



Support Vector Machine Prediction a Man in the Middle Attack on Traffic Networking

Nahla Ibraheem Jabbar

Authors affiliations:

1) Department Chemical Engineering, University of Babylon, Iraq.
eng.nahla.ibraheem@uobabylon.edu.iq

Paper History:

Received: 1st Aug. 2024

Revised: 23th Oct. 2024

Accepted: 15th Jan. 2025

Abstract

The goal of the study is to predict the Man in the Middle attack in the packets of Wireshark program by using Support Vector Machines (SVM). In the time of using the internet, it has become a tool targeted by attackers and hackers; it is a serious threat to the devices. A uniqueness of an attack that appears in multiple identities for legitimate agencies. It is very necessary to know the behavior attack and predict the possible actions of an attacker. In this research a detection of Man in the Middle attack by monitoring the Wireshark program and recording any changes can be recognized in packet information. The classification of packets is divided into two categories (normal and abnormal). The proposed model is designed in many stages: loading data, processing data, training data, and testing data. The detection of SVM based on abnormal network packet through movement packets in the Wireshark program that needs to deal with current packets to recognize a new attack that one does not have prior knowledge of its detection, and there is a need for an intelligent way to separate network packets that represent normal. The proposed approach achieved an accuracy of 97.34% in detecting attacks. The results show that the proposed model effectively visualizes attacker behavior from data that represents abnormal network attackers. Research achieves successful accuracy in predicting abnormalities.

Keywords: Computer Network, Clustering, Support Vector Machine, Man in the Middle attack.

دعم التنبؤ بألة المتجهات لرجل في منتصف الهجوم على شبكات المرور
نهلة ابراهيم جبار

الخلاصة:

إن الهدف من الدراسة هو التنبؤ بهجوم Man in the Middle باستخدام Support Vector Machines. في زمن استخدام الإنترنت، أصبحت أداة يستهدفها المهاجمون والمتسللون؛ فهو يشكل تهديدا خطيرا للأجهزة. تفرّد الهجوم الذي يظهر في هويات متعددة للوكالات الشرعية. من الضروري جدًا معرفة سلوك الهجوم والتنبؤ بالتصرفات المحتملة للمهاجم. في هذا البحث تم الكشف عن هجوم Man in the Middle من خلال مراقبة برنامج Wireshark. يمكن التعرف على أي تغييرات في معلومات الحزمة. قمنا بتصنيف الحزم إلى فئتين (عادية وغير طبيعية). تم تصميم النموذج المقترح على عدة مراحل: تحميل البيانات، معالجة البيانات، بيانات التدريب، واختبار البيانات باستخدام أجهزة ناقل الدعم (SVM). يتم اكتشاف SVM بناءً على حزم الشبكة غير الطبيعية من خلال حزم الحركة في برنامج Wireshark الذي يحتاج إلى التعامل مع الحزم الحالية للتعرف على هجوم جديد ليس لدى المرء معرفة مسبقة باكتشافه، وهناك حاجة إلى طريقة ذكية للفصل حزم الشبكة التي تمثل وضعها الطبيعي. حقق الأسلوب المقترح دقة بلغت 97.34% في كشف الهجمات. أظهرت النتائج أن النموذج المقترح يتصور بشكل فعال سلوك المهاجم من البيانات التي تمثل مهاجمين غير طبيعيين على الشبكة. يحقق البحث دقة ناجحة في التنبؤ بالتشوهات.

