases to 99.86% when the packet size is 1024 bytes. When packet size increases to 2048 bytes and 4096 bytes, the packet delivery ratio is constant at 99.75%. In conclusion, PMIPv6 performed better than MIPv6 in terms of packet delivery ratio.

This is the reason why MIPv6 is used in global networks while PMIPv6 is used in localized networks. In global networks, the handoff procedure is not efficient and causes large latency. The packet drops when it reaches its timeout in TCP transmission. In localized networks where PMIPv6 is implemented, the limited topology contributes to minimal handoff signaling delays, low latency, and lower probability for packet drop due to the timeout. Hence, this explains the packet delivery ratio of MIPv6 is lower compared with PMIPv6.

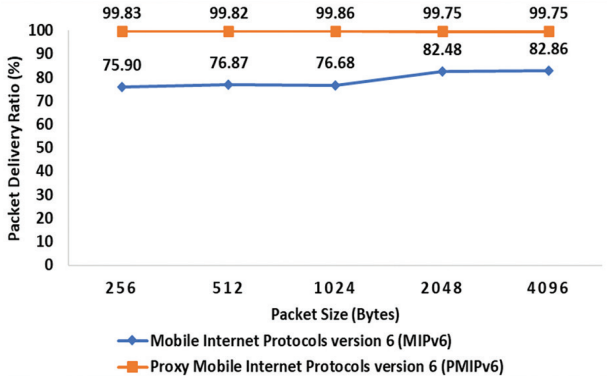



**Fig. 7.** A comparison of MIPv6 and PMIPv6 in terms of aspect packet delivery ratio (%)

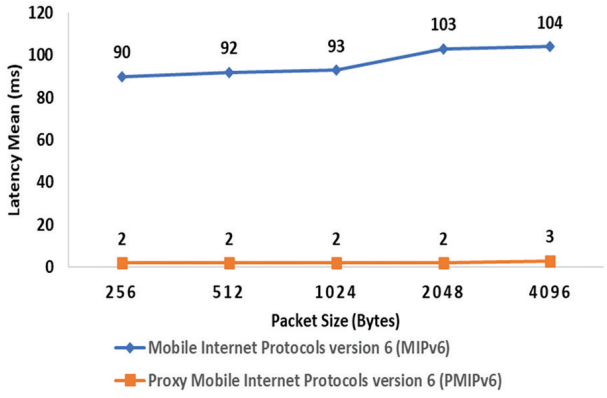



**Fig. 8.** A comparison of MIPv6 and PMIPv6 in terms of aspect latency mean (ms)

Figure 8 shows that the latency of the MIPv6 is very high compared with PMIPv6. For MIPv6, the latency reaches 90 ms when the packet size is 256 bytes; this increases to 92 ms, 93 ms, 103 ms, and a peak of 104 ms when the packet size increases from 512 bytes, 1024 bytes, 2048 bytes, and 4096 bytes respectively. As for PMIPv6, the latency stays constant at 2 ms when the packet size increases from 256 bytes, 512 bytes, 1024 bytes, and 2048 bytes. The highest latency for PMIPv6 is 3 ms when the packet size reaches 4096 bytes.

The high latency of MIPv6 is caused by the handoff procedure. The handoff latency of MIPv6 mainly comes from the process of localized routing, modification of Mobile Node (MN), and return routing. These steps or processes in MIPv6 are excluded from PMIPv6. Thus, the latency of PMIPv6 is lower as compared to MIPv6.

## 6. CONCLUSION

In this research, a comparison between MIPv6 and PMIPv6 was made, and PMIPv6 has been shown to outperform MIPv6. The evaluation shows that PMIPv6 offers better throughput, lower latency, and higher PDR. The basic MIPv6 has various problems which are not able to fulfill the huge demand of mobile users. Hence, PMIPv6 is developed and implemented to overcome the drawback of MIPv6. For future works, the researcher suggests various modified handover methods to overcome this bottleneck. In future, the HMIPv6, FMIPv6, and FHMIPv6 with route optimization schemes should be compared with each other. The proposed methods are believed to outperform MIPv6, and are able to be used in preparation for the wireless network to 5G network technologies.

## 7. REFERENCES:

[1] A. B. Waluyo, W. Rahayu, D. Taniar, B. Scrinivasan, "A Novel Structure and Access Mechanism for Mobile Data Broadcast in Digital Ecosystems", IEEE Transactions on Industrial Electronics, Vol. 58, No. 6, 2011, pp. 2173-2182.

[2] C. Perkins, J. Arkko, D. Johnson, "Mobility Support in IPv6", IETF RFC 3775, 2004.

[3] R. Koodli, "Fast Handovers for Mobile IPv6", IETF RFC 4068, 2005.

[4] H. Soliman, C. Castelluccia, K. Elmalki, L. Bellier, "Hierarchical mobile IPv6 mobility management (HMIPv6)", RFC 4140, p. 5380, 2008.

[5] K. Chowdhury, K. Leung, B. Patil, V. Devarapalli, S. Gundavelli, "Proxy Mobile IPv6", RFC 5213, 2008.

[6] H. Yokota, K. Chowdhury, R. Koodli, B. Patil, F. Xia, "Fast handovers for proxy mobile IPv6", RFC 5949, 2010.

[7] M. Skořepa, R. Klügl, "Analytical method for L3 handover latency evaluation", Proceedings of the European conference of systems, and European conference of circuits technology and devices,

and European conference of communications, and European conference on Computer science, Tenerife, Spain, 30 November 2010, pp. 342–347.

[8] A. Ahmad, D. Sasidharan, "Handover efficiency improvement in Proxy Mobile IPv6 (PMIPv6) networks", Procedia Computer Science, Vol. 46, 2015, pp. 1064-1071.

[9] Y. Zhang, H. Bi, "The simulation of Hierarchical Mobile IPv6 with fast handover using NS2", Procedia Engineering, Vol. 37, 2012, pp. 214–217.

[10] W. K. Jia, "A unified MIPv6 and PMIPv6 route optimization scheme for heterogeneous mobility management domains", Computer Networks., Vol. 75, 2014, pp. 160–176.

[11] S. Muthut, B.-L. Ong, N. A. H. Zahri, R. B. Ahmad, "An overview of performance enhancement of FHMIPv6 on wireless mesh network", International Journal of Future Computer and Communication, Vol. 4, No. 3, 2015, pp. 160–164.

[12] A. Yadav, A. Singh, "Performance analysis and optimization Of Hmipv6 and Fmipv6 handoff management protocol", International Journal of Engineering Research, Vol. 5, No. 3, 2014, pp. 305–308.

[13] W. S. Hoh, S. Muthut, B.-L. Ong, M. Elshaikh, M. N. M. Warip, R. B. Ahmad, "A survey of mobility management protocols", ARPN Journal of Engineering and Applied Sciences, Vol 10, No. 19, 2015, pp. 9015-9019.

[14] W. S. Hoh, B.-L. Ong, R. B. Ahmad, H. Ahmad, "Consolidation of Host Based Mobility Management Protocols with Wireless Mesh Network", Innovative Computing, Optimization and Its Applications, pp. 111-129, Springer, 2018.

[15] X. Wu, G. Nie, "Comparison of different mobility management schemes for reducing handover latency in Mobile IPv6", Proceedings of the International Conference on Industrial Mechatronics and Automation, Chengdu, China, 15-16 May 2009, pp. 256-259.

[16] S. Shayma, I. Mahamod, J. Kasmiran, "A comparison of mobile node 's handoff between mobile IPv6 and fast handover protocol", The Institution of Engineers Malaysia, 2008.

[17] S. Y. Ren, R. Chai, L. Tang, Q. B. Chen, "Proxy Mobile IPv6 based inter-domain mobility management approach and performance analysis", Application Research of Computers, Vol 27, No. 3, 2010, pp.1118-1121.

[18] N. Neumann, J. Lei, X. Fu, G. Zhang, "I-PMIP: an inter-domain mobility extension for proxy-mobile IP", Proceedings of the International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly, Leipzig, Germany, 21-24 June 2009, pp. 994-999.

[19] Z. L. Lee, B.-L. Ong, A. Amir, W. S. Hoh, "A survey of session initiation protocol in Wireless Mesh Network", Proceedings of the IEEE 15th Student Conference on Research and Development, Wilayah Persekutuan Putrajaya, Malaysia, 13-14 December 2017, pp. 286-290.

[20] S. Muthut, B.-L. Ong, W. S. Hoh, R. B. Ahmad, "Integration of fast handover and hierarchical mobile internet protocol with wireless mesh network", Australian Journal of Basic and Applied Sciences, Vol 9, No. 25, 2015, pp. 72-78.

# SIEM Network Behaviour Monitoring Framework using Deep Learning Approach for Campus Network Infrastructure

Original scientific paper

**Mohd Azmi Bin Mustafa Sulaiman**
National Defence University of Malaysia
Faculty of Defence Science and Technology,
Computer Science Department
Sungai Besi Camp, Kuala Lumpur, Malaysia
azmi@upnm.edu.my

**Mohammad Adib Khairuddin**
National Defence University of Malaysia
Faculty of Defence Science and Technology,
Computer Science Department
Sungai Besi Camp, Kuala Lumpur, Malaysia
adib@upnm.edu.my

**Mohd Rizal Mohd Isa**
National Defence University of Malaysia
Faculty of Defence Science and Technology,
Computer Science Department
Sungai Besi Camp, Kuala Lumpur, Malaysia
rizal@upnm.edu.my

**Mohd Nazri Ismail**
National Defence University of Malaysia
Faculty of Defence Science and Technology,
Computer Science Department
Sungai Besi Camp, Kuala Lumpur, Malaysia
m.nazri@upnm.edu.my

**Mohd Afizi Mohd Shukran**
National Defence University of Malaysia
Faculty of Defence Science and Technology,
Computer Science Department
Sungai Besi Camp, Kuala Lumpur, Malaysia
afizi@upnm.edu.my

**Aznida Abu Bakar Sajak**
University Kuala Lumpur,
MIIT, Computer Engineering Technology
aznida@unikl.edu.my

***Abstract*** *– One major problem faced by network users is an attack on the security of the network especially if the network is vulnerable due to poor security policies. Network security is largely an exercise to protect not only the network itself but most importantly, the data. This exercise involves hardware and software technology. Secure and effective access management falls under the purview of network security. It focuses on threats both internally and externally, intending to protect and stop the threats from entering or spreading into the network. A specialized collection of physical devices, such as routers, firewalls, and anti-malware tools, is required to address and ensure a secure network. Almost all agencies and businesses employ highly qualified information security analysts to execute security policies and validate the policies' effectiveness on regular basis. This research paper presents a significant and flexible way of providing centralized log analysis between network devices. Moreover, this paper proposes a novel method for compiling and displaying all potential threats and alert information in a single dashboard using a deep learning approach for campus network infrastructure.*

***Keywords****: SIEM, Network Behaviour Monitoring, Campus Network Infrastructure*

## 1. INTRODUCTION

Network security plays a critical part in Information Technology. It is still difficult for organizations to meet security standards. Identity attacks, intrusions, and hacking have been the most common security threats to the public and have also highlighted the importance of information security [6]. By focusing on threats of both internal and external of the network, network security can secure and stop the threat from entering and spreading on the network. A secure network involves a complex connection of hardware devices such as firewalls, routers, and anti-malware tools.

In the campus network, all system and server equipment depends on the admin to collect logs of network equipment and servers, and also to monitor and notify the system status to users. Therefore, it is important to have comprehensive centralized log management in a campus network. It uses to analyze events that occur from thousands of nodes to several dedicated servers where central analysis is carried out. When the analyses are obtained in a real-time process, the safety events can be identified from future events through event correlation and other advanced surveillance techniques. Moreover, it also can be an offline forensic activity,

where the past events are investigated to understand the occurrence of security that has taken place.

Aggregate the data generated from multiple sources, identify specific threats and take appropriate action are the basic principles of each analysis of network and security reporting system. For instance, the system can take additional log information, generate alerts and ensure that all security controls can be monitored and prevented when such issues are detected. The networking, software, hardware, and media used to produce, distribute, store, analyze, and erase log data are referred to as log management infrastructure. Almost every organizations have one or more log management infrastructures.

Most organizations or businesses use SIEM (Security Information and Event Management) tool. This tool is designed to simplify company compliance reporting through the usage of a centralized logging solution. Each host that is in use must have a log security record included in the report and can pass log data to the SIEM server. Single SIEM servers can collect log data from as many devices as they need and can produce a detailed report and manage all security events of each log they receive. In the current situation, each system needs to be able to manually retrieve data from each device regularly and to ensure that a central configuration of configuration can be generated to produce a report.

The SIEM system server is a tool for detecting unidentified events. Almost most of the equipment used does not comply with safety regulations and cannot track events or logs more deeply. Although such tools can identify and monitor events and produce audit log entries, they cannot analyze logins to detect unacceptable activities. Best of all, tools such as personal computers and laptops can alarm users when an event occurs. SIEM equipment can also perform higher detection by linking the events or logs of the equipment used. By collecting the events or logs of the linked equipment, the SIEM system can see attacks that have different angles on each of the different devices and can therefore record events or logs to decide if the attack is of nature and if it works.

SIEM equipment is used to improve the ability to manage any future accidents that can save time and money for incident handlers. The ability to deal with accidents rapidly and effectively will speed up the delay of occurrence, thus reducing the safety risk that cannot be followed by security events. SIEM equipment can also increase performance, mainly by offering a single report and review to display all security log data from many of the devices connected to it.

This research will provide a significant and flexible way of providing centralized log analysis between the security and network devices and how to display all threats alert information in a single dashboard. The system can assist the IT administrator in collecting, analyzing, storing, investigating, and reporting on the logs and other data for forensics, incident response, and regulatory compliance purposes and as well as analyzing real-time event data to aid the early detection of advanced threats, data breaches, and targeted attacks. Hence, the proposed framework could provide an effective way of presenting the log file to the management.

This research paper is divided into four sections and begins with the introduction. The second section is on the literature review on SIEM concepts as well as the current research relevant to this research paper. It is then followed by the third section on the proposed network behavior monitoring framework based on SIEM concepts using deep learning analysis. In the fifth section, the case study evaluation is presented. The sixth and last section is on the conclusion and future works.

## 2. LITERATURE REVIEW

### 2.1 CURRENT LOG MANAGEMENT ISSUES

#### 2.1.1 Logs are scattered

It is very hard to compile and view each event in the campus network and therefore, all logs in the campus network have been stored individually on their system. Few tools for log management, rather than performance and capabilities, are listed in a random order [1]. Although threat detection platforms such as SIEM are significantly effective based on the recent reports that Sayed [5] found.

#### 2.1.2 High number of false positive

Based on Filkins [4], the network administrator and the company network infrastructure monitoring are facing numerous tools which are not integrated. Open standards are developed and maintained through a collaborative process that consensus-driven to facilitate interoperability and exchange of information between different products and services.

#### 2.1.3 Logs are scattered

For analysts, a solution needs to be created. It will not be meaningful if Syslog only pulls from various data sources. While it is not difficult to preserve the data collected with traditional methods such as hacking, it is an enormous challenge in an IoT environment to preserve the scene [2].

#### 2.1.4 Lack of support & expertise

Some logs sometimes are a massive difficulty. As a result, the agency needs to recruit dedicated staff to support the collection, analysis, correlation, and normalization of all the logs collected, or to retain time for the current team. The rapid growth of the campus network and several IT staff help with the challenging size of data. Monitoring, maintaining, and expanding IT budgets 24/7. This means that the campus network

must recruit professional staff or reserve the time of the current team to support the collection of data to detect, analyze, correlate and normalize all the logs collected. Crowley and Pescatore [3] discovered that the most often reported reasons for existing SOC failures to reach excellence are a lack of competent employees, a lack of resources, and effective automation.

## 2.2 SECURITY AND EVENT MANAGEMENT INFORMATION (SIEM)

### 2.2.1 SIEM Components

Organizations need to protect themselves regularly from the daily growing number of cyber-attacks. Security and Event Management Information (SIEM) is a security system that is widely used by various organizations to protect their networks from cyber-attacks. A SIEM solution consists of several components to assist security teams to identify data violations and malicious activities through constant monitoring and analyzing network devices and events. According to Chikonga [7], the SIEM component includes a collection from various systems, network devices, and applications collected, filtering, aggregation, normalization, and correlation of event messages. These features are a significant improvement from SIEM event log collection and storage.

Sayed [5] reports that SIEM operates hierarchically by deploying different agents. SIEM also gathers security data for specific safety equipment and tools including intrusion detection systems, firewalls, and antivirus, as well as event information from end-user devices, network equipment, and systems servers. The gathered data is sent to the centralized control and administration console. On the central console, additional logs and anomalies analysis are performed. SIEM product roles are collected, consolidated, correlated, communicated, and controlled generally. Initially, log data from various devices and applications are collected. The data is then added and standardized, as a process called consolidation. Afterward, the log data is analyzed and linked. This step enhances the usefulness of contextual network information and common threats. The data is initially kept locally in an organization's network until being transferred for analysis and archiving to a central area.

### 2.2.2 SIEM Architecture

Based on Figure 1, the SIEM architecture is explained as below:

- Server: It represents the core part of the entire deployment that collects and processes for the correlation engine of the logs from the external world.
- DataBase: It stores all the data to analyze the SIEM itself and to set the runtime (asset tables, taxonomies, basic modules configuration, etc.).
- FrontEnd: A console that offers a server's user interface. It provides a visual panel for the security administrator to both controls the individual component configuration and analyzes the system security under control with specific dashboards.
- Probes: A collection of sensors used within the infrastructure monitored. Probes typical examples include firewall and intrusion prevention, perimeter protection, host sensors, and security applications, for example like host IDSs.
- Agents: probes that are integrated into a server and can convert heterogeneous logs generated by various samples into logs with a syntax and a semantic.

The probe can be used to retrieve the information from IT components like routers, firewalls, web servers, anti-virus systems, and intrusion detectors so that it can generate analyzable logs. Probes usually work in two modes: active and passive. The controlled IT component cannot generate logs in the active mode, thus the sample must retrieve information actively through specific queries. For the passive mode, the component monitored is capable of generating and sending logs to the sensor in order not to require ad-hoc queries. Upon retrieval of logs, each probe may carry out preliminary security analysis by using such information. When a security issue arises, a probe alert will be generated and an information log will be sent to the respective agent representing the entry point of the SIEM architecture.

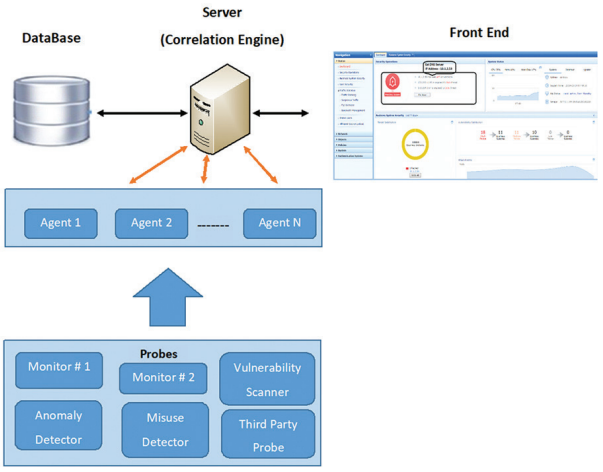



**Fig. 1.** A classical architecture of a SIEM system [9]

### 2.2.3 Benefits of SIEM

The advantage of using a SIEM tool is it serves to streamline a centralized logging solution for reporting business conformity. The host used must contain a log security record and can transfer log data to the SIEM server. Single SIEM servers can receive log data from as many devices as needed, can generate a comprehensive report, and handle every log security event it receives. It is unlikely that the efficient central logging capacity of

agencies or companies without SIEM can generate fast and concise reports, as required for their reports to comply. In the current environment, the reporting of each device is essential to periodically recover data manually from each device and to ensure a central configuration for the generation of a report can be created.

The unknown events can be detected by the SIEM system server. SIEM tool can also perform better detection by connecting used equipment events or logs. The SIEM system can see attacks from various angles to each of the various equipment by collecting the events or logs of the connected equipment, and can thus record events or record logs to determine whether the attack works and its nature. SIEM tool is used to increase the ability to handle events that can save incident handlers time and resources. The ability to manage incidents fast and efficiently can accelerate the delay, thereby reducing the security risk that security events cannot be followed. The SIEM tool can also enhance efficiency mainly by reporting and analyzing each security log dataset from many connected devices.

#### 2.2.4 SIEM Security Analysis Techniques

##### 2.2.4.1 Event Normalization

Collecting data in the scope of SIEM is almost impossible for any person to process the data in its raw state. Through a process called event normalization, it can be performed by peculiar parsers and requires a method to modify logs and alerts delivered by probes to present homogeneous data formats to the server [9]. The process involves splitting each field of a raw event into variables and then combining them into views that are important to security administrators. This process is very important for seeking significance in often isolated and heterogeneous events.

SIEM systems can normalize logs to allow efficient analysis of data from various sources and event correlation. This normalization method includes processing logs in a readable and standardized format, extracting important data from them, and mapping the various fields that they contain.

##### 2.2.4.2 Event Correlation

SIEM system using correlation rules to detect any potential suspicious events in the presence of encrypted real-time traffic [9]. The amount of recorded data is huge in environments. Even small to medium-sized companies are likely to send tens of GB of data every day. Sorting one-by-one authentication logs is a complicated task. Thus, using correlation rules a SIEM solution can solve the problem.

This allows administrators to see anomalies like login attempts from suspicious locations, network scans, and simultaneous user authentication attempts from various locations. The SIEM also monitors network traffic using this rule for better detection of threats and unusual activity [11]. It also can automatically extract this important information into a report or diagram that allows us to visualize activities from many sources. Events are generated using raw data to search for patterns, map them to known expressions, and assign unique identifiers. If the SIEM meets an unknown log or data type, it can define an event by the editor and allocate variables, such as name, severity, and facility.

The correlation rule is being a specific sequence of events that could be indicative of a branch in security. It combines multiple normalized events from different sources into a single correlated event [8].

##### 2.2.4.3 Mining Process

There are millions of data and database need to be processed for useful information. In IT infrastructure, event logs provide an input of all the activities for any organization. The raw data acts as a SIEM input, providing security alerts and output reports. All raw data can be processed through a data mining technique. According to Zope et al., [12], this technique can quickly be implemented on existing software platforms and hardware platforms for enhancing the value of existing information resources. It is the process of examining data from several angles and synthesizing it into meaningful knowledge. The procedure allows people to comprehend the content of data relations. It identifies hidden patterns and trends in the data. As the data mining scope is applied to all event logs created by various networking devices, systems, and application servers, the performance of corporate security may be improved.


**Fig. 2.** Data mining architecture [12]

Based on Figure 2, the data layer might be either a database or a data storage system. An interface is a layer that connects all data sources. The findings of data mining are saved in a data layer so that they may be displayed to the user as reports or another type of visualization. To obtain database data, the application layer of the data mining is employed. There is a transformation code here that will turn the data into the necessary format. The data is subsequently processed using various data mining methods. This layer provides an easy-to-use user interface that allows end-users to engage with the data mining technology. The findings are displayed to the user via a visualization form on the front end layer.

##### 2.2.4.4 Attack Graphs

Attack graphs represent a system's exposure. This section aims to view the approach used for calculating security measurements almost in real time. This

approach should enable new security information and events in the network operating process to be taken into account and security metrics to be recalculated appropriately. To do that, it needs to develop a characterization of metrics that takes into account the following aspects. In a recent survey in the area of security metrics, modeling of attacker steps is as attack graphs. It uses known and adopted techniques to calculate security metrics. Based on these metrics, it identifies the current security situation, including attack existence, skill and position of attackers, potential previous attackers, and future attack targets.

#### 2.2.5 How SIEM work

##### 2.2.5.1 Collection

SIEM system collects event and logs data generated throughout a company's infrastructure from host systems, applications, and security devices such as antivirus filters and firewalls. The information that is gathered is sent to a centralized platform. The SIEM identifies and classifies data into successful and failed logins, the activity of malware, and other possibly malicious activities. So the threats can be detected and security alerts can be created. The customized dashboards and event management system of SIEM enhance investigative efficiency and reduces waste of time on false-positive elements. Real-time monitoring and incident management can be performed by SIEM for security-related events which are collected from the network, security devices, system, and applications.

##### 2.2.5.2 Consolidation or Normalisation and Aggregation

Consolidation happens when all the collected log data from various devices and applications are then aggregated and normalized. The SIEM consolidates logs, parsers every log, and classifies them into event types, including successful and unsuccessful logins, exploits attempts, malware activities and, port scans. Peculiar parsers are used to normalize and involve in a process for manipulating logs and alerts provided by probes, to represent homogenous data formats to the server. These types of events are then run with the outline rules to decide whether illegal traffic exists. If a rule is triggered, an alert will be created. This step enhances the usefulness of contextual information concerning a network and common threats. The collected data will be stored locally in the organization's network first before it is transferred to the analysis and archiving central area. With this step, the data security violations can be studied as closely as necessary with the data classified.

Aggregation is described as a collection of large amounts of data in one place from different applications and databases. SIEM aggregates device data and interprets key attributes related to the identification of security incidents or problems. Devices generate event logs that are submitted for analysis to the SIEM. SIEM tools form an important part of the ecosystem for data security. It aggregates data from several systems and analyzes data to capture abnormal behavior or potential cyber-attacks. SIEM tools play an important role in the collection of events and alerts.

##### 2.2.5.3 Correlation and Contextual Information

The core of SIEM architecture for correlation combines several normalized events from various sources into one correlated event. The log storage is used to store log volumes for retention purposes and historical queries. This correlation is done when the presence of relevant patterns of events is detected. The correlation method is used as a means for performing detection at multiple layers. The analysis of the data collected through the introduction of a set of correlation rules that detect potential suspicious events as encrypted real-time traffic. To add additional data and filter duplicate events, the correlation between data and external services is preferred. If there is a feedback mechanism, a consumer of data should use the mechanism to provide providers with information to improve the quality of their service.

Every correlation was shown to gain greater insight, eliminate false-positive effects, or detect replicates for even simple cyber incidents. The process of comparing incidents is the identification of patterns and relationships to determine events from multiple sensors and data sources that are the result of an attack or a general indicator of malicious activities. It enables an improved understanding of the nature of an event, reduces the workload required to deal with incidents, and automates the classification and forwarding of incidents that are only relevant to a specific constituency. Correlation is useful both for the processing of data on a monitored network from multiple tools and for the use of multiple external services which supply incident data.

This step is decided to make more helpful by contextual information on a network and common threats. In the network organization, the data is first stored locally before they are transferred for analysis and archiving to a central area.

##### 2.2.5.4 Communication or Alerting/Reporting

An evaluation of three elements for the basic assessment parts of the SIEM system. Firstly, the central console, secondly, the monitoring entity, and, lastly, the process of communication between the control entity and the central console. For a SIEM to operate successfully, its design and development shall provide complete, integrated information to the central console for the supervisory entity and the communication process. The core role of the SIEM solution involving the analysis and detection of system-related incidents can be compromised through attacks on communication channels.

A proper evaluation of the SIEM solution aids in preventing an attacker from evading the system. A proper

SIEM evaluation should consist of three major steps. Firstly, to guarantee that the SIEM solution identifies the greatest number of threats while producing the fewest false events and alerts, it should first assess the entities that gather, aggregate, correlate and analyze audit log data from the monitoring entity. Secondly, the audit data collector or the Agent should be assessed independently to guarantee that all information and data acquired by the Agent is valid and truthful. Finally, all communication between SIEM entities must be evaluated and ensured that no attacks such as packet injections and packet alterations occur in the channels.

#### 2.2.5.5 Control or Storage

The usage of SIEM in IT security has been shown to improve security professional's capacity to monitor security risks. When monitoring the threat of outgoing traffic, it may be used to examine traffic going for external locations with a high-risk rating based on their IP addresses. While this traffic was created manually and only from a few log sources, the availability of this type of event data would be useful for security data in monitoring network security risks using a centralized SIEM integrated with threat intelligence services such as IP.

In a study of botnet detection, the researchers similarly describe methods for botnet detection based on output traffic monitors to potential botnet control and control centers. The monitoring of outbound traffic to enrich threats and linked to user-related event data, for example, would allow the security professionals to see more clearly both the internal source of suspect traffic and the destination of the traffic. Insider threats are more and more recognized by the need to monitor user activity as just as damaging as external security threats. The correlation of events that might be part of a composite security threat, while reducing false positives that were an issue in security systems such as IDs, can improve threat detection in an environment with a large number of log sources [7].

As with log management solutions, SIEM technology guarantees that gathered event logs from the application, network, and system components distributed inside an IT infrastructure are consolidated. SIEM, on the other hand, offers more sophisticated features. These capabilities include advanced filtering, aggregation, normalization and correlation, alerting, and reporting. The SIEM should allow an analysis of events in real or close to real-time. This research paper focuses on the use of the SIEM framework to enhance IT security management.

### 2.2.6 Current Research on SIEM

Here are a few examples of SIEM research conducted by other authors that are related to this research paper. In their study, K. Agrawal and H. Makwana [1] examined a few log management tools based on criteria such as data input type, primary application area, SDK accessible for languages, dashboard design capabilities, pricing, real-time supportability, online interface, etc. On the other hand, Sayed [5], discussed on SIEMs and advanced evasion techniques. The paper examined the most frequent AETs as well as the tools used to carry out such attacks. An Adaptive learning system based on an Artificial Immune System (AIS) has been proposed by A. Majeed et al. [11]. The authors use the near-miss situation based on visual analysis for the SIEM rules. L. Coppolino, et al [8] in their research paper discussed the comparison study between OSSIM and GET Data. They concluded that data collecting is the primary role of SIEM design, Indeed, by combining information from several data sources, SIEMs may provide diverse viewpoints on security incidents that occur throughout the system.

## 3. THE PROPOSED FRAMEWORK

According to Agrawal and Makwana [1], the best functional log management tools must consist of the following components; Log management, Log analysis, and Event Management (Figure 3). The Campus network environment required a SIEM solution which provides all in one feature to identify, analyze, correlate, normalize and security logs from multiple data source in the campus network. The complete security features must provide essential security capabilities in a collectible platform controlled by a single management console.

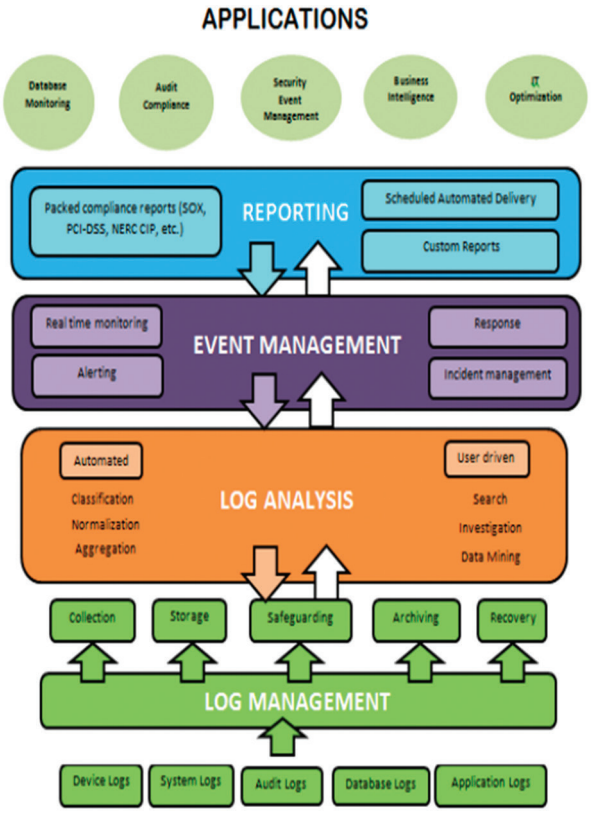



**Fig. 3.** Flowchart of working of log management tools [1]

Based on [1] framework, we proposed our network behavior monitoring framework to suited the campus network environment (see Figure 4). The proposed framework is divided into three components namely i. Log management, ii. Log analysis and iii Event management.

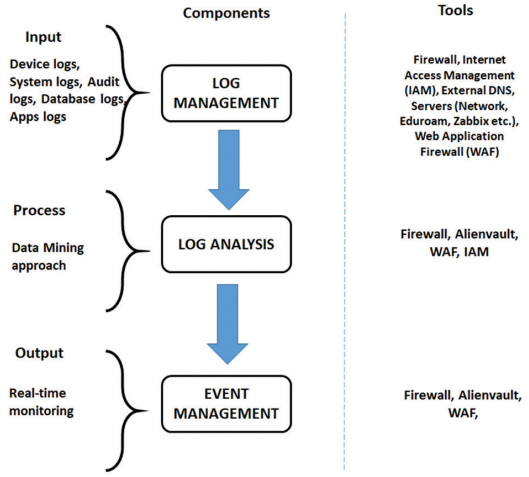



**Fig. 4.** The proposed framework of network behavior for campus network

### 3.1 COMPONENT 1 – LOG MANAGEMENT

Comprehensive centralized log management is a must to simplify the campus network. Centralized event log management is the most important component of monitoring network security and forensics, as it can analyze events that occur from thousands of nodes to several dedicated servers where central analysis is carried out. The analyses obtained can be a real-time process, where safety events can be identified from future events through event correlation and other advanced surveillance techniques; it can also be an offline forensic activity, where past events are investigated to investigate the occurrence of security that has taken place.

The underlying principles of every analysis of network & security reporting system is to aggregate data generated from different sources, identifying relevant threats, and taking appropriate action. For example, when certain problems are identified, the system will take additional log information, generate warnings and ensure other security controls can be controlled and stop them.

Log management infrastructure is a component covering networking, software, hardware, and media used to store, transmit, generate, analyze and delete log data. Almost all organizations have one or more log management infrastructure.

The main functions of a log management tool are:

- Identify and collect all logs of events involving involved software such as operating system, Syslog, flat file, database, or application
- Ensure all logs are stored integrally, scalable and secure
- All logs can be obtained quickly, fast and flexible
- Logs can be retrieved and stored for a long time
- Systems, databases, applications, databases, and devices are available in real-time.
- Normalization, aggregation, log classification, and correlation can be automated more efficiently.

### 3.2 COMPONENT 2 – LOG ANALYSIS

The log analysis describes how the security and traffic threats, the complete security log analysis obtained provides critical network intelligence for attempts to violate security and attacks such as viruses, trojans, service denials, and others. From the log report analysis obtained from the NGFW and WAF, security administrators and networks will be able to interpret network threatening activities and plan their strategies to protect and address threats that occur.

For servers (Eduroam and Sybase) include operating systems involving Windows, Linux, and Unix systems and other Syslog support devices, and applications such as IIS, MS SQL event log analysis is recorded for secure security. Important logs and security events are recorded and generated on equipment within the network, this SIEM system can collect important information, perform log analysis, and display all logs and events on SIEM Dashboard, in real-time and concise real-time.

Internet Access Management (IAM) is used to analyze traffic logs that provide detailed and valuable information about bandwidth usage, employee internet usage, web page confusion broadband, and smart interface traffic. From the firewall analysis report, network/security administrators will monitor the fair use of broadband to reduce existing traffic safety and data security to plan for future broadband capacity requirements.

This framework was proposed using the deep learning approach for log evaluation. Deep learning is a complex element of machine learning inspired by the function of interconnecting neurons in the human brain. The evolution of Machine Learning and an element of AI teaches itself to make more accurate and faster predictions by observing, processing and analyzing massive amounts of data. Over time, Deep Learning teaches itself every time it is executed, resulting in the identification of many previously undetected malicious domain names.

### 3.3 COMPONENT 3 – EVENT MANAGEMENT

In the Security Assessment System, the part that implements the proposed security assessment technique is based on attack graphs. Figure 5 shows the architecture of the component. The main component is the suite of vulnerability assessment algorithms for the computation of metrics. Mapper, another important subcomponent, allows an attacker to detect lo-

cation and attack structure based on security events. The security assessment component receives data from several sources, including an attack graph generator which creates reports from network analysis. Dependency graphs showing graphical dependencies between network services are also given. Output data has several security parameters, according to the proposed categorization. Additional output data would be received from the visualization system.

## 4. THE PROPOSED METHODOLOGY

Figure 6 shows the methodology which is split into four main steps:

1. Data Acquisition,
2. Data Extraction and Enrichment,
3. Reporting, Alerting, and Monitoring, and
4. Dashboard, Forms, and Integration.

### 4.1 DATA ACQUISITION (STEP 1)

In the log management, all log sources can be found. The log sources include application logs, audit logs, device logs, system logs, and database logs. In [5] article said that the logging process should be automated, precise, and visible to have a secure custody chain. These logs collected using several tools. Examples of tools used are External Domain Name Servers (Ext DNS), Intelligent Management Center (IMC). Dynamic Host Configuration Protocol (DHCP) server and firewalls like Web Application Firewall (WAF) and Next-Generation Firewall (NGFW).

