Attack Detection And Classification In Iot Network Using Machine Learning

Priya Tamrakar[1], Prof. Swati Soni[2]
[2] Prof,
[1, 2] Takshila Institute of Engineering & Technology, Jabalpur, M.P.

Abstract- *With the rapid development of internet of things (iot) devices, the frequency and intensity of cyber attacks is increasing. Recently, Denial of Service (DoS) and distributed denial of service (DDoS) attacks have been reported to be the most common attacks against iot networks. Firewalls, intrusion detection systems, and traditional security solutions cannot detect DoS and DDoS attacks because they usually filter both and block traffic according to the rules listed first. But these solutions can be effective and efficient when combined with artificial intelligence-based technology.*

In recent years, deep learning models, especially neural networks, have received great attention due to their excellent performance in image processing. The capability of this convolutional neural network (rnn) model can be used to identify complex DoS and DDoS and other attacks. Therefore, in this study, we propose a method for Improvement of attack detection performance on the internet of things network with Multilayer Perceptron and analyzing network data containing negative data and training the MLP state model.

In case of dual deployment, the plan achieves 98.37% accuracy in DoS and DDoS detection. In addition the proposed method achieved lower values for RMSE in identifying various DoS and DDoS attack patterns, which is lower compared to the latest technology.

Keywords- DoS, DDoS, IoT, IoT Security, IoT Devices, Machine Learning, Neural Network, RNN.

I. INTRODUCTION

Ss The Internet of Things (IoT) is a network of interconnected sensors and actuators is connected to the internet allowing them to send and receive data from the authenticated devices. Some major advantages of using IoT is we can empower our computers to gather information about environment without depending on human and by processing the data received we can decrease the effort, loss, and cost. The Internet of Things allows for communication between the physical world and the digital world. The digital world interacts with the physical world through sensors and actuators. These sensors collect data that must be stored and processed in secure environment. The Data processing can take place at the edge of the network or at a remote server or cloud. Internet of Things offers various benefits to organizations which encourage companies to have different approaches to their business and gives them necessary tools to business growth. It is most widely used in manufacturing, transportation, home automation industries, agriculture, infrastructure and .IoT provides in field of agriculture by measuring physical quantity like humidity, soil content, and temperature, automate forming techniques with the help of sensors. IoT can be used in home automation to manipulate and monitor mechanical and electrical systems in a building which would help to reduce waste and energy consumption on a broader scale in cities. IoT can be used in various industries, including businesses within healthcare, finance, retail and manufacturing. These are the benefits which lead to mass consumption of IoT device in the world right now. In IoT threats, security requirements, challenges, and the attack vectors pertinent to IoT networks [1]. This proposed architecture involves the creation of vulnerability scanner tool for IoT device which can be run in the system to detect vulnerability. The configuration of the IoT devices , which can lead to Brute force attack and with the help of attacks script we are able to find out if there are any open ports which could be used to exploit the system and also perform Dos & DDoS attack to check if the system is vulnerable for it. Once the scan is complete a detailed report is sent to the authenticated user Email.

The principle of the Internet of things depends on the abundance of many vital components, which are smart devices supporting the Web to harness them to process data, sensors, and communication devices of various kinds to collect data from the private environment and transmit it to its beneficiaries. Connect all internet devices to individual sensors to attract data and web-supported tools. Possible to track the movement of the individual daily over the Internet and find out what he is doing. Moreover, obstruct his work and get information through DDOS attacks. The simple example of this method is to continue pressing the ENTER button on a terminal that has not yet logged in to the login network to a particular type of IWAN or workstations. The reason that this method cans a denial-of-service attack method

is that the input button often initiates a routine to identify the tool within the operating system. This habit is usually of high implementation priority. By continuing to press this button, there is a high demand for the processing needed to identify the tool (the keyboard in this case), resulting in 100% of the processor's power be consumed and unable to receive additional processing requests. This causes paralysis in the operating system, which does not usually have the intelligence to distinguish between legitimate entry requests and abusive entry requests. In this case, there is no mechanism to respond to this attack. Another method of this type of attack is the targeting of other fixed resources in infrastructure, such as SYN dumping attacks.