1. Introduction

Computer networks are systems of interconnected devices that allow users to share resources, such as files and printers, and enable communication via email or

messaging [1][2]. Data is transferred across a network via packets, which are small units of information sent and received by devices. Using the Internet and exchanging information is subject to attack and transmission by viruses. Network threats and



intrusions pose major concerns for daily activities in a computer network [3][4]. It is considered one of the most difficult tasks for analysts and security officials within local networks. Many systems and programmers are deployed to eliminate hacks. When the system provides the following security characteristics: confidentiality, integrity, and availability, it is classified as a secure system. Artificial Intelligence (AI) creates computer systems and applications that perform tasks in different ways for computer networks [5]. Artificial Intelligence (AI) creates computer systems and applications that perform tasks typically associated with mental intelligence. Machine Learning (ML) is one of the most common applications of AI in classification traffic [6]. Many researchers have developed using AI to detect attacks. In [7] researchers applied SVM and ANN to classify the normal and abnormal packets in the traffic network they achieved successful classification in the rate of 94.02%. Deep Learning were applied in detecting DDoS attack [8]. In [9] [10] attacked are recognized by SVM and DL. The achievement occurs in reduced dimension of data base by features extraction that improvement in time process. AI plays an important role in the cybercrime that attacks an organization's network. Cybercriminals can spend a lot of time in a business network without being noticed. They can launch ransomware, eavesdrop on meetings, propagate malware and establish privileged accounts to access other systems, and/or attack larger corporations during this time. Due of AI's capacity to learn and make predictions about the past, present, and future, cybercriminals can use it to their advantage. Support Vector Machine (SVM) for short is a technique used in many applications within the fields of Artificial Intelligence and machine learning. SVM is a technique used in many applications within the fields of [11]. SVM is used in neural computer systems as a tool to determine the separation of data in a given data set. When new data comes in, the SVM method can be used to identify it and share it between two different classes. The researchers aim to develop SVM models that can be used in facial recognition, vital signs, disease treatment, and prediction on datasets. According to a study presented [12], we have to clarify how to detect intrusion resulting from a network attack [13][14]. A machine-learning system has been developed to classify network traffic. The cloud is considered one of the most important things to provide security for due to the fears that threaten it. One of the most important developments in detecting attacks in networks are SVM and ML [15][16]. They are used to detect the most common types of cloud attacks. In [17], support vector machines are used to detect fraud in the network by encrypting its secure execution. It was a safety data cloud for bank systems. Myo Myint Oo et al [18] applied SVM to detect Distributed Denial of Service (DDoS) in software-defined networking (SDN), and the accuracy of detection reached 97%. Through the use of the Wireshark program [19], our research attempts to detect Man in the Middle attacks (MIMT) using Support Vector Machines (SVM) on traffic networks. Various features are taken from Wireshark and

employed in prediction, based on SVM. In clustering technology testing and training divided into groups of normal and abnormal. The algorithm is capable of detecting attacks and distinguish Man in the Middle attacks in a real time system. In section 4 explains more details about the research like collection data, processing and classification. In section 5 are explained results of the research.

2. Man-In-The-Middle attack

The term cyber-attack or electronic attack refers to the hacker's attempts to intentionally cause damage and harm. The electronic attack is launched by a group of professionals or amateurs in the world of programming to enter site and make a group of changes or modifications, cause damage, or hack it. The goal of the attack varies depending on the goal of the hacker. There are different types of cyber or electronic attacks, through which attackers seek to achieve special goals that differ from one attacker to another and according to the nature of the attack and the targeted party. The following is a list of the most prominent types of cyber-attacks: Phishing, Malware and Man- in- the- Middle [20]. When someone positions themselves between two computers such as a laptop and a distant server and intercepts the traffic, it's known as a man-in-the-middle (MITM) attack [21]. This person can then eavesdrop on or even intercept the conversations between the two devices and steal the data. Man-in-the-middle attacks are a major security risk, but the most obvious way for anyone to get one is to connect to an encrypted public Wi-Fi network, such those at coffee shops or airports. With a free program like Wireshark, an attacker can be record every packet sent across network. After that, they can examine and pinpoint facts that might be helpful. Network- If successful, DNS spoofing is a similar technique that involves mapping the victim's hardware A Media Access Control (MAC) address to another person's IP address.