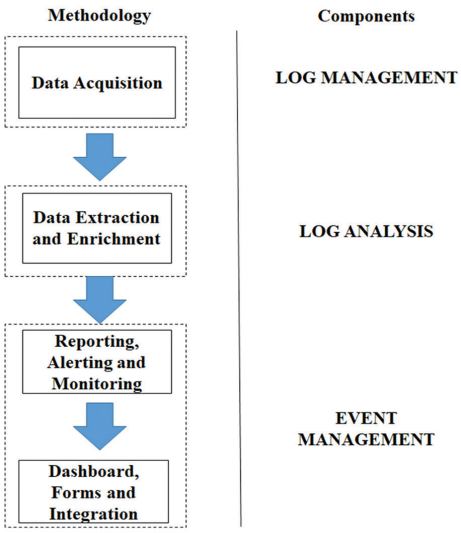



**Fig. 6.** Proposed Methodology

There are also fewer standard data sources including Internet Access Management (IAM) systems and all other tools that can provide relevant information to better identify organizational safety cases. The proposed system collects data from different security and network devices in the form of original logs and converts the collected logs into events through normalization and parsers. As a data exchange format for event data between different hardware and equipment and regulation engines became standard format is needed. It should differentiate between formats for content and formats for series. The structure of the data to be transferred (e.g. Event data includes two fields as "event timestamp" and "information about acknowledged attacks") is descriptive to content format [14].

### 4.2 DATA EXTRACTION AND ENRICHMENT (STEP 2)

In the proposed framework, data extraction is the method of taking raw data input and extracting only related fields, which is known as normalization. The standard event format can be created from all sources of data that allow the consistent comparison and collection of information across the network [15]. According to [7], the context allows a more comprehensive analysis by providing further information related to the relevant event.

A firewall is a key module that makes up the proposed framework. The information from the firewall constitutes the basic source of the log and event data of the system. Currently, vendors are producing and providing very sophisticated firewalls that can detect and prevent malicious activities from attackers. The firewall filter and inspect the traffic. Monitor the process by allowing all devices to collect audit data such as system logs and firewall alerts. Such data will be sent to the central console in the proposed system, where it will be aggregated, correlated, analyzed, and reported to prevent abnormalities. Based on [16] the study found two potential sources of error in the information enrichment process. The first is the unstructured original data and the other is the accuracy of the methods used for machine learning to obtain data and information extraction. All the logs that are related to this NGFW firewall in the network will go through the outbound and inbound of the firewall.

The process of association analysis can usually be divided into three parts that are,

- filter redundant information and format safety information,
- match association rules, and
- generate security incidents.

The first two techniques can be used in the proposed system to detect anomalies. It is also known as an abnormal association rule. This rule is valid only when defining the threshold. By comparing the rules of the normal category data set with the rules of the actual traffic category data set based on the similarity measure, anomaly detection can be calculated. If the similarity result is higher than the user threshold, it means that the data set is not intrusive and vice versa. The general category data set is reference data and should not contain intrusions. To apply this technique, the data must first be converted into a data set.

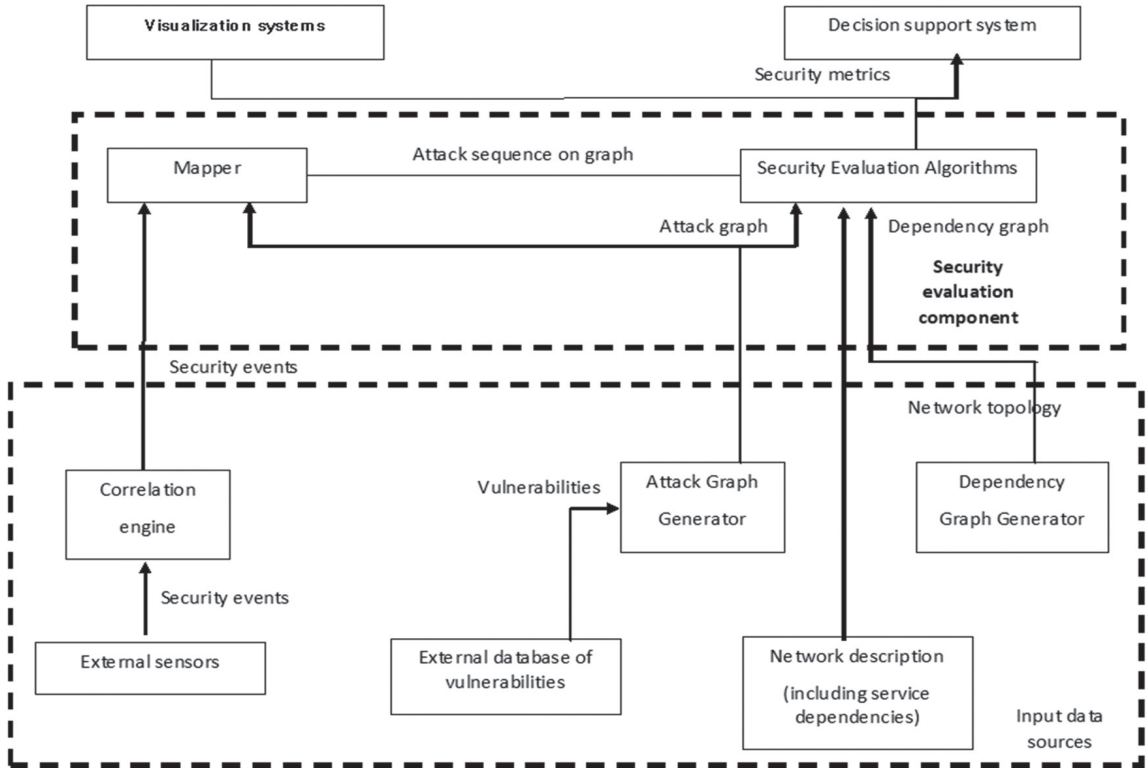



**Fig. 5.** Security evaluation component architecture [10]

### 4.3 REPORTING, ALERTING, AND MONITORING (STEP 3)

In the actual SIEM deployment environment, the monitoring requirements determined are usually mapped from specific business monitoring requirements. Monitoring requirements can be seen as business or IT security requirements for active tracking, alerting, and reporting. It is recommended to be as specific as possible when specifying monitoring requirements [7]. The proposed system will present the results and findings in a way that depends on the user's role after analyzing the log data. For all log sources, this information is displayed in a user-readable and understandable format. The reporting function allows users to configure log information files and only capture relevant information about their tasks. In addition to real-time tracking, log data analysis, and interactive reports for visual records, the proposed solution also provides compliance reporting functions, which can provide detailed and achievable audit logs.

In addition, the proposed solution has compliance reporting features, which generate a detailed and operational audit log record, in addition to real-time monitoring, log analysis, and interactive visual information reporting. During the audit of the organization, the auditor may review the records, the information reports, and other regulatory-specific content, to ensure compliance with the regulations. In short, the data will give in statistical format, such as reports, graphs.

#### 4.3.1 Alerting threats

Network and system protections have developed as the threats grow over the years. Distributed denial of services (DDoS) and other types of attacks affect organizations worldwide. Another major network threat is unauthorized access. It occurs when a malicious user invades an account and uses it to change permissions, gain access to resources or information, and other malicious activities. This is the route for hackers to execute APT and is a common problem in large organizations, which usually keep confidential information and other data with high business value. To secure the network, AlienVault generated alerts that provide response procedures. Of course, the data itself will be passed on when received for real-time analysis and monitoring. Analysis of the data from the devices used can show patterns and how conducive activities are to the fingerprints of individuals or groups of threats, they are detected throughout the sensor network through a trace of scattered data [17].

#### 4.3.2 Monitoring logs

The proposed systems will identify and warn a network in real-time if an incident and an important protection problem are detected. SIEM's key functions within the network of an organization are to track, capture and archive log data in a central console. The log data must be analyzed, warnings filtered and correlation rules developed. Both file origins are included in the log management. The log sources include network-based applications, systems, and computers. Based on

[18], effective log monitoring requires active data log analysis due to the size and quantity of the log file.

### 4.4 DASHBOARD, FORMS, AND INTEGRATION (STEP 4)

The infographic of network security provides for the visual depiction of security data, which may help users understand complicated technical information and security aspects [19]. The visual representation also enables users to search for specific types of graphical chart methods, appropriate visualization methods, classification of security data visualizations, etc., allowing us to meet the growing demand for network security monitoring.

## 5. CASE STUDY INVESTIGATION – EVALUATION OF THE PROPOSED FRAMEWORK

Unauthorized access introduces serious security problems. The SIEM system is of great significance in dealing with the security issues of critical infrastructure. The proposed solution based on the SIEM framework can monitor the network at any time to detect and issue alerts when incidents and serious security issues are discovered. Figure 7 shows the experimental setup for this study.

### 5.1 UNAUTHORIZED ACCESS – AN EXTERNAL DNS SERVER (HACKED)

Figure 8 and Figure 9 shows an example for unauthorized access from an external DNS server from NGFW where it can detect the security operations found three IP that threatened the external DNS Server and the graph and other analysis.

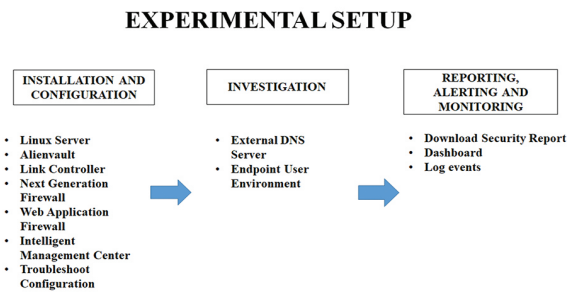



**Fig. 7.** Experimental Setup for this project

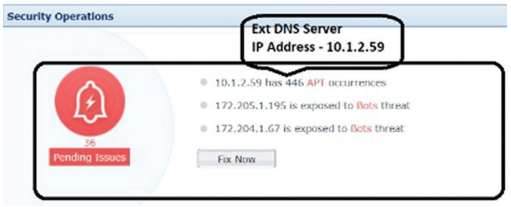



**Fig. 8.** The security operations found three IP that threatened the external DNS Server (Ext DNS server)

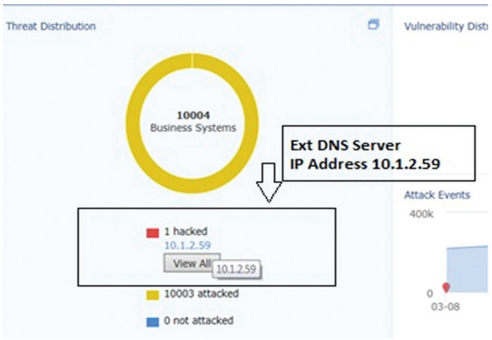



**Fig. 9.** The graph and other analysis


**Fig. 10.** The detail of the threat

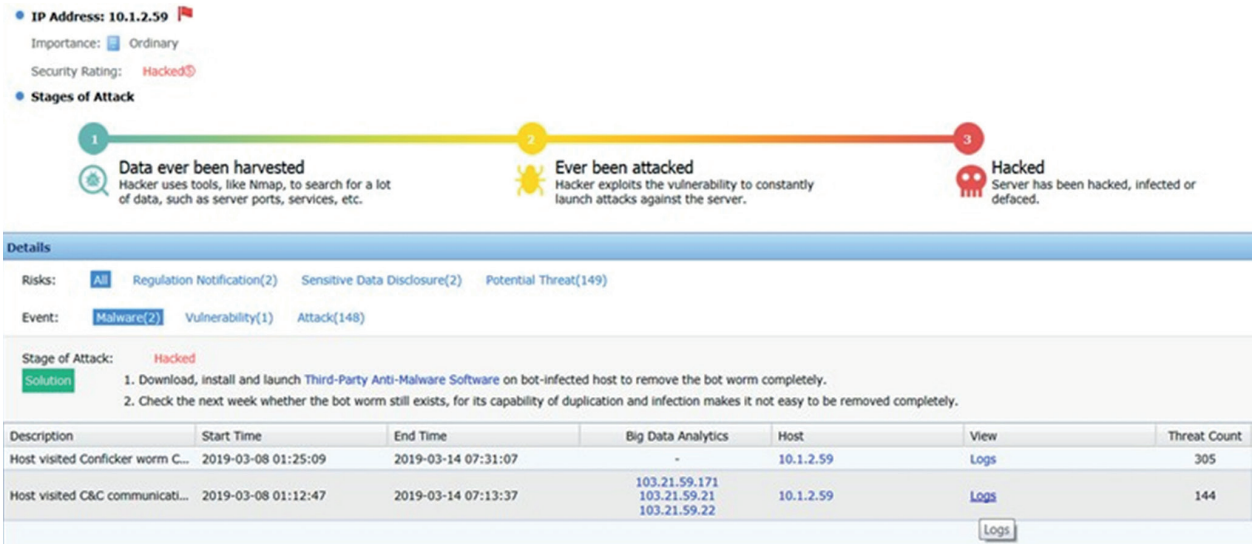



**Fig. 11.** The summary of the attacks

| No. | Date | Type | Source IP/User | Dst IP | Dst Location | Threat... | Acti... | Description | Data... | Threat... | Det... | Whitelist | Locked |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2019-03-14 07:13:37 | Botnet | 10.1.2.59 | 103.21.59.21 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 2 | 2019-03-14 06:17:24 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 3 | 2019-03-14 05:34:01 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 4 | 2019-03-14 05:22:47 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 5 | 2019-03-14 04:29:47 | Botnet | 10.1.2.59 | 111.118.215.77 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 6 | 2019-03-14 04:18:12 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 7 | 2019-03-14 04:08:54 | Botnet | 10.1.2.59 | 103.21.59.171 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 8 | 2019-03-14 02:05:02 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 9 | 2019-03-14 01:42:51 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 10 | 2019-03-14 01:12:40 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 11 | 2019-03-14 00:13:28 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 12 | 2019-03-14 00:02:31 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 13 | 2019-03-13 23:55:11 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 14 | 2019-03-13 23:23:31 | Botnet | 10.1.2.59 | 103.21.59.21 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 15 | 2019-03-13 22:30:11 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 16 | 2019-03-13 22:08:31 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 17 | 2019-03-13 21:48:01 | Botnet | 10.1.2.59 | 35.187.36.248 | United States | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 18 | 2019-03-13 21:41:51 | Botnet | 10.1.2.59 | 103.21.59.171 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 19 | 2019-03-13 21:14:31 | Botnet | 10.1.2.59 | 111.118.215.77 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 20 | 2019-03-13 20:55:58 | Botnet | 10.1.2.59 | 123.30.109.9 | Vietnam | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 21 | 2019-03-13 20:23:20 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 22 | 2019-03-13 20:06:55 | Botnet | 10.1.2.59 | 103.21.59.21 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 23 | 2019-03-13 19:45:25 | Botnet | 10.1.2.59 | 103.21.59.22 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 24 | 2019-03-13 19:41:06 | Botnet | 10.1.2.59 | 103.21.59.21 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |
| 25 | 2019-03-13 19:30:53 | Botnet | 10.1.2.59 | 111.118.215.77 | India | High | Deny | Host attempted to communic... | View | View | View | Add | Add |

**Fig. 12.** APT logs 1

| | |
|---|---|
| **No. 1** | |
| **Date:** | 2019-03-14 07:13:37 |
| **Type:** | Botnet |
| **Protocol:** | UDP |
| **URL/Directory:** | - |
| **Src Zone:** | DMZ |
| **Source IP/User:** | 10.1.2.59 |
| **Group:** | / |
| **Src Port:** | 58651 |
| **Dst Zone:** | Untrust |
| **Dst IP:** | 103.21.59.21 |
| **Dst Location:** | India |
| **Dst Port:** | 53 |
| **Rule ID:** | - |
| **Policy Name:** | APT |
| **Threat Level:** | High |
| **Action:** | Deny |
| **Description:** | Host attempted to communicate with the botnet C&C server (103.21.59.21) |

**Fig. 13.** APT logs 2

While, Figure 10, Figure 11, Figure 12, and Figure 13 show an example of unauthorized access from an external DNS Server from NGFW where it can detect the detail of the threat, the summary of the attacks come from, APT logs 1 and logs 2 actions.

Lastly, Figure 14 and Figure 15 shows the sample attack and map event of the attacker location.

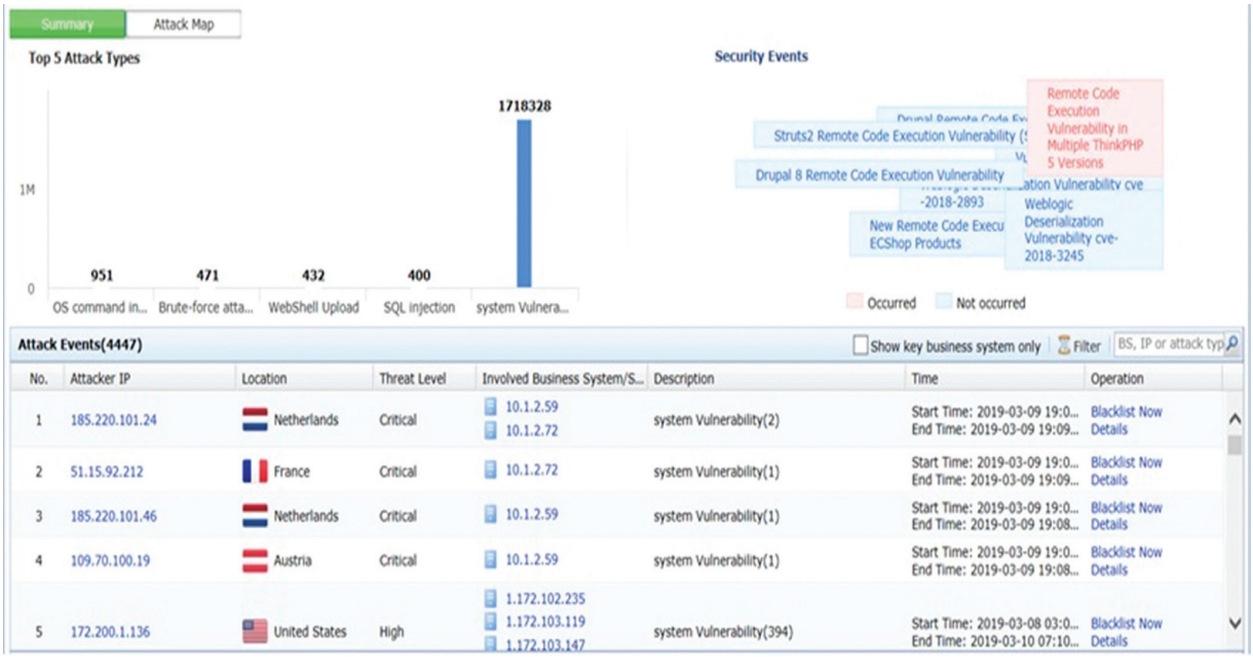



**Fig. 14.** Attack events


**Fig. 15.** The map of attack events

## 6. CONCLUSION AND FUTURE WORKS

Nowadays, High Education Institution (HEI) is working towards the deployment of SIEM as a solution to provide control across the whole network by gathering logs from both security and network equipment. Throughout the framework, logs may be correlated to include reliable and appropriate threat warnings. As a result, this research paper outlines the proposed network behavior monitoring architecture based on SIEM principles using a deep learning analysis. The case study to evaluate the proposed framework is also presented. Future studies will concentrate on evaluating a deep learning approach to DDOS attacks to determine the detection accuracy rate.

## 7. REFERENCES:

[1] K. Agrawal, H. Makwana, "Data Analysis and Reporting using Different Log Management Tools", International Journal of Computer Science and Mobile Computing, Vol. 47, No. 7, 2015, pp. 224-229.

[2] M. Conti, A. Dehghantanha, K. Franke, S. Watson, "Internet of Things security and forensics: Challenges and opportunities", Future Generation Computer Systems, Vol. 78, 2018, pp. 544-546.

[3] C. Crowley, J. Pescatore. "Common and Best Practices for Security Operations Centers: Results of the 2019 SOC Survey", https://www.dflabs.com/lp/thank-you-for-downloading-the-sans-2019-soc-surveyreport/?__s=aivfmwwer9oqny8sxwyf&drip_email=jack.whitter-jones%40southwales.ac.uk&drip_subscriber_id=aivfmwwer9oqny8sxwyf (accessed: 2021)

[4] B. Filkins, "2019 SANS Automation & Integration Survey", https://www.sans.org/media/vendor/Automation-and-Integration-Survey.pdf (accessed: 2021)

[5] M. Z. Seyed, "Analysis of Security Information and Event Management (SIEM) – Evasion and Detection Methods", Tallinn University of Technology, Faculty of Information Technology, Tallinn, Estonia, Master Thesis, 2016.

[6] A. Khan, R. Khan, F. Nisar, "Novice threat model using SIEM System for Threat Assessment", Proceedings of the 2th International Conference on Communication Technologies, Rawalpindi, Pakistan, 19-21 April 2017, pp. 72-77.

[7] M. Chikonga, "Exploring the Applicability of SIEM Technology in IT Security", Auckland University of Technology, Auckland, New Zealand, Master Thesis, 2014.

[8] L. Coppolino, S. D'Antonio, V. Formicola, L. Romano, "A framework for mastering heterogeneity in multi-layer security information and event correlation", Journal of Systems Architecture, Vol. 62, 2016, pp. 78-88.

[9] M. Di Mauro, C. Di Sarno, "Improving SIEM capabilities through an enhanced probe for encrypted Skype traffic detection", Journal of Information Security and Applications, Vol. 38, 2018, pp. 85-95.

[10] I. Kotenko, E. Doynikova, "Security assessment of computer networks based on attack graphs and security events", Lecture Notes in Computer Science, 8407, 2014, pp. 462-471.

[11] A. Majeed, R. ur Rasool, F. Ahmad, M. Alam, N. Javaid, "Near-miss situation based visual analysis of SIEM rules for real time network security monitoring", Journal of Ambient Intelligence and Humanized Computing, Vol. 10, No. 4, 2019, pp. 1509-1526.

[12] A. R. Zope, A. Vidhate, N. Harale, "Data Mining Approach in Security Information and Event Management", International Journal of Future Computer and Communication, Vol. 2, No. 2, 2013, pp. 80-84.

[13] K. O. Detken, M. Jahnke, C. Kleiner, M. Rohde, "Combining Network Access Control (NAC) and SIEM functionality based on open source", Proceedings of the IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Bucharest, Romania, 21-23 September 2017, pp. 300–305.

[14] H. A. Khan, "Advancing Security Information and Event Management Frameworks in Managed Enterprises using GeoLocation", University of Cape Town, Faculty of Science, Department of Computer Science, Master Thesis, 2014.

[15] P. Andruszkiewicz, H. Rybinski (2018), "Data Acquisition and Information Extraction for Scientific Knowledge Base Building". Proceedings of the 12th IEEE International Conference on Semantic Computing, Laguna Hills, CA, USA, 31 January - 2 February 2018, pp. 256–259.

[16] B. D. Bryant, & H. Saiedian, "Improving SIEM alert metadata aggregation with a novel kill-chain based classification model", Computers and Security, Vol. 94, 2020, p. 101817.

[17] I. Yagoub, M.A. Khan, L. Jiyun, "IT Equipment Monitoring and Analyzing System for Forecasting and Detecting Anomalies in Log Files Utilizing Machine Learning Techniques", 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems, Durban, South Africa, 6-7 August 2018, pp. 1-6.

[18] A. Majeed, R. ur Rasool, F. Ahmad, M. Alam, N. Javaid, "Near-miss situation based visual analysis of SIEM rules for real time network security monitoring", Journal of Ambient Intelligence and Humanized Computing, Vol. 10, No. 4, 2019, pp. 1509–1526.

# Pixel Value Graphical Password Scheme: K-Means as Graphical Password Fault Tolerance

Original Scientific Papers


**Mohd Afizi Mohd Shukran**
Professor, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia
afizi@upnm.edu.my

**Mohd Sidek Fadhil Mohd Yunus**
Senior Lecturer, Department of Computing,
Faculty of Arts, Computing and Industry Creative,
Sultan Idris Education University
Tanjung Malim, Malaysia
msidek@fskik.upsi.edu.my

**Mohd Rizal Mohd Isa**
Senior Lecturer, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Fatimah Ahmad**
Professor, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Muhammad Naim Abdullah**
Lecturer, Department of Computing, Academic Affairs,
University Malaysia of Computer Science and Engineering (UNIMY),
Selangor, Malaysia

**Syed Muzzameer Syed Zulkiplee**
Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Mohammad Adib Khairuddin**
Senior Lecturer, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Mohd Nazri Ismail**
Professor, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Mohd Fahmi Mohamad Amran**
Senior Lecturer, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Norshahriah Wahab**
Senior Lecturer, Department of Computer Science,
Faculty of Defence Science and Technology, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia

**Nur Adnin Ahmad Zaidi**
Department of Computer Science,
Faculty of Medicine and Defence Health, National Defence University of Malaysia
Kem Sungai Besi, Kuala Lumpur, Malaysia



***Abstract*** *– Pixel value access control (PVAC) was introduced to deliver a secure and simple graphical password method where it requires users to load their image as their password. PVAC extracts the image to obtain a three-octet 8-bits Red-Green-Blue (RGB) value as its password to authenticate a user. The pixel value must be matched with the record stored in the database or otherwise, the user is failed to authenticate. However, users which prefer to store images on cloud storage would unintentionally alter and as well as the pixel value due to media compression and caused faulty pixels. Thus, the K-Means clustering algorithm is adapted to fix the issue where the faulty pixel value would be recognized as having the same pixel value cluster as the original. However, most of K-Means algorithm works were mainly developed for content-based image retrieval (CBIR) which having opposite characteristics from PVAC. Thus, this study was aimed to investigate the crucial criteria of PVAC and its compatibility with the K-Means algorithm for the problem. The theoretical analysis is used for this study where the suitable characteristics of K-Means are analyze based on PVAC requirements. The compliance analysis might become a referencing work for digital image clustering techniques adaptation on security system such as image filtering, image recognition, and object detection since most of image clustering works was focused on less sensitive image retrieval.*



***Keywords****: Cybersecurity, PVAC, Pixel Value, Graphical Password, Clustering, K-Means*

## 1. INTRODUCTION

The pixel value access control (PVAC) is a graphical password method that utilizing pixel value extracted from a digital image, referred to as Password Pixel or *PassPix*, to authenticate a username. PVAC was produced through the design and development of Pixel Value Graphical Password [1] idea in 2014 was motivated to solve the human memory burden on memorizing a strong and challenging password. PVAC is suited to be implemented as the guardian mechanism for server access which would operate in a cloud-based environment which is demanding nowadays especially during the current pandemic outbreak. This trend would bring the same way as the way PVAC users storing their *PassPix* where unfortunately would altering the pixel value unintentionally and caused users failed to authenticate and depict the concept of cloud facility.

As a study conducted in 2019 [2], an image that is transferred and accessible through WhatsApp [3] application would change an image attribute where the dimension is reduced and caused pixel value is transformed. Therefore, theoretically by adapting the K-Means [4] clustering algorithm would solve this issue theoretically by allowing PVAC to identify the faulty pixel value was residing in the same cluster as the pixel value it's supposed to be [5]. The idea is, when PVAC encounters a fail pixel value query during the authentication process, the Query By Example [6-7] or QBE method will be activated to query the specific similarity index (SI) range using the K-Means clustering algorithm.

However, most of K-Means-related works were based on how efficient the algorithm performed on content-based image retrieval system (CBIRS) [8] rather than access control system such as PVAC where both kinds of systems were having contrast criteria. Thus the adaptation of the K-Means clustering algorithm for PVAC must be involving a comparison analysis and algorithm fitting works.

## 2. THE PVAC

The current PVAC model [9] described that a user requires enrolling to the server first to create an access record with a unique username and upload an image as *PassPix*. As the provided username does not exist in the database, the PVAC will extract the pixel value from the fed *PassPix* as the password for the username and store them in the database. As for the authentication (Log-in) process, a registered user must repeat the same procedure as the enrolment (sign-up) process where the username and the *PassPix* are being fed again to the PVAC. PVAC will extract the newly fed *PassPix* for its pixel value again and perform the record query from the database. As the fed data and the database record are a match, PVAC will grant access to the user to the server. The process shows likelihood as the common log-in process as well as the user interface for PVAC by replacing the alphanumeric password with the *PassPix* as shown in (Fig. 1).


**Fig. 1.** The user interface for PVAC log-in

The pixel value extraction module plays a vital role in PVAC where it computes the RGB value that is presented as three octet 8-bits numbers (0,0,0 to 255,255,255). In general, a digital image is constructed by number of pixels that arranged in two-dimension resolution with X-axis and Y-axis which presented as an image attribute as Sum of X-axis pixel by Sum of Y-axis pixel (example: 200px * 200px, 1080px * 750px, or 2400px * 1030px). For example, the attribute resolution is 850px * 480px means the image is constructed 850 pixels on X-axis and 480 pixels on Y-axis that make 408,000 pixels in total ($\sum_{px}$). Every single pixel is holding the RGB color combination that made up the color of the pixel. The RGB color presentation is the standard color presentation as it widely applies for digital images, web colors, OS settings, and others where it is commonly referred to as RGB color-wheel [10]. The color of a pixel is built up by combining the hue and saturation (also referred to as strength for each color) of red, green, and blue that also create other colors as well such as yellow, violet, cyan, and others.

### 2.1. THE PIXEL VALUE

In a digital image, the colored pixel is arranged in two-dimension based where the combination of all colored pixels creates a region area and as well as to object visible on screen. The pixel value of a digital image is calculated by total up the strength value of red, green, and blue from each value and divide it with total pixels to get the average strength of red, green, and blue color (referred to as a color histogram) which can be calculated as:

$$\frac{\sum(R,G,B)}{\Sigma px} \tag{1}$$

In PVAC, before the extraction module performing the pixel value extraction process, the *PassPix* fed by users is divided into the logical grid where the dimension of the logical grids is a variable that the value could be modified by the system provider. As in a study on fake *PassPix* attempts [11], the logical grid is a fea-

ture for PVAC to extending the password space where it enables a *PassPix* could produce a single pixel value to an extremely long pixel value. The study also shows that it requires tremendous works to reconstructing the *PassPix* as the fake *PassPix* (as in case the pixel value is sniffed or leak from the server), compared to utilizing a single pixel value as the password. Currently, the fake *PassPix* reconstruction is a failed attempt that proved that the logical grid is a proven method. For example, if a developer decided to apply for 8 grids on the X-axis and 3 grids on the Y-axis, the total dimension (d) is 24. The pixel value on every grid is extracted which makes the total pixel value utilize on the PVAC is 24-pixel values. By using the same equation (1), the pixel value is calculated within the grid size which the $\sum_{px}$ is the sum of pixels in the respective grid as well as the ∑(R, G, B) value. The pixel value is arranged into a single string object through the array process as the password is kept in the database. In other words, the extraction module on PVAC is performing three processes:

i. Logical grid process,
ii. Grid pixel value extraction process and
iii. Password array process.

The size of the logical grid (Gpx) is computed as:

$$G_{px} = \frac{x_{px}}{x_n} * \frac{y_{Px}}{y_n} \quad (2)$$

Where,
$Xp_x$ is image dimension on $x$ axis
$Yp_x$ is image dimension on $y$ axis
$x_n$ is number of grids determined on $x$ axis
$y_n$ is number of grids determined on $y$ axis

Then, PVAC is computing the pixel value as:

$$\frac{\Sigma_g(R,G,B)}{G_{px}} \begin{bmatrix} x_1 y_1 & x_n y_1 \\ x_1 y_n & x_n y_n \end{bmatrix} \quad (3)$$

Where,
$\sum_g$ is the sum of RGB value in a grid
$x_n$ is the last grid column on $x$ axis
$y_n$ is the last grid row on $y$ axis

### 2.2. FAULTY *PASSPIX* PROBLEM

As reported by Widjaja et al. [12] in 2018, almost 2 billion cloud storage active users recorded globally with a countless amount of file size stored, and this trend is forecast to increase over the next few years. Since the world is having a global COVID19 pandemic outbreak, the trend is not just extremely increasing, but also demanding. In addition, Wu et al. [13] mentioned, with cloud storage practices, users do not have to worry about data losses due to the failure of physical storage, data breach, or even Ransomware where all of the risks were mitigated to the service provider. This trend makes no exception for the PVAC users where the *PassPix* is stored cloudily for the above-mentioned reasons.

However, the cloud storage conditions and settings are cloudy where the effect on files is uncertain as found by Li et al. [14] where service providers might apply some compression to prevent certain deficiencies such as data processing consumption, storage insufficiency, and service delays. Such conditions would alter the file 8-bit attribution, especially multimedia files (digital images, digital sound, and digital video) as well as *PassPix* files. That will cause the PVAC to simply discarded the damaged *PassPix* and the affected users unable to authenticate.

Conceptually, a digital image clustering algorithm is an admissible method to integrate with PVAC since both methods are processing the digital image computation that involving the process of pixel value extraction. Despite those promising conceptual ideas, as digital image clustering was commonly designed for CBIR, the adaptation of digital image clustering is necessarily compliant with PVAC rules and regulations.

### 2.3. CRITERIA FOR PVAC

As the PVAC is meant to apply for client-server access control (user authentication security), there are a few requirements and rules that the digital image clustering method is strictly needed to comply with. Based on that, PVAC security-sensitive requirements are the major factor that regulated the digital image clustering algorithm selection. PVAC requires a digital image clustering algorithm that can compute the pixel value differently as specific as possible to prevent the PVAC could be tolerable to the fake *PassPix* authentication. In other words, the recognition ranges between the faulty *PassPix* and the original *PassPix* are subject to limit from the fake *PassPix*. By concept, during the query process, only *PassPix* that resides within the tolerable range is recognized as authentic *PassPix*. All of the PVAC criteria needed as fault tolerance mechanism are concluded in Table 1.

**Table 1.** The required PVAC rules and regulation

| Requirements | | Descriptions |
|---|---|---|
| Security | Tolerable Range | To reduce the vulnerability risk coming from fake PassPix |
| Features | Feature Extraction | PVAC is employing the pixel-based extraction method. |
| | Color-Space | PVAC is working on RGB colour-space. |
| | Logical-Grids Extractions | To preserve the password space strength, the clustering data is analyzed in two-dimension data. |

## 3. KMEANS CLUSTERING ALGORITHM

In Partition-Based DIC strategy, there a variety of algorithms that performed the clustering process in many ways. Among them, the K-Means algorithm gains the most attention among researchers [15-17]. Nayini et al. [18] stated that K-Means is the most interesting DIC algorithm because of its capability to clustering large

datasets with low computational complexity. Slamet et al. [15] suggested that the K-Means algorithm is a great help for efficiently understanding a complex dataset. In other words, most researchers are interested in K-Means due to their computational simplicity.

Basic K-Means at first was introduced by Macqueen [4] to classify a set of objects into predetermined groups. When the dataset receives a new object, he suggests the iteration process predetermine the group properties such as partition and centroid position. K-Means was equipped with the pixel value extraction function that translates a digital image into a string of pixel values which is similar to the PVAC pixel value extraction function as presented in equation (1). Then, the extracted object in the form of pixel value data is assigned to a cluster through the original K-Means algorithm.

Number ($n$) of the object ($P$) is the number of data ($x$) extracted using equation ($1$) and the Centroid seed ($S$) is the initial user to determine the centroid ($\mu$) point for placing the centroids. The $S$ value is randomly picked by a user could be any number and normally not greater than the $n$ value.

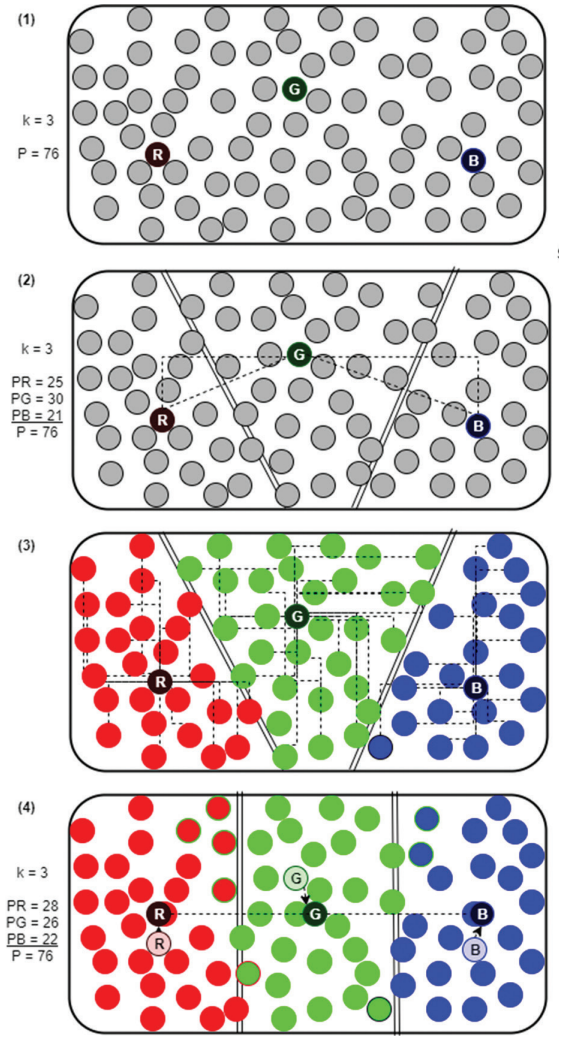



**Fig. 2.** 3 steps K-Means as in object matrix

For example, if a dataset containing 2,000 objects ($n$) to group into 20 clusters ($k$), the 20 $S$ value is picked from 1 to 2,000. The number of seed must be equal to number or cluster ($k$), also a value that determined by users. However, as emphasized by Jin & Han [17] there is no absolute fail-proof framework to handpick the seed value.