TABLE 1 includes some obvious vulnerability for IOT layer.

IOT layer	Vulnerability
Physical Layer.	Malicious code injection.
Software layer	DDOS attack
Network layer	Traffic Analysis

TABLE 1: The most common types of vulnerability IOT layer [2].

II. LITERATURE REVIEW

2.1 DDoS Attack

Denial of service (DoS) attacks and distributed denial of service (DDoS) attacks have been reported as the most common attacks on IoT devices and network [3]. A DoS attack is a malicious attempt done by an attacker using a single source to make a service or network resources inaccessible to legitimate users. When a DoS attack is launched using multiple distributed sources, it is called a DDoS attack. The DoS and DDoS attacks are increasing rapidly both in frequency and intensity with an average of 28.7K attacks per day [4], [5]. Recently, Neustar's report of cyber threats and trends [6] revealed that the DDoS attacks have been increased 200% in frequency while 73% increased in volume during the first six months of 2019 as compared to the same period in 2018 [7]. Fig. 1 depicts the surging trend of DDoS attacks as anticipated in Cisco's annual Internet report, 2018–2023 [8]. It can be observed that by 2023, the total count of DDoS attacks would become double, i.e., 15.4 million as compared to 2018. Hence, there is a crucial need for developing such solutions which can effectively detect and devastate the DoS and DDoS attempts. So far, firewalls, intrusion detection systems (IDS) and intrusion prevention systems (IPS) are used as major security shields to protect the IoT devices and network from the cyber attacks. However, the traditional firewalls, IDS and IPS cannot defend against the complex DDoS attacks, [9]–[11] as most of them filter the normal and suspicious traffic based upon the static predefined rules. However, the IDS and IPS that filter the intrusive attempts using artificial intelligence (AI) techniques are more reliable and effective as compared to the static predefined rules.

The traditional IDS use signatures or deep packet inspection (DPI) techniques for detecting malicious activities in the network. These techniques filter the packets based upon the packet content and header information. Unfortunately, such techniques have poor performance and become a bottleneck when deployed on high bandwidth and high-speed backbone links [9]. Moreover, these techniques fail to check packet contents when the encrypted traffic flows over the network [12]. Although many machine learning (ML) based solutions have been proposed for IoT attack detection, the prediction power of a well-tuned deep learning model especially convolutional neural network (CNN) is much better and effective as compared to the ML models [13]. During the past few years, deep residual network (ResNet) drastically captivated the attention of researchers due to its tremendous performance.

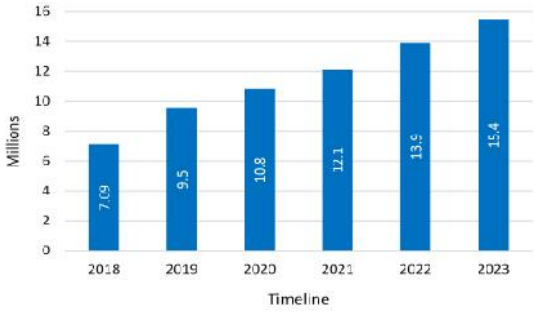

Figure 2.1: IoT Network Attacks.

No matter, the deep learning models especially CNN models have achieved high significance to their efficient performance in image processing and computer vision field. However, these CNN models are also being used for detecting the network attacks. Liu et al. [12], proposed a CNN-based approach to detect the malicious traffic from NetFlow data. The authors first encoded the features then applied feature correlation and converted the data into images through surrounding correlation matrix. Finally, they fed these generated images to the deep learning models. Among these models, residual network (ResNet) [14] outperformed the other models. Likewise, Salman et al. [15] devised a framework for IoT device identification and attack detection. The authors used a self-generated dataset of seven IoT devices and evaluated the processed framework using two machine learning and three deep learning networks out of which a machine learning model, i.e., Random Forest outperformed. The authors in [15] revealed that ResNet [14] is prone to

overfit in case of low dimensional and small size dataset due to which ResNetbased IDS do not perform well. To combat this challenge, the authors reconstructed the ResNet model by simplifying the residual block. The experiments proved that the simplified ResNet performed better as compared to the actual ResNet for low dimensional data.