3. Support Vector Machine

Support Vector Machine (SVM): After analyzing the dataset, SVM is a kind of supervised machine learning method that is used to categories all the data in the dataset [22]. The support vector machine algorithm is based on the idea of converting original data into a new pattern in a way that distinguishes the data in space higher than its original space, through which we can separate and classify data. Categorization issues in which a maximally separated hyperplane is built. A hyperplane is a linear pattern whose maximum margin determines how far apart the decision classes can be from one another [23]. given data set $\{x^i, y^i\}^N$, where N number of samples, $x_i \in RD$ denotes the feature vectors from sample- i , y_i denotes the class labels, and D is the number of features (dimension). The multiclass classification issue, where $y_i \in 1, 2, \dots, k$, where k is the number of classes, is different from the two-class classification problem, where $y_i \in \{-1, +1\}$. Finding the optimal hyperplane is the primary goal of SVM. To find the nearest point on the hyperplane to the origin, maximizes x as it appears on the hyperplane.

We have a similar situation for the other side points. The total distance from the dividing hyperplane to the nearest points is obtained by solving and subtracting the two distances. Maximum Margin ($M/||w|| = 2$).

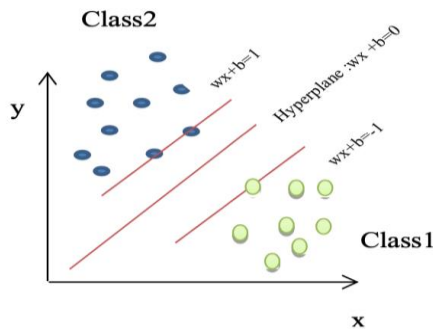


Figure (1): SVM separation two class1 and class2

Efficiency applications of SVM are succeed in image retrieval of face recognition [24] or classification text assignment, Natural Language Processing (NLP) and others types of classifiers. SVM as the machine learning technique that performs best in multi-label classification [25].

4. Methodology

In a computer network, devices are connected via wired or wireless connections, such as Ethernet, Wi-Fi, or cellular data cables. Data is transferred across a network via packets. Packets are defined as information sent and received by devices. These packets are routed across the network to their destination, where they are reassembled and processed. These packets are subject to hacking or attack during their passage. Man-In- The- Middle is one of the attacks infected packets. In our method, we employ the Wireshark program which is used in packet analysis or packet sniffing and is carried out by a packet sniffer to detect MIMT in SVM. It captures raw network traffic as it passes network, it contains the attacking web servers. The following fundamental elements form the basis of the research:-

4.1 Data Collection

It is a library of information provided by the Wireshark program that is collected from the Wi-Fi of café which distribute data between different customers. Information about the Wireshark as shown in Figure.2. Packet extracted from the program. We focus on a frame as in the following Figure.3

We find the Ethernet which lead to source and destination MAC address. There is more than option to inform MAC address by using filter. Data are distributed between different customers. The general important information includes IP addresses source and destination, type protocol, and time. We focused our research on the IP.

Addresses of the source and destination, as well as the MAC address of the IP in the type of protocol. Filters can be used by taking advantage of Wireshark's program. one of the most filter utilized in the packages that were captured and includes the selected IP address with the MAC address of the destination. It is suitable for individuals who wish to focus on a single MAC address in the entire traffic. For a period of time,

whose files data are taken to capture network packets in real- time, extract the descriptors or attributes contained within the packets and store data in collection of files that ready for analysis.

4.2 Improved Data Set

It performs the analysis and decision making process and represents a predictive model composed of the SVM method in the first stage to select features or attributes according to the type of data traffic in the network. It is a stage of removing huge amount of duplicate data and performing a preprocessing data by deleting duplicate and incomplete data. This stage of performs a modelling process to prepare data in appropriate measurement of the data that has a significant impact on the prediction variable. As Internet protocol a result of this stage, we produce a modified data set (scaled data set) that contains numeric values.

