Reviewing the effectiveness of artificial intelligence techniques against cyber security risks

Mohammed. I. Alghamdi
Department of Computer Science, Al-Baha University, Al-Baha City, Kingdom of Saudi Arabia

ABSTRACT

The rapid increase in malicious cyber-criminal activities has made the field of cybersecurity a crucial research discipline. Over the areas, the advancement in information technology has enabled cybercriminals to launch increasingly sophisticated attacks that can endanger cybersecurity. Due to this, traditional cybersecurity solutions have become ineffective against emerging cyberattacks. However, the advent of Artificial Intelligence (AI) – particularly Machine Learning (ML) and Deep Learning (DL) – and cryptographic techniques have shown promising results in countering the evolving cyber threats caused by adversaries. Therefore, in this study, AI's potential in enhancing cybersecurity solutions is discussed. Additionally, the study has provided an in-depth analysis of different AI-based techniques that can detect, analyse, and prevent cyber threats. In the end, the present study has also discussed future research opportunities that are linked with the development of AI systems in the field of cybersecurity.

Keywords: Artificial Intelligence; cybersecurity, cyber threats; cyber-attack, machine learning

Corresponding Author:

Author Name, Mohammed. I. Alghamdi
Departement, Department of Computer Science
University, Al-Baha University
Address, Al-Baha City, Kingdom of Saudi Arabia
E-mail: mialmushilah@bu.edu.sa

1. Introduction

Technological evolution has resulted in the development of such a system that can think and act like a human. In particular, with the introduction of new technologies, such as Artificial Intelligence (AI) and Machine Learning (ML), significant advancements have been made to ensure fool-proof security of the IT systems against cybersecurity risks [1]. Cybersecurity can be understood as the set of systems, human behaviours, and processes that assists in protecting online and electronic resources. Additionally, it is a fast-evolving discipline that has grasped the attention of practitioners because of the increment in the number of cyber-attacks and threats that endanger sensitive data. However, despite the effectiveness of AI and ML in cybersecurity, cybercriminals have also been using AI to launch increasingly complex cyberattacks while simultaneously hiding their tracks [2, 3]. Nonetheless, the fact remains that AI techniques in cybersecurity have the potential to alter the existing asymmetric adversary-versus-defender balance. Moreover, with the integration of AI on the defensive side, certain forms of defence, such as malware and spam detection could be improved [4]. Besides, it is also established that traditional solutions of cybersecurity are becoming ineffective, in terms of adequately identifying and mitigating cyberattacks. However, advancements in AI and cryptographic techniques are found to provide promising results to cybersecurity experts, in terms of countering cyber threats. To fully comprehend the potential of AI, this paper aims at reviewing different AI-based techniques that can be used to counter cybersecurity risks.

2. Related work

The recent increase in identity theft and cyber-attacks have made the internet a daunting place. Such attacks can severely affect society as today's communication and economic infrastructure heavily depend on IT and computer networks. Over the decades, researchers have discovered various types of cyber threats that could potentially jeopardize information security. According to a recent report [5], denial of service (DoS), eavesdropping, and malware attacks are among the most dangerous cybersecurity threats that can compromise digital security. In a DoS attack, the adversary attempts to clog the computing resources of a victim's computer by sending a substantial number of requests. According to [6], such attacks can be conducted in different ways. For example, one single attacker machine can overwhelm a victim's machine by transmitting a substantial number of network packets – which often appears to be legitimate. This activity is usually done to bypass network security. On the other hand, an adversary can also use multiple machines to launch simultaneous attacks in a distributed-style. Likewise, eavesdropping sniff through the network communication line to misuse obtained data[7, 8]. Malicious attackers either listens to the message transmission to detect information, which is often referred to as passive eavesdropping. Besides, attackers might also actively collect information by sending multiple friendly queries, which is called active eavesdropping. Another cyber threat that can jeopardize the cybersecurity is the malware attack. Studies like [2, 9] and [10]examined the phenomenon and proclaims that malware attacks spread automatically through the network by exploiting unknown or known vulnerabilities. Such attack not only threatens the confidentiality and integrity of individual computers, but it is also capable of bringing down any server via Distributed Denial of Service (DDoS) attack. These attacks severely affect the safety and security of the digital realm. To counter these security threats, multiple AI-based solutions have been proposed to classify and intelligently analyse cyber-attacks automatically. Some of the fruitful techniques are thoroughly reviewed and discussed in the proceeding manuscript.