The second step is assigning membership where each object ($x$) is assigned to the nearest centroid ($\mu$) using the Euclidean distance ($\in$) that formulate as:

$$\epsilon = \sqrt{\sum_{k=1}^{n}(x-\mu)^2} \tag{4}$$

A partition or cluster border is built perpendicular to the mean point of Euclidean distance between centroids and objects that reside between centroid and cluster border is converged into a cluster. The next step is the iteration process where the centroid in every cluster is repositioned to the mean point of every object. When the centroid moved, the convergence process as in step 2 is performed again until no object is transferring to a different cluster. The iteration is a repetition process that happens every object transition in the object matrix. This 3-steps K-means object matrix is exemplified in (Fig. 2) for 76 objects that need to divide into 3 groups.

## 4. COMPLIANCE ANALYSIS OF K-MEANS ON PVAC

Referring to equation (3), PVAC is extracting the pixel value in the two-dimension logical grid to extend the password space and password strength as well. The multigrid pixel value is arranged into a string object through an array function. The efficiency of clustering for such data has been performed by Kumari et al. [19] on network packets using the K-Means algorithm where each packet is encapsulating the 8-bit data and multiple packets are forming informative data. The authors referring the technique as multiset clustering which every object is referred to as a subset in which every subset is holding series of packets. The term multiset clustering is applied as two-dimensional data as extracted from default PVAC pixel value extraction. The password derived from the array process on PVAC pixel value password creation is fit as the multiset structure object is an advantage. PVAC is avoiding the risk of adopting a harsh and complex multi-dimension algorithm that excessively consume computational resource.

In this study, the analysis is based on the result produced by the prototype of K-Means patch PVAC which was set up as an exhibit in (Fig.3).

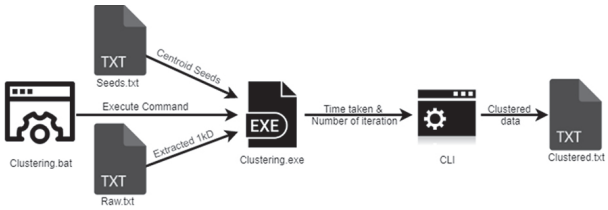



**Fig. 3.** Clustering experiment setup

There are 1,000 images use in this study which is randomly picked from various research dataset and gathered as one uncategorized dataset. To investigate the K-Means algorithm compliance to the PVAC feature requirement, the modified PVAC prototype is a challenge to perform clustering on the 1,000 images. The Clustering.exe is developed with the multiset clustering function which is required as the simulation of PVAC style data as discussed theoretically. Technically, the multiset clustering that is set on Clustering.exe is a set of Euclidean distance for every pixel value in each grid where every object is returning the set of Euclidean distance of every pixel value. The pixel value from every grid is locked to the object centroid or mean point for all grids to avoid the grid is scattered over the dataset. Then the RGB value in every grid is calculated to get the RGB distance different from the cluster centroid ($\mu c$) and the Euclidean distance of every grid pixel value becomes the attribute for the object.

### 4.1. FEATURE REQUIREMENT FOR PVAC

The desired observation for this study is the multiset clustering functionality that determined the feature requirement compliance and the accuracy of clustering result which determined the security requirement for PVAC. The data used for clustering is produced by the original or default PVAC *PassPix* extraction module without any additional coding. The *PassPix* extracted by PVAC in multiset clustering work as a subset and containing series of Euclidean distance of every grids pixel value that refers as Euclidean of a subset ($\in\subset$) where the whole object becomes one cluster. The $\in\subset$ calculation for every object is derived from equation (4) that the *k* value is removed from the equation since the object's grids are not segmented and the object is calculated as in a matrix form. The $\in\subset$ is calculated as:

$$\in\subset = \sqrt{\sum_{\subset}^{n}([G_1\ G_n] - \mu_c)^2} \quad (5)$$

Where,
$n$ is the sum of grids extracted by PVAC
$G_1$ is the first grid
$G_n$ is the last grid
$\subset$ is the subset or object in 1kD
$\mu_c$ is the centroid of the object cluster

The multiset clustering is working as every grid is calculated as an object. Then, all of the grid is arranged in a dimensional sequence to form a set of grids to become an object before being calculated with default Euclidean distance ($\in$). By using equation (5), the K-Means algorithm can perform the clustering for 2 dimension data as extracted by default PVAC PassPix extraction module.

As a result, by functionality, the modeled K-Means clustering algorithm is accepted to perform the clustering process as required by PVAC as listed in table 1 previously. The ability of the K-Means clustering algorithm to calculate multiset object and cluster assignment task resulting K-Means clustering algorithm is fulfilling the all feature requirement of PVAC as concluded in table 2.

**Table 2.** The PVAC features requirement complied with K-Means multiset clustering

| Feature Requirements | K-Means multiset clustering |
|---|---|
| **Feature Extraction** | K-Means process the data extracted by the default PVAC PassPix extraction mechanism using pixel-based extraction |
| **Color-Space** | Same as feature extraction, K-Means process the data extracted by the default PVAC PassPix extraction mechanism that produced the RGB color-space |
| **Logical-Grids Extractions** | Clustering.exe is patched with multiset clustering that worked on 2 Dimension data extracted by PVAC |

### 4.2. SECURITY REQUIREMENT FOR PVAC

The second element is the query accuracy experiment; the result from the experiment is the key factor to determine the security requirements for the PVAC pixel fault tolerance mechanism. The query-based experiment is a way to investigate whether the K-Means algorithm would be able to query for similarity of faulty pixel *PassPix* with the clustered original *PassPix* data.

By default, when the query is performed, it will group all objects that are similar to the queried image as in CBIR. However, for this study, the desired query output was the closest distance between the queried image and the clustered database. The accuracy would prevent the fake *PassPix* is unable to bypass the fault tolerance mechanism and create the PVAC vulnerability.

As a simulation for the unintentionally images pixel value altering, all original *PassPix* dataset is transferred to WhatsApp repository and download it to local storage media as a compressed dataset. K-Means were challenged to query all of the compressed images with the clustered database produced from the previous experiment. This experiment is aimed to obtain the accuracies rate of all tested DIC algorithms where the accuracy is determined by the ability to return the queried compressed object with the rightful clustered object.

Observation from the result shows that K-Means with a higher number of clusters produce a higher accuracy rate as concluded in table 3.

**Table 3.** Query accuracy result

| Number of *k* | Accuracy rate |
|---|---|
| *k* = 10 | 78.2 % |
| *k* = 20 | 78.3 % |

The K-Means with the *k* parameter set to 20 is producing a 0.1% higher accuracy rate than 10 clusters. This can be concluded that the number of *k*'s is affected the accuracy rates where more *ks* produce more accurate query result. To estimate the effect of more k settings, a line graph is plotted as shown in (Fig. 4).

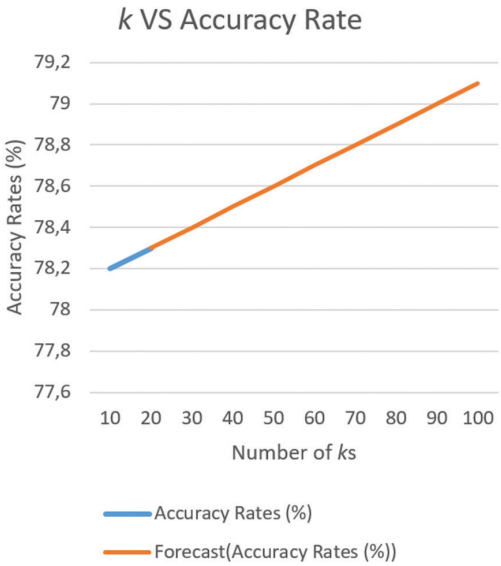



**Fig. 4.** *k* VS Accuracy rates

The k vs Accuracy rates graph shows that, estimated about 79.1% accuracy rate scored if the K-Means was set with 100 ks based on 10 *ks* and 20 *ks* accuracy rates. That proves that, for the 1,000 objects dataset, more ks are required to be set to increase the accuracy rate.

However, the number of *k* also affected the time taken by K-Means to perform the clustering process for such a dataset. K-Means with 10 *ks* setting taking 10 seconds to complete while K-Means with 20 ks taking 22 seconds. To gain more accurate results, more ks are required which would cause more time consumption as projected in (Fig. 5).

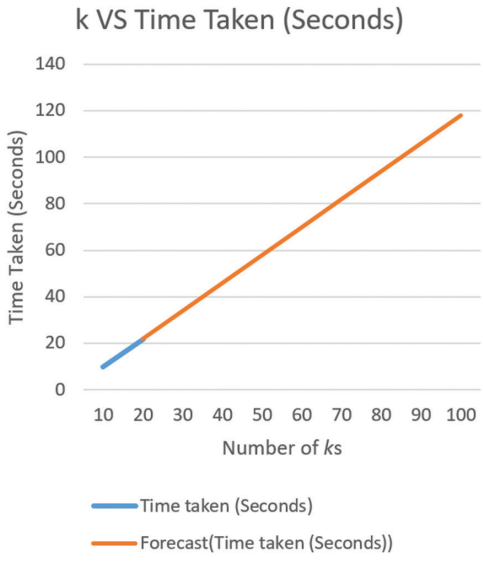



**Fig. 4.** *k* VS Time Taken

From the graph that forecasts based on $k = 10$ and $k = 20$ data, it can be concluded that the accuracy rate and the time consumption are both have to trade-off where to more time is needed to get a higher accuracy rate. It required about 2 minutes to complete the 100 ks clustering. Further works and effort on this issue might solve the issue as there is still more room for improvement.

## 5. CONCLUSION

Through this study, we can conclude that the K-Means clustering algorithm is a suitable algorithm to be adapted as a fault tolerance mechanism for PVAC where the multiset clustering enhancement and PVAC feature extraction can accomplish the clustering process.

This study also shows that, by applying more k, the accuracy of querying clustered databases is increased. This finding will reduce the possibility of depending on the similarity range module that would cause the range to exploit and be vulnerable. Besides, since the K-Means clustering algorithm is the most interesting algorithm among researchers, there are several variants derived from the K-Means algorithm which would enhance the performance and accuracy of the PVAC fault tolerance mechanism as well.

## 6. REFERENCES:


[1] M. A. M. Shukran, M. S. F. M. Yunus, "Method and System For Authenticating User Using Graphical Password For Access Control", Malaysia Patent MY-167835-A, 2018.

[2] M. A. M. Shukran, M. S. F. M. Yunus, M. N. Abdullah, M. N. Ismail, M. R. M. Isa, "Pixel Value Graphical Password: A PassPix Clustering Technique For Password Fault Tolerance", International Journal of Recent Technology and Engineering, Vol. 8, No. 3, 2019, pp. 2973-2975.

[3] WhatsApp Inc. WhatsApp Features, https://www.whatsapp.com/features/, (Accessed: 2019).

[4] J. MacQueen, "Some methods for classification and analysis of multivariate observations", Proceedings of the 5th Berkeley symposium on mathematical statistics and probability,1967.

[5] M. S. F. M. Yunus, "A Novel Graphical Password Clustering Method for Fault Tolerance Mechanism", National Defence University of Malaysia, Faculty of Defense Science and Technology, Kuala Lumpur, Malaysia, PhD Thesis, 2020.

[6] M. M. Zloof, "Query by example", Proceedings of the national computer conference and exposition, 19-22 May 1975, pp. 431-438.

[7] H. Kamper, A. Anastassiou, K. Livescu, "Semantic query-by-example speech search using visual grounding", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12-17 May 2019, pp. 7120-7124.

[8] M. Azam, N. Bouguila, "Bounded Laplace Mixture Model with Applications to Image Clustering and Content Based Image Retrieval", Proceedings of 17th IEEE International Conference on Machine Learning and Applications, Orlando, FL, USA, 17-20 December 2018.

[9] M. S. F. M. Yunus, M. A. M. Shukran, M. N. Abdullah, "Pixel-based Graphical Password Scheme: Password from Digital Image File", UPNM Press, Kuala Lumpur, 2019.

[10] M. A. M. Shukran, N. M. S. Ahmad, S. Ramli, F. Rahmat, "Melanoma Cancer Diagnosis Device Using Image Processing Techniques", International Journal of Recent Technology and Engineering, Vol. 7, No. 5S7, 2019, pp.490-494.

[11] M. A. M. Shukran, M. S. F. M. Yunus, "Pixel Value Graphical Password Scheme: Fake Passpix Attempt on Hexadecimal Password Style", International Journal of Information and Communication Sciences,Vol. 3, No. 3, 2018, p.104.

[12] A. E. Widjaja, J. V. Chen, B. M. Sukoco, Q. A. Ha, "Understanding users' willingness to put their personal information on the personal cloud-based storage applications: An empirical study", Computers in Human Behavior, Vol. 91, 2019, pp. 167-185.

[13] K. Wu, J. Vassileva, Y. Zhao, "Understanding users' intention to switch personal cloud storage services: Evidence from the Chinese market", Computers in Human Behavior, Vol. 68, 2017, pp. 300-314.

[14] C. Li, J. Bai, C. Yi Y. Luo, "Resource and Replica Management Strategy for Optimizing Financial Cost and User Experience in Edge Cloud Computing System", in Information Sciences, 2019.

[15] C. Slamet, A. Rahman, M. A. Ramdhani, W. Darmalaksana, "Clustering the Verses of the Holy Qur'an using K-Means Algorithm", Asian Journal of Information Technology, Vol. 15, No. 24, 2016, pp. 5159-5162.

[16] S. Irfan, G. Dwivedi, S. Ghosh, "Optimization of K-means clustering using genetic algorithm", Proceedings of the International Conference on Computing and Communication Technologies for Smart Nation, Gurgaon, India, 12-14 October 2017.

[17] X. Jin, J. Han, "K-means clustering", Encyclopedia of Machine Learning and Data Mining, 2017, pp. 695-697.

[18] S. E. Y. Nayini, S. Geravand, A. Maroosi, "A novel threshold-based clustering method to solve K-means weaknesses", Proceedings of the International Conference on Energy, Communication, Data Analytics and Soft Computing, Chennai, India, 1-2 August 2017.

[19] R. Kumari, M. K. Singh, R. Jha, N. K. Singh, "Anomaly detection in network traffic using K-mean clustering", Proceedings of the 3rd International Conference on Recent Advances in Information Technology, Dhanbad, India, 3-5 March 2016.

# Damage Cost/Value Clustering in Timber Harvesting Decision Making for Sustainable Forest Management

Original Scientific Paper

**Hana Munira Muhd Mukhtar**
Applied Statistics and Data Science Research Cluster,
Universiti Kuala Lumpur,
Malaysian Institute of Information Technology,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur.
hanamunira@unikl.edu.my

**Yasmin Yahya**
Applied Statistics and Data Science Research Cluster,
Universiti Kuala Lumpur,
Malaysian Institute of Information Technology,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur.
yasmin@unikl.edu.my

**Azizah Rahmat**
Service and Information Science Research Cluster,
Universiti Kuala Lumpur,
Malaysian Institute of Information Technology,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur.
azizah@unikl.edu.my

**Roslan Ismail**
Applied Statistics and Data Science Research Cluster,
Universiti Kuala Lumpur,
Malaysian Institute of Information Technology,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur.
drroslan@unikl.edu.my

***Abstract*** *– The most important factor to ensure forest regrowth strongly relies on minimizing damage as well as maintaining an adequate quantity and quality of residual stands. Currently, most of the Malaysian concessions are applying the Selective Management System (SMS). The SMS had been introduced about 40 years ago and various studies discovered that it contributes a negative impact on the forest. Thus, revision and adoption of an appropriate harvesting method are required. The main objective of this study is to propose a new method that promotes forest regrowth and reduces damages due to logging activities for Sustainable Forest Management (SFM). The two primary elements introduce in this new method are 1) to determine the minimum damage cost/value to the residual trees according to tree felling direction and 2) on the division of logging area into clusters where only certain clusters will be affected in a logging operation and the rest are conserved. The overall results of this study proven that the analysis of potential logged value, productions, damage value, and damage volume by dividing forest into clusters able to minimize damage and maintain forest regeneration.*

***Keywords****: damage cost/value, tree felling direction, forest clustering algorithm*

## 1. INTRODUCTION

Forests are crucial in terms of biodiversity and ecosystem, it gives numerous benefits to humans as well as timber products and biodiversity conservation. About 300 to 350 million people are directly and indirectly dependent on forests [1]. Tropical deforestation and forest degradation are some of the world's most urgent environmental problems. It contributes to biodiversity loss, accounts for approximately 17% of total global carbon emissions, and has adverse socio-economic consequences for forest-dependent people (e.g. EU, 2016). In addition, tropical forest degradation is one of the significant factors of carbon dioxide (CO2) emission [2]–[4] approximately 2.1 billion tons of CO2 yearly [2]. To address this issue, one of the crucial decisions to make in forest planning and forest management is to determine the best logging operation to increase timber harvesting productivity that reduces damages and promotes forest regrowth for sustainable forest management. In this paper, we describe the newly proposed methods to promote forest regrowth and reduce damages due to logging activities for sustainable forest management. By understanding the calculation of minimum damage based on tree felling direction, we can expect to have significant results in implementing the new methods and algorithms.

This paper is organized as follows; in the next section, the related work regarding the implementation and the limitations of the Selective Management System (SMS) are presented. Then followed by Section 3 where the newly proposed algorithms on how to determine the minimum damage cost to residual trees and also forest clustering for preservation are thoroughly described. Section 4 combines the results from the selection of the best-felling direction and decision-making on the minimum damage. Finally, Section 5 and 6 presents our conclusion and acknowledgment respectively.

## 2. RELATED WORK

Since 1978, the Selective Management System (SMS) was implemented for timber harvesting in Malaysia. This commercial logging system mainly targets dipterocarp species. Selective Management System (SMS), the current logging system is; a year before felling, commercially viable trees are marked for felling, the harvestable trees applied are >45 cm dbh for non-dipterocarp, and >50 cm dbh for dipterocarp species. Later, arrows are painted on trees to indicate the direction of felling to avoid damaging other valuable trees. Then, the system calculates the damage volume of residual trees [5], [6]. However, this event only takes place 10% of the pre-felling inventory as sampling for the rest of the felling area [7]. These decisions play an important role in maintaining the species composition and structure of the forest [8], [9].

According to various researchers [7], [8], [10]–[13], the most important factor to ensure forest regrowth strongly relies on minimizing damage as well as maintaining an adequate quantity and quality of residual stands. These studies discovered that the current selective logging contributes negative impacts to the forest; such as frequency distribution of gap area was strongly skewed, a low recovery rate of forest conditions after logging, tree volume of non-dipterocarp species higher than dipterocarp species, and absence of large-sized mammals. Therefore, a revision of current forest management in Peninsular Malaysia, mitigation actions, and the adoption of an appropriate harvesting plan for sustainable forest management are needed.

The purpose of this study is to propose a solution that has the potential to mitigate the stated current issues corresponding to sustainable forest management practices by dividing the forest into clusters, determine potential trees to log according to clusters with minimum damage value and damage volume to the surrounding trees. The analysis of these minimum damages provides a significant impact on forest preservation. The study will produce an output of the analysis that could be used by the government for timber harvesting decision-making.

## 3. MATERIALS AND METHOD

In response to this challenge, the main objective of this study is to propose a new method that promotes forest regrowth and reduces damages due to logging activities for Sustainable Forest Management (SFM). The two primary elements introduce in this new method are 1) to determine the minimum damage cost/value to the residual trees and 2) forest clustering to retain areas of unlogged forest for preservation. Preserve unlogged forest is critically important to safeguard species biodiversity of the tropical rainforest [7], [14], [15].

Selective Management System (SMS) is the current method that has been implemented by the majority of Malaysian concessions. Although this method is based on SFM practices, there are some negative side effects to the forest after more than 40 years of practicing.

For the logging activities, the SMS can be categorized into 3 stages for the logging activities under the SMS system. Table 1 describes the Selective Management System (SMS) that has been employing in a Malaysian forest. One of the limitations of this practice; these activities only take place on 10% of the whole logging area. From the accuracies point of view, this 10% sampling is no longer practical. Therefore, a new method is required.

The stages and activities according to the current practice are well illustrated.

**Table 1.** The SMS activities

| Stage | Year | Activities |
|---|---|---|
| Pre-Harvesting | n-2 years to n-1 years | Pre-felling forest inventory of 10% sampling intensity using systematic-line plots to determine appropriate cutting regimes (>45 cm dbh for non-dipterocarp and >50 cm dbh for dipterocarp). |
| | n-1 year to n | Tree marking incorporating directional felling. |
| Harvesting | n | Felling all marked trees. |
| Post-Harvesting | n + ¼ year to n ½ year | Forest survey to determine fines on trees unfelled, royalty on short logs and tops, and damage residual stands. |
| | n + 2 year to n + 5 year | Post-felling inventory of 10% inventory using systematic-line plots to determine residual stocking and appropriate silvicultural treatments. |
| | n+10 years | Forest inventory of regenerated forest to determine the status of the forest. |

The newly proposed solution has the potential to improve the current method. Table 2 is the general algorithm of the new method.

**Table 2.** The Algorithm of the New Proposed Method.

| **Calculates Volume and Value of each Tree in Logging Area** |
|---|
| Step 1: Read ***idno, speciesName, speciesGroup, dbh, height*** from the tree mapping pre-felling table. |

Step 2: Calculate the volume.

$$volume = \pi(dbh/2)^2 \, height;$$

Step 3: Calculate the value.

```
priceValue=
        "select priceValue from treeValue
        where speciesName=' sN' ";
value= volume .priceValue;
```

**Calculates Threshold Value (The Maximum Allowable Harvest)**

Step 4: Prompt and get the logging area.

Step 5: Calculate the maximum allowable harvest.

$threshold = logArea .30m^3;$

**Divides Forest into Clusters**

Step 6: Determine the length (x) of the logging area and divide it into clusters.

```
begin_x="select x-coor from
        preFelling order by x-coor asc limit 1"
end_x="select x-coor from
        preFelling order by x-coor desc limit 1"
length_x = end_x -begin_x
num_x=(length_x)/50
clus_x=(length_x)/(num_x)
```

Step 7: Determine the length (y) of the logging area and divide it into clusters.

```
begin_y="select y-coor from
        preFelling order by y-coor asc limit 1"
end_y="select y-coor from
        preFelling order by y-coor desc limit 1"
length_y = end_y -begin_y
num_y=(length_y)/50
clus_y=(length_y)/(num_y)
```

Step 8: Determine the clusters.

```
set cno= 0
set begin_x =begin_x
set next_x=begin_x+clus_x
foreach num_x increment by 1
        set begin_y =begin_y
        set next_y =begin_y +clus_y
        foreach num_y increment by 1
            cno=cno+1
            cno(begin_x,next_x,begin_y,next_y)
            begin_y  = next_y
            next_y = begin_y +clus_y
        begin_x =next_x
        next_x  =begin_x +clus_x
```

Step 9: Calculate the number of trees, total value, total volume, and total damage volume based on 20 sets of cutting regimes for every cluster.

```
read and write 20 set of cutting regime to database
foreach set of cutting regime [nonDip,dip]
        create a table:earlyprediction_[nonDip][dip]
        foreach cno
            insert into earlyprediction_[nonDip][dip]
            select count(tree),sumValue,sumVolume,
            sumDamage(0.43(sumResidual))
            where (dbh_G3 || dbh_G4= nonDip)
            && (dbh_G1 || dbh_G2= dip)
```

Step 10: Calculate the AVERAGE of number of trees, total value, total volume, and total damage volume by cluster

```
foreach cno
    read 20 set of cutting regime
    foreach set of cutting regime [nonDip,dip]
        record="select cno,tree,value,volume,damage
        from earlyprediction_[nonDip,dip]
        where cluster=cno"
        foreach data in record
            accumulate tree,value,volume,damage
            determine the tree_average ,value_average ,
                volume_average ,damage_average
            insert into calculatedcluster (cno,tree_average,
                value_average,volume_average,damage_average)
```

Step 11: Sort and sum the records in table: calculatedcluster.

```
sorted="select * from calculatedcluster
order by damage_average asc"
        foreach record of sorted
        calculate Σtree_average,Σvalue_average,Σvolume_average,
        Σdamage_average
```

Step 12: Determine cluster to log and to retain based on the threshold value.

```
update status='L' where
Σvolume_average.calculatecluster≤thresholdValue
update status='R' where status IS NULL
```

Step 13: Determine trees to log based on the cutting regime for each harvestable cluster.

```
cutting regime="insert into treeCutingRegime
select cno,nondip,dip from calculatecluster
where damage=damage_minimum
& status='L'
trees to log="insert into treesToLog
select * from preFelling pf
inner join treeCuttingRegime ct
on ct.treeNo= pf.treeNo
retain trees ="insert into retainTrees
where not exists
(select*from treesToLog)
```

Step 14: Determine minimum damage value and minimum damage volume to residual trees.

```
trees ="select tree_coordinate from treesToLog"
foreach record in tree
determine the residual trees of felling_direction P1
P1_volume= Σ Volume_residualTrees
P1_value= Σ Value_residualTrees
determine the residual trees of felling_direction P2
P2_volume= Σ Volume_residualTrees
P2_value= Σ Value_residualTrees
determine the residual trees of felling_direction P3
P3_volume= Σ Volume_residualTrees
P3_value= Σ Value_residualTrees
determine the residual trees of felling_direction P4
P4_volume= Σ Volume_residualTrees
P4_value= Σ Value_residualTrees
decide tree felling direction:
minimum_value=minimum(P1_value,P2_value,P3_value,P4_value)
```

Firstly, the proposed solution requires a tree mapping pre-felling database that contains detailed information on trees in tropical forests of logging areas. The tree mapping pre-felling database stores the position of each tree (x,y coordinate), DBH (diameter-breast-height), the tree height, species group, and species name. There are about 7650 trees in 9 hectares of forest. The following Fig. 1 displays random of 20 tree records of pre-felling data.

| idno | speciesname | g | dbh | height | x | y |
|---|---|---|---|---|---|---|
| 4991 | Kelandan | G2 | 2.99 | 1.61 | 222.68 | 40.53 |
| 3623 | Pepauh | G4 | 23.19 | 4.87 | 162.39 | 129.96 |
| 2447 | Kelandan | G2 | 23.97 | 4.11 | 127.11 | 114.13 |
| 6170 | Mengkulang jari bulu | G3 | 1.93 | 2.34 | 60.35 | 216.9 |
| 2535 | Pulai basong | G4 | 1.35 | 1.57 | 226.72 | 214.63 |
| 919 | Giam rambai | G1 | 2.71 | 2.64 | 12.2 | 82.08 |
| 671 | Otak udang | G4 | 30.18 | 17.81 | 32.62 | 234.66 |
| 3569 | Jongkong | G5 | 1.34 | 1.65 | 99.34 | 10.36 |
| 3821 | Asam pupoi | G3 | 14.85 | 5.28 | 222.09 | 280.6 |
| 6531 | Kelumpang | G4 | 4.82 | 2.9 | 15.9 | 220.16 |
| 4693 | Balau tembaga | G1 | 3.85 | 2.4 | 254.57 | 220.09 |
| 820 | Jongkong | G5 | 41.56 | 16.31 | 90.15 | 6.01 |
| 2369 | Buluh | G5 | 3.23 | 3 | 181.8 | 24.15 |
| 7580 | Damar hitam bulu | G1 | 4.91 | 2.39 | 25.4 | 30.06 |
| 6137 | Durian tupai | G3 | 3.6 | 1.74 | 248.49 | 146.96 |
| 6681 | Kelandan | G2 | 4.11 | 2.65 | 27.96 | 234.91 |
| 1339 | Keruing gombang | G1 | 1.08 | 2.86 | 225.36 | 232.87 |
| 6880 | Merawan daun bulat | G1 | 3.48 | 2.99 | 88.78 | 59.66 |
| 1101 | Gapis | G4 | 3.86 | 1.29 | 270.49 | 96.28 |
| 3312 | Babai | G5 | 3.41 | 2.42 | 254.54 | 74.09 |

**Fig. 1.** Trees pre-felling data.

Compares to the SMS; that only takes 10% sampling. This study records each tree in the logging area. Therefore, there are various constructive calculations, simulations, and analyses that can be performed using this tree mapping pre-felling records.

At first, before the division of the logging area into clusters, the volume (in $m^3$) and value (in RM) of each tree are calculated. Next, the threshold value is determined according to the size of the logging area. The maximum allowable harvest is $30m^3$ per hectare [16] and the threshold value for 9 hectares is $270m^3$.

Provided with the tree coordinates; the algorithm is designed and executed to determine the logging area which later divides it into clusters. Fig. 2 shows that 9 hectares of forest with a width of 300m and length of 300m are divided into 36 clusters. The size of a cluster is 50m in width and 50m in length [17]. While Fig. 3 presents the detailed position of each cluster.

It is hard to do a comparison between tree to tree of voluminous forest data. Therefore, this study takes into consideration dividing the forest into standard clusters/plots. It appeared that it is more practical and relevant to assess and analyze when data is group and divided accordingly.

Step 9 describes that the algorithm read 20 sets of cutting regimes as shown in Table 3. The algorithm creates 20 tables for 20 cutting regimes and each table consists of 36 records for 36 clusters. Each record in the table store the cluster number, total number of trees, total volume (production), total value, and total damage of residual trees (*there are about 43% of residual trees damaged after harvest [18]*) of potential trees to be harvested according to the specific cutting regime.

Later, the algorithm checks and summarize across all 36 clusters from the 20 tables in Step 9 and calculates the average, accumulate, and sort: number of trees, tree volume, tree value, and damages of residual trees. These records are stored in a dedicated table (*calculatedcluster*).

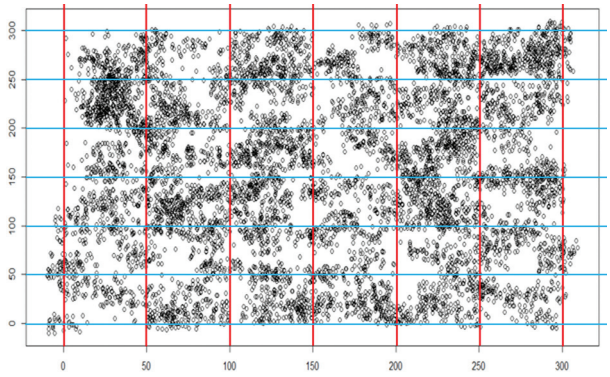



**Fig. 2.** 9 hectares of logging area is divided into 36 clusters.

| cno | begin_x | end_x | begin_y | end_y |
|---|---|---|---|---|
| 1 | -9.64 | 43.52 | -9.89 | 43.31 |
| 2 | -9.64 | 43.52 | 43.32 | 96.52 |
| 3 | -9.64 | 43.52 | 96.53 | 149.73 |
| 4 | -9.64 | 43.52 | 149.74 | 202.94 |
| 5 | -9.64 | 43.52 | 202.95 | 256.15 |
| 6 | -9.64 | 43.52 | 256.16 | 309.36 |
| 7 | 43.53 | 96.69 | -9.89 | 43.31 |
| 8 | 43.53 | 96.69 | 43.32 | 96.52 |
| 9 | 43.53 | 96.69 | 96.53 | 149.73 |
| 10 | 43.53 | 96.69 | 149.74 | 202.94 |
| 11 | 43.53 | 96.69 | 202.95 | 256.15 |
| 12 | 43.53 | 96.69 | 256.16 | 309.36 |
| 13 | 96.7 | 149.86 | -9.89 | 43.31 |
| 14 | 96.7 | 149.86 | 43.32 | 96.52 |
| 15 | 96.7 | 149.86 | 96.53 | 149.73 |
| 16 | 96.7 | 149.86 | 149.74 | 202.94 |
| 17 | 96.7 | 149.86 | 202.95 | 256.15 |
| 18 | 96.7 | 149.86 | 256.16 | 309.36 |
| 19 | 149.87 | 203.03 | -9.89 | 43.31 |
| 20 | 149.87 | 203.03 | 43.32 | 96.52 |
| 21 | 149.87 | 203.03 | 96.53 | 149.73 |
| 22 | 149.87 | 203.03 | 149.74 | 202.94 |
| 23 | 149.87 | 203.03 | 202.95 | 256.15 |
| 24 | 149.87 | 203.03 | 256.16 | 309.36 |
| 25 | 203.04 | 256.2 | -9.89 | 43.31 |
| 26 | 203.04 | 256.2 | 43.32 | 96.52 |
| 27 | 203.04 | 256.2 | 96.53 | 149.73 |
| 28 | 203.04 | 256.2 | 149.74 | 202.94 |
| 29 | 203.04 | 256.2 | 202.95 | 256.15 |
| 30 | 203.04 | 256.2 | 256.16 | 309.36 |
| 31 | 256.21 | 309.37 | -9.89 | 43.31 |
| 32 | 256.21 | 309.37 | 43.32 | 96.52 |
| 33 | 256.21 | 309.37 | 96.53 | 149.73 |
| 34 | 256.21 | 309.37 | 149.74 | 202.94 |
| 35 | 256.21 | 309.37 | 202.95 | 256.15 |
| 36 | 256.21 | 309.37 | 256.16 | 309.36 |

**Fig. 3.** Cluster number and the detail positions.

The *calculatedcluster* table shows the simplified records in a form of the total and average; the number of trees, value, volume, and damage of each cluster of overall 20

sets cutting regime. Referring to this table; potential harvestable clusters are determined based on the threshold value or the maximum harvestable volume (production) as shown in the algorithm of Step 12. Meanwhile, the clusters to be preserved are updated to status = 'R'.

**Table 3.** 20 sets of Cutting Regime

| No. | Non-Dipterocarp (dbh in cm) [nonDip] | Dipterocarp (dbh in cm) [dip] |
|---|---|---|
| 1 | 45 | 50 |
| 2 | 45 | 55 |
| 3 | 45 | 60 |
| 4 | 45 | 65 |
| 5 | 50 | 50 |
| 6 | 50 | 55 |
| 7 | 50 | 60 |
| 8 | 50 | 65 |
| 9 | 55 | 50 |
| 10 | 55 | 55 |
| 11 | 55 | 60 |
| 12 | 55 | 65 |
| 13 | 60 | 50 |
| 14 | 60 | 55 |
| 15 | 60 | 60 |
| 16 | 60 | 65 |
| 17 | 65 | 50 |
| 18 | 65 | 55 |
| 19 | 65 | 60 |
| 20 | 65 | 65 |

Once the potential clusters to be harvested are finalized. The algorithm selects the best cutting regime for each potential cluster. The selection is based on the cutting regime which yields the minimum damage to residual trees. Then, in Step 13 the algorithm able to identify the potential harvestable trees according to the selected cutting regime based on the clusters to be harvested which is determined in Step 12.

In addition to the series of steps in deciding the potential harvestable trees, the new timber harvesting techniques that we introduce also determine the direction of the felling tree which yields the minimum damage volume and the minimum damage value to the residual trees. Apart from the calculation on damage volume, calculation on damage value (in monetary value) is also included.

This study able to determine which felling direction yields the minimum damage cost and minimum damage volume to the surrounding trees. This method takes into consideration of various tree species, tree value, and tree volume of the trees that surround the potential tree to be harvested. There are only a few studies that produce and analyze the damages in terms of monetary values to the residual trees.