4.3 Analysis Data set

The research aims to use fewer features in Wireshark program to detect attacks. All features are mentioned above about IP addresses of source and destination, check the source and destination fields from the traffic you have captured, and then match the physical address of them. Depending on this information to see which frames computer delivered or received. Wireshark is monitored for MIMT by the MAC address, which has two different IP addresses. This is a contradiction in the result, leading to suspicion of the existence of a MIMT attack. After extracting important features that were specified in each file uploaded to the SVM, the training model is used to form a training dataset. Loading the data set (the training dataset) into the classifier (SVM) that was chosen to classify the data and load the testing dataset into the prediction method.

5. Results

We represented an approach to detecting Man- in- the- Middle attack in the network packets based on an support vector machine (SVM) algorithm as a method for learning and extracting information available in Wireshark features. Abnormal packets or victim packets are one of the main issues while using and installing Wireshark. These can arise from a variety of attacks on Wireshark packets, which could explain why so many packets appear. Strange anomalies in packets, illustrates the abnormality through the attributes of the attack type in Table 1.

Table (1): Describe how the data set and anomaly

Data set	Abnormal packets
10000	Double-tagging attack
25000	Direct Attacks
40000	IP unreachable attack
50000	DDoS attack

We omit other anomalous packets since we have focused on our research on the identification of a MIMT attack that centralizes MAC addresses. The scaled data set is divided into two databases namely the training data set with a data size of 70% and the testing data set with a data size of 30%. To train and test the model, Finding the link between dependent and



independent variables is done using the training set, and analysis is done using the testing set. A model's performance group of experiments done on packets of data belonging to normal and abnormal. The experiment includes 1000 packets from Wireshark which obtained accuracy detection of MITM is very weak as shown in Figure (4). Increasing data set for training and testing with a data size of 55000 records, the accuracy improved to reachable 78% that shown in Figure.5. From 60000 to 75000 an obvious jump in the accuracy of prediction, it is reached to 97.3% in

Figure.6. The training process for two types of (normal and abnormal) in this method of learning possible to test within real-time systems and known the time of detection time. The rate of normal and abnormal packets is different according to normal packets and abnormal. Normal packets are easily recolonized from other types of abnormal. We chose special type of abnormal not mentioned others types abnormal in the research. Among the normal approaches, the proposed approach achieved an accuracy of 97.34% in detecting attacks.