2.1. Machine learning

Conventionally, ML techniques can be categorised into two types: unsupervised and supervised learning. In unsupervised learning, no training or data labelling is required. Instead, the algorithm, according to [2], automatically determines the degree of dispersion/coherence among the sample data while systematically creating classes. However, in supervised learning data samples are labelled according to different classes, e.g., legitimate or malicious. It is important to note that data labelling or training is manually performed and require humans to detect and analyse data patterns. The use of ML techniques is highly effective in detecting malicious activities. In this account[11], proposed a hybrid Support Vector Machine (SVM) approach for developing intrusion detection system (IDS). The study utilised two feature techniques of reducing security attacks: BPSO and Information Gain to detect DoS attack. The study further reported 99.4% classification performance on the DoS attack. Likewise, studies like [12] and [13] used Index Partial Distance Search k-nearest Neighbour (IPDS), and K-Means clustering and KNN classification to detect various types of attacks. The studies report 99.6% and 99.8% accuracy in NSI-KDD dataset. Galal et al. [14], on the other hand, used three supervised learning algorithms, including Decision Tree, SVM, and Random Forest to detect malware from the given dataset. The study proclaims that their proposed model achieved an accuracy of 97.19% in detecting malware attacks. This means that ML methods are powerful techniques that can be used to ensure robust security against cyberattacks.

2.2. Artificial neural network

Another technique that makes use of AI is the Artificial Neural Network (ANN). ANNs can be defined as statistical models that imitate the function and structure of the human brain. ANN has the capability of learning problems, particularly in environments where rules and algorithms for solving a problem are either unknown or difficult to express. In cybersecurity, ANNs can be used for observing network traffic [15]. As shown in Figure 1, ANNs analyse the past network activities to detect an actual attack during the delivery phase. As compared to traditional cybersecurity techniques, the great benefit of ANNs is their ability to learn the cybersecurity environment. This could be understood by considering the study of Chandrasekhar and Raghuveer [16, 17] who integrated SVM, ANN, and fuzzy C-means clustering to develop an intrusion detection model that predicts network intrusion. The findings of the study indicate that the proposed algorithm achieved the detection accuracy of 98.99% for PROBE, 99.66% for DoS, 98.81% for U2R, and 98.99% for R2L attacks. Shenfield et al. [18], on the other hand, used ANN for deep packet inspection-based IDS. The study proclaims that their model achieved an accuracy of 98% in distinguishing between malicious and benign network traffic. Dahl et al. [19] used random projections to train neural networks in detecting automatically generated malware. With this model, the authors were able to achieve classification result with an error rate of 0.49%. This indicates that ANN is a powerful AI technique that is widely used for detecting cyber-attacks.

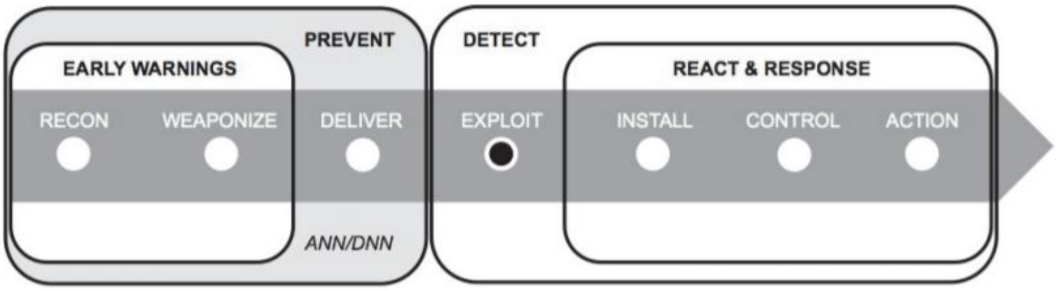

Figure 1. ANNs preventing network attack [15]