To determine the minimum damage cost/value and minimum volume of the residual trees. Total values and volumes of all affected residual trees due to the felling direction of a harvestable tree are calculated. In this study, there are 4 options for the felling direction of a harvestable tree. Fig. 5 shows the Part 1, Part 2, Part 3, and Part 4 felling directions of each harvestable tree. The algorithm is designed and executed to verify which felling direction that produces the minimum damage cost and minimum damage volume to the surrounding of the harvestable tree.

| cno | no | x | y | dbh | volume | value |
|---|---|---|---|---|---|---|
| 9 | 10 | 57.12 | 149.28 | 61.45 | 7.79 | 3129.6 |
| 9 | 26 | 54.42 | 144.11 | 38.1 | 3.27 | 1313.08 |
| 9 | 33 | 75.04 | 114.7 | 38.74 | 1.54 | 837.69 |
| 9 | 50 | 78.27 | 122.26 | 40.4 | 1.89 | 1031.88 |
| 9 | 52 | 68.29 | 143.3 | 42.22 | 3.13 | 1257.29 |
| 9 | 101 | 89.21 | 128.23 | 44.89 | 1.79 | 974.68 |
| 9 | 121 | 67.78 | 145.73 | 82.45 | 8.61 | 6924.07 |
| 9 | 151 | 87.62 | 100.98 | 84.61 | 4.97 | 2708.83 |
| 9 | 178 | 61.32 | 118.64 | 81.57 | 12.31 | 4949.4 |
| 9 | 194 | 81.81 | 121.71 | 34.06 | 1.74 | 698.85 |
| 9 | 293 | 67.92 | 125.46 | 40.62 | 2.92 | 1174.22 |
| 9 | 302 | 96.17 | 144.31 | 54.46 | 2.97 | 1195.81 |
| 9 | 362 | 48.4 | 135.77 | 31.82 | 1.95 | 783.22 |
| 9 | 365 | 56.11 | 129.79 | 86.39 | 8.48 | 3409.66 |
| 9 | 377 | 65.34 | 114.73 | 80.35 | 10.65 | 4280.62 |
| 9 | 396 | 51.14 | 99.05 | 51.09 | 1.97 | 792.8 |
| 9 | 397 | 67.81 | 123.98 | 66.38 | 6.11 | 2456.86 |
| 9 | 406 | 68.09 | 128.88 | 61.17 | 4.32 | 1735.46 |
| 9 | 446 | 90.2 | 99.39 | 78.75 | 6.54 | 2627.66 |
| 9 | 456 | 51.08 | 129.67 | 35.28 | 1.17 | 471.97 |
| 9 | 474 | 55.07 | 120.86 | 41.78 | 3.39 | 1379.89 |
| 9 | 511 | 85.96 | 127.58 | 49.24 | 2.34 | 940.81 |
| 9 | 596 | 90.74 | 112.13 | 58.18 | 3.89 | 1565.67 |
| 9 | 618 | 50.74 | 124.03 | 82.06 | 12.42 | 5056.27 |
| 9 | 624 | 55.58 | 108.94 | 70.61 | 9.07 | 3645.75 |
| 9 | 659 | 62.52 | 101.59 | 32.45 | 1.05 | 423.56 |
| 9 | 662 | 87.76 | 133.37 | 33.95 | 0.76 | 307.14 |

**Fig. 4.** Cluster No. 9 trees volume and value.

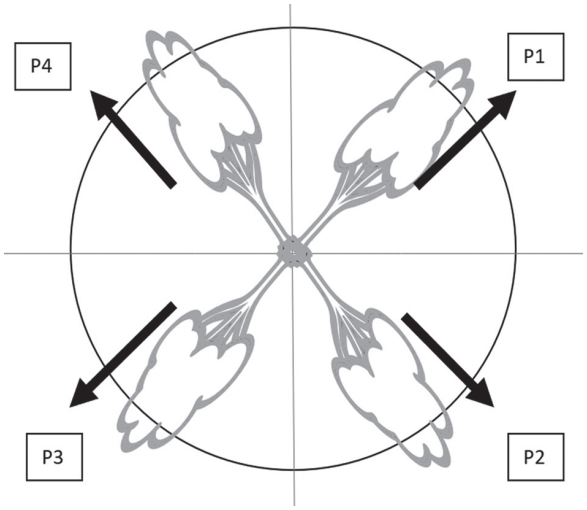



**Fig. 5.** The 4 options of felling.

| no | speciesname | value | speciesgroup |
|---|---|---|---|
| 1 | Giam hantu | 407 | G1 |
| 2 | Kapur | 549 | G1 |
| 3 | Keruing bukit | 545 | G1 |
| 4 | Meranti batu | 402 | G1 |
| 5 | Balau membatu jantan | 407 | G1 |
| 6 | Merawan batu | 402 | G1 |
| 7 | Keruing kertas | 545 | G1 |
| 8 | Balau bukit | 407 | G1 |
| 9 | Merawan penak | 402 | G1 |
| 10 | Meranti merah muda | 732 | G1 |
| 11 | Chengal | 1182 | G1 |
| 12 | Meranti paya | 402 | G1 |
| 13 | Mersawa durian | 626 | G1 |
| 14 | Keruing ropol | 545 | G1 |
| 15 | Merawan mata kucing bukit | 402 | G1 |

**Fig. 6.** The tree species name, value (in RM per 1m$^3$), and tree group.

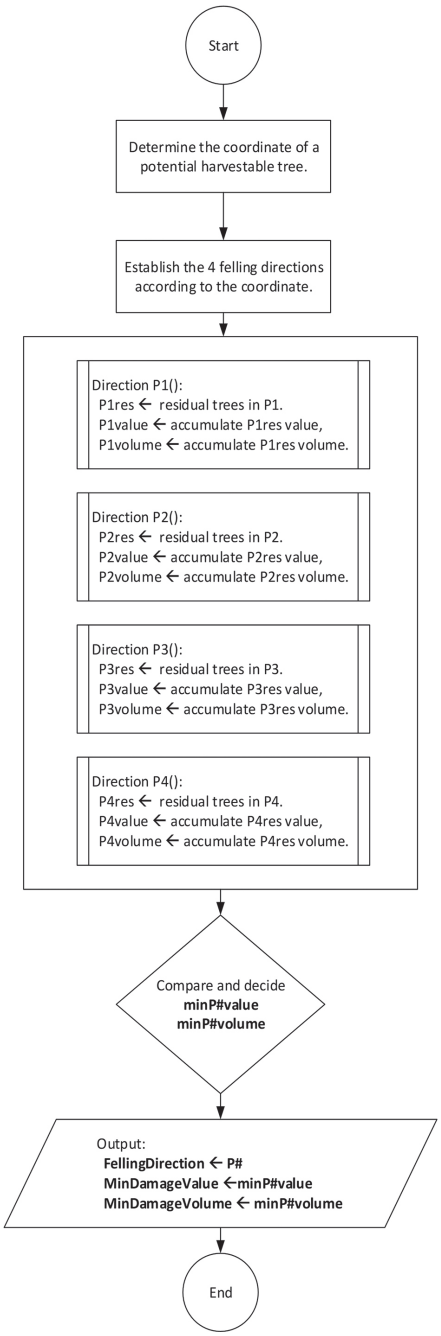



**Fig. 7.** Flowchart to determine minimum damage value and minimum damage volume.

For the damages calculation, at first, the algorithm will determine surrounding trees and verify them into Part 1, Part 2, Part 3, and Part 4 based on their coordinates. Later for each Part or felling direction, the algorithm calculates the total volume and value of residual trees included in it. The felling direction of the harvestable tree is based on the most minimum damage value between Part 1, Part 2, Part 3, and Part 4. The flowchart in Fig. 7 demonstrates the process of this calculation.

## 4. RESULT AND DISCUSSION

In this study, 7650 trees in 9 hectares had been recorded to determine the minimum damage value and minimum damage volume of the potential harvestable trees. The system has to select the best-felling direction before forming the calculation of minimum damage value and minimum damage volume. Fig. 8 shows that the system determines the felling direction of tree number 38; tree species name *Giam Rambai* is P3 (Part 3). The decision-making is based on the minimum damage volume and minimum damage value of the residual trees. This technique takes into account the various tree species, volumes, and values of surrounding affected trees.

| no | x | y | name | dbh | vol | val | dmgvol | dmgval | direct |
|---|---|---|---|---|---|---|---|---|---|
| 38 | 5.78 | 92.02 | Giam rambai | 54.08 | 2.97 | 1207.87 | 0.18 | 77.46 | P3 |

mysql> select * from realexecution_fellingdirection;

| fellno | speciesname | speciesgroup | x | y | dbh | height | volume | value | status |
|---|---|---|---|---|---|---|---|---|---|
| 38 | Giam rambai | G1 | 5.78 | 92.02 | 54.08 | 12.92 | 2.97 | 1207.87 | M30 |
| 367 | Kelandan | G2 | 7.15 | 92.88 | 58.88 | 22.54 | 6.14 | 2467.21 | M30 |
| 501 | Penarahan arang | G3 | 13.44 | 82.58 | 64.37 | 22.59 | 7.35 | 2955.28 | M30 |
| 767 | Jelawai mentalun | G4 | 13.2 | 93.81 | 67.09 | 8.08 | 2.86 | 1148.27 | M30 |
| 803 | Pulai penipu paya | G4 | 1.08 | 102.42 | 33.85 | 16.9 | 1.52 | 611.39 | M30 |
| 901 | Giam rambai | G1 | -3.11 | 94.51 | 1.76 | 1.26 | 0 | 0.12 | SE |
| 902 | Giam rambai | G1 | 10.94 | 93.36 | 3.99 | 1.4 | 0 | 0.71 | SE |
| 903 | Giam rambai | G1 | 14.05 | 100.65 | 3.97 | 2.66 | 0 | 1.34 | SE |
| 904 | Giam rambai | G1 | 11.8 | 84.59 | 1.77 | 2.81 | 0 | 0.28 | SE |
| 905 | Giam rambai | G1 | -0.86 | 82.21 | 4.78 | 2.11 | 0 | 1.54 | SE |
| 906 | Giam rambai | G1 | 13.65 | 96.82 | 1.75 | 2.92 | 0 | 0.29 | SE |
| 907 | Giam rambai | G1 | -0.11 | 85.13 | 4.64 | 1.34 | 0 | 0.92 | SE |
| 908 | Giam rambai | G1 | 1.54 | 98.93 | 2.04 | 1.61 | 0 | 0.21 | SE |
| 909 | Giam rambai | G1 | 1.72 | 92.38 | 3.42 | 1.77 | 0 | 0.66 | SE |
| 910 | Giam rambai | G1 | 12.33 | 101.63 | 2.97 | 1.46 | 0 | 0.41 | SE |
| 911 | Giam rambai | G1 | -2.09 | 84 | 1.4 | 2.42 | 0 | 0.15 | SE |
| 912 | Giam rambai | G1 | 7.62 | 101.79 | 4.76 | 2.27 | 0 | 1.64 | SE |
| 913 | Giam rambai | G1 | -1.41 | 96.23 | 4.66 | 1.83 | 0 | 1.27 | SE |
| 914 | Giam rambai | G1 | 10.65 | 95.82 | 2.18 | 2.38 | 0 | 0.36 | SE |
| 915 | Giam rambai | G1 | 15.68 | 93.5 | 1.85 | 1.85 | 0 | 0.2 | SE |
| 916 | Giam rambai | G1 | 14.44 | 95.89 | 3.98 | 1.58 | 0 | 0.8 | SE |
| 917 | Giam rambai | G1 | 9.93 | 85.23 | 3.84 | 2.05 | 0 | 0.97 | SE |
| 918 | Giam rambai | G1 | 11.77 | 100.9 | 4.9 | 1.72 | 0 | 1.32 | SE |
| 919 | Giam rambai | G1 | 12.2 | 82.08 | 2.71 | 2.64 | 0 | 0.62 | SE |
| 920 | Giam rambai | G1 | 1.39 | 91.11 | 3.62 | 1.73 | 0 | 0.72 | SE |

**Fig. 8.** Felling direction of tree number 38.

An iteration of this algorithm is designed to calculate and determine the minimum damage volume and minimum damage value of each harvestable tree to decide

on the tree felling direction. Next, the algorithm produces the total number of felling trees, the total value of felling trees, the total damage volume (production), and the total damage value (damage) by the cluster as shown in Fig. 9.

| cno | treenum | value | volume | damage |
|---|---|---|---|---|
| 1 | 9 | 22352.13 | 51.44 | 12.1 |
| 2 | 13 | 31356.86 | 74.48 | 8.18 |
| 3 | 7 | 18741.01 | 45.67 | 8 |
| 4 | 12 | 38162.93 | 88.53 | 7.19 |
| 5 | 13 | 48521.95 | 100.99 | 12.13 |
| 6 | 13 | 40712.49 | 89.41 | 8.7 |
| 7 | 11 | 25655.23 | 59.59 | 10.58 |
| 8 | 10 | 24175.57 | 59.86 | 8.73 |
| 9 | 19 | 59140.05 | 136.57 | 18.43 |
| 10 | 16 | 39798.17 | 97.46 | 11.27 |
| 11 | 14 | 29958.67 | 71.91 | 13.39 |
| 12 | 10 | 23871.5 | 56.69 | 8.8 |
| 13 | 9 | 24482.18 | 59.84 | 9.96 |
| 14 | 12 | 34669.07 | 86.14 | 11.08 |
| 15 | 13 | 33382.47 | 71.82 | 16.56 |
| 16 | 9 | 32342.46 | 77.48 | 17.81 |
| 17 | 15 | 35833.61 | 88.89 | 13.8 |
| 18 | 10 | 22916.71 | 54.77 | 10.53 |
| 19 | 13 | 35708.76 | 87.4 | 9.92 |
| 20 | 15 | 51275.39 | 121.26 | 6.94 |
| 21 | 16 | 41875.33 | 100.35 | 13.65 |
| 22 | 9 | 16952.78 | 42.07 | 10.58 |
| 23 | 11 | 35415.85 | 77.47 | 9.88 |
| 24 | 10 | 36676.15 | 79.11 | 11.08 |
| 25 | 9 | 26227.82 | 60.12 | 10.95 |
| 26 | 10 | 27743.13 | 66.95 | 7.36 |
| 27 | 18 | 58442.13 | 142.29 | 11.79 |
| 28 | 10 | 22862.43 | 56.76 | 22.31 |
| 29 | 9 | 22379.6 | 55.46 | 12.41 |
| 30 | 15 | 40788.53 | 91.43 | 10.96 |
| 31 | 10 | 18415.63 | 43.51 | 5.35 |
| 32 | 8 | 22583.45 | 47.58 | 6.05 |
| 33 | 9 | 22859.98 | 54 | 5.11 |
| 34 | 7 | 15868.92 | 38.48 | 13.36 |
| 35 | 10 | 24030.57 | 52.51 | 17.08 |
| 36 | 17 | 52571.27 | 122.28 | 17.64 |

**Fig. 9.** Calculated the total number of felling trees, the total value of felling trees, the total damage volume, and the total damage value by clusters.

The newly proposed method introduces the logging area divided into clusters. Based on records shown in Fig. 9, the system descending sorts according to its value, production and ascending sort to its damage volume, and damage value by clusters. Then, the system accumulates those values as stated in Fig. 10.

This study proposed a new method by dividing the logging area into clusters and able to determine which tree to fell is based on the value of timber and minimum damage to the residual tree. With this new method, certain clusters will be preserved to maintain forest regeneration.

Referring to the threshold value (*maximum allowable harvest = 270m³ for 9-hectare forest*), the system then calculated and only trees within 4 clusters are affected for a logging operation as shown in Fig. 11.

| cno | totaltree | totalvalue | totalvolume | totaldamage |
|---|---|---|---|---|
| 33 | 9 | 22859.98 | 54 | 5.11 |
| 31 | 19 | 41275.61 | 97.51 | 10.46 |
| 32 | 27 | 63859.06 | 145.09 | 16.51 |
| 20 | 42 | 115134.45 | 266.35 | 23.45 |
| 4 | 54 | 153297.38 | 354.88 | 30.64 |
| 26 | 64 | 181040.51 | 421.83 | 38 |
| 3 | 71 | 199781.52 | 467.5 | 46 |
| 2 | 84 | 231138.38 | 541.98 | 54.18 |
| 6 | 97 | 271850.87 | 631.39 | 62.88 |
| 8 | 107 | 296026.44 | 691.25 | 71.61 |
| 12 | 117 | 319897.94 | 747.94 | 80.41 |
| 23 | 128 | 355313.79 | 825.41 | 90.29 |
| 19 | 141 | 391022.55 | 912.81 | 100.21 |
| 13 | 150 | 415504.73 | 972.65 | 110.17 |
| 18 | 160 | 438421.44 | 1027.42 | 120.7 |
| 7 | 171 | 464076.67 | 1087.01 | 131.28 |
| 22 | 180 | 481029.45 | 1129.08 | 141.86 |
| 25 | 189 | 507257.27 | 1189.2 | 152.81 |
| 30 | 204 | 548045.8 | 1280.63 | 163.77 |
| 24 | 214 | 584721.95 | 1359.74 | 174.85 |
| 14 | 226 | 619391.02 | 1445.88 | 185.93 |
| 10 | 242 | 659189.19 | 1543.34 | 197.2 |
| 27 | 260 | 717631.32 | 1685.63 | 208.99 |
| 1 | 269 | 739983.45 | 1737.07 | 221.09 |
| 5 | 282 | 788505.4 | 1838.06 | 233.22 |
| 29 | 291 | 810885 | 1893.52 | 245.63 |
| 34 | 298 | 826753.92 | 1932 | 258.99 |
| 11 | 312 | 856712.59 | 2003.91 | 272.38 |
| 21 | 328 | 898587.92 | 2104.26 | 286.03 |
| 17 | 343 | 934421.53 | 2193.15 | 299.83 |
| 15 | 356 | 967804 | 2264.97 | 316.39 |
| 35 | 366 | 991834.57 | 2317.48 | 333.47 |
| 36 | 383 | 1044405.84 | 2439.76 | 351.11 |
| 16 | 392 | 1076748.3 | 2517.24 | 368.92 |
| 9 | 411 | 1135888.35 | 2653.81 | 387.35 |
| 28 | 421 | 1158750.78 | 2710.57 | 409.66 |

**Fig. 10.** Descending sort and accumulates trees, value, volume, and damage.

| cno | treenum | value | production |
|---|---|---|---|
| 20 | 15 | 51275.39 | 121.26 |
| 31 | 10 | 18415.63 | 43.51 |
| 32 | 8 | 22583.45 | 47.58 |
| 33 | 9 | 22859.98 | 54.00 |

**Fig. 11.** The affected clusters for logging operations.

## 5. CONCLUSION

The research introduces two new elements to be included in timber harvesting pre-felling analysis is to ensure forest regrowth which able to minimize damage as well as maintaining an adequate quantity and quality of residual stands. The first element that this research highlighted is to determine the minimum damage cost/value to the residual trees according to tree felling direction.

In addition, to retain areas of unlogged forest for preservation; this research focused on the division of logging area into clusters where only certain clusters will be affected in a logging operation and the rest are conserved.

## 6. ACKNOWLEDGMENT

The authors would like to thank the members of the Interest Group on Research in Intelligent System (IGRIS) research group for their ideas and recommendations throughout the paper writing. Our thanks to Universiti Kuala Lumpur for full support in this research.

## 7. REFERENCES

[1] S. Chao, "Forest-Peoples-Numbers-Across-World-Final 0", For. Peoples Program., Vol. 1, 2012.

[2] T. R. H. Pearson, S. Brown, L. Murray, G. Sidman, "Greenhouse gas emissions from tropical forest degradation: An underestimated source", Carbon Balance Management, Vol. 12, No. 1, 2017.

[3] A. Ahmad, Q. J. Liu, S. M. Nizami, A. Mannan, S. Saeed, "Carbon emission from deforestation, forest degradation and wood harvest in the temperate region of Hindukush Himalaya, Pakistan between 1994 and 2016", Land use policy, Vol. 78, 2018, pp. 781–790.

[4] B. Bernal, L. T. Murray, T. R. H. Pearson, "Global carbon dioxide removal rates from forest landscape restoration activities", Carbon Balance Management, Vol. 13, No. 1, 2018.

[5] F. H. Susanty, "Study of recovery rates of natural forest stands after logging in East Kalimantan", IOP Conference Series: Earth and Environmental Science, Vol. 533, No. 1, 2020.

[6] E. Cedamon, G. Paudel, M. Basyal, I. Nuberg, K. K. Shrestha, "Applications of single-tree selection guideline following a DBq approach on Nepal's community forests", Banko Janakari, No. 4, 2018, pp. 104–112.

[7] J. Jamhuri et al., "Selective logging causes the decline of large-sized mammals including those in unlogged patches surrounded by logged and agricultural areas", Biological Conservation, Vol. 227, 2018, pp. 40–47.

[8] I. Saiful, A. Latiff, "Canopy gap dynamics, effects of selective logging: A study in a primary hill dipterocarp forest in Malaysia", Journal of Tropical Forest Science, Vol. 31, No. 2, 2019, pp. 175–188.

[9] M. Demies, H. Samejima, A. K. Sayok, G. T. Noweg, "Tree diversity, forest structure, species composition in a logged-over mixed dipterocarp forest, Bintulu, Sarawak, Malaysia", Transitions on Science and Technology, Vol. 6, No. 2, 2019, pp. 102–110.

[10] N. J. Naim Jemali, M. F. Abd Rhani, M. Muhammad, N. K. S. Abdul Majid, "Forest Growth Analysis of Ulu Sat Forest Reserve", IOP Conference Series: Earth and Environmental Science Sci., Vol. 549, No. 1, 2020.

[11] M. Azian et al., "Carbon emission assessment from different logging activities in production forest of Pahang, Malaysia", Journal of Tropical Forest Science, Vol. 31, No. 3, 2019, pp. 304–311.

[12] A. H. Atiqah, P. P. Rhyma, J. Jamhuri, A. W. Zulfa, M. S. Samsinar, K. Norizah, "Using Google earth imagery to detect distribution of forest cover change-Is the technique practical for Malaysian forests?", Malaysian Forester, Vol. 83, No. 1, 2020, pp. 1–15.

[13] R. Pillay, F. Hua, B. A. Loiselle, H. Bernard, R. J. Fletcher, "Multiple stages of tree seedling recruitment are altered in tropical forests degraded by selective logging", Ecology and Evolution, Vol. 8, No. 16, 2018, pp. 8231–8242.

[14] D. B. Lindenmayer, "Integrating forest biodiversity conservation, restoration ecology principles to recover natural forest ecosystems", New Forests, Vol. 50, No. 2, 2019, pp. 169–181.

[15] G. R. Cerullo, D. P. Edwards, "Actively restoring resilience in selectively logged tropical forests", Journal of Applied Ecology, Vol. 56, No. 1, 2019, pp. 107–118.

[16] R. B. de Lima et al., "Accurate Estimation of Commercial Volume in Tropical Forests", Forest Science, Vol. 67, No. 1, 2021, pp. 14–21.

[17] H. Omar, M. Hasmadi Ismail, M. Hakimi Abu Hassan, "Optimal Plot Size for Sampling Biomass in Natural, Logged Tropical Forests", Proceedings of the Conference on Forestry and Forest Products, Sunway Putra, Kuala Lumpur, 11-12 November 2013.

[18] E. van der Werf, Y. Indrajaya, F. Mohren, E. C. van Ierland, "Logging damage, injured tree mortality in tropical forest management", Natural Resource Modeling, Vol. 32, No. 4, 2019, pp. 1–20.

# Review of Ad Hoc Networks scenarios and challenges in years 2015-2019

Review paper

**Amna Saad**
Universiti Kuala Lumpur,
Malaysia Institute of Information Technology
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250 Kuala Lumpur,
Wilayah Persekutuan Kuala Lumpur, Malaysia
amna@unikl.edu.my

**Husna Osman**
Universiti Kuala Lumpur,
Malaysia Institute of Information Technology
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250 Kuala Lumpur,
Wilayah Persekutuan Kuala Lumpur, Malaysia
husna@unikl.edu.my

**Mufind Mukaz Ebedon**
Universiti Kuala Lumpur,
Malaysia Institute of Information Technology
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250 Kuala Lumpur,
Wilayah Persekutuan Kuala Lumpur, Malaysia
mufind.mukaz@s.unikl.edu.my

***Abstract*** *– A Mobile Ad-hoc Network (MANET) protocol performance analysis depends on the type of simulation tools, mobility models, and metrics used. These parameters' choice is crucial to researchers because it may produce an inaccurate result if it is not well chosen. The challenges researcher is facing are on the choice of these four parameters. Our survey shows an inclination to used Ad-hoc On-Demand Distance Vector routing (AODV) for performance comparison and enhancement of it by the researcher. Network simulation 2 (NS2) was the most selected tool, but we observe a decline in its utilization in recent years. Random Waypoint Mobility model (RWPM) was the most used mobility model. We have found a high percentage of the published article did not mention the mobility models use; this will make the result difficult for performance comparison with other works. Packet Delivery Ratio (PDR), End to End Delay (E2ED) were the most used metrics. Some authors have self-developed their simulation tools; the authors have also used new metrics and protocols to get a particular result based on their research objective. However, some criteria of choosing a protocol, metrics, mobility model, and simulation tool were not described, decreasing the credibility of their papers' results. Improvement needs to be done in the Ad-hoc network in terms of benchmark, acceptable scenario parameters. This survey will give the best practice to be used and some recommendations to the Ad-hoc network community.*

***Keywords****: MANET, Routing protocols, Mobility Model, Metric, Simulation tool, Performance analysis*

## 1. INTRODUCTION

MANET can be implemented without any centralized administration. Mobile nodes in the Ad-hoc network form a network on the go; each mobile node can be a forwarding node and sending node simultaneously.

Ad-hoc networks can be classified into Mobile Ad-hoc Networks (MANET), Vehicular Ad-hoc Network (VANET), Wireless Sensor Network (WSN), and Flying Ad-hoc Network (FANET). Ad-hoc networks need a protocol that adapts to these parameters: topology change, updating of a new route path, and energy-efficient of the node.

Designing a protocol that will include these criteria and give a better result in a different scenario is difficult to achieve. Various proto- cols have been proposed in the literature, such as Ad- hoc On-demand Distance Vector (AODV) [1], Dynamic Source Routing (DSR) [2], Dynamic Manet On-demand (DYMO) [3], Ad-hoc On-demand Multipath Distance Vector (AOMDV) [4], Multicast Ad-hoc On-demand Distance Vector (MAODV) [5], Destination Sequenced Distance Vector (DSDV) [6],

Optimized Link State Routing (OLSR) [7], Zone Routing Protocol (ZRP) [8] and Low-energy Adaptive Clustering Hierarchy (LEACH) [9].

The mobile nodes in ad-hoc networks move arbitrarily, and the topology of the network change frequently. The mobile node's arbitrary movement makes it hard for a researcher to find the mobility models close to the reality (human movement) for implementation in the simulation tools.

Here are some well-known simulation tools in Adhoc network : Network simulator 2 (NS2) [10], Network Simulator 3 (NS3) [11], Objective Modular Network Testbed in C++ (OMNet++) [12], Optimized Network Engineering Tool (OPNET) [13], Global Mobile Information System Simulator (GloMoSim) [14], Matrix Laboratory (MATLAB) [15], Castalia [16] and EXATA Cyber [17].

The mobility model in Ad-Hoc Networks shows how the mobile nodes move in the network. There are different mobility models based on their movement patterns, such as Random Way Point Mobility Model (RWPM), Random Walk Mobility Models (RWM), Reference Point Group Mobility Models (RPGM), Gauss-Markov Mobility Models (GMM) [18].

Open Simulation of Urban Mobility (SUMO) combines with a Mobility generator for Wireless Networks (MOVE) [19, 20], and MOBISIM [21, 22]. Metrics play a significant role in data analysis; some well-known metrics are Packet Delivery Ratio (PDR), Average Overhead, Throughput, End-to-End Delay E2ED, Energy consumption, and Jitter.

The review analyses 169 papers retrieved from 5 Scopus journals selected based on their focus on mobile ad-hoc networks and other criteria from 2015 to 2019. The result shows that 45.6% of the researchers used Adhoc On-Demand Distance Vector routing (AODV) as a protocol for their performance comparison, 37.9% used the Random Waypoint Mobility Model (RWPM) in their simulation scenarios. 52.1% used Packet Delivery Ratio (PDR) as their metrics, and 60.9% used Network Simulation 2 (NS2) as the simulation tool. We have observed that NS2 utilization is declining in recent years because it does not support new technology (IoT, 5G). Some papers did not mention the mobility model used, which can create difficulty for the researcher to compare with other works. The researcher will face some challenges when selecting the simulation tools, the protocol for comparison or enhancement, choice of mobility model, and metrics to use for a better view of the networks. Our review shows what parameters have been used as a reference for future researchers. We have proposed some best practices and some recommendations to the ad hoc network community for improvement.

The rest of this review is organized as follows: Section 2 reviews existing literature, Section 3 gives detail of the review methodology, the result is provided in Section 4. Section 5 discusses the result and recommendation. Finally, Section 6 contains the conclusion.

## 2. LITERATURE REVIEW

Yoon et al. [23] studied the random waypoint and found out that it is not a good mobility model to use because unreliable results can be obtained using this model. They observe that if the simulation time increases, some metrics show a drop in performance. They proposed a new modified Random mobility model.

Kurkowski et al. [24] surveyed proceeding paper in Association for Computing Machinery (ACM) on MANET from 2000-2005. They extracted simulation parameters from those surveyed papers. They found out that NS2 was the most simulation tool used by authors and RWPM as a mobility model. Some missing parameters were observed, and the unavailability of code was missing for self-developed simulation tools.

Hiranandani et al. [25] conducted a review of published papers in the ACM conference between 2006 to 2010. They found out that missing parameters are still observed. Default parameters for simulators have been used by research, which raises a question on the result's credibility. They observed that since the study in [24], the current mobile ad hoc simulation practices are still not progressing. Still, some custom simulators were more used compared to existing simulator tools.

Naicken et al. [26] surveyed 280 papers in the peer-to-peer network to see what simulation tools were used. They found out that custom simulators surpass the well-known simulators in terms of usage.

Kurkowski et al. [27] proposed a simulation standard by using two metrics for characterization of the simulation scenarios, such as the average shortest path hop count and the average amount of network partitioning. Enhancement of Kurkowski works was done in [28], which adds a new metric average neighbour count.

Andel et al. [29] focused on the credibility of manet simulation tools. By analyzing multiple review papers published, they found out that most authors do not specify their simulator's version, and missing simulation parameters can be observed. They proposed some solutions to improve the credibility of the simulation tools.

Ahmad et al. [30] proposed a comprehensive comparison of AODV, DSDV, and ZRP protocol by reviewing some related works in terms of routing used, simulator tool, metrics, network type, and qualitative analysis was done.

Sanchez et al. [31] surveyed Unmanned aerial and aquatic vehicles, and the focus was on their communication, application, and tools for the evaluation.

Garcia et al. [32] proposed a methodology to help the researcher conduct good simulation practice in their scenarios in VANET. Different simulation scenarios were proposed in NS2 to determine an excellent method to use.

Other research focuses on the performance comparison of different simulation tools [33-36].

To the best of our knowledge, no review paper has covered the literature on performance evaluation or analyses of articles in Ad-Hoc networks from 2015 to 2019 in MANET, WSN, FANET, and VANET based on different parameters. Therefore, this study aims is to conduct a literature review on the performance evaluation to:

a) Identify the mobility model, metrics, simulation tools, and routing protocols used in MANET, VANET, WSN, and FANET with a synthesis of empirical evidence

b) To analyze the result and present our finding

## 3. REVIEW METHODS

This review will show the process of formulating the research questions, which include the search processes that represent the keyword and the Inclusion-Exclusion criteria for selecting the papers. The following are the formulating research questions for our review:

a) RQ1: Which of the simulation tools, mobility models, metrics, and routing protocol are the most used in the performance analysis of MANET, VANET, WSN, and FANET?

b) RQ2: What are the lessons learned and best practices in the performance analysis of ad hoc networks?

### 3.1. SEARCH PROCESSES

The literature has been searched from selected Scopus journals, which are focusing on ad-hoc networks IEEE Access (IEEE), IEEE Transactions on communication (IEEE), Wireless Networks (Springer), Wireless Personal Communication (Springer), and Ad-hoc networks (Elsevier). The criteria for selecting these journals were based on the higher number of published articles in ad hoc networks compared to other journals.

The keyword search in the article and the abstract was: "performance comparison" and "mobility model" and "simulation" and "ad-hoc networks" or secondary keyword "performance evaluation" and "mobility model" and "routing protocols" and "ad-hoc networks" in those five journals ranging from 2015 to 2019. Retrieving paper after using the keyword was analyzed first by screening the title to see if it is relevant to our objective; if yes abstract was read to double confirm if the paper can be selected or not. A full reading of the selected paper was done to see if the paper contains these parameters (metrics, simulation tools, protocol, mobility models). For protocol, we included the article, which details the protocol used for the implementation of the scheme or algorithm, and those who compared new protocol with existing protocols. We need to clarify that the compulsory parameters for inclusion were the Simulation tools, metrics, and protocols. The articles that do not show the mobility model but have all the other parameters were selected in our review.

### 3.2. INCLUSION-EXCLUSION CRITERIA

The selection of the articles was conducted based on the inclusion and exclusion criteria. The articles which fulfil the criteria in Table 1 were selected, and those that do not were excluded in Table 2.

**Table 1.** Inclusion criteria for the selection of articles

| Inclusion criteria |
|---|
| Publication date 2015-2019 |
| English |
| Any geographical location |
| The articles published in IEEE Access, IEEE Transactions on communication, Wireless Networks, Wireless Personal communication, and Ad-hoc networks. |
| Performance comparison, routing protocols, and mobility model in ad-hoc networks must be the primary topic or secondary topic of the publication |
| The article should report the simulation parameters table or a simulation parameter description and should include the protocol, simulation tools |

**Table 2.** Exclusion criteria for the selection of articles

| Exclusion Criteria |
|---|
| Published pre-2015 |
| Non-English |
| Proceedings and peer-reviewed papers, and articles published in other journals, patent. |
| The proposal, lectures notes, A summary of conference Keynote, Dissertation/Thesis, Doctoral Workshop, and tutorial |
| Articles did not present the simulation parameters table or a simulation parameter description such as (protocol, metrics, simulation tools) |
| An article which presents algorithm and schemes without comparing it to the new protocol or without describing in what protocol it was implemented were excluded |

## 4. RESULTS

The flow diagram in Figure 1 shows the articles' selection process for the review. The following number of articles from 5 Scopus journals have been retrieved from 3 databases (IEEE, Springer, Elsevier): From IEEE two journals; IEEE Access (429), IEEE Transaction on communication (127), from Springer 2 journal; Wireless Networks (306) and Wireless Personal communication (384) and finally from Elsevier Ad-hoc networks (400). A total of 1646 articles were identified. After the title and abstract filtering, 796 articles were excluded because they were not relevant to the topic and did not focus on ad-hoc networks. We have conducted a full article review on the remaining 850 articles, and 672 were excluded due to missing some parameters listed in the Exclusion criteria, as illustrated in Table 2. The remaining 169 were selected for the review.

## 5. DISCUSSION

*A. Which of the simulation tools, mobility models, metrics, and routing protocol are the most used in the performance analysis of MANET, VANET, WSN, and FANET?*

To answer RQ1, we have extracted 169 articles with the following information. For a summary result see Figure 4.

### 5.1. SIMULATION TOOLS

Figure 2 shows the network simulator tools used in all the articles we have reviewed from 2015 to 2019. We can observe an inclination toward free simulator tools than the paid ones. Network simulator 2 (NS2) was the most used with 103 (60.9%) out of 169 articles, 13 (7.7%) NS3, 14 (8.3%) MATLAB, 7 (4.1%) used OMNET++ and OPNET Modeler, 5(3%) GloMoSim and Qualnet and 3 (1.8%) EXATA/Cyber. Figure 3 shows that 70 out of 99 (77.7%) in MANET, 15 out of 32 (46.8%) in VANET and 18 out of 37 (48.6%) in WSN. MATLAB and OPNET surpass NS3 in terms of usage in MANET. OMNET++ and NS3 have similar articles used in VANET.

NS2 has shown a high percentage of usage in our review because it is a free license simulator. NS2 has multiple models, many protocols are implemented in it, the source code is available for free, and the documentation can be found on the NS website and other pages [37]. The negative side is that it does not contain new features that can support other research areas like the Internet of Things (IoT), 5G. Another disadvantage is that NS2 code cannot be implemented directly to a real system due to multiple languages (TCL, C++). Comparing our result with the result in [24] we can observe similarities in terms of utilization of NS2. Our survey shows 60.9% out of 169 papers and 43.8% out of 80 for the previous review. The result in [25] shows that a custom simulator was more used, but fewer authors used a custom simulator in our research. This review has observed a decrease in the utilization of NS2 in the current year, and improvement has been seen since the work in [25] that standard simulation is more used than custom simulation tools, which can help researchers to repeat the work.

We can conclude that there are no predefined simulator tools to use in a particular research area, as long as the code's availability is there, and the evaluation methodology is well designed. New simulators are available, which can bring new features and models to the research area. Still, the challenges are the source code's availability, implementation of various models, and documentation.