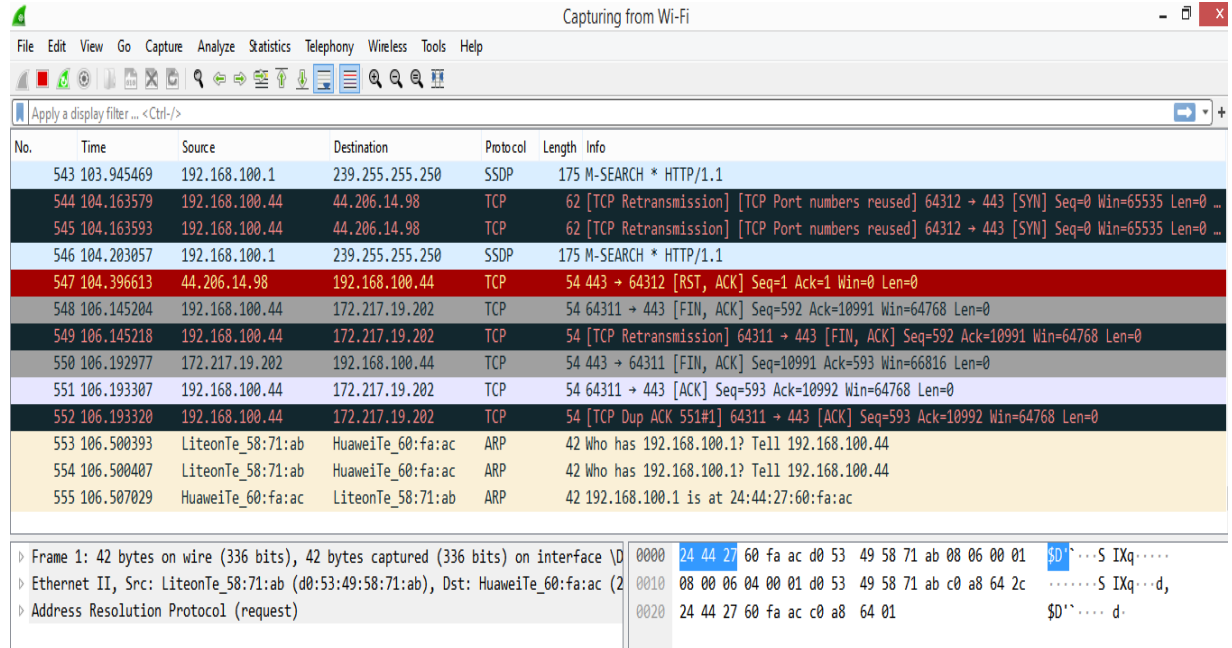


Figure (2): Wireshark program

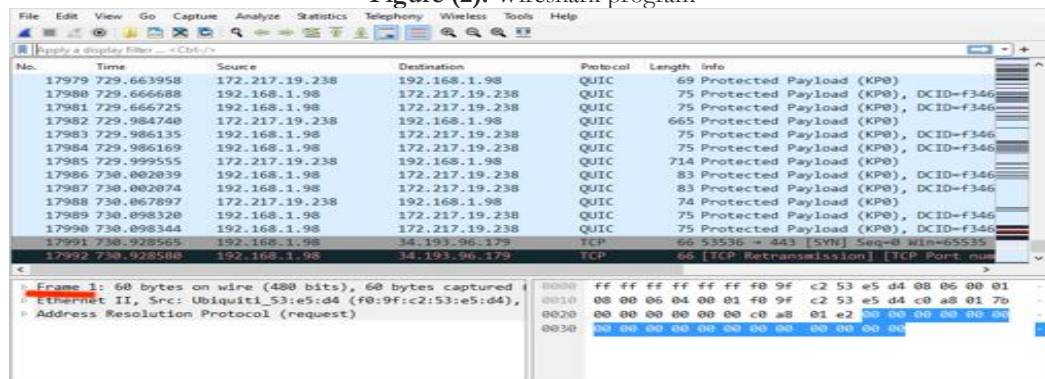


Figure (3): Shows Frame in Wireshark

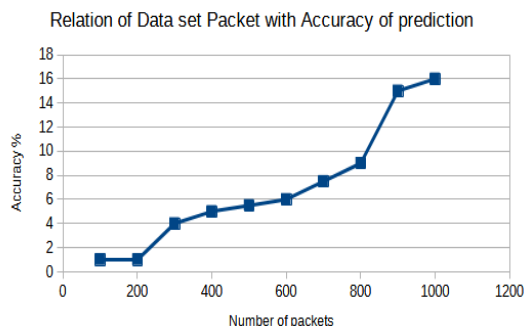


Figure (4): Relation number of Data (1000) packets with Accuracy

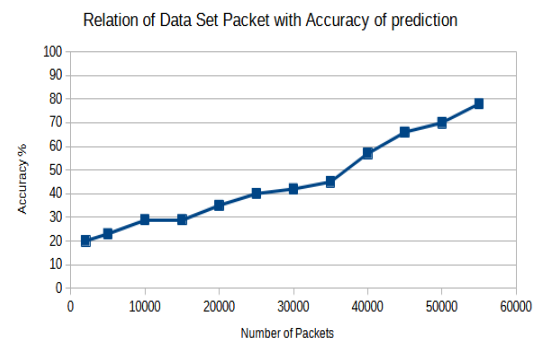


Figure (5): Relation number of Data (5500) packets with Accuracy

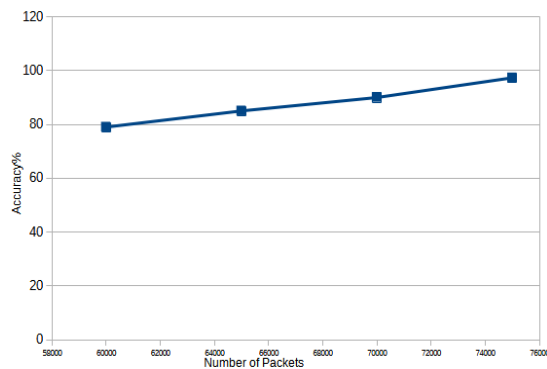


Figure (6): A Relation number of Data (6000-7500) packets with Accuracy

6. Conclusions

In this work, a detection Man in The Middle attack system is based on a Support Vector Machine that operates in real-time data information from a Wireshark program. The packets series consists of a local network that operates in the manner of Wi-Fi, which leads to a high speed of packets response, in addition to the fact that the data is analyzed locally within the Wireshark program. Descriptors were specified in each file unfolding in a specific number depending on the type or nature of network data traffic, and a decision is made about the availability of an attack or not at the level of each packet, which detects the attack on the network, and the results of the analysis are sent to files only using a special Internet protocol with a different MAC address. The process Prediction of the packets across variant types of MAC addresses. The proposed approach also has the ability to predict dynamically without affecting the operation of the Wireshark program, by detecting within any new network that is added to the network topology through the use of the framework, which contributes to knowing the attack in the process of detection, analysis, and decision-making Man In the Middle attack into (normal and abnormal). Two important measures should be taken in the extraction of packets: At the first is waiting time for capture packet infects and a scheduling of the values of IP. At the second is according to the experiments presented, which are used in the classification process at the level of training data using the support beam of the SVM machine, it is possible to test it within real-time systems to know the detection time; and finally, the proposed approach achieved an accuracy of 97.34% in detecting attacks. Preprocessing data collection achieves successful accuracy in prediction. In future works will extend this work by proposing Deep Learning in prediction types of attack occurs in the current packets of network.

7. References:

- [1] S. Nabilah, S. Novita, N. Isra', A. Maria, and S. Dahlan, "Computer networking," *Int. J. Comput. Netw. Inf. Secur.*, Jan. 2023.
- [2] A. Sumit and A. Anshul, "An introduction to computer networking," *Int. J. Comput. Sci. Inf. Technol. Res.*, vol. 2, no. 2, pp. 373–377, Apr.–Jun. 2014.