2.3. Deep belief network

Deep Belief Network (DBN) can be defined as a probabilistic generative model that make use of multiple layers of hidden and stochastic variables. This method is effective in detecting cybersecurity threats because it can effectively work in a random environment and a random network. Ding et al. [20], in their study, used DBNs to detect malware attacks. According to the study, DBNs utilised unsupervised learning to find multiple features layers and then use it in a feed-forward neural network (NN) fashion to fine-tune, particularly for the optimisation of discrimination. As a result of the unsupervised pre-training algorithm, DBNs were found to be less susceptible to over-fitting. This is also apparent in the study, as DBNs achieved an accuracy of 96.1%. Nadeem et al. [21] used NN with semi-supervised learning for obtaining optimal accuracy. The study used KDD Cup 99 dataset to detect the non-labelled data via Ladder Network and then used DBN for data classification. The study proclaims an average accuracy of 99.18%, which is similar to supervised learning. Zhao et al. [22], on the other hand, proposed an IDS based on probabilistic neural network (PNN) and DBN to solve an existing problem in traditional IDS, including huge data volumes, long-term training, and redundant information. The study used KDD CUP 99 dataset for testing the performance and obtained an accuracy of 99.14%. This means that DBN is an efficient AI model that is capable of detecting malicious activities within a network.

2.4. Convolutional neural network

Convolutional Neural Network (CNN) is a widely used AI-based technique whose network structure is similar to that of the biological neural network [23]. As a deep learning (DL) architecture, CNN is a multi-layered sensor, as presented in Figure 2. It is particularly designed to identify two-dimensional shapes, which are highly invariant to tilting, scaling, translation, or other deformation forms. To utilise the full effectiveness of CNN against cybersecurity risk, the study [24, 25]designed a dilated convolutional autoencoders (DCAEs) that make use of CNNs and stacked auto-encoders. In essence, the proposed model has the ability to learn crucial features from more-varied and large-scale unlabelled raw traffic network data consisting of web-based malware, botnets, advanced persistent threats (APTs), exploits, and scans. The study proclaims that their model achieved the detection accuracy of 99.59%. Likewise, [26] proposed a one-dimensional CNN based on end-to-end encrypted traffic classification technique. The method combines feature classifier, feature selection, as well as feature extraction into an end-to-end framework that intelligently learns the non-linear connection between the expected output and the original input. The study proclaims that their proposed method shown excellent performance in 2-class classification with 99% and 100% accuracy for VPN and non-VPN traffic, respectively. Thus, CNN is capable of detecting potential cybersecurity threats with great precision and accuracy.

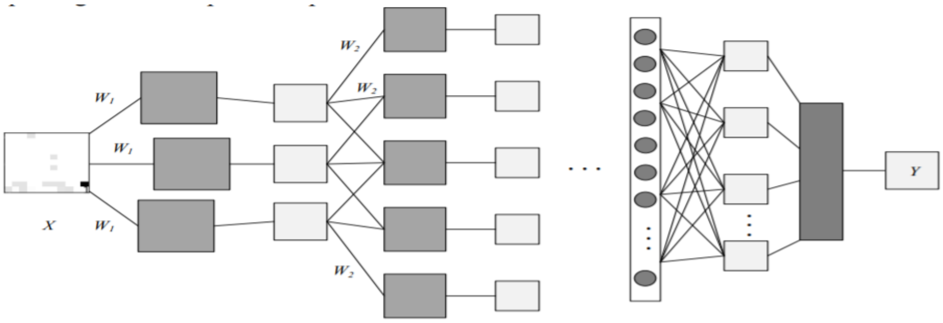

Figure 2. Convolutional neural network (CNN) structure [19]

2.5. Decision tree

The decision tree is a method that uses training data samples to create a set of rules. In the structure of a decision tree, each branch represents a test output while each node and leaf node represent a test on one property and a category, respectively [2]. As depicted in Figure 3, the decision tree categorises the data samples via the training conditions; thus, it has a better detection accuracy for cyberattacks. The study [27] proposed decision tree-based IDS for NSL-KDD dataset that uses Correlation Feature Selection (CFS) for evaluating features. Using this algorithm, the study reported an accuracy of 96.65% in detecting cyberattacks. Another study [28]proposed an IDS based on the genetic algorithm and C4.5 decision tree. This system not only reduces the overall false-positive rate but also contributes to improving the classification accuracy and resolves the small separation problem of the decision tree. The results show better performance when compared to renowned cybersecurity systems, such as Reptree, Naïve Bayes, and Random tree with an accuracy of 99.89%. Thus, it is apparent that AI-based techniques are highly effective against cybersecurity risks.