The authors in [16] claimed that CNN best performs on images while the network traffic datasets are in nonimage form. In order to efficiently use the potential CNNs for detecting the network intrusions, the authors proposed a methodology to convert the network traffic into a three dimensional (3D) image. For this, the authors used a publicly available dataset, i.e., NSL-KDD dataset, applied fast Fourier transformation (FFT) onto it, converted it into 3D images and then passed it to a state-of-the-art CNN model to detect the network intrusions. Likewise, Li et al. [17] converted NSL-KDD dataset feature values into binary vectors using a binary encoding scheme then transformed these vectors into images. These images were fed into two deep neural networks. The authors concluded that CNN models show better performance as compared to the machine learning methods. Although the potential of CNN models is being used for developing intrusion detection systems, these CNN models do not perform efficiently when trained on non-image dataset. Hence, there is a need for developing such a mechanism that transforms the network traffic into a representable form on which CNN models perform efficiently. Usually, the network traffic datasets are in low dimensional form, i.e., either in .pcap format or in .csv or .txt format. While the CNN models are designed and widely known for solving image processing and computer vision problems.

2.2 Machine Learning and Cyber Security

The more devices are connected, the more devices can be attacked and used by botnets or other threats. With the increasing amount of connected devices in a network, it is very difficult for a non-technical user to determine the level of security of the network [18]. Especially with Ambient Assisted Living (AAL) devices and digital assistants, security is extremely important, because highly personal data are collected by such devices. In private households and especially bathrooms are connected sensors that help older people or detect if they fell down [19]. Assistants like Google Home Mini [20] are used to make life easier and control other devices with voice input. The microphones are active all the time to receive voice commands. However, this can be also used to monitor third parties, such as visitors. To improve the security of the networks, security software like firewalls is needed. Current firewalls are getting extended with intelligent algorithms to keep up with the increasing development of attacks. But there are still new, growing botnets, like Ares [21]. To improve the security level further, Intrusion Detection Systems (IDS) are used. Network based IDS can detect attacks without any additional software on single devices. However, these systems cannot detect every attack. With current artificial intelligence (AI) algorithms, the detection rates can be improved above eighty to ninety percent. Without AI they are just detecting below seventy to eighty-five percent [22]. This difference shows the importance of IDS with AI.

AI and machine learning (ML) algorithms are part of many software and research projects. Therefore, a lot of approaches for IDS with different kind of AI integration can be found, too. The methods in [23] and [24] are both using ML algorithms to improve the detection rate of their IDS. Autonomous machine learning and deep learning algorithms improve the detection rate. However, we are trying to get no false positive results. To achieve this, we need to combine more approaches. There are existing hybrid methods, like the hybrid IDS from [25]. They are using this approach, because of the high false alarm rate of the neural network. The rule-based component should reduce this rate. Our goal is quite similar, but we are using different AI algorithms. One algorithm for a low false positive rate and the other for the classification of the attack, combined with the classic components. We found no similar combination of AI algorithms and rule based components for our zero false positive goals, but a lot of work, evaluating single AI algorithms for IDS, e.g. [26].

The Software-Defined Networking (SDN) paradigm provides a way to control IoT devices securely. For the IoT paradigm, K. M. Shayshab Azad et al. [27] suggested a general system for detecting and mitigating Distributed Denial-of-Service (DDoS) attacks using an SDN. The proposed architecture consists of a pool of controllers comprising SDN controllers, IoT gateway-integrated. Also, we have offered an IoT DDoS attack detection and mitigation algorithm attached to the proposed SDN IoT platform. Finally, the proposed algorithm shows the experimental results that have improved performance and the proposed architecture adapts to heterogeneous and fragile devices to enhance IoT security.