- [3] M. Mimi and L. Satishkumar, "Mitigation of DDOS and MITM attacks using belief based secure correlation approach in SDN-based IoT networks," *Int. J. Comput. Netw. Inf. Secur.*, vol. 1, pp. 52–68, 2022.
- [4] K. Ruzaina and H. Mohammad, "Network threats attacks and security measures: A review," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 8, Sep.–Oct. 2017.
- [5] J. Tanya, "Artificial intelligence in computer networks," *Period. Eng. Nat. Sci.*, vol. 10, no. 1, pp. 309–322, Jan. 2022.
- [6] N. Evgenii and V. Vitalii, "Application of network traffic using machine learning methods," *Int. Sci. Pract. Conf.*, Dec. 2023.
- [7] U. Pranjal and S. Perina, "Classification network attack detection using machine learning," *Int. Res. J. Eng. Technol.*, vol. 8, Apr. 2021.
- [8] A. Mahmood, Al-Shareeda, M. Selvakumar, and S. Murtaja, "DDoS attacks detection using machine learning and deep learning techniques: Analysis and comparison," *Bull. Electr. Eng. Inform.*, vol. 12, no. 2, pp. 930–939, Apr. 2023.
- [9] A. Mohammad, H. Russul, T. Shatha, A. Ahmed, M. Mostafa, and S. Tole, "Distributed denial of service attack defense system-based auto machine learning algorithm," *Bull. Electr. Eng. Inform.*, vol. 12, no. 1, pp. 544–551, Feb. 2023.
- [10] G. Toufik, B. Mohamed, and P. Purnendu, "Automated diagnosis of attacks in internet of things using machine learning and frequency distribution techniques," *Bull. Electr. Eng. Inform.*, vol. 10, no. 2, pp. 950–961, Apr. 2021.
- [11] J. Weston, S. Mukherjee, and L. Pontil, "Feature selection for SVMs," *Adv. Neural Inf. Process. Syst.*, vol. 13, pp. 668–674, 2000.
- [12] M. Zerina, K. Dino, D. Nejdert, and Kemal, "Flood attack detection in cloud computing using support vector machine," *TEM J.*, vol. 6, pp. 752–759, Nov. 2017.
- [13] S. Muhammad, N. Shah, and Y. Xiangzhan, "Identification of attack traffic using machine learning in smart IoT networks," *Secur. Commun. Netw.*, Apr. 2022, doi: 10.1155/2022/9804596.
- [14] N. Vivine and M. Zuriani, "Improving sentiment reviews classification performance using support vector machine-fuzzy matching algorithm," *Bull. Electr. Eng. Inform.*, vol. 12, no. 3, pp. 1817–1824, Jun. 2023, doi: 10.11591/eei.v12i3.4830.
- [15] N. Mukrimah, A. Amiza, Y. Naimah, and L. Ong, "Effective and efficient network anomaly detection system using machine learning algorithm," *Bull. Electr. Eng. Inform.*, vol. 8, no. 1, pp. 46–51, Mar. 2019, doi: 10.11591/eei.v8i1.1387.
- [16] A. Raveendra and G. Gurumoorthi, "Cloud-based machine learning algorithms for anomalies detection," *Indones. J. Electr. Eng. Comput. Sci.*,