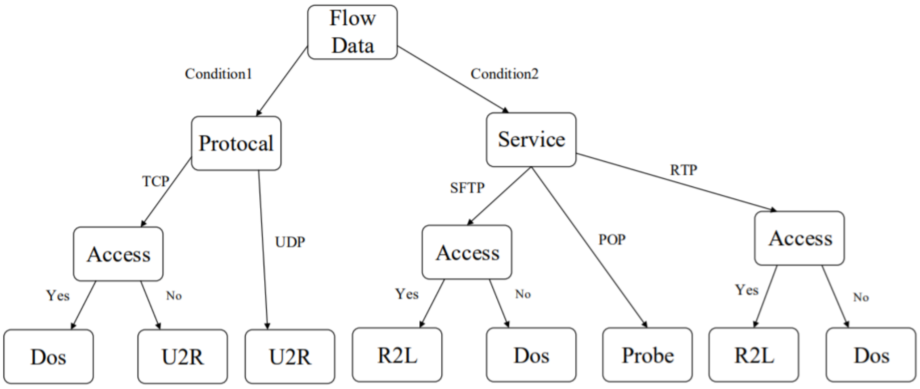

Figure 3. Decision tree example [19]

3. Results and discussion

Based on the existing body of literature, AI-based techniques, such as ML, decision tree, DBN, ANN, and CNN are capable of detecting various cybersecurity threats, including eavesdropping, malware, DoS, and DDoS attacks. While reviewing the effectiveness of AI techniques, it has been identified that ML-based techniques such as SVM, K-nearest neighbour, and decision tree are widely used in the development of IDS. Studies like [29]and [30]used SVM to detect cyber-attacks like DoS and achieved an accuracy of 99.8% and 99.3%, respectively. This accuracy is consistent in studies [11]and [16] which, similar to previous studies, used SVM to achieve a detection accuracy of 99.4% and 99.66% on the DoS attacks, respectively. The results from [13] and [31] provide evidence that k-nearest neighbour is equally capable of analysing cybersecurity threats with an accuracy of 99% and 94.17%, respectively. While ML algorithms are highly effective, DL-based algorithms, such as CNN are also consistent in threat detection. Studies like [32] and [33] used CNN to design an algorithm that detects malware attacks. According to the studies, their proposed methods achieved 93% precision and 99.41% accuracy, respectively. This indicates that AI techniques improve security practices that can defend against cyberattacks and cybercriminals while protecting valuable assets.

4. Discussion

AI has the potential to change the dynamics of cybersecurity. Initially, cyberattacks were recognized with rule-based systems, which uses attack signatures to identify the attack. However, with new technological trends, cyberattack strategies have become increasingly sophisticated, and therefore a need for advanced technologies and tools which can help investigate, detect, and make optimal decisions have increased [2]. Today, AI-based cybersecurity solutions are being utilised to automate the attack detection while improving their capability to detect over time. Besides, AI techniques like decision trees use network flow rate, duration, and size to identify DoS attacks [34]. Furthermore, different ML-based techniques such as SVM can categorise multiclass data into the attack and legitimate class. These features make AI an attractive technology that can be used to examine network traffic patterns for potential cybersecurity risks. Moreover, in cybersecurity, ANN can recognize a zero-day attack as it learns from the patterns of recent incidents. For instance, traffic network patterns obtained from recent DoS or eavesdrop attacks can be fed to ANNs as the training data, so that neurons can modify their weight to detect future DoS attacks [35]. This indicates that AI techniques are suitable for cybersecurity applications and can learn from previous attack patterns to predict future incidents.