To ensure that the information system can provide services for users normally, it is important to detect the occurrence of DDoS attacks quickly and accurately. Therefore, keeping in view the author's H. -C. Chu and C. -Y. Yan [28] proposes a system based on packet continuity to detect DDoS attacks. On average, it only takes a few milliseconds to collect a certain number of consecutive packets, and then DDoS attacks can be detected. Experimental

results show that the accuracy of detecting DDoS attacks based on packet continuity is higher than 99.9% and the system response time is about 5 milliseconds.

It can be seen that the accuracy of the DoS detection systems are still low. This study [29] aims to provide a solution to the above problems by proposing an Intrusion Detection System based on Artificial Intelligence (AI, AdaBoost) for IoT system. The method used in this study is supervised learning which measures the accuracy of predictions in detecting DoS on IoT network data. The experiments have been carried out on 130223 DoS attack data and 130284 normal data. The detection accuracy of the DoS detection is 95.84 % and the F1-Score is 95.72 %. Recall and precision have achieved 93.28% and 98.29%, respectively.

A Distributed Denial of Service (DDoS) attack is a lethal threat to web-based services and applications. These attacks can cripple down these services in no time and deny legitimate users from using these services. The problem has further prevailed with the massive usage of unsecured Internet of Things (IoT) devices across the Internet. Moreover, many existing rule-based detection systems are easily vulnerable to attacks. In paper [30] A. Chopra et al. performed a comparative analysis of Machine Learning (ML) algorithms to detect and classify DDoS attacks. As part of the work, various machine learning algorithms such as Naive Bayes, J48, Random Forest ML classifiers are compared. Principal Component Analysis (PCA) method has been used to select the optimal number of features. WEKA tool has been used to implement ML algorithms.

IV. CONCLUSION

From PROPOSED WORK

The proposed methodology consists of four key steps which include: data acquisition, data cleaning, data conversion and attack pattern recognition. Figure 4.1 below provides an overview of the proposed methodology.

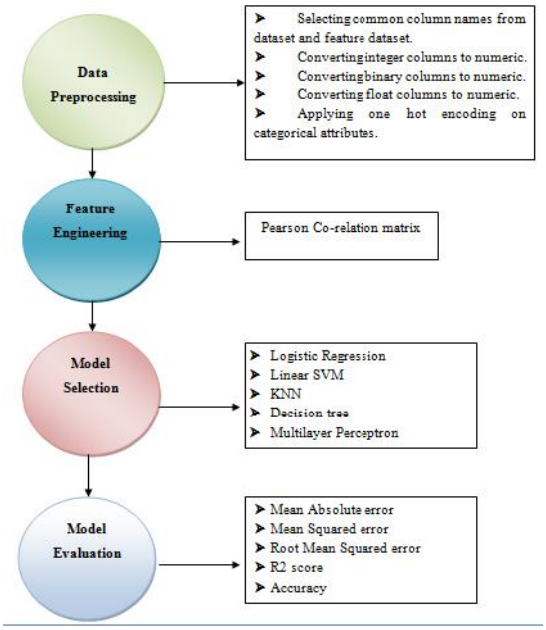

Figure 4.1: Architecture of Proposed Model.

A. Data Pre-processing: This is the first step that we perform for the cleaning of the dataset. In the data processing different methods are implemented. The methods that are implemented are shown above in the figure. The methods that are implemented are:

Following pre-processing steps are applied:

- Selecting common column names from dataset and feature dataset.
- Converting integer columns to numeric.
- Converting binary columns to numeric.
- Converting float columns to numeric.
- Applying one hot encoding on categorical attributes.

B. Feature Engineering: In this step we apply feature selection techniques for selecting important features. Feature selection is done on the basis of Pearson Co-relation matrix. The value of Pearson correlation matrix is used to calculate the value of features that are more important for the matrix to be decided.