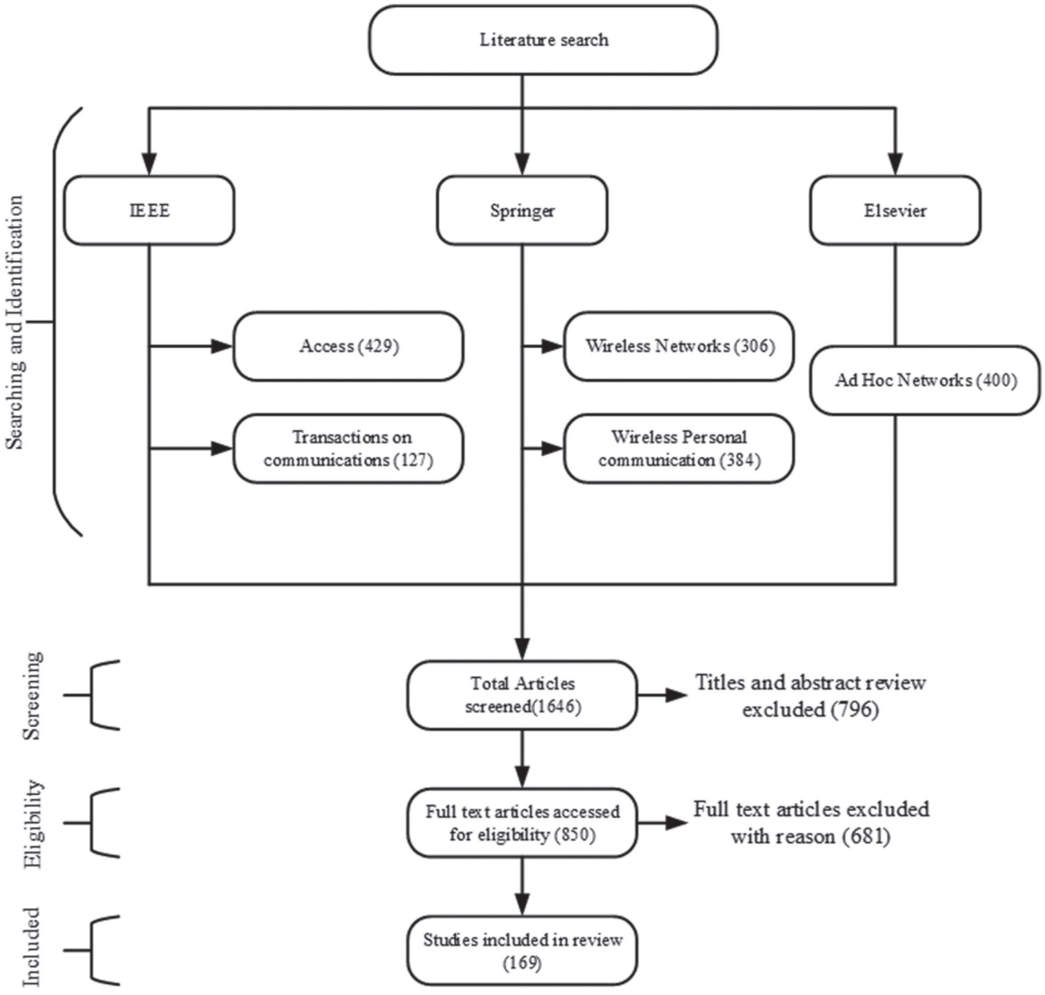



**Fig. 1.** The flow diagram of the article for the literature review

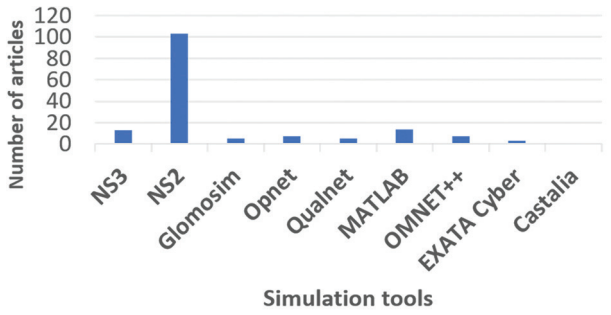



**Fig. 2.** Simulation tools usage

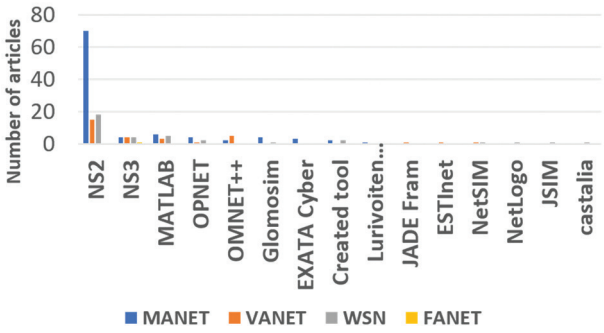



**Fig. 3.** Simulation tools usage base on MANET, VANET, WSN, FANET

### 5.2. PROTOCOLS

The studies of protocols are very important before deciding on what protocol to enhance or compare with when designing an evaluation study. The non-availability of some protocol source code makes it difficult for the researcher to enhance or compare with their new protocol. Figure 5, shows 77 (45.6% ) out of 169 articles used AODV, 27 (16% ) DSR, 17 (10.1%) OLSR, 9 (5.3%) LEACH, 8 (4.7%) DSDV, 6 (3.6%) DYMO, 9 (5.3%) AOMDV, 5 (3%) MAODV and 3 (1.8%)ZRP. Table 3 shows protocols that have a low percentage of utilization. This study's trend indicates that reactive protocol is preferred for performance comparison with the new one. Figure 6 shows the most protocol used in ad-hoc networks, 50 out of 99 (55.5%) articles used AODV in MANET, in VANET 12 out of 32 (37.5%), WSN 14 out of 37 (37.8%), and 1 out of 1 article in FANET, the next protocol which was most used is DSR. LEACH was only used in WSN 8 out of 37 (21.6%).

The challenge facing is on what protocol to compare with and in what scenarios. Comparing an existing protocol with a new one helps to view the new one's positive and negative aspects. The question is, in what situation should a protocol be chosen for comparison. Research in [38] compared AODV and ZRP with their proposed protocol GeoZRP. Why not compare it Directly to ZRP or compare it with protocol base on their focus, e.g., security, energy, multipath. Those are the challenges researchers face in the selection of protocol to analyze or enhance.

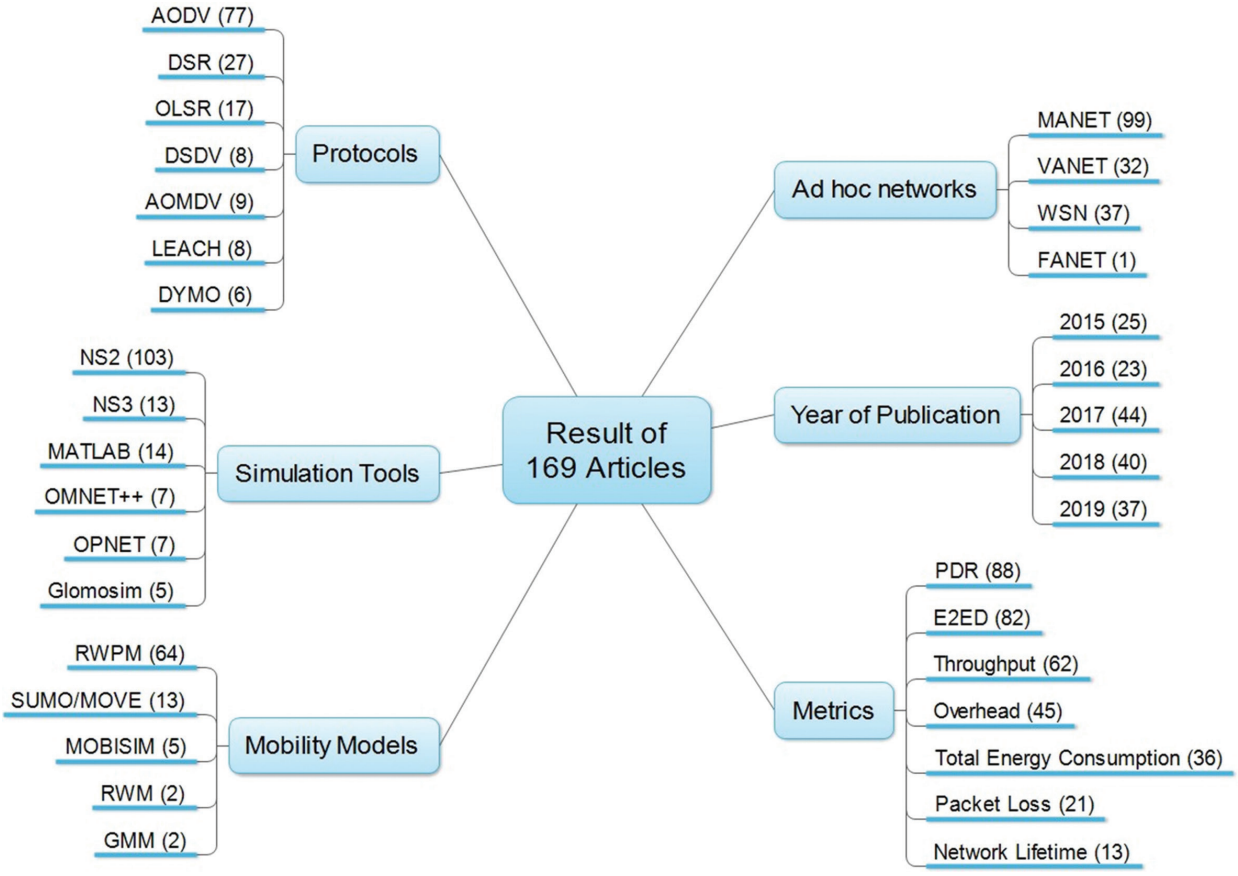



**Fig. 4.** Summary result of 169 articles selected in the review

We can conclude that a solution to solve this problem is for research to implement new protocols in well-known simulators tools and make the source code available. It will enable future researchers in ad-hoc networks to have the possibility to modify, reproducing, and confirming their results with the new existing protocol, rather than comparing it with an old protocol because of its availability of the source code.

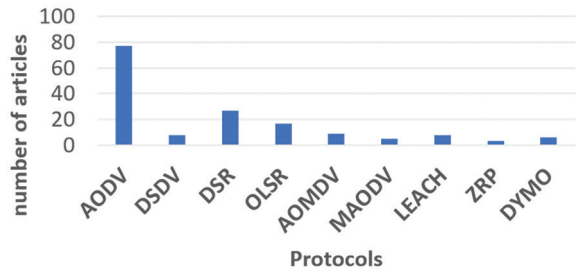



**Fig. 5.** Routing protocol choice by authors for the performance comparison

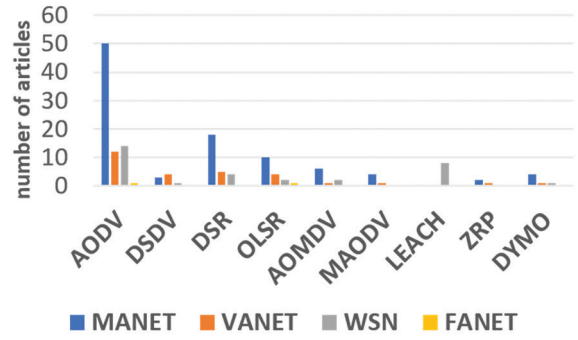



**Fig. 6.** Routing protocols usage based on MANET,VANET, WSN, FANET

**Table 3.** Protocols that have a low percentage of utilization

| Ref | Protocol | Ref | Protocol | Ref | Protocol |
|---|---|---|---|---|---|
| [39] | CR-EAOMDV | [40] | LEACH-M | [41, 42] | STFDR |
| [43] | EPC-AODV | [44] | LEACH-C | [45] | NCLR |
| [39] | AODMV-MR | [46] | EACHP | [47] | RIP |
| [39] | CAODV | [40] | EPCR | [48] | LAR |
| [49] | DAODV | [50] | FDCRP | [51] | WECRR |
| [52] | PMT-AODV | [51] | DFCR | [53] | A-CAR |
| [54] | FTDSR | [53] | IVD-CAGR | [55] | EAR |
| [56] | QoS-UMDSR | [53] | CSR | [20, 57] | GSR |
| [56] | QoS-UDSR | [55] | CLB | | |
| [20] | SC-OLSR | [55] | DGLB | | |

### 5.3. MOBILITY MODELS

Mobility models represent how the mobile nodes move inside the mobile ad-hoc network based on a specific pattern, position, and speed changes. Change in speed, position and pattern will result in the displacement of the mobile nodes in a particular region. In our review, different mobility models were selected by authors. Figure 7 shows that RWPM was the most used mobility model with 64 (37.9%), SUMO/MOVE 13 (7.7%), MOBISIM 5 (3%), and 2 (1.2%) for GMM and RWM, KRAUSS, RPGM, Bezier curves Mobility (BCM), Semi-random circular movement (SCRM). SUMO/MOVE was used to create the movement of the vehicle in VANET. VANET can also use Vanet Mobisim for mobility models such as IDM (Intelligent Driver Model), IDM-LC(Intelligent Driver Model with lane changes), IDM-IM (intelligent driver model with intersection management), and FTM (fluid traffic model) [22]. We have observed that RWPM was the most used mobility model in our review. SUMO/MOVE was most used to create the movement of the vehicle in VANET.

The use of one mobility model is useful to show the node's movement in a particular scenario for performance analysis, but which mobility model is the right one to choose for the performance analysis?. We can observe that an inclination is toward using only RWPM but using it alone will not guarantee a good result; it can give unreliable results, as shown in [23]. Comparing our result with the work in [24], RWPM is still the researcher's preferred mobility model. Another observation is that 49.7% of the paper did not mention the mobility models used. No mentioning the mobility model will lead to a work that cannot be repeatable or compared with others' works.

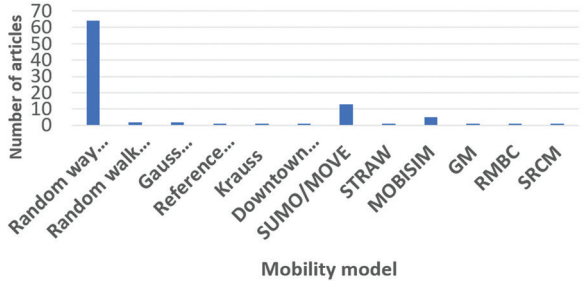



**Fig. 7.** Mobility model used by authors

### 5.4. METRICS

The metrics are essential for performance analysis. It gives you a big picture of the network performance in different ways. The choice of it is crucial, and our review shows in Figure 8 that 88 (52.1%) out of 169 articles used Packet Delivery Ratio, 82 (48.5%) End to End delay, 62 (36.7%) Throughput, 45 (26.6%) Overhead,36 (21.3%) Energy consumption, 21 (12.4%) Packet loss and Network Lifetime, 13 (7.7%) Network lifetime, 7 (4.1%) on Jitter and 5 (3%) Latency. It can be observed that PDR, E2ED, Throughput, Overhead was the most chosen metric in our review. These metrics give the network's general performance, but adding other specific metrics can give the researcher more details for a particular aspect of the network's performance. Figure 9 shows the metric choice base on the field of study. In MANET 57 out of 90 (63.3%) articles used PDR, 55 (61.1%) E2ED, 41 (45.5%)

Throughput, 31 (34.4%) Overhead, 17 (18.8%) Total Energy consumption and Packet Loss, 6 (6.6%). In VANET, 19 out of 32 (59.3%) articles used PDR, 14 (43.7%) E2ED, 8 (25%) Throughput, and Overhead. In WSN, 11 out of 37 (29.7%) articles used PDR, 13 (35.1%) E2ED, 12 (32.4%) Throughput, 5 (13.5%) Overhead, 18 (48.6%) Total energy consumption, and 12 (32.4%) Network Lifetime. Form these results we can observe that Energy consumption was most used in WSN and MANET. Other choices of metrics that were less used can be found in Table 4.

In conclusion, challenges are there in terms of what metrics to choose for the performance analysis, as the trend is on using the old metrics (PDR, E2ED, Throughput, Overhead). Our comment is to use for a general view of the performance PDR, E2ED, Throughput, Overhead metrics. For detailed information on a particular aspect of a protocol's performance analysis, specific metrics can be used.

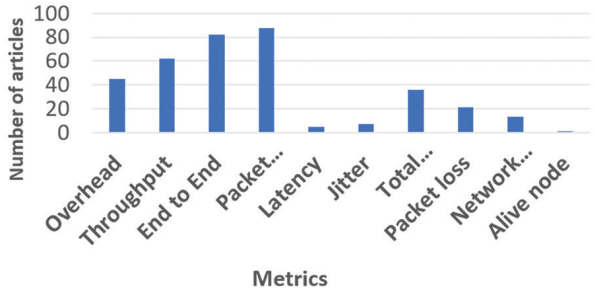



**Fig. 8.** Metric used by authors

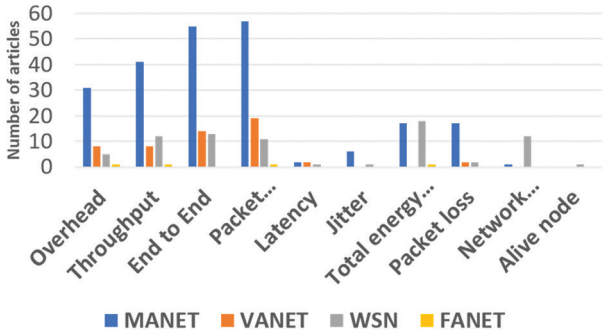



**Fig. 9.** Metric choice base on based on MANET,VANET, WSN, FANET

*B. What are the lessons learned and best practices in the performance analysis of ad hoc networks?*

In section 5, we have identified some challenges the researcher faces in choosing a simulator tool, a protocol to compare with or enhance, metrics, and the mobility model. The performance analysis of a protocol depends on these four parameters, which are essential for improving the result and future research in the ad hoc network. We have seen a significant improvement in ad hoc networks over the five years. New protocols have been created, enhancement of existing ones, and performance comparison with the existing ones. New metrics were created and apply in different scenarios with different mobility models. Nevertheless, there is room for improvement.

1. Best practices

Here we propose the best practices the researcher can use as a guide in terms of the simulator's choice, the protocol for comparison, mobility model, and metrics. The guide will help researchers and designers of software to improve the research quality in the Ad Hoc Networks.

- We recommended NS2 due to the availability of multiple models. Many protocols are implemented in it. The source code is available for free; the documentation and example can be found on the NS website and other pages (big user group) even though it has not been supported since 2010. Apart from NS2, which is our first choice, we also recommend NS3, MATLAB, and OMNET++ because these tools are activity updates and well documented.
- For selecting a specific protocol to compare with or to use for the enhancement, we recommend using the protocol that is closer in all aspects, e.g., security protocol with a security protocol. For example, the protocol should be recent, four to five years ago, for a better comparison. The same criteria should also be followed for the enhancement of a particular protocol.
- A selection of one mobility model can be accepted but not encourageable because it will not give the node movement's overall result. We encourage to use 2 or 4 mobility models for an excellent study of the performance analyses of a protocol. The best should be to use more mobility models to help other researchers see the protocol's strengths and weaknesses in different mobility models used.
- For a general view of a specific protocol's performance, we recommend using PDR, E2ED, Overhead, and Throughput, but for an in-depth analysis, the use of other metrics will be the right choice. Our recommendation is to add to the general metrics mention earlier one or two metrics; the best choice should be to use only specific metrics.

2. Recommendation

Here is our recommendation to the developer of simulator tools and the research community in ad hoc networks

**Shared source code:** The availability of code is essential for enhancing future research in ad hoc networks, but most published articles do not contain the source code and not even a link to a shared open free source code website. It is difficult for the researcher to compare the new protocol with the existing one or enhance it. We recommend the publisher to ask the author who voluntarily wants to share their source code with other researchers.

**Table 4.** Other selective metrics used

| Ref | Metric | Ref | Metric | Ref | Metric |
|---|---|---|---|---|---|
| [46] | Received message | [58, 59] | Route discovery | [60] | Average SIP |
| [46] | Alive sensor | [61] | Distributed key generation | [62] | Average travel time |
| [9, 63-65] | Accuracy | [59, 65] | Reachability | [66] | Decoded frame rate |
| [67] | Hello message | [68] | Data dissemination | [69] | Link expiration time |
| [70-72] | Bandwidth | [68] | Frequency of event | [53] | Broadcast problem |
| [73] | Variance | [74] | Average number of route broken | [75] | Connectivity rate |
| [73] | Traffic load | [76] | Wormhole Detection Ratio | [77] | Node trust value |
| [3] | Average node weight | [9] | Location accuracy | [78] | FIFO packets dropped |
| [3] | Number of accusations | [54] | Selfish nodes | [17, 79, 80] | Route failure |

**Benchmarks and acceptable scenario parameter**: The result's creditability should be compared to a benchmark to prove that the result is acceptable, but Ad-Hoc Networks does not have a benchmark. The majority of authors always compare their results with other author's results. Parameters like the number of nodes, topology size, speed, and packet size are chosen arbitrarily by the researcher because there are no standard parameters acceptable to use. The lack of standard parameters can make the result doubtful. We recommend to the community to work on making standard benchmarks and scenarios parameter for ad hoc networks.

**Simulation tools and Documentation**: Simulation tools should contain more examples for new users to adapt to them. In our observation, the lack of more examples makes it difficult for the researcher to choose new simulation tools. New researchers preferred to use simulation tools that have more users so that it is very easy to find existing code and get advice from others researcher. New simulation tools are the most favourable tools the researcher must use because they have new features, but it is time-consuming for new researchers to be familiarized with them. A well-documented simulation tool with more examples and has helpful community can facilitate the use of new tools

3. Limitation of the study

This review used a selected journal in some digital libraries to analyze the performance of ad hoc networks. We used a limited search string to retrieve the articles in those selected journals, and articles published before 2015 were not included in the survey.

## 6. CONCLUSION

The performance analysis depends on several parameters included in the simulation tools; those parameters will influence the result if not well chosen. This survey analyzes articles based on these parameters; protocol, mobility models, metrics, and simulation tools. The statistic shows an inclination to used Ad-hoc On-Demand Distance Vector routing (AODV) for performance comparison and the researcher's enhancement. Network simulation 2 (NS2) was the most selected tool, but we observe a decline in its utilization in recent years. Random Waypoint Mobility model (RWPM) was the most used mobility model. We have found a high percentage of the published article did not mention the mobility models use; this will make the result difficult for performance comparison with other works. Packet Delivery Ratio (PDR), End to End Delay (E2ED) were the most used metrics. The survey explains some lessons learned in the study and proposes best practices and recommendations to the researcher and Ah Hoc Community.

## 7. REFERENCES


[1] M. R. Belgaum, S. Musa, M. MohdSu'ud, M. Alam, S. Soomro, Z. Alansari, "Secured Approach towards Reactive Routing Protocols Using Triple Factor in Mobile Ad Hoc Networks", Annals of Emerging Technologies in Computing, Vol. 3, 2019, pp. 32-40.

[2] M. R. Belgaum, S. Soomro, Z. Alansari, M. Alam, "Ideal Node Enquiry Search Algorithm (INESH) in MANETS", in Annals of Emerging Technologies in Computing Vol. 1, 2017, pp. 26-33.

[3] M. Masdari, M. Bidaki, F. Naghiloo, "Comprehensive Evaluation of the Localized Certificate Revocation in Mobile Ad Hoc Network", Wireless Personal Communications, Vol. 94, 2017, pp. 977-1001.

[4] G. M. Borkar, A. R. Mahajan, "A secure and trust based on-demand multipath routing scheme for self-organized mobile ad-hoc networks", Wireless Networks, Vol. 23, 2017, pp. 2455-2472.

[5] M. P. Arthur and K. Kannan, "Cross-layer based multiclass intrusion detection system for secure multicast communication of MANET in military networks", Wireless Networks, Vol. 22, 2016, pp. 1035-1059.

[6] N. Abbani, H. Artail, "Protecting data flow anonymity in mobile ad hoc networks that employ cooperative caching", Ad Hoc Networks, Vol. 26, 2015, pp. 69-87.

[7] S. Tan, X. Li, Q. Dong, "Trust based routing mechanism for securing OSLR-based MANET", Ad Hoc Networks, Vol. 30, 2015, pp. 84-98.

[8] P. Kaur, D. Kaur, R. Mahajan, "Simulation based comparative study of routing protocols under wormhole attack in manet", Wireless Personal Communications, Vol. 96, No. 1, 2017, pp. 47-63.

[9] T. Karthikeyan, V. Brindha, P. Manimegalai, "Investigation on Maximizing Packet Delivery Rate in WSN Using Cluster Approach", Wireless Personal Communications, Vol. 103, 2018, pp. 3025-3039.

[10] T. Issariyakul, E. Hossain, "Introduction to network simulator 2 (NS2)", in Introduction to network simulator NS2, Springer, 2009, pp. 1-18.

[11] G. F. Riley, T. R. Henderson, "The ns-3 network simulator", in Modeling and tools for network simulation, Springer, 2010, pp. 15-34.

[12] A. Varga, R. Hornig, "An overview of the OMNeT++ simulation environment", Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops, Marseille, France, 3-7 March 2008, pp. 1-10.

[13] Z. Lu, H. Yang, Unlocking the power of OPNET modeler. Cambridge University Press, 2012.

[14] X. Zeng, R. Bagrodia, M. Gerla, "GloMoSim: a library for parallel simulation of large-scale wireless networks", Proceedings. 12th Workshop on Parallel and Distributed Simulation, Banff, AB, Canada, 29-29 May 1998, pp. 154-161.

[15] B. Hahn, D. Valentine, Essential MATLAB for engineers and scientists. Academic Press, 2016.

[16] A. Benzerbadj, B. Kechar, A. Bounceur, B. Pottier, "Cross-Layer Greedy position-based routing for multihop wireless sensor networks in a real environment", Ad Hoc Networks, Vol. 71, 2018, pp. 135-146.

[17] G. Singal, V. Laxmi, M. S. Gaur, V. Rao, "Moralism: mobility prediction with link stability based multicast routing protocol in MANETs", Wireless Networks, Vol. 23, 2017, pp. 663-679.

[18] F. Bai, A. Helmy, "A survey of mobility models", Wireless Adhoc Networks, Vol. 206, 2004, p. 147.

[19] M. Oche, A. B. Tambuwal, C. Chemebe, R. M. Noor, S. Distefano, "VANETs QoS-based routing protocols based on multi-constrained ability to support ITS infotainment services", Wireless Networks, Vol. 26, No. 3, 2020, pp. 1685-1715.

[20] Y. Hernafi, M. B. Ahmed, M. Bouhorma, "ACO and PSO algorithms for developing a new communication model for VANET applications in smart cities", Wireless Personal Communications, Vol. 96, No. 2, 2017, pp. 2039-2075.

[21] T. Li, J. Ma, C. Sun, "SRDPV: secure route discovery and privacy-preserving verification in MANETs", Wireless Networks, Vol. 25, No. 4, 2019, pp. 1731-1747.

[22] I. Zaimi, Z. S. Houssaini, A. Boushaba, M. Oumsis, D. Aboutajdine, "An evaluation of routing protocols for vehicular ad-hoc network considering the video stream", Wireless Personal Communications, Vol. 98, No. 1, 2018, pp. 945-981.

[23] J. Yoon, M. Liu, B. Noble, "Random waypoint considered harmful", Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, San Francisco, CA, USA, 30 March-3 April 2003, pp. 1312-1321.

[24] S. Kurkowski, T. Camp, M. Colagrosso, "MANET simulation studies: the incredibles", ACM SIGMOBILE Mobile Computing and Communications Review, Vol. 9, No. 4, 2005, pp. 50-61.

[25] D. Hiranandani, K. Obraczka, J. Garcia-Luna-Aceves, "MANET protocol simulations considered harmful: the case for benchmarking", IEEE Wireless Communications, Vol. 20, No. 4, 2013, pp. 82-90.

[26] S. Naicken, B. Livingston, A. Basu, S. Rodhetbhai, I. Wakeman, D. Chalmers, "The state of peer-to-peer simulators and simulations", ACM SIGCOMM Computer Communication Review, Vol. 37, No. 2, 2007, pp. 95-98.

[27] S. Kurkowski, W. Navidi, T. Camp, "Constructing manet simulation scenarios that meet standards", Proceedings of the IEEE International Conference on Mobile Adhoc and Sensor Systems, Pisa, Italy, 8-11 October 2007, pp. 1-9.

[28] A. Munjal, T. Camp, W. C. Navidi, "Constructing rigorous MANET simulation scenarios with realistic

mobility", Proceedings of the European Wireless Conference, Lucca, Italy, 12-15 April 2010, pp. 817-824.

[29] T. R. Andel, A. Yasinsac, "On the credibility of manet simulations", Computer, Vol. 39, no. 7, 2006, pp. 48-54.

[30] I. Ahmad, U. Ashraf, A. Ghafoor, "A comparative QoS survey of mobile ad hoc network routing protocols", Journal of the Chinese institute of engineers, Vol. 39, No. 5, 2016, pp. 585-592.

[31] J. Sánchez-García, J. García-Campos, M. Arzamendia, D. G. Reina, S. Toral, D. Gregor, "A survey on unmanned aerial and aquatic vehicle multi-hop networks: Wireless communications, evaluation tools and applications", Computer Communications, Vol. 119, 2018, pp. 43-65.

[32] J. M. García-Campos, J. Sánchez-García, D. Reina, S. Toral, F. Barrero, "An evaluation methodology for reliable simulation based studies of routing protocols in VANETs", Simulation modelling practice and theory, Vol. 66, 2016, pp. 139-165.

[33] P. P. Garrido, M. P. Malumbres, C. T. Calafate, "ns-2 vs. OPNET: a comparative study of the IEEE 802.11 e technology on MANET environments", Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops, Marseille, France, 3-7 March 2008: Citeseer, pp. 1-10.

[34] E. Schoch, M. Feiri, F. Kargl, M. Weber, "Simulation of ad hoc networks: ns-2 compared to jist/swans", Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops, Marseille, France, 3-7 March 2008, pp. 1-8.

[35] F. Kargl, E. Schoch, "Simulation of MANETs: a qualitative comparison between JiST/SWANS and ns-2", Proceedings of the 1st international workshop on System evaluation for mobile platforms, San Juan, Puerto Rico, 11 June 2007, pp. 41-46.

[36] E. Weingartner, H. Vom Lehn, K. Wehrle, "A performance comparison of recent network simulators", Proceedings of the IEEE International Conference on Communications, Dresden, Germany, 14-18 June 2009, pp. 1-5.

[37] L. Begg, W. Liu, K. Pawlikowski, S. Perera, H. Sirisena, "Survey of simulators of next generation networks for studying service availability and resilience", Technical report, University of Canterbury, 2006.

[38] T. M. T. Pham, T. T. Nguyen, D. S. Kim, "Geographical awareness hybrid routing protocol in Mobile Ad Hoc Networks", Wireless Networks, Vol. 23, 2017, pp. 1-13.

[39] Y. H. Xu, Y. Wu, J. Song, "Joint Channel Assignment and Routing Protocol for Cognitive Radio Wireless Sensor Networks", Wireless Personal Communications, Vol. 97, 2017, pp. 41-62.

[40] N. Ismat, R. Qureshi, S. Mumtaz ul Imam, "Adaptive Power Control Scheme for Mobile Wireless Sensor Networks", Wireless Personal Communications, Vol. 106, 2019, pp. 2195-2210.

[41] Y. H. Robinson, R. S. Krishnan, E. G. Julie, R. Kumar, L. H. Son, P. H. Thong, "Neighbor Knowledge-based Rebroadcast algorithm for minimizing the routing overhead in Mobile Ad-hoc Networks", Ad Hoc Networks, Vol. 93, 2019, p. 101896.

[42] J. Sathiamoorthy, B. Ramakrishnan, M. Usha, "STF-DR: Architecture of Competent Protocol for Efficient Route Discovery and Reliable Transmission in CEAACK MANETs", Wireless Personal Communications, Vol. 97, 2017, pp. 5817-5839.

[43] S. Vidhya, T. Sasilatha, "Secure Data Transfer Using Multi Layer Security Protocol with Energy Power Consumption AODV in Wireless Sensor Networks", Wireless Personal Communications, Vol. 103, 2018, pp. 3055-3077.

[44] G. Dhand, S. S. Tyagi, "SMEER: Secure Multi-tier Energy Efficient Routing Protocol for Hierarchical Wireless Sensor Networks", Wireless Personal Communications, Vol. 105, 2019, pp. 17-35.

[45] N. Kaur, R. Singhai, "Analysis of Traffic Impact on Proposed Congestion Control Scheme in AODV", Wireless Personal Communications, Vol. 109, 2019, pp. 1395-1418.

[46] H. Barati, A. Movaghar, A. M. Rahmani, "EACHP: Energy Aware Clustering Hierarchy Protocol for Large Scale Wireless Sensor Networks", Wireless Personal Communications, Vol. 85, 2015, pp. 765-789.

[47] V. K. Verma, S. Singh, N. P. Pathak, "Optimized Battery Models Observations for Static, Distance Vector and On-Demand Based Routing Protocols Over 802.11 Enabled Wireless Sensor Networks", Wireless Personal Communications, Vol. 81, 2015, pp. 503-517.

[48] K. K. Rana, S. Tripathi, R. S. Raw, "Analytical analysis of improved directional location added routing protocol for VANETS", Wireless Personal Communications, Vol. 98, 2018, pp. 2403-2426.

[49] M. Rajesh Babu, G. Usha, "A Novel Honeypot Based Detection and Isolation Approach (NHBADI) To Detect and Isolate Black Hole Attacks in MANET", Wireless Personal Communications, Vol. 90, 2016, pp. 831-845.

[50] S. Jayaraman, R. Bhagavathiperumal, U. Mohanakrishnan, "A Three Layered Peer-to-Peer Energy Efficient Protocol for Reliable and Secure Data Transmission in EAACK MANETs", Wireless Personal Communications, Vol. 102, 2018, pp. 201-227.

[51] K. Haseeb, K. A. Bakar, A. Ahmed, T. Darwish, I. Ahmed, "WECRR: Weighted Energy-Efficient Clustering with Robust Routing for Wireless Sensor Networks", Wireless Personal Communications, Vol. 97, 2017, pp. 695-721.

[52] X. Anita, M. A. Bhagyaveni, J. Martin Leo Manickam, "Collaborative Lightweight Trust Management Scheme for Wireless Sensor Networks", Wireless Personal Communications, Vol. 80, 2015, pp. 117-140.

[53] M. U. Hassan, M. H. Rehmani, Y. Faheem, "Performance evaluation of broadcasting strategies in cognitive radio networks", Wireless Networks, Vol. 25, 2019, pp. 999-1016.

[54] S. A. Thorat, P. J. Kulkarni, "Opportunistic Routing in Presence of Selfish Nodes for MANET", Wireless Personal Communications, Vol. 82, 2015, pp. 689-708.

[55] T. M. Rajeh, A. I. Saleh, L. M. Labib, "A New Cooperative Balancing Routing (CBR) Protocol to Enhance the Lifetime of Wireless Sensor Networks", Wireless Personal Communications, Vol. 98, 2018, pp. 2623-2656.

[56] V. V. Mandhare, R. R. Manthalkar, V. R. Thool, "Novel Approach for Cache Update on Multipath DSR Protocol in MANET for QoS Support", Wireless Personal Communications, Vol. 98, 2018, pp. 505-519.

[57] S. K. Bhoi, P. M. Khilar, "Self soft fault detection based routing protocol for vehicular ad hoc network in city environment", Wireless Networks, Vol. 22, 2016, pp. 285-305.

[58] S. Tabatabaei, M. Teshnehlab, S. J. Mirabedini, "Fuzzy-Based Routing Protocol to Increase Throughput in Mobile Ad Hoc Networks", Wireless Personal Communications, Vol. 84, 2015, pp. 2307-2325.

[59] S. S. Basurra, M. De Vos, J. Padget, Y. Ji, T. Lewis, S. Armour, "Energy efficient zone based routing protocol for MANETs", Ad Hoc Networks, Vol. 25, 2015, pp. 16-37.

[60] F. Alshahwan, M. Alshamrani, A. A. Amer, "Dynamic Novel Cross-Layer Performance Enhancement Approach for SIP over OLSR", IEEE Access, Vol. 6, 2018, pp. 71947-71964.

[61] H. Kojima, N. Yanai, J. P. Cruz, "ISDSR+: Improving the Security and Availability of Secure Routing Protocol", IEEE Access, Vol. 7, 2019, pp. 74849-74868.

[62] J. S. Pan, I. S. Popa, C. Borcea, "DIVERT: A distributed vehicular traffic re-routing system for congestion avoidance", IEEE Transactions on Mobile Computing, Vol. 16, 2017, pp. 58-72.