5. Conclusion

As the sophistication and speed of attack increases, artificial intelligence has become an essential technology in the cybersecurity domain. The present research article has highlighted how cyber threats have evolved, increased in complexity, and using modern technology to jeopardize security. The research has provided in-depth insight into the types of cyber threats that increases cybersecurity risk. Besides, with the advancement of technology, cyberattacks will continue to rise, even if the cybersecurity community develops solutions to counter these threats. In the present research, different AI-based techniques, including ML, ANN, DBN, CNN, and decision trees, have been reviewed. The present study also provided evidence that supports the use of these technologies. In contrast to the traditional rule-based cybersecurity solutions, AI-based frameworks are robust, flexible, and adaptable; thereby, contributes to improving security execution and protection against an increasing number of cyber threats.

6. Future work

The present research examined a significant number of academic cybersecurity studies based on AI. However, almost all studies have given less consideration to actual deployment efficiency and most experiments were conducted in the lab, which does not provide an actual picture of detection efficiency in the actual network. Therefore, it is important to consider real-world statistics and performance metrics in future research. Moreover, current datasets contain an unbalanced number of categories, redundant information, and old data [19]. Thus, establishing balanced attack categories, network intrusion detection (ID) datasets with a substantial amount of data, and dynamic type coverage becomes a top priority in future work. In addition, while access to the modern, high-tech computing infrastructure will help AI experts to solve AI-related problems with efficacy and efficiency, the fact remains that analysing the substantial amount of data requires state-of-the-art computing platforms and resources [2]. This problem, however, can be resolved by using cluster computing solutions, including Hadoop and Apache Sparke. In the future, the introduction of quantum computing will also ease solving sophisticated computing problems with great precision and accuracy; thereby, making it a potential future research domain. Future work should also explore the possibility of hybrid solutions, as it can potentially increase the reliability of AI techniques.

Refrences

[1] V. D. J. A. a. S. Soni, "Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA," 2020.

[2] S. Zeadally, E. Adi, Z. Baig, and I. A. J. I. A. Khan, "Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity," vol. 8, pp. 23817-23837, 2020.

[3] I. A. Aljazaery, H. T. S. Alrikabi, and M. R. J. i. Aziz, "Combination of Hiding and Encryption for Data Security," vol. 14, no. 9, p. 35, 2020.

[4] M. Brundage *et al.*, "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation," 2018.

[5] I. Al-Barazanchi, S. A. Shawkat, M. H. Hameed, and K. S. L. Al-Badri, "Modified RSA-based algorithm: A double secure approach," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 6, pp. 2818–2825, 2019.

[6] T. Mahjabin, Y. Xiao, G. Sun, and W. J. I. J. o. D. S. N. Jiang, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques," vol. 13, no. 12, p. 1550147717741463, 2017.

[7] I. Abdulshaheed, H. R., Yaseen, Z. T., Salman, A. M., & Al_Barazanchi, "An Evaluation study of WiMAX and WiFi on Vehicular Ad-Hoc Networks (VANETs)," *IOP Conf. Ser. Mater. Sci. Eng. Pap.*, vol. 3, no. 12, pp. 1–7, 2012.

[8] O. H. Yahya, H. Alrikabi, I. A. J. I. J. o. O. Aljazaery, and B. Engineering, "Reducing the Data Rate in Internet of Things Applications by Using Wireless Sensor Network," vol. 16, no. 03, pp. 107-116, 2020.

[9] N. S. Alseelawi, E. K. Adnan, H. T. Hazim, H. Alrikabi, and K. Nasser, "Design and Implementation of an E-learning Platform Using N-Tier Architecture," 2020.

[10] L. J. J. o. I. S. E. Jaramillo and Management, "Malware Detection and Mitigation Techniques: Lessons Learned from Mirai DDOS Attack," vol. 3, no. 3, p. 19, 2018.

[11] H. Saxena and V. J. I. J. o. C. A. Richariya, "Intrusion detection in KDD99 dataset using SVM-PSO and feature reduction with information gain," vol. 98, no. 6, 2014.

[12] R. J. E. o. I. S. Rao and Technology, "Wavelet transforms," 2002.

[13] H. Shapoorifard and P. J. I. J. C. A. Shamsinejad, "Intrusion detection using a novel hybrid method incorporating an improved KNN," vol. 173, no. 1, pp. 5-9, 2017.

[14] H. S. Galal, Y. B. Mahdy, M. A. J. J. o. C. V. Atiea, and H. Techniques, "Behavior-based features model for malware detection," vol. 12, no. 2, pp. 59-67, 2016.