After that, we figured out the features which were either duplicated or entirely had a constant value in case of all labels. Such constant features are not useful for discriminating the attack or normal traffic and may decrease the performance of the machine learning model, if included in the training set. Therefore, we also dropped constant features from the training set. On the other hand, the duplicate features are those which have similar values but have a different name. In the case of duplicate features, we keep the first original feature and

dropped its duplicate feature. Finally, after the cleaning the data we left with 61 features which were unique and important.

C. Attack Detection Model: The acquired network traffic data was in .csv format which includes more than 80 flow features. In order to better train our model for attack pattern detection, we removed the unwanted features from the data set which are not useful for classifying the attack and normal traffic.

These features include Flow ID, Source IP, Source Port, Destination IP, Destination Port, Protocol and Timestamp. As based upon these static features, one cannot decide whether a certain flowid, srcIP, etc., whenever found will always generate malicious or normal packets. That's why we dropped such unwanted features and excluded them from our training set.

Thereafter, we analyzed the whole dataset in order to deal with missing or malformed data. For this purpose, we first checked that which samples contain missing values, which samples have inadequate values like nan, -inf, +inf, etc. As we had a large number of samples in the dataset, so we dropped all those samples which comprise of missing or malformed values.

Table 4.1 below shows the Accuracy comparison of all the algorithms.

Algorithm used	**Accuracy in %**
Linear Regression	97.81
Linear SVM	97.85
KNN	98.31
Decision Tree	98.10
Multilayer Perceptron	98.37

Table 4.1: Accuracy comparison.

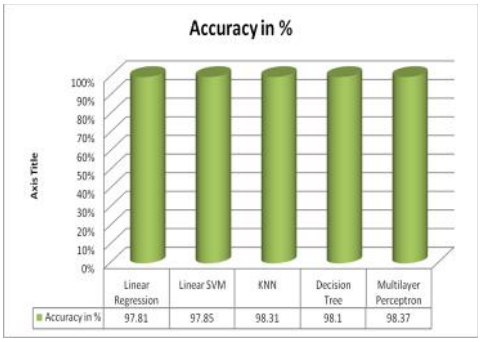

Figure 4.1: Accuracy Chart.

From the result it is clear that MLP algorithm outperforms the other models.

Table 4.2: RMSE Score Comparison.

Algorithm used	**RMSE in %**
Linear Regression	14.88
Linear SVM	14.66
KNN	13.01
Decision Tree	13.81
Multilayer Perceptron	12.77

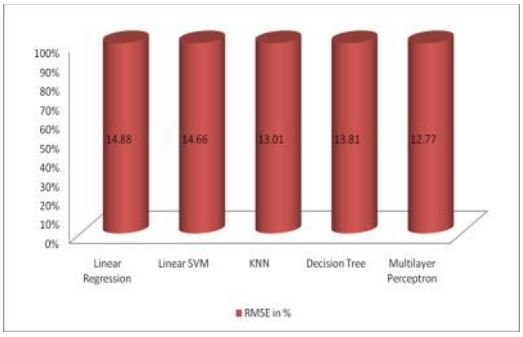

Figure 4.2: RMSE Values.

- In data science, RMSE has a double purpose:
- To serve as a heuristic for training models
- To evaluate trained models for usefulness.
- From results it is clear that the MLP Classifier has the lower values of RMSE.

V. CONCLUSION

We evaluated the proposed methodology for DoS and DDoS attack detection based on the above-mentioned parameters during the training and testing phase for both detecting and recognizing DoS and DDoS attacks in IoT networks.

For classification, the proposed methodology achieved 98.37% accuracy for detecting the DoS and DDoS attacks.

Furthermore, it achieves lower RMSE values as compared to the state-of-the art solution which was proposed on the same dataset.

REFERENCES

[1] C. Wagner, G. Wagener, and A. Dulaunoy, "SDBF : Smart DNS Brute-Force," pp. 1001–1007.

[2] D. Yuvaraj, M. Sivaram, A. M. U. Ahamed, and S. Nageswari, "Some investigation on DDOS attack models in mobile networks," Int. J. Interact. Mob. Technol., vol.