- vol. 35, no. 1, pp. 156–164, Jul. 2024, doi: 10.11591/ijeecs.v35.i1.pp156-164.
- [17] Vazquez-Saavedra, J. Jimenez, Loureiro-Acuna, Fernandez-Veiga, and Pedrouzo-Ulloa, “Homomorphic SVM inference for fraud detection,” *Ongoing Research*, 2019.
- [18] T. Mohammed, “Advanced support vector machine (ASVM) based detection for distributed denial of service (DDoS) attack,” *Softw. Defined Netw.*, 2022, doi: 10.56294/dm202272.
- [19] J. Vinit, *Wireshark Fundamentals: A Network Engineer’s Handbook to Analyzing Network Traffic*, Apress, 2022, doi: 10.1007/978-1-4842-8002-7.
- [20] B. Andreea, “Cyber-attacks – trends, patterns and security countermeasures,” *Procedia Econ. Finance*, vol. 28, pp. 24–31, Dec. 2015, doi: 10.1016/S2212-5671(15)01077-1.
- [21] B. Bhushanm, G. Sahoo, and A. Raj, “Man-in-the-middle attack in wireless and computer networking: A review,” *Int. Conf. Adv. Comput. Commun. Autom.*, Sep. 2017.
- [22] P. Silvana, S. Phil, Z. Valentina, and P. Aleksandar, “Predicting bidding price in construction using support vector machine,” *TEM J.*, vol. 5, no. 2, pp. 143–151, May 2016, doi: 10.18421/TEM52-04.
- [23] S. Pongsametrey and T. Nguonly, “Support vector machine (SVM) based classifier for Khmer printed character-set recognition,” *APSIPA Annu. Summit Conf.*, 2014, doi: 10.1109/APSIPA.2014.7041823.
- [24] S. Anton, U. Wendi, B. Arif, and H. Khalid, “Content based image retrieval and support vector machine methods for face recognition,” *TEM J.*, vol. 8, no. 2, pp. 389–395, May 2019, doi: 10.18421/TEM82-10.
- [25] C. Yange, M. Qinyu, W. Baocang, P. Duan, Z. Benyu, and H. Zhiyong, “Privacy-preserving multi-class support vector machine model on medical diagnosis,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 7, pp. 3342–3353, Jul. 2022, doi: 10.1109/JBHI.2022.3157592.