[63] K. Raja, A. Deivasigamani, V. Ravi, "A Reliant Certificate Revocation of Malicious Nodes in MANETs", Wireless Personal Communications, Vol. 90, 2016, pp. 435-455.

[64] A. Pal, P. Dutta, A. Chakrabarti, J. P. Singh, S. Sadhu, "Biogeographic-Based Temporal Prediction of Link Stability in Mobile Ad Hoc Networks", Wireless Personal Communications, Vol. 104, 2019, pp. 217-233.

[65] A. E. Hilal, A. B. MacKenzie, "A distributed coalition game model for cooperation in MANETs", Ad Hoc Networks, Vol. 85, 2019, pp. 46-59.

[66] S. González, W. Castellanos, P. Guzmán, P. Arce, J. C. Guerri, "Simulation and experimental testbed for adaptive video streaming in ad hoc networks", Ad Hoc Networks, Vol. 52, 2016, pp. 89-105.

[67] D. S. Sakkari, T. G. Basavaraju, "GCCT: A Graph-Based Coverage and Connectivity Technique for Enhanced Quality of Service in WSN", Wireless Personal Communications, Vol. 85, 2015, pp. 1295-1315.

[68] M. Guerroumi, A. S. K. Pathan, "Hybrid data dissemination protocol (HDDP) for wireless sensor networks", Wireless Networks, Vol. 24, 2018, pp. 1739-1754.

[69] A. Naushad, G. Abbas, Z. H. Abbas, A. Pagourtzis, "Novel strategies for path stability estimation under topology change using Hello messaging in MANETs", Ad Hoc Networks, Vol. 87, 2019, pp. 76-99.

[70] S. R. Malwe, N. Taneja, G. P. Biswas, "Enhancement of DSR and AODV Protocols Using Link Availability Prediction", Wireless Personal Communications, Vol. 97, 2017, pp. 4451-4466.

[71] M. Malathi, S. Jayashri, "Modified Bi-directional Routing with Best Afford Path (MBRBAP) for Routing Optimization in MANET", Wireless Personal Communications, Vol. 90, 2016, pp. 861-873.

[72] J. K. Jayabarathan, A. Sivanantharaja, S. Robinson, "Quality of Service Enhancement of Mobile Adhoc Networks Using Priority Aware Mechanism in AODV Protocol", Wireless Personal Communications, Vol. 96, 2017, pp. 5897-5909.

[73] Y. Qin, L. Li, X. Zhong, Y. Yang, C. L. Gwee, "A Cross-Layer QoS Design with Energy and Traffic Balance Aware for Different Types of Traffic in MANETs", Wireless Personal Communications, Vol. 85, 2015, pp. 1429-1449.

[74] C. Lal, V. Laxmi, M. S. Gaur, M. Conti, "Enhancing QoE for video streaming in MANETs via multi-constraint routing", Wireless Networks, Vol. 24, 2018, pp. 235-256.

[75] W. Wang, J. Wang, M. Wang, B. Wang, W. Zhang, "A realistic mobility model with irregular obstacle constraints for mobile ad hoc networks", Wireless Networks, Vol. 25, 2019, pp. 487-506.

[76] T. T. Vo, N. T. Luong, D. Hoang, "MLAMAN: a novel multi-level authentication model and protocol for preventing wormhole attack in mobile ad hoc network", Wireless Networks, Vol. 0123456789, 2018.

[77] A. M. Shabut, K. P. Dahal, S. K. Bista, I. U. Awan, "Recommendation Based Trust Model with an Effective Defence Scheme for MANETs", IEEE Transactions on Mobile Computing, Vol. 14, 2015, pp. 2101-2115.

[78] W. A. Jabbar, W. K. Saad, M. Ismail, "MEQSA-OLSRv2: A multicriteria-based hybrid multipath protocol for energy-efficient and QoS-aware data routing in MANET-WSN convergence scenarios of IoT", IEEE Access, Vol. 6, 2018, pp. 76546-76572.

[79] O. S. Gnana Prakasi, P. Varalakshmi, "Decision Tree Based Routing Protocol (DTRP) for Reliable Path in MANET", Wireless Personal Communications, Vol. 109, 2019, pp. 257-270.

[80] L. Sayad, L. Bouallouche-Medjkoune, D. Aissani, "IWDRP: An Intelligent Water Drops Inspired Routing Protocol for Mobile Ad Hoc Networks", Wireless Personal Communications, Vol. 94, 2017,pp. 2561-2581.

# Deep Learning Approach for cognitive competency assessment in Computer Programming subject

Review Paper

**Shahidatul Arfah Baharudin**
Malaysian Institute of Information Technology,
Universiti Kuala Lumpur,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur, Malaysia
shahidatularfah@unikl.edu.my

**Adidah Lajis**
Malaysian Institute of Information Technology,
Universiti Kuala Lumpur,
1016, Jalan Sultan Ismail, 50250 Kuala Lumpur, Malaysia
adidahl@unikl.edu.my

***Abstract*** *– This research examines the competencies that are essential for an lecturer or instructor to evaluate the student based on automated assessments. The competencies are the skills, knowledge, abilities and behavior that are required to perform the task given, whether in a learning or a working environment. The significance of this research is that it will assist students who are having difficulty learning a Computer Programming Language course to identify their flaws using a Deep Learning Approach. As a result, higher education institutions have a problem with assessing students based on their competency level because; they still use manual assessment to mark the assessment. In order to measure intelligence, it is necessary to identify the cluster of abilities or skills of the type in which intelligence expresses itself. This grouping of skills and abilities referred to as "competency". Then, an automated assessment is a problem-solving activity in which the student and the computer interact with no other human intervention. This review focuses on collecting different techniques that have been used. In addition, the review finding shows the main gap that exists within the context of the studied areas, which contributes to our key research topic of interest.*

***Keywords****: Cognitive competency, deep learning, automated assessment, Bloom's Taxonomy, computer programming*

## 1. INTRODUCTION

Assessment is a core and critical requirement in an educational system since it contributes to the great extent affects of students' learning [1][2]. Therefore, deep learning approach is applies to enrich and enhance the lecturer assessment. A good assessment is where the lecturers are well understanding on the assessment principles includes assessment terminology, development and use of assessment methodologies and techniques, assessment quality standards and any alternative to traditional measurements of learning. Because of that, it does require integration in assessment practices, theories, philosophies to support teaching and learning in education.

The lecturer is also served as facilitator, a mentor or a coach to guide the process of students' learning. The student is responsible for his or her own learning. The students should also get trainings for developing various competencies, including cognitive, meta-cognitive, social and affective competencies, for the success of their future. Therefore, with deep learning, the learning process is internally motivated and is associated with an intention to understand, rather than to simply pass an assessment task. In this connection, assessment has been identified as a powerful aid to engage students into a more in-depth learning process and transform them into reflective practitioners.

By applying deep learning in education, the research on pedagogies of assessment education can be enriched. Moreover, deep learning can be applied either in face-to-face teaching or online learning or blended learning itself. Therefore, the structure of the paper is as follows: Section II describes the main problem in computer programming subject . Section III discuss on Competency-Based Education and Bloom Taxonomy. Section IV discuss on review of cognitive competency assessment techniques. Section V, explains about the deep learning in the assessment of cognitive competency. Section VI, discuss on to conclude the paper's finding.

## 2. THE MAIN PROBLEM IN COMPUTER PROGRAMMING SUBJECT

In 2019, Bennedsen and Caspersen [4] has conducted a worldwide survey for the research on failure rate for programming course. The total respondent during 2019 is 170 response to their survey. Fig. 1 shows the number of respondents per continent is Africa, Asia, Australia, Europe , North America and South America.

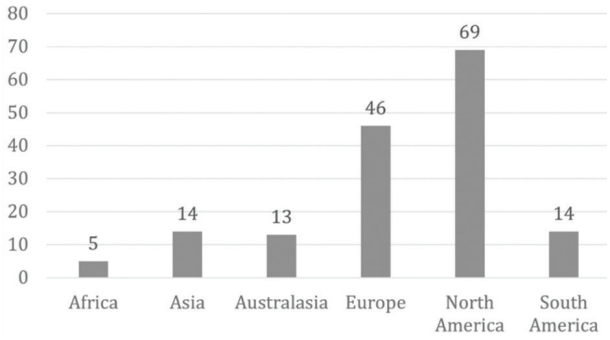



**Fig. 1.** Number of respondents per continent

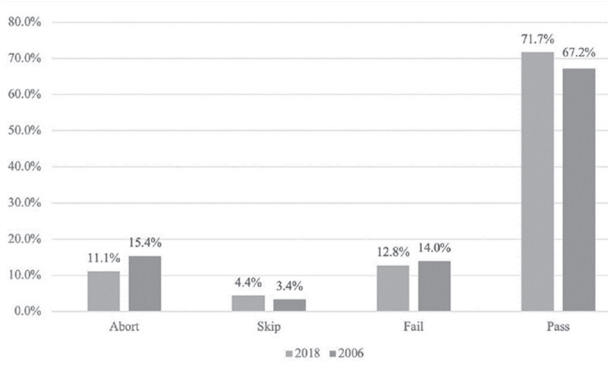



**Fig. 2.** Pass, Fail, Abort and Skip Rate: aggregate

Based on Fig. 2, each of the respondents were asked to give a number for abort which is the number of student aborting the course before final exam, skip is the number of students not showing up for the final exam but was allowed to, fail is the number of students who failed the course and pass is the number of students who passed the course. As a results , they found that the average failure is 28% based on Fig. 2. The main reason they were uninterested in programming. Students believe that learning to program is difficult. They struggle to understand the program code and write a simple program. This research is also agreed by other researchers, while the student also already has a comfort level, difficulties in understanding the course content, time management issues and expectations and perfections of not getting enough help from their lecturer [5][6]. Other than that, the student also demotivated the student to learn programming [7]. This reason also agrees with Nurul Farahin et al. where the major problem in computer programming is lack of problem-solving skills, no prior knowledge, low motivation, poor mathematical knowledge, peer influence and lack of future expectation [8].

## 3. COMPETENCY-BASED EDUCATION AND BLOOM'S TAXONOMY

Computer-Based Education (CBE) is one of the concept to reflect human competency motor, intellectual and emotional competency. Using CBE, it can measure learning progress by the student. CBE also is the smaller concept of outcome-based leaning (OBE). The competence student is those who can and want to interact effectively three kinds of environment presented by the socially ascribed, self-selected and self-developed roles [9].

OBE is defined as an education system that focuses on learning outcomes rather than educational curriculum content. Learning outcomes, for example, quantified in terms of information, abilities, attitudes gained during the learning process. [10]. It covers three learning domains, which are the Psychomotor, Cognitive and Affective domains. They have implemented these three learning domains in various ways.

Cognitive domain is the one where the student's cognitive activities are structured. Starting with the knowledge level and ending with the evaluation level of Bloom's Taxonomy [11]. There also some evidence that cognitive training able to improve cognitive function, which potentially slow cognitive decline and able to help the student. Cognitive domain deals with how a student acquires processes and utilizes the knowledge. For Affective domain, it is focused on attitude, motivation, willingness to take part, valuing what is being learn and discipline values into real life. The last one, psychomotor domain focuses on performing sequences of motor activities to a specific level of accuracy, smoothness, rapid or force. Underlying the motor activities is cognitive understanding [12]. Evidence of outcome is required to fulfill the shortage of the soft skill of an employee in the workplace [13].

Bloom's taxonomy of Educational Objectives is a classification system by an educational psychologist Benjamin Bloom who creates in year 1956. It focuses on developing thinking ability, which involves simple information acquisition to a more complex process [14]. Bloom's taxonomy contains six categories of cognitive skills ranging from lower-order skills that require less cognitive processing to higher-order skills that require deeper leaning and a greater degree of cognitive processing. Though in year 2001, the Bloom's Taxonomy has been revised [15]. Refer to Fig. 3. The differentiations into categories of higher-order and lower-order skills arose later.

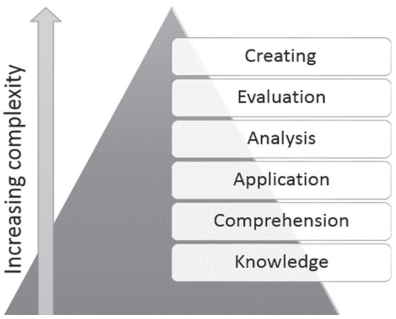



**Fig. 3.** Revised Bloom's Taxonomy

For the older version of Bloom's Taxonomy, it consists Knowledge, Comprehension, Application, Analyzing, Evaluation and Creating. *Knowledge* is the foundational cognitive skill and refers to the retention of specific, discrete pieces of information like facts and definitions or methodology for example as the sequence of events in a step-by-step process. *Comprehension* is the meaning of the information that they encounter by paraphrasing it in their own words, classifying items in groups, comparing and contrasting items with other similar entities of explaining a principle to others. For comprehension, it is requires more cognitive processing than simply remembering information and learning objectives that address comprehension will help learners to incorporate knowledge into their existing cognitive schemas by which they understand the world [16]. Based on comprehension, it will allow student to learn how to use the knowledge, skills and techniques in a new situation via *application*, which is the third level of Bloom's taxonomy. For the higher level of Bloom's taxonomy is *analysis*. Analysis can break down a material into its constituent parts in order to comprehend its organizational structure. It is also where skills such as critical thinking come into play. Distinguish between facts and opinions and identify the claims that underlie the analysis. Following the analysis, the next level is synthesis. *Synthesis* entails creating a novel product in a specific situation. Its behavior is recombine the parts created during analysis to form a new entity where is differs from the original one. Finally, *evaluation* is the pinnacle of Bloom's Taxonomy. *Evaluation* is also an important aspect of critical thinking skills. It will show the student's ability to assess the worth of a material for a specific purpose using predetermined criteria and rationale. [17]. For the revised Bloom's taxonomy, refer to Table 1.

**Table 1.** Revised Bloom's Taxonomy

| | |
|---|---|
| **Creating** | Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions. |
| **Evaluation** | Present and defend opinions by making judgments about information, validity of ideas, or quality of work based on a set of criteria. |
| **Analyzing** | Examine and break information into parts by identifying motives or causes. Make interences and find evidence to support generalizations. |
| **Applying** | Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way. |
| **Understanding** | Demonstate understanding of facts and ideas by organizing, comparng, translating, interpreting, giving desriptions, stating main ideas. |
| **Remembering** | Retrieve relevant knowledge by recalling facts, terms, basic concept and answer from long-term memory. |

In year 2019, [18] has used Bloom's Taxonomy as a scale for preparing the assessment questions, it quantified the competency level based on that. The results show that Bloom's Taxonomy is a beneficial tool for learning and assessing computer-programming subject.

## 4. 4. REVIEW OF COGNITIVE COMPETENCY ASSESSMENT TECHNIQUE

Cognitive competency defined as critical thinking and creative thinking skills which effective problem solving, decision making, learning and development [19]. These criteria are important for the student to learn Computer Programming. Therefore, cognitive competency assessment follow by guideline from the Bloom's Taxonomy in the cognitive domain using automated assessment.

Table 2 illustrates the summarized of technique cognitive competency assessment which came from the previous research.

**Table 2.** Summarized of technique cognitive competency assessment (C-Competency, NM-Not mentioned)

| Researchers | Subjects | Level / Age | Techniqu-es | Focus |
|---|---|---|---|---|
| [20] | Comp. Science | Undergrad and Schools | NLP and info. theory | C |
| [21] [22] | Introductory course in Computer Literacy | Undergrad Students | LSA | NM |
| [23] [24] [25] | Introductory data structure course | Undergrad students | Text Similarity | NM |
| [13] | Introduction to programming | Undergrad students | Assessment Framework based on Bloom's Taxonomy | C |
| [26] | C++ Programming | Undergrad Students | Semi-automated assessment | NM |
| [18] [27] | Computer Programming | Undergrad students | Rule-Based Method | C |
| [28] | Parallel Programming | Undergrads students | Code Evaluation and Debugging | NM |
| [29] | Computer Methods | Undergrads students | Computer Adaptive Testing Tools | Non-cognitive |
| [30] | Computer Programming | Undergrads students | Flexible and systematic teaching framework | C |
| [31] | Programming | Undergrads students | Online EasyHPC Tool | NM |

| Researchers | Subjects | Level / Age | Techniqu-es | Focus |
|---|---|---|---|---|
| [32] | Programming | Undergrads students | Collaborative Scenario-TASystem tools | NM |
| [33] | UML | Undergrads students | Comprehensive Approach | NM |
| [34] | Computer Programming | Undergrads students | Online Learning system | C |
| [35] | Computer Programming | Undergrads students | Integration Automated Test Data Generation and programming assessment | NM |
| [36] | Computer Programming | Undergrads students | Software Testing technique | NM |
| [37] | Programming | Undergrads students | Classical Test Theory (CTT) or Item Response Theory (IRT) - SIETTE | NM |
| [38] | Computer Programming | Undergrads students | 2TSW – testing-based approach | NM |
| [39] | Chinese Subject | Undergrads students | NLP | NM |
| [40] | Computer Programming | Undergrads students | Metacognitive Support | C |
| [41] | Object-Oriented Programming | Undergrads students And High School | Competency Structure Model -COMOOP | C |

Based on data in Table 3, it is possible to propose the cognitive competency assessment can still assist the student. Several studies have been used to identify the cognitive problem. Other than that, the research did not mention the focus on cognitive. Some of the research is using the Mobile Learning application to investigate the factors that influence student's learning performance and evaluate the effectiveness of mobile learning to use the Learn C application in programming subjects. The findings shows that a variety factors that affect the student learning. There are misunderstanding, lack of practices, poor logical thinking and problem solving [8] . However, there is no mention of them being able to determine their cognitive level. Only one research using the Bloom's Taxonomy as a benchmark to evaluate the student. Thus, the study concludes that the cognitive level of Bloom's Taxonomy as a tool for the assessing a student's competency in programming is appropriate and cable of reducing the high failure rate among student enrolled in Computer Programming subjects [13].

## 5. DEEP LEARNING IN THE ASSESSMENT OF COGNITIVE COMPETENCY

Deep learning is a class of machine learning algorithms that employ multiple layers to represent various levels of abstraction. It comprises of an input layer, an output layer and a few hidden layers. It showed this assessment in Table 3.

**Table 3**. Summarized of deep Learning in the cognitive competency assessment (C-Competency, NM-Not mentioned)

| Researchers | Subjects | Level / Age | Techniqu-es | Focus |
|---|---|---|---|---|
| [42] | Kaggle ASAP | 7th to 10th grader | LSTM classification and regression task | NM |
| [43] | Kaggle ASAP | 7th to 10th grader | LSTM- CNN-attention- based | NM |
| [44] | IT, Engineering, Management | Higher education | NLP, CNN | C |
| [45] | IT , Medical Engineering, Management | Higher education | LSTM | C |

From the observation in Table 3, Kaggle Automated Student Assessment Prize (ASAP) conducted a competition dataset. It is sponsored by the William and Flora Hewlett Foundation (Hewlett). Their variables are used to test their scoring capabilities using neural network techniques such as Long Short-Term Memory (LSTM) and Convolutional Neural Network. Dimitrios et al [46] also mention that a deep neural network is capable of using as automatic text scoring using a neural network.

In 2019, Tiliza [45] developed the rule-based Long Short Term Memory (LSTM) classification to assess higher level of cognitive competency via short text answers. This study analysed short free text assessment answers which fell under three criteria. Table 4 shows the three criteria for this study.

**Table 4.** Criteria for Rule-Based LSTM classification

| No | Criteria |
|---|---|
| 1 | The word count of the assessment answer must be in the range of 0 to 200 words each. |
| 2 | The model questions are composed of Bloom's Taxonomy higher order cognitive process dimentions. |
| 3 | The scope of assessments answers are from three academic domains namely Information Technology (IT), Medical Engineering (ME) and Management. |

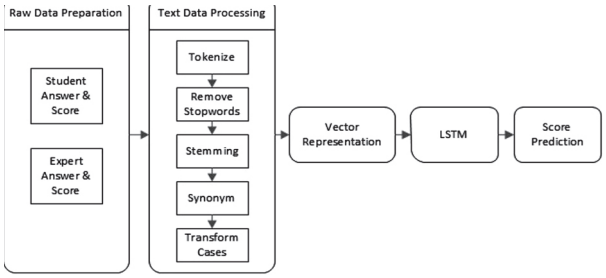



**Fig. 4.** Technique using Rule-based Long Short Term Memory (LSTM) [47]

As shown in Fig. 4, the technique uses Rule-Based Long Short Term Memory classification to assess higher-level cognitive competency via short free text answers. The results of the study are that the rule-based LSTM classification achieved a mean correlation of 0.80, 0.88, 0.85 against test materials sets, Bloom's Taxonomy levels and the academic domain, respectively. Whereas the benchmark results, Latent Semantic Analysis (LSA), show a mean correlation of 0.32, 0.332 and 039 against test material sets, Bloom's Taxonomy levels, academic domain respectively. This is one example of how the Deep learning approach can reduce student failure rates in computer programming.

## 6. CONCLUSION

In conclusion, based on the review of cognitive competency assessment techniques and deep learning in cognitive competency assessment, this research is workable to help reduce of failure rate in the subject.

As to date, there is no research in this field that applied deep learning. For now, there is only research on short free text answers [45].This study will concentrate on the Cognitive domain of Bloom's Taxonomy, where we will evaluate student's programming exercise in C language programming. This research also covers both Higher-Order Thinking Skills and Lower-Order Thinking Skills. With this guideline, we can identify the student's weakness and motivate the student to learn computer programming. By this research, the student and lecturer will able to identify the weaknesses at the early stage and this will help student to pass the subject.

## 7. REFERENCES


[1] Ministry of Education Malaysia (MoE), "Malaysia Education Blueprint 2015-2025 (Higher Education)", Vol. 2025, p. 40, 2015.

[2] C. Douce, D. Livingstone, J. Orwell, "Automatic test-based assessment of programming: a review", Journal on Educational Resources in Computing, Vol. 5, No. 3, 2005, pp. 1–13.

[3] B. C. Surve, B. R. Londhe, "Artificial Intelligence based assessment and development of student's Non-cognitive skills in Professional Education through an online Learning Management System", Proceedings of the 4th International Conference on Inventive Systems and Control, Coimbatore, India, 8-10 January 2020, pp. 329–336.

[4] J. Bennedsen, M. E. Caspersen, "Failure rates in introductory programming - 12 years later", ACM Inroads, Vol. 10, No. 2, 2019, pp. 30–35.

[5] Simon et al., "Pass rates in introductory programming and in other STEM disciplines", Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education, December 2019, pp. 53–71.

[6] R. Gao, "Reforming to improve the teaching quality of computer programming language", Proceedings of the 6th International Conference on Computer Science & Education, Singapore, 3-5 August 2011 pp. 1267–1269.

[7] D. Jaithavil, N. Kuptasthien, "An adaptive algorithm for learning computer programming course", Proceedings of the 15th Int. CDIO Conference Aarhus University Aarhus, Denmark, July 2019.

[8] N. Farahah Abdul Halim, D. Nincarean Eh Phon, "Mobile Learning Application Impact Towards Student Performance in Programming Subject", IOP Conference Series. Materials Science and Engineering, Vol. 769, No. 1, 2020.

[9] J. H. Block, "The 'C' in CBE", Educational Research, 1978.

[10] W. G. Spady, K. J. Marshall, "Beyond traditional outcome-based education", Educational Leadership, Vol. 49, No. 2, 1991, pp. 67–72.

[11] A. R. M. Zaghloul, "Assessment of lab work: A three-domain model; Cognitive, affective, psychomotor", Proceedings of the American Society for Engineering Education Annual Conference & Expositionpp, 2001, pp. 2279–2285.

[12] G. Kasilingam, M. Ramalingam, E. Chinnavan, "Assessment of learning domains to improve student's learning in higher education", Journal of Young Pharmacists, Vol. 6, No. 1, 2014, pp. 27–33.

[13] A. Lajis, H. Md Nasir, N. A. Aziz, "Proposed assessment framework based on bloom taxonomy cognitive competency: Introduction to programming", Proceedings of the 7th International Conference on Software and Computer Applications, February 2018, pp. 97–101.

[14] J. Conklin, "Book review of: "A Taxonomy for Learning , Teaching , Assessing : A Revision of Bloom' s Taxonomy of Educational Objectives", Vol. 18345, 2005, pp. 22–25.

[15] L. W. Anderson et al., Taxonomy for_ Assessing a Revision 0F Bl00M'S Tax0N0My 0F Educati0Nal Objectives. 2001.

[16] R. E. Mayer, "A taxonomy for computer-based assessment of problem solving", Computers in Human Behavior, Vol. 18, No. 6, 2002, pp. 623–632.

[17] R. Tong, B. P. Lim, N. F. Chen, B. Ma, H. Li, "Subspace Gaussian mixture model for computer-assisted language learning", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, Italy, 4-9 May 2014, pp. 5347–5351.

[18] Z. Ullah, A. Lajis, M. Jamjoom, A. Altalhi, F. Saleem, "Bloom's taxonomy: A beneficial tool for learning and assessing students' competency levels in computer programming using empirical analysis", Computer Applications in Engineering Education, 2020, pp. 1–13.

[19] R. C. F. Sun, E. K. P. Hui, "Cognitive competence as a positive youth development construct: A conceptual review", The Scientific World Journal, Vol. 2012, 2012, pp. 21–23.

[20] A. Lajis, N. A. Aziz, "NL scoring and bloom competency test: An experimental result", Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication, February 2012, pp. 1–5.

[21] B. P. Mitchell Tom, Russel Terry, "Towards robust computerised marking of free-text responses", 2002.

[22] P. Wiemer-Hastings, K. Wiemer-Hastings, A. C. Graesser, "Improving an intelligent tutor's comprehension of students with Latent Semantic Analysis", Proceedings of Artificial Intelligence Education, 1999, pp. 535–542.

[23] F. A. G. Wael, "Short Answer Grading Using String Similarity And Corpus-Based Similarity", International Journal of Advanced Computer Science and Applications, Vol. 3, No. 11, 2012.

[24] M. Mohler, R. Bunescu, R. Mihalcea, "Learning to grade short answer questions using semantic similarity measures and dependency graph alignments", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Vol. 1, 2011, pp. 752–762.

[25] M. Mohler, R. Mihalcea, "Text-to-text semantic similarity for automatic short answer grading", Proceedings of the 12th Conference of the European Chapter of the ACL, Athens, Greece, March 2009, pp. 567–575.

[26] S. Buyrukoglu, F. Batmaz, R. Lock, "A new marking technique in semi-Automated assessment", Proceedings of the 12th International Conference on Computer Science and Education, Houston, TX, USA, 22-25 August 2017 pp. 545–550.

[27] Z. Ullah, A. Lajis, M. Jamjoom, A. H. Altalhi, J. Shah, F. Saleem, "A rule-based method for cognitive competency assessment in computer programming using bloom's taxonomy", IEEE Access, 2019.

[28] Y. Zhang, J. Li, D. Wu, Y. Du, "Improving Student Skills on Parallel Programming via Code Evaluation and Feedback Debugging", Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering, Wollongong, NSW, Australia, 4-7 December 2018, pp. 1069–1073.

[29] P. R. G. M. S. P. Molins-Ruano, C. González-Sacristán, F. Díez, "Adaptive Model for Computer-Assisted Assessment in Programming Skills", International Journal of Engineering Education, 2000.

[30] R. Gacitua, M. Dieguez, J. Diaz, S. Sepulveda, "A flexible and systematic teaching framework to develop cognitive skills through programming courses", Proceedings of the 38th International Conference of the Chilean Computer Science Society, Concepcion, Chile, 4-9 November 2019.

[31] Z. Zou, Y. Zhang, J. Li, X. Hei, Y. Du, D. Wu, "EasyHPC: An online programming platform for learning high performance computing", Proceedings of the 6th International Conference on Teaching, Assessment, and Learning for Engineering, pp. 432–435.

[32] L. Echeverría, R. Cobos, L. Machuca, I. Claros, "Using collaborative learning scenarios to teach programming to non-CS majors", Computer Applications of Engineering Education, Vol. 25, No. 5, 2017, pp. 719–731.

[33] H. Cheers, M. Javed, Y. Lin, S. Smith, "Exploring a Comprehensive Approach for the Automated Assessment of UML", Proceedings of the 8th International Congress on Advanced Applied Informatics, Toyama, Japan, 7-11 July 2019, pp. 133–139.

[34] P. E. Robinson, J. Carroll, "An online learning platform for teaching, learning, assessment of programming", Proceedings of the IEEE Global Engineering Education Conference, Athens, Greece, 25-28 April 2017, pp. 547–556.

[35] R. Romli, S. Sulaiman, K. Z. Zamli, "Test data generation framework for Automatic Programming Assessment", Proceedings of the 8th Malaysian Software Engineering Conference, 23-24 September 2014, pp. 84–89.

[36] D. Galan, R. Heradio, H. Vargas, I. Abad, J. A. Cerrada, "Automated Assessment of Computer Programming Practices: The 8-Years UNED Experience", IEEE Access, Vol. 7, 2019, pp. 130113–130119.

[37] R. Conejo, B. Barros, M. F. Bertoa, "Automated Assessment of Complex Programming Tasks Using SIETTE", IEEE Transactions on Learning Technologies, Vol. 12, No. 4, 2019, pp. 470–484.

[38] G. Polito, M. Temperini, A. Sterbini, "2TSW: Automated assessment of computer programming assignments, in a gamified web based system", Proceedings of the 18th International Conference on Information Technology Based Higher Education and Training, Magdeburg, Germany, 26-27 September 2019, pp. 1–9.

[39] R. Li, Y. Zhu, Z. Wu, "A new algorithm to the automated assessment of the Chinese subjective answer", Proceedings of the International Conference on Information Technology and Applications, Chengdu, China, 16-17 November 2013, pp. 228–231.

[40] M. A. I. Situ Nurulain Mohd Rum, "Metocognitive Support Accelerates Computer Assisted Learning for Novice Programmers", Educational Technology & Society, Vol. 20, No. 3, 2017, pp. 170–181.

[41] M. Kramer, P. Hubwieser, T. Brinda, "A competency structure model of object-oriented programming", Proceedings of the International Conference on Learning and Teaching in Computing and Engineering, Mumbai, India, 31 March-3 April 2016, pp. 1–8.

[42] K. Taghipour, H. T. Ng, "A neural approach to automated essay scoring", Proceedings of the Conference on Empirical Methods in Natural Language Processing, Austin, Texas, USA, November 2016, pp. 1882–1891.

[43] F. Dong, Y. Zhang, J. Yang, "Attention-based recurrent convolutional neural network for automatic essay scoring", Proceedings of the 21st Conference on Computational Natural Language Learning, Vancouver, Canada, August 2017, pp. 153–162.

[44] J. Z. Sukkarieh, J. Blackmore, "C-rater: Automatic content scoring for short constructed responses", Proceedings of the 22nd International Florida Artificial Intelligence Research Society Confenrence, 2009, pp. 290–295.

[45] T. A. Mat, A. Lajis, H. Nasir, "Text Data Preparation in RapidMiner for Short Free Text Answer in Assisted Assessment", Proceedings of the IEEE 5th International Conference on Smart Instrumentation, Measurement and Application, Songkhla, Thailand, 28-30 November 2018, pp. 28–30.

[46] D. Alikaniotis, H. Yannakoudakis, M. Rei, "Automatic text scoring using neural networks", Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, August 2016, pp. 715–725.

[47] T. A. Mat, "Rule-based LSTM classification to assess higher level cognitive competency via short free text answer", 2019.

# Post Acceptance Model for Online Teleconsultation services: An Empirical Study in Malaysia

Case study

**Abdulaziz Aborujiah**
Universiti Kuala Lumpur
Malaysian Institute of Information Technology (MIIT)
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250,
Kuala Lumpur, Malaysia
abdulazizsaleh@unikl.edu.my

**Rasheed Mohammad Nassr**
Universiti Kuala Lumpur
Malaysian Institute of Information Technology (MIIT)
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250,
Kuala Lumpur, Malaysia
rasheed@unikl.edu.my

**Abdulaleem Al- Othmani**
De montfort university
Cyber Technology Institute
Gateway House, Leicester LE1 9BH, UK
aleem.al-othmani@dmu.ac.uk

**Zalizah Awang Long**
Universiti Kuala Lumpur
Malaysian Institute of Information Technology (MIIT)
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250,
Kuala Lumpur, Malaysia
zalizah@unikl.edu.my

**Mohd Nizam Husen**
Universiti Kuala Lumpur
Malaysian Institute of Information Technology (MIIT)
1016, Jalan Sultan Ismail, Bandar Wawasan, 50250,
Kuala Lumpur, Malaysia
mnizam@unikl.edu.my

***Abstract*** *– Most nations across the world are actively pursuing equal access to healthcare services. Teleconsultation technology is a substantial improvement in terms of an effective framework for the provision of healthcare services. However, a lack of understanding of people's willingness towards the use of this technology has been observed. The goal of this study is to investigate the factors affecting the post-acceptance of teleconsultation services in Malaysia. This study developed a theoretical model which involves the combination of the second generation of Unified Theory of Acceptance and Use of Technology (UTAUT2) and Expectation Confirmation Theory (ECT), with the inclusion of several other constructs. An online survey was used to collect data from 154 university students and partial least squares (PLS) approach was used for data analysis. The research findings indicate that confirmation, performance, effort expectancy, usefulness, and satisfaction were the key factors that affect the post-acceptance of teleconsultation services. Furthermore, actual use, ease of use, technology readiness, and facilitating conditions did not impact participants' post intention in the continuous usage of teleconsultation facilities.*

***Keywords****: Teleconsultation Technology, UTAUT2, ECT, Post Acceptance model, COVID19, PLS_SEM*

## 1. INTRODUCTION

The new Information and Communication Technology (ICT) age have radically transformed human life, econom-ic processes, and culture into a new era of application that makes life easier [1]. The expansion of the Internet has contributed to the popularization of various virtual networks of online services [2] such as online learning and telehealth services. Telehealth services are described as health services that allow patients to receive therapy within their day-to-day life through one or more medical specialists [3]. Researchers have found that telehealth has steadily become the leading ICT service with an impres-sive impact on conventional health mechanisms [4]. Globally, telehealth programs boost doctors' efficacy, reduce medical costs, and increase access to healthcare [5][6]. They also provide services of medical practitioners consisting of tracking, diagnosis, and care provision over long distances using telecommunications. Previous stud-ies have proposed telehealth as a potential option for treating multiple health conditions including high blood pressure, obesity, diabetes, and cancer [6]. It is crucial to analyze the factors that influence end users' perception of adopting telehealth services [7]. so more studies have explored the key fac-

tors that motivate users to adopt such applications [12-14]. However, despite the huge number of emerging health applications, only a small number of apps (such as Noom Diet, Nike+, and Lose It) have been successful across the entire mHealth market. Although health apps are extremely useful in helping individuals to manage their health effectively, their usage often lasts a short while. This indicates a lack of under-standing of people's actions after installing health apps on their smartphones [3]. The aim of this study is, therefore, to fill the research gap and develop a research model based on UTAUT2 and ECT theories to discover the most influential factors that affect future intentions to use tele-health systems. As for the remainder of the paper, Sec-tion 2 describes the research background of factors that influence the intention of respondents to use telehealth services. Section 3 highlights the evolution of the hy-potheses while Section 4 discusses the research method-ology. Section 5 presents the results from the data analy-sis. Section 6 highlights the discussion of data analytics while Section 7 presents the conclusions and recom-mendations drawn from the research.

## 2. RELATED STUDIES

The novel coronavirus disease 2019 (COVID-19) had spread to Malaysia via Singapore on 24th January 2020. The pandemic has set a huge challenge to the delivery of neurosurgical services including the transfer of patients. Patients are triaged depending on their urgent needs for surgery or transferred to a neurosurgical center and man-aged accordingly. All patients are screened for the poten-tial risk of contracting COVID-19 before any surgery [2]. General surgery departments in Malaysia are part of Ma-laysia's tertiary centers that treat COVID-19 patients. The core highlights of these strategies during this pandemic are (1) surgery ward and clinic decongestions; (2) defer-ment of elective surgeries; (3) restructuring of medical personnel; (4) utilization of online applications for tele-communication; (5) operating room adjustments and patient screening; and (6) continuous learning and up-date practices in terms of COVID-19. These adaptations are important for the continuation of emergency surgery services, prevention of transmission of COVID-19 amongst healthcare workers, and optimization of the medical personnel workforce in times of a global pan-demic [3]. Patients are evaluated by a psychiatrist in the COVID-19 wards where they are hospitalized. The con-sultants wearing personal protective equipment provided for them enter the rooms of patients with COVID-19 to reduce their risks of exposure [4]. As the novel corona-virus SARS-CoV-2 (COVID-19) outbreak is highly conta-gious, there has been an urgent need to devise and iden-tify new models of delivering healthcare to avoid 'face-to-face' consultation between clinician and patient, thus reducing the risk of disease transmission [5]. In the ab-sence of high-tech communication facilities, resuming healthcare services during ongoing lockdown is highly demanding for related healthcare facilities in the country [6]. M-health may be a valuable strategy for expanding health coverage and empowering people to track their health, as well as potentially lowering medical costs [11]. Malaysia, as a developing nation with a strong technology market, should benefit from the use of m-health services due to its high Internet and broadband penetration rates, as well as its high smartphone penetration rate of 144.8%, showing that the majority of Malaysians own multiple mobile devices [1]. Teleconsultation is an exam-ple of m-health where patients communicate with a healthcare specialist via video chat or online platforms that provide videos of physical activities based on a physi-cal therapist's training programs. This technology can be valuable only when people start using it, given its known benefits. Consequently, end users' general attitude towards embracing telehealth services may play an important role [7]. However, most people are hesitant in using such technology. Hence, there is a need to explore to what extent the patients trust such systems [8]. There is a need to study the aspects that influence people's acceptance of teleconsultation in Malaysia.