[15] M. Taddeo, T. McCutcheon, and L. J. N. M. I. Floridi, "Trusting artificial intelligence in cybersecurity is a double-edged sword," pp. 1-4, 2019.

[16] V. Chamola, V. Hassija, V. Gupta, and M. J. I. A. Guizani, "A Comprehensive Review of the COVID-19 Pandemic and the Role of IoT, Drones, AI, Blockchain, and 5G in Managing its Impact," vol. 8, pp. 90225-90265, 2020.

[17] H. T. Alrikabi, A. H. M. Alaidi, A. S. Abdalrada, and F. T. J. I. J. o. E. T. i. L. Abed, "Analysis the Efficient Energy Prediction for 5G Wireless Communication Technologies," vol. 14, no. 08, pp. 23-37, 2019.

[18] S. A. Shawkat, K. S. L. Al-Badri, and I. Al Barazanchi, “Three band absorber design and optimization by neural network algorithm,” *J. Phys. Conf. Ser.*, vol. 1530, no. 1, 2020.

[19] G. E. Dahl, J. W. Stokes, L. Deng, and D. Yu, "Large-scale malware classification using random projections and neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 3422-3426: IEEE.

[20] Y. Ding, S. Chen, and J. Xu, "Application of deep belief networks for opcode based malware detection," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 3901-3908: IEEE.

[21] M. Nadeem, O. Marshall, S. Singh, X. Fang, and X. Yuan, "Semi-supervised deep neural network for network intrusion detection," 2016.

[22] G. Zhao, C. Zhang, and L. Zheng, "Intrusion detection using deep belief network and probabilistic neural network," in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, 2017, vol. 1, pp. 639-642: IEEE.

[23] Y. Xin *et al.*, "Machine learning and deep learning methods for cybersecurity," vol. 6, pp. 35365-35381, 2018.

[24] X. Yu, J.-C. Shen, J. Zhang, and K. B. J. I. J. o. S. T. i. S. P. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," vol. 10, no. 3, pp. 485-500, 2016.

[25] H. T. S. Al-Rikabi, *Enhancement of the MIMO-OFDM Technologies*. California State University, Fullerton, 2013.

[26] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 2017, pp. 43-48: IEEE.

[27] B. Ingre, A. Yadav, and A. K. Soni, "Decision tree based intrusion detection system for NSL-KDD dataset," in *International Conference on Information and Communication Technology for Intelligent Systems*, 2017, pp. 207-218: Springer.

[28] C. Azad, V. K. J. I. J. o. C. N. Jha, and I. Security, "Genetic algorithm to solve the problem of small disjunct in the decision tree based intrusion detection system," vol. 7, no. 8, pp. 56-71, 2015.

[29] E. Hodo, X. Bellekens, E. Iorkyase, A. Hamilton, C. Tachtatzis, and R. Atkinson, "Machine learning approach for detection of nontor traffic," in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, 2017, pp. 1-6.

[30] M. V. Kotpalliwar and R. Wajgi, "Classification of Attacks Using Support Vector Machine (SVM) on KDDCUP'99 IDS Database," in *2015 Fifth International Conference on Communication Systems and Network Technologies*, 2015, pp. 987-990: IEEE.

[31] I. Al Barazanchi, H. R. Abdulshaheed, M. Safiah, and B. Sidek, “Innovative technologies of wireless sensor network : The applications of WBAN system and environment,” *Sustain. Eng. Innov.*, vol. 1, no. 2, pp. 98–105, 2020.

[32] B. Kolosnjaji, G. Eraisha, G. Webster, A. Zarras, and C. Eckert, "Empowering convolutional networks for malware classification and analysis," in *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 3838-3845: IEEE.

[33] W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, "Malware traffic classification using convolutional neural network for representation learning," in *2017 International Conference on Information Networking (ICOIN)*, 2017, pp. 712-717: IEEE.

[34] S. P. Sahu and S. Verma, "Secured and authentic communication by combined approach of digital watermarking and steganography," in *International Conference on Advances in Communication, Network, and Computing*, 2011, pp. 595-599: Springer.

[35] A. Saied, R. E. Overill, and T. J. N. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks," vol. 172, pp. 385-393, 2016.