### *A. Expectation-Confirmation*

The Expectation Confirmation Theory (ECT), presented by Oliver in 1980, describes consumer satisfaction because of the disaffirmation of desires and aspirations. Using ECT, Oliver argued that the shift in mood and intention of the customer is caused by satisfaction [15-17] Bhattacherjee subsequently proposed in 2001 to provide information systems (IS) consistency with ECT. Bhattacherjee pro-posed that the decision of IS users to continue is like the decision of customers to buy back, as both are based on initial knowledge of IS or product use. Therefore, both are closely related to customer satisfaction [18]. The more expectations people have on technology, the more de-sire they must use it. The continuous intention of using information systems for compulsory use was investigated and the value of user satisfaction was found by Sorebo and Eikebrokk [19]. IT uses were described by Rai, Lang, and Welker in 2002 as an undemanding, but not neces-sarily voluntary, system-based usage due to social pres-sure and environmental subjective norms [20]. ECT is widely used in different post-adoption contexts. It ends with the assumption that the extent of user confirmation and perceived usefulness are the main determinants of user satisfaction. Hence, confirmation and satisfaction are linked favorably. Usefulness and satisfaction also affect individuals' continuous intent to use technology [21]. In this context, the following hypothesis is proposed:

*H1: Performance is positively associated with users' confir-mation of continuous use of telehealth technology.*

### *B. Actual use*

In terms of better work results (effectiveness), fast com-pletion (efficiency), and a positive attitude to a job (en-gagement), technical usefulness is the product of task success [22-24]. The expectations of technical usefulness and satisfaction [25-29] have been related to

improved task efficiency. For example, satisfaction and prior experi-ence regulate the desire to proceed with Internet-based learning technology [30]. Furthermore, a strong relation-ship between perceived usefulness, confirmation, and satisfaction has been developed [31]. Prior experience of using technology plays a main role in usage continuation. For example, if a user perceives better organizational and technological support and ICT services, the more the e-learning program is used. The present study defined the qualities that trigger actual use and ongoing use of e-learning systems to be considerably beneficial [57]. In this context, the following hypothesis is proposed:

*H2: Actual use is positively associated with users' continuous intention of telehealth technology usage.*

### *C. Ease of use*

Technology has to be seen as a useful tool to assist peo-ple in doing their jobs easily [32]. The ease of learning and user-intuitiveness of the system can be measured by the users revisiting the technology and not having to re-learn the tools to effectively perform a task [33-35]. If the system is easy to learn, effective, and efficient, people will be more interested to use it. Minimizing errors that may exist in the technology plays a main role to attract more people to use it [36]. For example, in e-learning education, the nature of effort expectation implies the extent of its acceptability and usage. Past studies of tech-nology adoption have shown that efforts are anticipated, both voluntary and involuntary, during the early steps of technology usage and are negligible over time for sus-tainable usage [37]. In this context, the following hy-pothesis is proposed:

*H3: Ease of technology use is positively associated with users' continuous intention of telehealth technology usage.*

### *D. Effort expectancy*

Previous researches have shown that whenever the effort to understand and learn new technology is lesser, users tend to have more intention of using the technology. Public relations professionals, for example, have been influenced by the simplicity and self-efficacy of the me-dia [38]. For example, in inpatient management, effort expectation is also projected as a key indicator of patients' likeability of using mobile systems [23]. Effort expectancy refers to the degree to which the systems are easy or difficult to be accepted and used. Previous studies on technology acceptance have shown that during the early stages of technology adoption, effort expectancy is significant, both voluntary and involuntary, and becomes insignificant over time for sustainable use [37]. In this context, the following hypothesis is proposed:

*H4: Effort expectancy is positively associated with users' continuous intention of telehealth technology usage.*

### *E. Performance*

Several studies have tested and validated the relationship between confirmation, usefulness, satisfaction, and con-tinuous intention [18], [40-42]. For example, the degree of confidence in Internet banking services has influenced the degree of perceived usefulness and satisfaction of the services [18]. Kim [18] showed that the relationship between perceived confirmation and usefulness upon goods purchased has a positive impact on the e-commerce satisfaction of consumers. The perceived amount of usefulness affects satisfaction and intention to use e-learning technology [41]. Students' performance expectancy refers to the degree to which the system allows the students to perform better in their curriculum. Preliminary studies have recognized that technology use in both voluntary and obligatory settings is strongly ex-pected to be employed by performance factors [43, 44] [39]. The expected performance has a significant impact on a system's continuous use in various studies. In this context, the following hypothesis is proposed:

*H5: Expected performance is positively associated with users' continuous intention of telehealth technology usage.*

### *F. Price*

Price is described as a cognitive trade-off of the consum-ers between the perceived advantages of apps and the monetary cost of using them [45][46]. There are three types of pricing schemes in the modern app market: free, paid, and freemium. Free apps are free to download and use while paid applications must be paid for by users be-fore they can be downloaded. Freemium schemes pro-vide users with the ability to test the application for free before agreeing to buy the premium features [47]. Con-sumers demand higher quality or improved services if they pay for them [48]. In this context, the following hy-pothesis is proposed:

*H6: Price is associated with users' continuous intention of telehealth technology usage.*

### *G. Technology readiness*

According to Parasuraman [49], technology readiness refers to one's propensity to embrace new technology to achieve goals in one's life at home and work. It is a multifaceted construction that has four dimensions: optimism (a positive perception of technology and a belief that it gives people greater control, flexibility, and efficiency in their lives), innovation (a tendency to become a technology pioneer), discomfort (a perceived lack of control over technology and a feeling of being overwhelmed by it), and insecurity (disruption). Optimism and innovativeness serve as the key drivers of technology readiness. They encourage people to use new technology and foster per-ceptions of safety and novelty [31]. Discomfort and inse-curity, on the other hand, are inhibitors of development readiness. They make customers hesitant to adopt new technology and create feelings of fear, confusion, and discomfort. Meanwhile, health-related information chan-nels and apps are seen as an innovative technology that can facilitate healthy behavior. Hence, the willingness of people to use applications affects their

efficacy, as some people are technology pessimists [50]. In this context, the following hypothesis is proposed:

*H7: Technology readiness is associated with users' contin-uous intention of telehealth technology usage.*

### *H. Usefulness*

The perceived service quality in the initial phase is de-scribed as the degree of an individual's expectation that the current system would improve the efficiency of a given task. Literature has shown the usefulness and im-portance of a particular technology for users' intention change [51]. As Bhattacherjee introduced ECT coupled with the technology acceptance model (TAM), it has been confirmed that perceived usefulness influences not just the implementation of IS, but also users' comfort and persistent desire to use it. Several studies have described the association between perceived usefulness, satisfac-tion, and continuous usage desire [52], [53]. Compared to the findings of earlier research, we predicted a beneficial impact of perceived usefulness on user satisfaction and continuous usage intention [54]. In this context, the fol-lowing hypothesis is proposed:

*H8: Usefulness is associated with users' continuous intention of telehealth technology usage.*

### *I. Facilitating conditions*

In a longitudinal study of Chen [7], facilitation of condi-tions refers to the degree to which people feel that the technical and institutional facilities are available to pro-mote the use of technology. The original UTAUT has shown that facilitating conditions only substantially im-pact actual use. Further studies, including a meta-analysis of 43 studies on technology acceptance, have shown that facilitating conditions also have positive effects on behav-ioral purposes [63]. For example, previous e-learning ac-ceptability studies have demonstrated that ease-of-use conditions have a positive effect on the intention to use [54]. The indication is that the better the students per-ceive the organizational and technical support, and ICT infrastructure, the more the e-learning system is used. The current study has hypothesized the facilitation of conditions that makes people keener to use telehealth. In this context, the following hypothesis is proposed:

*H9: Facilitating conditions are associated with users' contin-uous intention of telehealth technology usage.*

### *J. Satisfaction*

Consumer satisfaction can be defined as consumer per-ception of the extent to which consumer requirements have been met [55]. Keiningham, Perkins-Munn, and Ev-ans confirmed the definition of satisfaction where con-sumer satisfaction has an impact on consumer behavior. In addition, high consumer satisfaction leads to higher consumer loyalty and buying intentions [56]. For exam-ple, market research has shown that the main reason for a consumer's decision to re-purchase or re-use a product is their level of satisfaction [17],[57,58]. Bhattacherjee had empirically shown that level of satisfaction is a critical factor in decision-making [18]. In this context, the follow-ing hypothesis is proposed:

*H10: Satisfaction is associated with users' continuous inten-tion of telehealth technology usage.*

## 3. METHODS AND MEASURES

This study aims to examine and investigate the causes that shape and influence the intention to use telehealth among Malaysians. Figure 1 represents the research model of this study. Telehealth intention was considered as a dependent variable. Students in Malaysian universi-ties were the targeted population for this study. To en-sure the validity of all measures, individual constructs of the determinants were adapted from previous research, provided in this paper.

**Table 1.** The proposed research hypotheses

| Factors | Abb* | Hypothesis |
|---|---|---|
| Actual Use | ATS | H1: Actual use positively influences continuous intention of use |
| Performance | PER | H2: Performance positively influences confirmation |
| Ease of use | EOS | H3: Ease of use positively influences continuous intention of use |
| Effort expectancy | EX | H4: Effort expectancy positively influences continuous intention |
| Confirmation | COF | H5: Confirmation positively influences user satisfaction |
| Price | PR | H6: Price influences continuous intention of use |
| Technology readiness | TR | H7: Technology readiness positively influences continuous intention of use |
| Usefulness | USF | H8: Usefulness positively influences continuous intention of use |
| Facilitating conditions | FC | H9: facilitating conditions positively influence the continuous intention of use |
| Satisfaction | SAT | H10: user satisfaction positively influences continuous intention of use |

***Abb:** Abbreviation

**Table 2.** Demographics information of the participants

| **Gender** | **N** | **%** | | | |
|---|---|---|---|---|---|
| Female | 75 | %49 | | | |
| Male | 79 | %51 | | | |
| **Study level** | **N** | **%** | | | |
| Undergraduate | 129 | %84 | | | |
| Postgraduate: | 25 | %16 | | | |
| **Online Health care services familiarity** | | | | | |
| I like to use websites to get health care services | Strongly disagree | Disagree | Neutral | Agree | Strongly Agree |
| | 5 | 7 | 40 | 50 | 52 |
| | 3.2% | 4.5% | 25.9% | 32.4% | 33.7% |
| **Living** | **N** | **%** | | | |
| Urban | 123 | 79.9% | | | |
| Rural | 31 | 20.1% | | | |
| **Age** | **N** | **%** | | | |
| | Between 21 and 30 154 | 100% | | | |

To test the multiple hypotheses, this study collected online survey data of full-time students from the University of Kuala Lumpur, Malaysia. The survey asked students to examine their acceptance of telehealth services. Two hundred and eight responses were initially received. Out of the 208 responses collected, 54 (i.e. incomplete, outlier) responses were dropped while 54 respondents were chosen. Table 1 shows the demographics information of the respondents. Eight responses were initially received.

The construction and research model in Figure 1 was developed based on a comprehensive literature review as described above. The independent constructs of the theoretical model consisted of Actual Use, Confirmation, Ease of Use, Effort Ex-pectancy, Performance, Price, Technology Readiness, Usefulness, Facilitating Conditions, and Satisfaction. Intent to Use was stated as the dependent variable. These constructs were derived from previous studies, with minor modifications in the lan-guage of the items used to capture the data. A 5-point Likert scale was used to capture the answers for each item, with 1 being strongly agreed and 5 as Strongly Disagree. Several demographic items that use various measurement scales, were also included in the questionnaire. We used partial least squares (PLS) for data analysis and research model testing. PLS path modeling is a variance-based structural equation modeling (SEM) technology that is widely implemented in business and social sciences. Its ability to simulate composites and factors makes it an effective computational method for new technology studies [59]. The advantages of SEM relative to first-generation statistical techniques include more robust assumptions where multicollinearity is partly enabled and less error of calculation is used with confirmatory factor analysis (CFA) [60]. We evaluated the model using the Smart PLS 3.0 bootstrapping methodology [60,61].

## 4. RESULTS

The details shown in Table 2 indicate that 51% are male interviewees while 49% are female. Most participants (75%) are Bachelor's degree holders. Appendix A shows the questions.

### *A. Measurement Model Assessment*

To measure the internal consistency of the hypothesized model, Cronbach's Alpha along with composite reliability and av-erage variance extracted (AVE) was used. Table 2 shows that the composite reliability values are between 93.3% and 78.3% which exceed the recommended threshold of 70% [62]. However, Cronbach's Alpha values are below, between 87.5% and 60.3%. Some items are below 70%, consisting of EX, EOS, SAT, TR, and Perceived USF. A low Cronbach's Alpha indicates a result of test length and dimensionality [62]. Therefore, all the indicators were considered reliable. Furthermore, the average variance extracted (AVE) method was used to measure the convergent validity of the selected items between 49.5 and 82.2. Table 2 shows that the AVE values of all the constructs are more than 0.5 except the TR construct, which assumed ade-quate convergent validity [62].

**Table 3.** Measurement Model Assessment

| Constructs | | Loading | AVE* | CR** | Alpha |
|---|---|---|---|---|---|
| ATS | ATS 1 | 0.919 | 0.808 | 0.894 | 0.765 |
| | ATS 2 | 0.879 | | | |
| COF | COF 1 | 0.702 | 0.616 | 0.906 | 0.875 |
| | COF 2 | 0.723 | | | |
| | COF 3 | 0.814 | | | |
| | COF 4 | 0.791 | | | |
| | COF 5 | 0.857 | | | |
| | COF 6 | 0.812 | | | |
| EX | EX 1 | 0.912 | 0.822 | 0.933 | 0.616 |
| | EX 2 | 0.923 | | | |
| | EX 3 | 0.884 | | | |
| FC | FC 1 | 0.814 | 0.581 | 0.847 | 0.762 |
| | FC 2 | 0.787 | | | |
| | FC 3 | 0.782 | | | |
| | FC 4 | 0.658 | | | |
| EoS | EoS 1 | 0.297 | 0.578 | 0.778 | 0.616 |
| | EoS 2 | 0.931 | | | |
| | EoS 3 | 0.882 | | | |
| PER | PER 1 | 0.858 | 0.757 | 0.903 | 0.84 |
| | PER 2 | 0.85 | | | |
| | PER 3 | 0.901 | | | |
| **Constructs** | | **Loading** | **AVE*** | **CR**** | **Alpha** |
| PR | PR 1 | 0.926 | 0.797 | 0.887 | 0.75 |
| | PR 1 | 0.858 | | | |
| SAT | SAT 1 | 0.903 | 0.575 | 0.783 | 0.603 |
| | SAT 2 | 0.885 | | | |
| | SAT 3 | 0.356 | | | |
| TR | TR 1 | 0.648 | 0.495 | 0.795 | 0.657 |
| | TR 2 | 0.712 | | | |
| | TR 3 | 0.815 | | | |
| | TR 4 | 0.623 | | | |
| CTU | CTU 1 | 0.641 | 0.623 | 0.92 | 0.898 |
| | CTU 2 | 0.809 | | | |
| | CTU 3 | 0.824 | | | |
| | CTU 4 | 0.782 | | | |
| | CTU 5 | 0.783 | | | |
| | CTU 6 | 0.861 | | | |
| | CTU 7 | 0.808 | | | |
| USF | USF 1 | 0.751 | 0.612 | 0.825 | 0.683 |
| | USF 2 | 0.848 | | | |
| | USF 3 | 0.744 | | | |

***AVE:** Average variance extracted
****CR:** Composite Reliability

**Table 4.** Correlation analysis

| | ATS | COF | CTU | EoS | EX | PER | PR | TR | USF | FC | SAT |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ATS** | 0.899 | | | | | | | | | | |
| **COF** | 0.485 | 0.785 | | | | | | | | | |
| **CTU** | 0.477 | 0.809 | 0.79 | | | | | | | | |
| **EoS** | 0.35 | 0.588 | 0.499 | 0.76 | | | | | | | |
| **EX** | 0.537 | 0.73 | 0.759 | 0.533 | 0.907 | | | | | | |
| **PER** | 0.453 | 0.835 | 0.778 | 0.569 | 0.647 | 0.87 | | | | | |
| **PR** | 0.287 | 0.566 | 0.636 | 0.336 | 0.519 | 0.496 | 0.892 | | | | |
| **TR** | 0.265 | 0.461 | 0.439 | 0.419 | 0.441 | 0.357 | 0.349 | 0.703 | | | |
| **USF** | 0.425 | 0.753 | 0.764 | 0.494 | 0.674 | 0.668 | 0.578 | 0.51 | 0.782 | | |
| **FC** | 0.362 | 0.608 | 0.639 | 0.482 | 0.642 | 0.478 | 0.553 | 0.541 | 0.683 | 0.762 | |
| **SAT** | 0.475 | 0.75 | 0.761 | 0.635 | 0.711 | 0.697 | 0.623 | 0.424 | 0.692 | 0.609 | 0.758 |

**Table 5.** Hypothesis testing

| | Original Sample (O)(β) | Sample Mean (M) | Standard Deviation (STDEV) | T Statistics (\|O/STDEV\|) | P Values | Conclusion |
|---|---|---|---|---|---|---|
| H1: ATS ->CTU | 0.04 | 0.055 | 0.051 | 0.787 | 0.432 | Not supported |
| H2: COF ->PER | 0.835 | 0.838 | 0.025 | 32.899 | 0 | supports |
| H3: EoS ->CTU | -0.03 | -0.037 | 0.082 | 0.36 | 0.719 | Not supports |
| H4: EX ->CTU | 0.291 | 0.279 | 0.081 | 3.588 | 0 | supports |
| H5: PER -> SAT | 0.697 | 0.697 | 0.056 | 12.339 | 0 | supports |
| H6: PR->CTU | 0.149 | 0.156 | 0.054 | 2.75 | 0.007 | supports |
| H7: TR ->CTU | -0.007 | -0.009 | 0.062 | 0.111 | 0.912 | Not supported |
| H8: USF ->CTU | 0.3 | 0.31 | 0.08 | 3.744 | 0 | supports |
| H9: FC ->CTU | 0.02 | 0.034 | 0.102 | 0.2 | 0.842 | Not supported |
| H10: SAT ->CTU | 0.244 | 0.228 | 0.088 | 2.782 | 0.006 | supports |

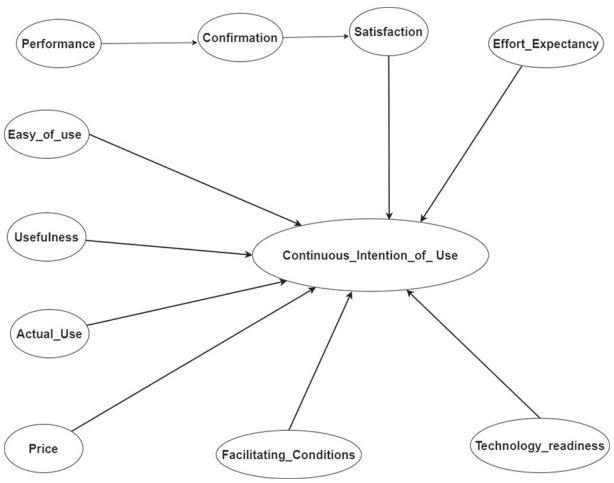



**Fig. 1.** Conceptual Research Model

### *B. DISCRIMINANT VALIDITY*

This research calculated the discriminant convergent va-lidity of the contracts by comparing the square root of AVE for each construct with its cross-correlation with oth-er constructs. The results showed that the square root of AVE was found to be higher than the off-diagonal ele-ments in the corresponding rows and columns which support the discriminant criteria set for all contracts. Table 3 summarizes the results. The accepted convergent validity of each construct must exceed the correlation it exhibits with other constructs [8]. In addition, the members in the matching columns and correlation matrix rows must be lower than the diagonal element [8]. Table 3 summarizes the data that con-firm convergent validity of all contracts which are be-tween 89.9% and 70.3%, indicating a minimum of 0.50 of AVE

exists for all constructs. In addition, the entire loadings were highly significant (t-statistics > (3.419), p < (0.001)) based on the output of SmartPLS, which demonstrated that the indicators represent noticeably different latent construction.

### *C. Limitations*

One drawback of this study is that the sample used did not include certain classes of individuals, such as school students who are strong consumers of video games for example, and academically challenged people who are more vulnerable to overuse of digital games. Future work is expected to extend the effects of this study to a specif-ic population, such as school pupils and uneducated peo-ple.

## 5. CONCLUSION

Public confidence in telehealth services is a highly under-researched field. A major antecedent of this technology adoption is to investigate decisions on the use of tele-health services. Now that telehealth services are more popular and large quantities of personal data are being collected, the public trust in telehealth services will be-come a more important feature. This research investigat-ed the drivers and obstacles that affect the willingness of people to utilize telehealth facilities. The findings pointed out that confirmation, performance, effort expectancy, usefulness, and satisfaction were the main drivers influ-encing the acceptance of telehealth ser-vices. Furthermore, actual use, ease of use, technology readiness, and facilitating conditions did not impact participants' confidence in the use of telehealth ser-vices. Despite the substantial influences of the constructs, educating workers and the public on how to use this technology by conducting special training programs at health care institutions is suggested to familiarize them with the technology. There is also a need to upgrade the current healthcare systems and make them compatible with telehealth technology requirements. The findings of this study contributed to the existing body of knowledge of adopting and implementing new healthcare systems such as telehealth.

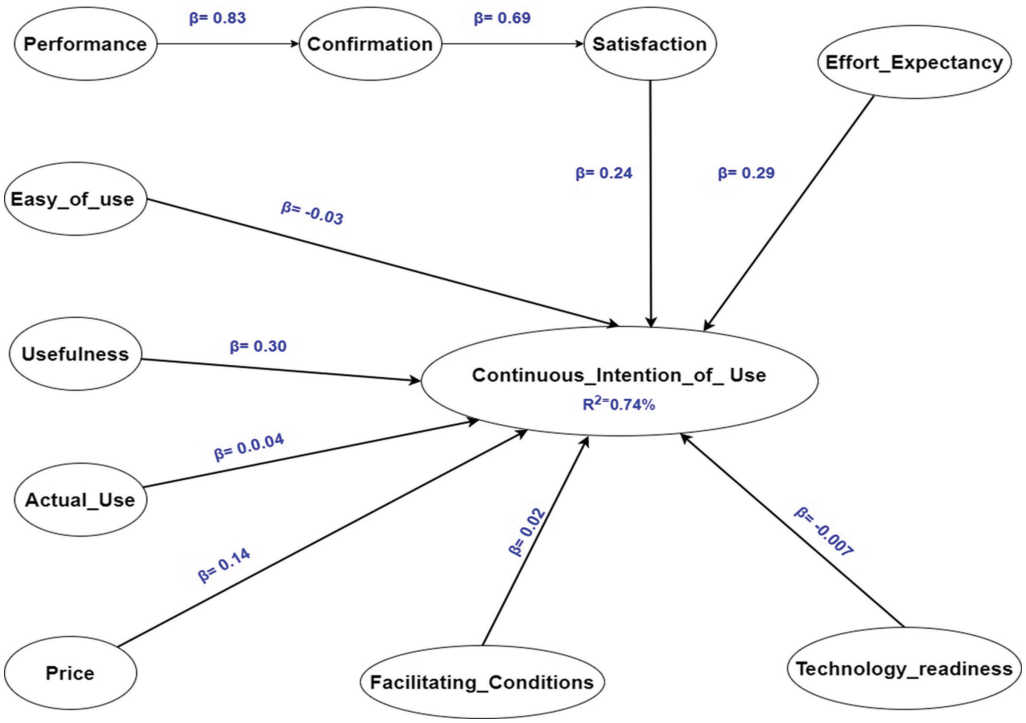



**Fig. 2.** Research mode

## 6. REFERENCES


[1] P. Tasca, T. Aste, L. Pelizzon, N. Perony, "Banking Beyond Banks and Money: A Guide to Banking Services in the Twenty-First Century", Springer, 2016.

[2] E. M. Al-Mukhaini, W. S. Al-Qayoudhi, A. H. Al-Badi, "Adoption of social networking in education: A study of the use of social networks by higher education students in Oman", International Journal of Educational Research, Vol. 10, No. 2, pp. 143–154, 2014.

[3] S. Kosterink, "The added value of telemedicine services for physical rehabilitation", Research UT, 2014, PhD Thesis.

[4] M. J. Rho, I. young Choi, J. Lee, "Predictive factors of telemedicine service acceptance, behavioral intention of physicians", International Journal of Medical Informatics, Vol. 83, No. 8, pp. 559–571, 2014.

[5] P. Esmaeilzadeh, M. Sambasivan, N. Kumar, "The challenges, issues regarding e-health, health in-

formation technology trends in the healthcare sector", Proceedings of the International Conference on E-business Technology, Strategy, Ottawa, Canada, 29-30 September 2010, pp. 23–37.

[6] Z. Jin, Y. Chen, "Telemedicine in the cloud era: Prospects, challenges", IEEE Pervasive Computing, Vol. 14, No. 1, 2015, pp. 54–61.

[7] S. A. Kamal, M. Shafiq, P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM)", Technology in Society, Vol. 60, 2020, p. 101212.

[8] L. van Velsen, M. Tabak, H. Hermens, "Measuring patient trust in telemedicine services: Development of a survey instrument, its validation for an anticoagulation web-service", International Journal of Medical Informatics, Vol. 97, 2017, pp. 52–58.

[9] D. D. Luxton, R. A. McCann, N. E. Bush, M. C. Mishkind, G. M. Reger, "MHealth for mental health: Integrating smartphone technology in behavioral healthcare", Professional Psychology: Research and Practice, Vol. 42, No. 6, 2011, pp. 505–512.

[10] E. Årsand et al., "Mobile health applications to assist patients with diabetes: lessons learned, design implications", Journal of Diabetes Science and Technology, Vol. 6, No. 5, 2012, pp. 1197–1206.

[11] M. C. Carter, V. J. Burley, C. Nykjaer, J. E. Cade, "Adherence to a smartphone application for weight loss compared to website, paper diary: pilot randomized controlled trial", Journal of Medical Internet Research, Vol. 15, No. 4, 2013, p. e32.

[12] J. L. Bender, R. Y. K. Yue, M. J. To, L. Deacken, A. R. Jadad, "A lot of action, but not in the right direction: Systematic review, content analysis of smartphone applications for the prevention, detection, management of cancer", Journal of Medical Internet Research, Vol. 15, No. 12, 2013, p. e287.

[13] J. Cho, M. M. Quinlan, D. Park, G.-Y. Noh, "Determinants of adoption of smartphone health apps among college students", American Journal of Health Behavior, Vol. 38, No. 6, 2014, pp. 860–870.

[14] J. Cho, H. E. Lee, S. J. Kim, D. Park, "Effects of body image on college students' attitudes toward diet/fitness apps on smartphones", Cyberpsychology, Cyberpsychology, Behavior, and Social Networking, Vol. 18, No. 1, 2015, pp. 41–45.

[15] R. Gulati, "Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances", Academy of Management Journal, Vol. 38, No. 1, 1995, pp. 85–112.

[16] R. L. Oliver, "A Cognitive Model of the Antecedents, Consequences of Satisfaction Decisions", J. Mark. Res., Vol. 17, No. 4, 1980, pp. 460–469.

[17] E. W. Anderson, M. W. Sullivan, "The Antecedents, Consequences of Customer Satisfaction for Firms", Marketing Science, Vol. 12, No. 2, 1993, pp. 125–143.

[18] A. Bhattacherjee, "Understanding Information Systems Continuance: An Expectation-Confirmation Model", MIS Quarterly, Vol. 25, No. 3, 2001, pp. 351–370.

[19] Ø. Sørebø, T. R. Eikebrokk, "Explaining IS continuance in environments where usage is mandatory", Computers in Human Behavior, Vol. 24, No. 5, pp. 2357–2371, 2008.

[20] A. Rai, S. S. Lang, R. B. Welker, "Assessing the validity of IS success models: An empirical test, theoretical analysis", Information Systems Research., Vol. 13, No. 1, 2002, pp. 50–69.

[21] S. H. Lim, D. J. Kim, Y. Hur, K. Park, "An empirical study of the impacts of perceived security, knowledge on continuous intention to use mobile fintech payment services", International Journal of Human–Computer Interaction, Vol. 35, No. 10, 2019, pp. 886–898.

[22] R. P. McDonald, M.-H. R. Ho, "Principles, practice in reporting structural equation analyses.", Psychological Methods, Vol. 7, No. 1, 2002, p. 64.

[23] J. Baldwin, "The fire next time", Vintage, 2013.

[24] A. Sonderegger, G. Zbinden, A. Uebelbacher, J. Sauer, "The influence of product aesthetics, usability over the course of time: a longitudinal field experiment", Ergonomics, Vol. 55, No. 7, 2012, pp. 713–730.

[25] D. R. Compeau, C. A. Higgins, "Computer self-efficacy: Development of a measure, initial test", MIS Quarterly, 1995, pp. 189–211.

[26] D. Green, J. M. Pearson, "Development of a web site usability instrument based on ISO 9241-11", Journal of Computing and Information Science in Engineering., Vol. 47, No. 1, 2006, pp. 66–72.

[27] D. T. Green, J. M. Pearson, "Integrating website usability with the electronic commerce acceptance model", Behaviour & Information Technology, Vol. 30, No. 2, 2011, pp. 181–199.

[28] S. Toleva-Stoimenova, D. Christozov, "Informing via Websites: Comparative Assessment of University Websites", in Proceedings of the Informing Science, Information Technology Education Conference, 2013, pp. 525–537.

[29] R. K.-J. Yeh, J. T. C. Teng, "Extended conceptualisation of perceived usefulness: empirical test in the context of information system use continuance", Behaviour & Information Technology, Vol. 31, No. 5, 2012, pp. 525–540.

[30] M. Limayem, C. M. K. Cheung, "Understanding information systems continuance: The case of Internet-based learning technologies", Information and Management, Vol. 45, No. 4, 2008, pp. 227–232.

[31] J. Wu, R. J. Tsai, C. C. Chen, Y. Wu, "An integrative model to predict the continuance use of electronic learning systems: hints for teaching", International Journal on E-Learning, Vol. 5, No. 2, 2006.

[32] L. Baker-Eveleth, R. W. Stone, "Usability, expectation, confirmation, continuance intentions to use electronic textbooks", Behaviour & Information Technology, Vol. 34, No. 10, 2015, pp. 992–1004.

[33] S. M. Z. Ahmed, "A comparison of usability techniques for evaluating information retrieval system interfaces", Performance Measurement and Metrics, 2008.

[34] R. P. Bringula, "Influence of faculty-and web portal design-related factors on web portal usability: A hierarchical regression analysis", Computers & Education, Vol. 68, 2013, pp. 187–198.

[35] U. Konradt, H. Wandke, B. Balazs, T. Christophersen, "Usability in online shops: scale construction, validation, the influence on the buyers' intention, decision", Behaviour & Information Technology, Vol. 22, No. 3, 2003, pp. 165–174.

[36] A. Chevalier, N. Fouquereau, J. Vanderdonckt, "The influence of a knowledge-based system on designers' cognitive activities: a study involving professional web designers", Behaviour & Information Technology, Vol. 28, No. 1, 2009, pp. 45–62.

[37] V. Venkatesh, F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies", Manage. Sci., Vol. 46, No. 2, 2000, pp. 186–204.

[38] L. Curtis et al., "Adoption of social media for public relations by nonprofit organizations", Public Relations Review, Vol. 36, No. 1, 2010, pp. 90–92.

[39] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, "User acceptance of information technology: Toward a unified view", MIS Quarterly, 2003, pp. 425–478.

[40] J.-H. Lee, "A literature review on security for internet of things in Korea based on IoT SPND-Se ecosystem model", Journal of Security Engineering, Vol. 12, No. 4, 2015, pp. 397–414.

[41] M.-C. Lee, "Explaining, predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model", Computers & Education, Vol. 54, No. 2, 2010, pp. 506–516.

[42] S. H. Lim, N. J. Cho, J. H. Whang, "Role of familiarity, perceived behavioral control in adoption, continual use of RFID services", Korea Logistic Review, Vol. 22, No. 3, 2012, pp. 5–25.

[43] Y. K. Dwivedi, N. P. Rana, H. Chen, M. D. Williams, "A Meta-analysis of the Unified Theory of Acceptance, Use of Technology (UTAUT)", Proceedings of the IFIP international working conference on governance, sustainability in information systems-managing the transfer, diffusion of it, Hamburg, Germany, 22-24 September2011, pp. 155–170.

[44] B. ŠUmak, M. HeričKo, M. PušNik, "A meta-analysis of e-learning technology acceptance: The role of user types, e-learning technology types", Computers in Human Behavior, Vol. 27, No. 6, 2011, pp. 2067–2077.

[45] V. Venkatesh, J. Y. L. Thong, X. Xu, "Consumer acceptance, use of information technology: extending the unified theory of acceptance, use of technology", MIS Quarterly, 2012, pp. 157–178.

[46] W. B. Dodds, K. B. Monroe, D. Grewal, "Effects of price, brand, store information on buyers' product evaluations", Journal of Marketing Research, Vol. 28, No. 3, 1991, pp. 307–319.

[47] J. H. West, P. C. Hall, C. L. Hanson, M. D. Barnes, C. Giraud-Carrier, J. Barrett, "There's an app for that: Content analysis of paid health, fitness apps", Journal of Medical Internet Research, Vol. 14, No. 3, 2012, p. e72.

[48] V. A. Zeithaml, "Consumer perceptions of price, quality, value: a means-end model, synthesis of evidence", Journal of Marketing, Vol. 52, No. 3, 1988, pp. 2–22.

[49] A. Parasuraman, "Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies", J. Serv. Res., Vol. 2, No. 4, 2000, pp. 307–320.

[50] M. L. Meuter, A. L. Ostrom, R. I. Roundtree, M. J. Bitner, "Self-service technologies: understanding customer satisfaction with technology-based service encounters", Journal of Marketing, Vol. 64, No. 3, 2000, pp. 50–64.

[51] P. C. Lai, "The literature review of technology adoption models, theories for the novelty technology", JISTEM-Journal Inf. Syst. Technol. Manag., Vol. 14, No. 1, 2017, pp. 21–38.

[52] R. Zhou, X. Wang, Y. Shi, R. Zhang, L. Zhang, H. Guo, "Measuring e-service quality, its importance to customer satisfaction, loyalty: an empirical study in a telecom setting", Electronic Commerce Research, Vol. 19, No. 3, 2019, pp. 477–499.

[53] C.-H. Hsiao, "The effects of post-adoption beliefs on the expectation–confirmation model in an electronics retail setting", Total Quality Management and Business Excellence, Vol. 29, No. 7–8, 2018, pp. 866–880.

[54] M. Park, J. Jun, H. Park, "Understanding Mobile Payment Service Continuous Use Intention: An Expectation-Confirmation Model, Inertia", Quality Innovation Prosperity, Vol. 21, No. 3, 2017, pp. 78–94.

[55] M. Horváth, A. Michalkova, "Monitoring Customer Satisfaction in Service Industry: A Cluster Analysis Approach", Quality Innovation Prosperity, Vol. 16, No. 1, 2012, pp. 49–54.

[56] T. L. Keiningham, T. Perkins-Munn, H. Evans, "The impact of customer satisfaction on share-of-wallet in a business-to-business environment", Journal of Service Research, Vol. 6, No. 1, 2003, pp. 37–50.

[57] P. Kotler, G. Armstrong, "Principles of marketing. Pearson education", 2010.

[58] T. Fernandes, R. Pedroso, "The effect of self-checkout quality on customer satisfaction, repatronage in a retail context", Service Business, Vol. 11, No. 1, pp. 69–92, 2017.

[59] J. Henseler, "Bridging design, behavioral research with variance-based structural equation modeling", Journal of Advertising, Vol. 46, No. 1, 2017, pp. 178–192.

[60] I. B. Hong, H. S. Cha, "The mediating role of consumer trust in an online merchant in predicting purchase intention", International Journal of Information Management, Vol. 33, No. 6, 2013, pp. 927–939.

[61] M. Wetzels, G. Odekerken-Schröder, C. Van Oppen, "Using PLS path modeling for assessing hierarchical construct models: Guidelines, empirical illustration", MIS Quarterly, 2009, pp. 177–195.

[62] C. Fornell, D. F. Larcker, "Evaluating structural equation models with unobservable variables, measurement error", Journal of Marketing Research, Vol. 18, No. 1, 1981, pp. 39–50.

[63] Y. K. Dwivedi, N. P. Rana, H. Chen, M. D. Williams, "A Meta-analysis of the Unified Theory of Ac-ceptance and Use of Technology (UTAUT)", Pro-ceedings of the IFIP international working con-ference on governance and sustainability in in-formation systems-managing the transfer and dif-fusion of it, Hamburg, Germany, 22-24 Septem-ber 2011, pp. 155-170.

## **Appendix A** SURVEY QUESTIONS

| Construct | Questions | Reference |
|---|---|---|
| **Satisfaction** | 1. How have you played video games for the last six months?<br>2. I feel positive about the health care services offered by the internet<br>3. I am satisfied with online health consultation that is delivered by the internet<br>4. I have a nice relationship with my family | [9] |
| **Confirmation** | 1. I can manage my health well by getting health treatment through the internet<br>2. I can improve the condition of my health by getting health treatment through the internet<br>3. My experience with having health treatment through the internet is better than what I expected.<br>4. level of health care services that are delivered through the internet is better than what I expected<br>5. Online health consultation meet my expectation<br>6. Obtaining health treatment via the internet can assist me to achieve better health | [8] |
| **Performance** | 1. Optioning health treatment via the Internet helps me to accomplish things more quickly<br>2. Optioning health treatment by the Internet increases my productivity.<br>3. Overall, most of my expectations from using the internet to deliver health care services are confirmed. | [8] |
| **Easy of use** | 1. I think that learning how to use health care services that are delivered through the internet will be easy<br>2. I think it will take longer to learn how to use health care services that are delivered through the internet<br>3. I think that it will be easy to use health care services that are delivered through the internet<br>4. I think that digital games overuse treatment will become easier if we use health care services that are delivered through the internet<br>5. Do you think that health care services that are delivered through the internet will be hard to use | [12] |
| **Perceived usefulness** | 1. Using online health consultations that are delivered by the internet effectively brings more energy to me.<br>2. I think that doctors and patients(addicted to digital games) will become closer using the internet<br>3. I find the use of online health consultations that are delivered by the internet makes me more knowledgeable. | [12] |
| **Continuance intention** | 1. I am willing to use online health consultations that are delivered by the internet<br>2. I intend to continue using online health consultations that are delivered by the internet than using any alternative means<br>3. I intend to continue using online health consultations that are delivered by the internet rather than discontinue their use.<br>4. I intend to continue using online health consultations that are delivered by the internet rather than discontinue their use.<br>5. I intend to continue using online health consultations that are delivered by the internet in the future.<br>6. I always try to use online health consultations that are delivered via the internet in my daily life<br>7. I will continue to use online health consultations that are delivered by the internet frequently | [12] |
| **Technology readiness** | 1. Technology gives me more freedom of mobility.<br>2. I often keep up with the latest technological development that I am interested in<br>3. I can figure out new high-tech products and services without any help<br>4. I am usually among the first in my circle of friends to acquire new technology | [13]<br>[11] |
| **Price value** | 1. Online health consultations that are delivered by the internet are good value for the price<br>2. Online health consultations that are delivered are reasonably priced<br>3. The current price of online health consultations that are delivered by the internet provides a good value | [12] |
| **Facilitating conditions** | 1. I have the resources necessary to use the internet for online health consultation<br>2. I have the resources necessary to use online health consultations that are delivered by the internet<br>3. I have a permanent connection to the internet<br>4. I know to use the internet to access health care websites | [13] |
| **Actual of Use** | 1. Currently, I am using the internet for online health consultation<br>2. I have used the internet for online health consultation before<br>3. I use the mobile phone to get health care services<br>4. I use some websites to get health care services<br>5. I use online services to get health care services | [13] |
| **Effort expectancy** | 1. It is easy for me to become skillful at using online treatment technology<br>2. Learning to use online treatment technology easy for me<br>3. My interaction with the internet to get health care treatment is clear and understandable | [13] |

# Evaluating Agile Information-Based Framework for Flood Management Utilizing Metadata Concept to Support Flood Operation Activities

Preliminary communication

**Mohamad Firdaus bin Mat Saad**
School of Information Technology
SEGi College Penang
Penang, Malaysia
firdaus.ms@outlook.com

**Aliza binti Abdul Latif**
Department of Information System, School of Computing and Informatic
Universiti Tenaga Nasional,
Selangor, Malaysia
aliza@uniten.edu,my

**Marini binti Othman**
Academic Consultant
Selangor, Malaysia

***Abstract*** *– Operational policies are established to handle natural hazards including floods to minimize the effect with differing degrees of effectiveness and increasing relaxation. Sometimes policies are time-consuming of rigid protocols that are inadequate in a dynamic and somewhat chaotic environment synonymous with the complexity of flood disaster. Hence, this research aimed at recommending the incorporation of agile concepts in flood control, which would offer stability and adaptability in the control of the complex flood situation. Extensive reviews on flood management and existing frameworks for disaster management were conducted to understand the problems and the potential solution to construct an agile framework. A grounded analysis was conducted to obtain insight into how the agility of standard operating procedures could be enhanced. The agile components have been defined by contrasting characteristics from other effective disciplines, including software development and health care, that share common complexity in management environments. Consequently, an Agile Information-Based for Flood Management Framework is proposed in previous publication. The validation component for agile key-values presented in the earlier article is, however, absent. This study therefore presents the validation component from earlier publication on the Agile Information Based Framework. A theoretical evaluation of the proposed key-values for the agile framework has been conducted using the metadata concept. The evaluation identified the similarity feature in the same area where the proposed framework was agreed to be implemented in tandem with electricity company emergency response plan to improve flood operations. The proposed key-values in the agile framework are required to be adopted and further strengthened by other significant variables.*

***Keywords****: Agile concept, Flood management, Flood operation, Metadata Model, Framework*

## 1. INTRODUCTION

Managing floods is a complex process since every disaster including flood is unique in nature. Effective communication between emergency relief organizations, flood victims control, media pressure, time constraints, operational demand, technical and facilities limitations are the key factors that contributed to the complexity of the management process [1]–[4]. In such a situation, effective management methods need to be considered to achieve disaster management (DM) goals and objectives to increase flexibility, time to respond, and satisfaction of stakeholders (authorities, non-governmental organization (NGO), victims). In the area of systems development known as an agile management approach in a complex environment has indeed been adopted [5]–[8]. This included constantly evolving customer demands, time pressures, and customer satisfaction, which need to be addressed from time to time while promoting corporate and operational priorities. Taking into account the similarities of the complex characteristics in both environments, the agile concept is seen as a new approach in

flood management. In fact, there are studies conducted by scholars to adapt the agile concept in disaster management [9]–[12]. Therefore, the AgiLe inFormation-bAsed (ALFA) Framework for Flood Management has been introduced in previous publications where it details the development process of the proposed frameworks [18]. The concept of agility is presented in that paper, promising the flexible approach in managing the complexities during chaotic situations. However, there is a missing validation aspect of the framework. Hence, this paper purposely to present the validation process of the AFLA Framework for flood management to ensure the correctness, completeness, and relevance of the proposed key-values inside the framework to be generalized to the actual implementation.

Validation has been carried out to ensure all key-value introduced in the ALFA framework satisfy the requirements of the analysis and are therefore applicable to the real-world application. The main purpose of validation is to guarantee each key-value discussed by the researcher meets each domain of the study conducted. It is to demonstrate that perhaps the agile model can be applied to any management system if well designed. While the complex environment can adopt agile methods in the implementation process, the traditional, systematic, and knowledge-based management should be passionate about doing so. This research followed theoretical evaluation approaches that explain the definition of metadata to improve the dimension of the results for the proposed key-values in the ALFA framework. This is because the assessment made is dependent on the outcomes of the analysis carried out and then reviewed with the current facts. This theory directs the evaluator by defining the key elements of the framework and describes how these elements are structured to be connected [13]. According to S. Cojocaru, this theory is beneficial as it directs the evaluator by defining the key elements of the framework and describing how these elements are structured to be connected [13]. Other researchers claimed that the objective of a theory-based evaluation is to evaluate a model that is hypothesized to describe the program and the mechanisms used to produce the expected outcomes [14].

The key-values incorporated into this ALFA framework are confirmed by the proof from previous research, documentation, and even actual application of the SOP utilizing this theoretical assessment process. In this analysis, the evidence-based approach used is the official documentation involving the implementation of the main framework used at the international level, namely the Hyogo and Sendai frameworks [28] and the framework developed by local researchers (referring to Governance of Flood Disaster Framework (GFD) [36]) by following the causal type of the theoretical-driven evaluation to define and underlies the causal relationship between those frameworks. Therefore, to verify the suggested key-values in the ALFA framework, these key values have been tested with existing Standard Operating Procedure (SOP) or frameworks (referred to Hyogo and Sendai) and another similar framework (referred to GFD) developed in the field of research (focused on the sense of Malaysia). Taking into consideration the relationship between key values implemented through the ALFA framework by the metadata approach, it is expected that there would be a connection between the key values and the accessible facts.

This paper is organized as follows: in the second section, an overview of the works related to agile management will be elaborated. The third section will discuss the methodology used in validating the ALFA framework and followed with the definition of ALFA framework. The next section will provide the result of the validation performed on the framework. Concluding, we discuss the framework adoption with a sample disaster in an electricity company as well as future research of the framework.

## 2. METHODOLOGY

In this study, theoretical validation has been adopted by considering metadata elements presented in the ALFA framework. Metadata, according to E. Brodie is structured content that identifies, defines, finds, or otherwise facilitates the collection, usage, or management of an information resource [15]. Metadata is often referred to as data about data or information about information. The use of the word metadata varies according to the discipline of study [15]. Some of them use it to refer to understandable machine information, while others only use it to identify electronic services in documents. The metadata helps to provide appropriate and accurate information to get a real picture of resource quality [16]. In this case, a descriptive metadata approach has been adopted by the researcher to define resources. This method was used to identify elements such as title, abstract, keyword, or any other item that would explain how the resources are used. In resource collection, metadata performs the same purposes as successful cataloging does by (1) allowing resources to be found by relevant criteria, (2) identifying resources, (3) bringing similar resources together, (4) distinguishing dissimilar resources, and (5) giving location information [15]. At any level of aggregation, metadata may describe resources. A selection, a specific commodity, or a part of a larger resource can be represented. Hence, the usage of metadata for theoretical-based validation will account for data and information accuracy by providing the highest degree of outcome based on the theoretical methods adopted.

The theory of highly intuitive metaphor from the court of law was introduced using the outline of a witness' oath to ensure the consistency of information offered. By considering the statement takes in the court when witness swears to "… tell the truth (the correctness), the whole truth (the completeness), and nothing but the truth (the relevance)" [16]. All three elements from the theory; correctness, completeness and relevance were applied to validate the quality of the metadata provided in the ALFA framework. The quality of the metadata

should emerge when the consistent data are associated with the domain (correctness), the data is provided with adequate information regarding data contents (completeness), and the ability to apply the data to the actual implementation (relevance).

Thus, to ensure uniformity of data introduced within the same context, which is disaster management specifically to the flood control, the identification of consistent data within the existing framework, firstly was applied, to the Hyogo and Sendai, the GFD, and the ALFA frameworks, for validation purposes. Secondly, to ensure the completeness of data, the integration and consideration of metadata information from the Hyogo and Sendai, and the GFD framework would be clarified in depth. Next, to demonstrate the relevance of the described data, the data should be clarified and implemented based on current practices concerning flood control. Thus, in the next sections, the presenter has presented the discussion and finding regards to the validation of the ALFA framework based on the theoretical evaluation using metadata methods.

## 3. DEFINING ALFA FRAMEWORK

The ALFA framework is a work-based framework on agile elements adapted from system development to disaster management. Due to the complex management features in system development, which has shifted from traditional management to agile management to allow the achievement of objectives promptly. While developing effective management, the agile approach is seen as a new benchmark for other areas of complex management including flood [17]. Many studies have been carried out on the acceptability of the agile concept in disaster management [9]–[12] by identifying the information of the agile concept adopted in disaster management. This research was additionally adopted observation, interviews, and formal training methodologies, to collect information relevant to agile management.

Literature reviews were conducted on system development, and health care to identify agile elements adapted in implementation in these two domains. Interviews with electricity company (EC) General Manager and the District and Land Office Director were performed to understand flood management practices and to define the agile concepts implemented in current flood management. Besides, observations have been made in many aspects of flood control, including emergency centers, evacuation centers, early-warning systems, water level monitoring systems, moorings, and flood management portals. Formal training has also been included in data collection to understand the overall flood disaster management practiced in Malaysia, particularly on the EC and District and Land Office operations as were presented in previous publication [18].

Based on the previous works, published in [18], seven agile elements have been identified from the study conducted. Among them are 1) Quick response, 2) Transformational leadership, 3) Small project management, 4) Technology and innovation, 5) Coordination and communication, 6) Community Engagement, and 7) Practice and training in various disaster management (DM) aspect especially flood management as shown in Figure 1 [18]. All seven elements were analyzed using thematic coding methods based on the similarity of data obtained during the data collection [18]. The thematic analysis is the qualitative approach focused on the pattern under which the gathered information is evaluated in conjunction with the theme. The thematic analysis allows for a more precise theme in the way the different activities give significant advantages [19]. Therefore, the ALFA framework has been developed to coordinate flood management based on the agile elements found during data collection.

The analysis has then been explained by two main recommendations on successful flood management as proposed in the ALFA framework [18].

1. Certain elements have to take into account complex circumstances. It recognizes that the situation and the disaster context are unique, can vary and change with time. Adaptive approaches are therefore expected during implementation. In the proposed framework, under the 'agile process,' these elements are defined.
2. The strict SOP must be preserved (where flexibility is not permitted) to ensure the effective and organized complication of certain activities, for instance, were shutting down a substation operation. A sleek and transparent SOP is needed in every disaster organization. A structured SOP for this specific type of disaster management is also required to regulate the uniformity of the system applied.

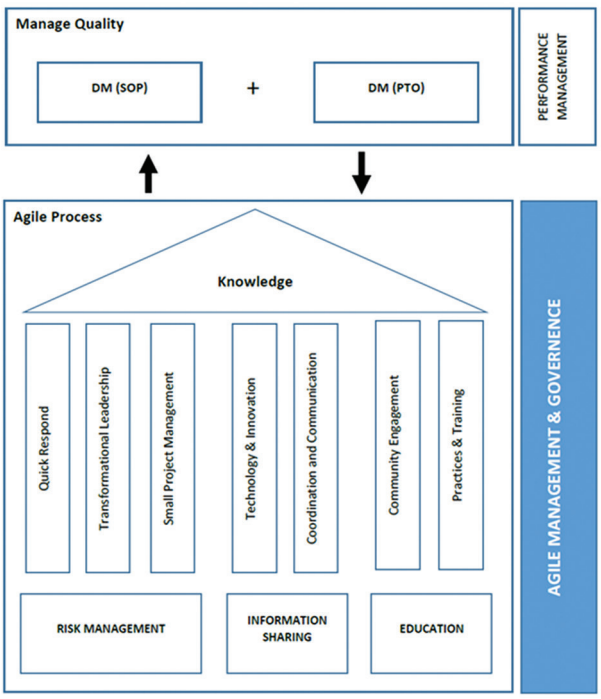



**Fig. 1.** Agile Information-Based Framework for flood management (The figure was reprinted from [18])

As a result, several aspects appear to be unproductive to an agile approach, for example, to define a Standard Operating Procedure (SOP) as a guideline in every management and administration practice. This rule is very specific to a limited part of the project. So, during implementation, these documentations could not adopt agility. However, there is a demand for post-evaluation or lessons learned from flood for future access and improvement to be documented and to use SOP for the smaller scope of activities for easy navigation. On the other aspect, agile key-values are introduced in the framework to manage flood, and the validation of the key-values would strengthen the finding. Hence the following section will present the validation phases of introduced agile key-values.

## 4. RESULT AND DISCUSSION

In this section, the researchers will present the validation process carried out on key-values established in the ALFA framework through the integration of the concept of metadata. Consistency of data in the actual implementation of the framework and management processes relates to disaster management are determined by the key-values. It is verified by the establishment of the relationship between the ALFA framework with the Hyogo & Sendai as well as the GFD framework. On top of that, the detailed information about the data has been explained by integrating the Hyogo, Sendai, and GFD framework to justify the data in ensuring the completeness of the introduced key-values. The importance of the agile key values to be widespread in real implementation has been explained to strengthen metadata quality identified using T.Margaritopoulos theory [16]. The merger of correctness, completeness, and relevance of the key-values is structured to ensure the consistency of the provided metadata is relevant within the research domain and has been presented in Agile information-based framework for flood management [18]. Table 1 below provides important details on how the Hyogo, Sendai, and GFD frameworks have a connection to the ALFA key-values. In this section, the following are significant.

**Table 1.** Mapping of Hyogo, Sendai, and GFD framework to the key-values proposed in the ALFA framework

| Key-Values introduced in ALFA Framework | Hyogo and Sendai Framework | GFD Framework | Relational to ALFA Framework |
|---|---|---|---|
| **SOP, Policy** | Both frameworks have adopted an emergency response plan. | GFD has emphasized designing and enforcing environmental management plans, initiatives, laws, and regulations for FD operations. | ALFA framework has introduced the use of SOP to be integrated with the ALFA framework. |
| **Knowledge** | Hyogo stressed the use of knowledge to construct a culture of protection and stability at all levels. | To enhance performance and competitiveness by maximizing capital, GFD applied resources management to people, competencies, and skills (refer to knowledge) in the managing of flood disasters. | The ALFA framework discusses the value of a strong understanding (knowledge) of emergency reduction. |
| **Quick Response** | Hyogo has highlighted DRR strategies integrated into priority 4 to reduce the underlying risk factor with adaptation to climate change. | None | The ALFA framework underlines the quick response, as it leads to disaster management's effectiveness. |
| **Transformational Leadership** | None | GFD emphasized the key point in which the public must know how to play an active role in the flood disaster. | The ALFA framework explains the value of transformational leadership to encourage more efficient disaster management. |
| **Small Project Management** | Hyogo highlighted raising the fundamental risk factor in balanced biodiversity and environmental management. | None | Throughout its more achievable form, small project management is highlighted in the ALFA Framework. |
| **Technology and Innovation** | Hyogo emphasized the advancement of technologies and the usage of technology (early warning, data sharing) as resources for crisis relief. | To promote interactions and knowledge exchange on flood control, GFD implemented nationwide flood system components to act as end-to-end IT infrastructures. | The ALFA framework involves technology and innovation because it has significance for improving emergency recovery plans. |
| **Coordination and Communication** | Hyogo and Sendai underlined the coordination, information sharing, and use of standard DRR terminology for effective response. | For efficient resource allocation in the FM Process, GFD has adopted the idea of collaboration and exchanging capital. | The ALFA framework emphasizes coordination and communication because of its beneficial qualities during the emergency recovery phase. |
| **Community Engagement** | Hyogo stressed community involvement as one of the important areas for developing a better DRR plan. | None | Community engagement in emergency recovery also leads to effective disaster management. The ALFA framework thus further outlines community engagement as a primary element. |
| **Practices and Training** | Hyogo stressed the importance of education, training, public awareness, review, and exercise plans in the area of preparedness and emergencies plan. | In cooperation with public and private organizations worldwide, the GFD introduced the idea of Education, Research, and Partnership (ERP). | The ALFA framework adapts practices and training owing to its capacity for effective disaster management. |

Table 1 presented the interconnection of the elements between Hyogo, Sendai, GFD, and ALFA frameworks. The consistency of the elements in the established (refer to Hyogo, Sendai) frameworks and similar (refer to GFD) framework, have strengthened the key-values proposed in the ALFA framework. The explanation below should clarify the completeness and relevance of key-values for actual flood operations.

The usage of SOP, and Policy in emergency recovery has been enforced formally by the primary disaster management agency, the National Security Council (NSC), which has now been carried up by the National Disaster Management Agency (NADMA) [20]. Directive No.20 is the primary emergency response guideline used in Malaysia [21], [22]. Every agency or organization has developed an internal SOP, in the meantime. EC has outlined the three SOPs for an emergency, including Corporate Emergency Respond Plan (CERP), and System and Non-system Emergency Respond Plan [23], [24]. For SOPs such as these, emergency recovery perhaps to be more easily managed. Nevertheless, it is argued whether the SOPs relative to the form of disaster should be more comprehensive, and the handling mechanism should also be narrower.

As important as SOP, and Policy implementation, it is critical to have a general understanding of an uncommon situation. Yet it is more important to provide a detailed understanding of the severity of issues. It is because knowledge has a significant part to play in deciding intervention. As a society, it is important to stress fundamental information regarding flood control, as it affects the living, properties, and everyday life. It helps the public to respond to issues that occur without waiting for orders or actions from the authorities, with the knowledge of flood control, because it impacts the life of the population. The enhancement of knowledge management and integration has been explored by many researchers in Malaysia in regards to disaster management [25]–[27]. Therefore, knowledge should be seen as an essential factor of flood control such that actions can be done following the issues occurring, the condition, the current circumstance, and the complexity of the issue that requires agile intervention.

Not to neglect, the importance of the quick response element in disaster management is viewed as a revolutionary approach that might have a successful influence on flood control. The Hyogo framework has used the concept since it was introduced in the year 2005 [27], [28]. However, the implementation remains disappeared. Adjusting the idea of a quick response when an incident occurs is viewed as an early step to mitigate the threats from expanding to certain regions. For some other cases, this quick response principle can be extended. Nonetheless, persons who can react effectively to a problem, have some main reasons, which is to save a life, to adjust rapidly to efficiently performed. A detailed understanding of the problem, experience in the decision-making phase, awareness about the data and information processing facilities, and other considerations that may influence the person to make a swift decision about the problem are some of the factors to be considered.

Another important aspect is the transformation leadership. This concept is an interactive viewpoint that encourages everyone to be a leader in flood control. This term is referred to as transformation leadership as it will affect flood control by offering experts who are not generally recognized (non-lead management staff) to handle specific flood operation activities. This approach enables everyone to manage the situations from their experience or knowledge in flood management. It gives both the local population and the middle or lower management the ability to serve as the leader in flood operations. This is not the concept of only top managers will make the decisions. The IC, often the general manager, is appointed as a director to take full responsibility for crisis management immediately in actual deployment cases of the EC. If the general manager is unavailable, the responsibilities go to the second higher-position workforce, followed by seniority.

Whereas the idea of small project management is to break the project into a limited size such that the operation can be efficiently carried out within a reasonable amount of time and resources. The complicated task can be easily dealt with provided it has been narrowed down to a reduced size of the operation [6]. EC defined three types of SOPs, including the Corporate Emergency Response Plan (CERP) and the System and Non-System Emergency Response Plan. The unique SOP designed by EC is one of the aspects to reduce operational activities related to flood management.

In any case, technology is a medium that enables certain items to be linked by technology-based services, such as discovering information by internet facilities, weather forecasting via the forecasting system, managing the modern aviation environment through global positioning systems, and many more. Technology is now used as a significant element for disaster management, including floods, as essential to management function. This will implicitly help the flood control mechanism and reduces the likelihood due to flood through the presence of technology. This can be demonstrated in recent findings by scholars on the usage of technologies to control disaster operations [29]–[31]. EC has used telephone and mobile phones for internal or external communications and the dissemination of information. The information should be conveyed by using the reachable devices as this is simpler in contrast to the usage of fax, message, and info blast. The same details would, therefore, be faxed to agencies as evidence that the knowledge is exchanged by EC and agencies. The purpose of the message and info blast is to remind the community (headman) of the current situation and the actions to be complied with by the potential victims (applied only at a certain station). Nevertheless, there are two independent views on the

involvement of social media to be used as a platform for channelling information to potential victims or agencies. However, the utilization of technology and innovation in channelling information, communicate with other stakeholders, and coordinate information and mission is the concept of agility where it may provide benefits to the flood control team.

Apart from the points discussed above, coordination and communication are the other important aspect of flood management as proposed in the ALFA framework. Coordination can be described as a collaboration between agencies or organizations in managing the activities related to the event, with the main goal of enhancing the quality of the operations [32]. To safeguard the effectiveness of flood control, cooperation between stakeholders is very critical. It involves cooperation between authorities, the public, and the victims of floods. Coordination of crisis management between the public and private sectors (locally, nationally, and internationally) is becoming increasingly relevant and effective. There are two kinds of coordination often mentioned: vertical and horizontal coordination. Vertical coordination is defined as the correlation between two or more organizations that share their duties, resources, and information to support similar end customers while horizontal coordination is based on internal communication, operational or group alliances, or mutual collaboration with competitors and non-competitors [32], [33]. On the other hand, the communication aspect during and after a disaster is a vital aspect of response and recovery, as it unites flood victims into touch with first responders, support systems, and other family members. For the survival of society, therefore, an accessible and reliable means of communication is very important. The consequences of the tragedy are not only for victims but also for families, friends, emergency responders, and care providers. The complexity of this significant incident has raised the concern of the need for more comprehensive and efficient communication and management approaches [34].

As important as the elements discussed above, a society is a group of people living or possessing the same characteristics in the same location. A society grows in depth when a collective of people who have specific characteristics relevant to social interaction, exchange common experiences and take cooperative action in an environment or geographical position [35]. To improve their capacity and capabilities to cope with disasters by themselves, society should be active and enable them to engage actively in any phase of a crisis management process. Also, community engagement in decision-making and active interest in developing an SOP is one means of strengthening emergency response strategies. This is because affected groups are the best markers of their vulnerability and can agree about their wellbeing adequately [35]. In the EC practices, local heroes were first formed to distribute information, expose disaster control activities and search and rescue, and all relevant emergency procedures to be taken out and carry out throughout the preparation phase in the context of an actual tragedy. The local heroes, whether the leader of the community or the interested representative (community participation) are listed. EC shall relay incident details through SMS or WhatsApp. Clear notice of potential disasters was issued to the local heroes to be vigilant or to act accordingly.

Lastly, practices and training are structured to ensure that everybody is completely educated and ready to respond as directions are given in any circumstance. Throughout the development of coordination, communication, and awareness between all rescue agencies, flood simulation is essentially needed. All rescue agencies have the same mission to save people from crises. This simulation is, therefore, necessary to let all organizations understand the operation of flood management, including local citizens' support. Even each organization has a different SOP, the key priority is to ensure that the emergency operations performed in the actual situation are carried out properly. This was intended that the flood program has been implemented and educated three times a year at EC-level, where EC is demanded to implement the appropriate emergency response plan protocol due to the critical operation during the flood involves the continuous supply of the electricity to the non-inundated area while off system to affected areas. However, the use of genset or mobile genset may serve as an alternative for the continuous operation of electricity in the inundated areas.

All the key principles presented in the agile sense have been inferred to be accepted frameworks. This is focused on metadata found from main frameworks used locally and internationally for disaster management (referred to in the Hyogo and Sendai) and also in the sense of flood governance (referred to in the GFD) developed by researchers in Malaysia. The plurality of core principles shared in an ALFA framework by scholars contains parallels from all areas of the frameworks. Nonetheless, key-values addressed in the suggested ALFA framework, which are not underlined by GFD are explored within the Hyogo and Sendai context and vice-versa.

## 5. CONCLUSION AND RECOMMENDATION

In a conclusion, the ALFA framework is not the same as the current framework that companies, governments, or academics in various fields have utilized or published. It is the emerging agile framework, a structure that can be tailored to any framework or SOPs for disaster management (refer to performance management in the AFLA framework) including the emergency response plan under EC and GDF framework. ALFA framework promises the operational aspects of flood management that can be adaptive in the application phase depending on the actual situation. As mentioned above where all the agile concepts introduced in the ALFA framework are tailored to the real flood management scenario. This is intended to offer an indi-

cation of incorporating the suggested ALFA framework into the actual implementation. The proposed framework also provides an agile management philosophy in implementation or operation activities either before, during, or after the flood, depending on the suitability to be adopted with the current SOP.

However, a thorough analysis of the management cycle of an agile dimension in other complex environments such as hospital management, airplane accident, building collapse, and some other relevant fields will strengthen the finding of agile key-values as proposed in the ALFA framework to be implemented as a directed process. On the other hand, the implementation of the ALFA framework over different types of geographical and topology will provide better information and improvement to the proposed framework. Hence, the framework demands the implementation from a variety of structures to improve the content and implementation.

## 6. REFERENCES:


[1] B. B. R. Turner, „Balancing Agility and Discipline: A Guide for the Perplexed", 7th ed. Uniten States: Pearson Education, Inc, 2009.

[2] M. . Dorasamy, M. . Raman, and M. . Kaliannan, "Evaluating CEMAS in simulated environment to support disaster management challenges", Proceedings of the11th International Conference on Information Systems in Crisis Response Management, 2014, pp. 444–453.

[3] C. Pathirage, K. Seneviratne, D. Amaratunga, R. Haigh, "KNOWLEDGE FACTORS AND ASSOCIATED CHALLENGES FOR SUCCESSFUL DISASTER KNOWLEDGE SHARING Prepared for the Global Assessment Report on Disaster Risk Reduction 2015", No. January 2014, 2014.

[4] M. Yu, C. Yang, Y. Li, "Big data in natural disaster management: A review", Geosciences, Vol. 8, No. 5, 2018.

[5] M. Hneif, S. Hock ow, "REVIEW OF AGILE METHODOLOGIES IN SOFTWARE", International Journal of Research and Reviews in Applied Sciences, Vol. 1, No. 1, 2009, pp. 1–8.

[6] K. N. Rao, G. K. Naidu, P. Chakka, "A Study of the Agile Software Development Methods , Applicability and Implications in Industry", International Journal of Software Engineering and Its Applications, Vol. 5, No. 2, 2011, pp. 35–46.

[7] H. Hajjdiab, A. S. Taleb, "Adopting Agile Software Development : Issues and Challenges", International Journal of Managing Value and Supply Chains, Vol. 2, No. 3, 2011, pp. 1–10.

[8] K. Pathak, A. Saha, "Review of Agile Software Development Methodologies", Advances in Computer Science and Information Technology, Vol. 3, No. 2, 2013, pp. 270–276.

[9] A. I. Nawaz, I. A. Zualkernan, "The Role of Agile Practices in Disaster Management and Recovery : A Case Study", Proceedings of the Conference of the Center for Advanced Studies on Collaborative Research, November 2009, pp. 164–173.

[10] K. A. Abdelouhab, D. Idoughi, C. Kolski, "Agile & user centric SOA based service design framework applied in disaster management", Proceedings of the 1st International Conference on Information and Communication Technologies for Disaster Management, Algiers, Algeria, 24-25 March 2014.

[11] L. L. Salvadó, M. Lauras, T. Comes, B. Van de Walle, "Towards More Relevant Research on Humanitarian Disaster Management Coordination", in Procedings of the ISCRAM, 2015.

[12] J. G. Brown, A. Chennamaneni, "Towards an Integrated Framework for Applying the Agile Project Methodology to Manage Task Uncertainty in Disaster Management", Proceedings of the Americas Conference on Information Systems, 2013, pp. 1–7.

[13] S. Cojocaru, "Clarifying the theory-based evaluation", Revista de Cercetare si Interventie Sociala, Vol. 26, 2009, pp. 76–86.

[14] G. Sharpe, N. Bay, "A Review of Program Theory and Theory-Based Evaluations 100 College Drive 1 . Purpose of the Paper 3 . When to develop a program theory 4 . Components of a program theory", American International Journal of Contemporary Research, Vol. 1, No. 3, 2011, pp. 1998–2001.

[15] E. Brodie, "Understanding M.E.", Nursing Times, Vol. 84, No. 31, 1988, pp. 48–49.

[16] T. Margaritopoulos, "A Conceptual Framework for Metadata Quality Assessment Merkourios Margaritopoulos", Proceedings of the International Conference on Dublin Core and Metadata Applications, 2008, pp. 104–113.

[17] G. Becker, D. Huitema, J. C. J. H. Aerts, "Prescriptions for adaptive comanagement: The case of flood management in the german rhine basin", Ecology and Society, Vol. 20, No. 3, 2015.

[18] M. Firdaus, M. Saad, A. A. Latif, M. Othman, "Agile information-based framework for flood management", Proceedings of the Annual Conference on Computer Science and Engineering Technology, , Medan, Indonesia, 23 September 2020.

[19] M. Maguire, B. Delahunt, "Doing a Thematic Analysis : A Practical , Step-by-Step Guide for Learning and Teaching Scholars",Reflections, Journeys and Case Studies, Vol. 3, No. 3, 2017.

[20] N. O. Chong, K. H. Kamarudin, "Disaster risk management in Malaysia: Issues and challenges from the persepctive of agencies", Planning Malaysia, Vol. 16, No. 1, 2018, pp. 105–117.

[21] Major Mohd Sakri Hussin, "Disaster Management (NSC Directive No.20) A Malaysian Perspective by National Security Division Prime Minister Department", 2015.

[22] A. R. Badruddin, "Issues of Disaster Management Preparedness : A Case Study of Directive 20 of National Security Council Malaysia", International Journal of Business and Social Science, Vol. 3, No. 5, 2012, pp. 85–92.

[23] Tenaga Nasional Berhad, "STATEMENT ON RISK MANAGEMENT This Statement on Risk Management and Internal Control has been Malaysia ' s Listing Requirements and Risk Management and Internal", 2013.

[24] N. H. Din, Datuk Ir. Baharin; Abd Kadir, Hjh Kamaliah; S., Parameswaran; Ibrahim, Ruslinda; MZ Halim, "After the Storm. Improved Safetr Measure in Hand, TNB Stads Ready for Future Floods", TenagaLink, 2015.

[25] M. N. Ahmad, M. Othman, N. H. Zakaria, M. Z. Mohd Rodzi, "Managing information and knowledge in Malaysia's flood management: Towards a new framework", Frontier in Artificial Intelligence Applications, Vol. 265, 2014, pp. 446–463.

[26] Y. Ali, A. Mohammad, N. Ahmad, N. Hidayati, "Knowledge Sharing Framework for Disaster Management", Journal of Information Systems Research and Innovation, Vol. 9, 2015, pp. 50–60.

[27] H. D. M. T. A. Wahab, "Malaysia - National progress report on the implementation of Hyogo Framework for Action (2009-2011)", Malaysia, 2011.

[28] UNISRD, "Hyogo Framework for Action 2005-2015 : *", 2005.

[29] E. Nations, ICT for Disaster Risk Reduction. 2010.

[30] J. Wilson, F. Wilson, J. Wilson, "The use of ICT in Disaster Risk Management : A Case Study of Nema Borno State", Journal of Remote Sensing GIS & Technology, Vol. 5, No. 1, 2018, pp. 44–66.

[31] H. Yuliandoko, Subono, S. H. Pramono, P. Siwindarto, "Innovation technology for disaster management", International Journal of Recent Technology and Engineering, Vol. 8, No. 1, 2019, pp. 396–398.

[32] M. Bahadori, H. R. Khankeh, R. Zaboli, I. Malmir, "Coordination in Disaster: A Narrative Review", International Journal of Medical Reviews, Vol. 2, No. 2, 2015.

[33] R. Kaynak, A. T. Tuğer, "Coordination and Collaboration Functions of Disaster Coordination Centers for Humanitarian Logistics", Procedia - Social and Behavioral Sciences, Vol. 109, 2014, pp. 432–437.

[34] R. Moorthy, G. Benny, S. S. Gill, "Disaster communication in managing vulnerabilities", Malaysian Journal of Communication, Vol. 34, No. 2, 2018, pp. 51–66.

[35] H. A. Rahman, "Community Based Approach Towards Disaster Management in Malaysia", Asian Journal of Environment, History and Heritage, Vol. 2, No. 2, 2018, pp. 55–66.

[36] S. Maidin, M. Othman, "Governance of the Flood Disaster Framework in Malaysia: A Way Forward in Enabling Information Technology Knowledge Sharing", Journal of Advanced Research in Dynamical and Control Systems, Vol. 11, No. 1, 2019, pp. 1533–1541.

# INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

## About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

## Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

## Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

## Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

## Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

### Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

## Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

## Copyright



## Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

## Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

# IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

**TITLE OF ARTICLE** (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

**Author/Authorized Agent** **Date**

**CONTACT**
**International Journal of Electrical and Computer Engineering Systems (IJECES)**
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr

International Journal of Electrical and Computer Engineering Systems
ISSN 1847-6996