13, no. 10, pp. 71–88, 2019, doi: 10.3991/ijim.v13i10.11304.

[3] T. G. Nguyen, T. V. Phan, B. T. Nguyen, C. So-In, Z. A. Baig, and S. Sanguanpong, "Search: A collaborative and intelligent nids architecture for sdn-based cloud iot networks," IEEE access, vol. 7, pp. 107 678–107 694, 2019.

[4] What Is a DoS Attack?, (accessed January 11, 2020). [Online]. Available: https://www.datto.com/library/what-is-a-dos-attack.

[5] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto, and A. Dainotti, "Millions of targets under attack: a macroscopic characterization of the dos ecosystem," in Proceedings of the 2017 Internet Measurement Conference. ACM, 2017, pp. 100–113.

[6] Neustar Cyber Threats and Trends Report Q1 2019, (accessed January 11, 2020). [Online]. Available: https://www.discover.neustar/rs/717-IIA-274/images/Neustar%20Cyber%20Threats%20and%20Trends%20Report%20Q1%202019%20-%20Web.pdf.

[7] DDoS attack statistics and facts for 2018-2019, (accessed January 11, 2020). [Online]. Available: https://www.comparitech.com/blog/ information-security/ddos-statistics-facts.

[8] Cisco Annual Internet Report (2018–2023) White Paper, (accessed June 11, 2020). [Online]. Available: https: //www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/ annual-internet-report/white-paper-c11-741490.html.

[9] B. A. Khalaf, S. A. Mostafa, A. Mustapha, M. A. Mohammed, and W. M. Abduallah, "Comprehensive review of artificial intelligence and statistical approaches in distributed denial of service attack and defense methods," IEEE Access, vol. 7, pp. 51 691–51 713, 2019.

[10] S. Ghazanfar, F. Hussain, A. U. Rehman, U. U. Fayyaz, F. Shahzad, and G. A. Shah, "Iot-flock: An open-source framework for iot traffic generation," in 2020 International Conference on Emerging Trends in Smart Technologies (ICETST). IEEE, 2020, pp. 1–6.

[11] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in 2019 International Carnahan Conference on Security Technology (ICCST). IEEE, 2019, pp. 1–8.

[12] X. Liu, Z. Tang, and B. Yang, "Predicting network attacks with cnn by constructing images from NetFlow data," in 2019 IEEE 5th Intl Conference on Big Data Security on Cloud (Bigdatasecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS). IEEE, 2019, pp. 61–66.

[13] Y. Xiao and X. Xiao, "An intrusion detection system based on a simplified residual network," Information, vol. 10, no. 11, p. 356, 2019.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[15] O. Salman, I. H. Elhajj, A. Chehab, and A. Kayssi, "A machine learning based framework for iot device identification and abnormal traffic detection," Transactions on Emerging Telecommunications Technologies, p. e3743, 2019.

[16] W. Liu, X. Liu, X. Di, and H. Qi, "A novel network intrusion detection algorithm based on fast Fourier transformation," in 2019 1st International Conference on Industrial Artificial Intelligence (IAI). IEEE, 2019, pp. 1–6.

[17] Z. Li, Z. Qin, K. Huang, X. Yang, and S. Ye, "Intrusion detection using convolutional neural networks for representation learning," in International Conference on Neural Information Processing. Springer, 2017, pp. 858–866.

[18] S. Fischer, K. Neubauer, L. Hinterberger, B. Weber and R. Hackenberg, "IoTAG: An Open Standard for IoT Device Identification and Recognition", The Thirteenth International Conference on Emerging Security Information, Systems and Technologies, 2019, in press.

[19] W. L. Zangler, P. Panek and M. Rauhala, Ambient assisted living systems - the conflicts between technology, acceptance, ethics and privacy. Dagstuhl Seminar Proceedings. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2008.

[20] Google Ireland Limited (2019, Oct.) Google Home Mini. [Online]. Available: https://store.google.com/product/google home mini.

[21] C. Cimpanu. (2019, Oct.) A new IOT botnet is infecting Android-based set-top boxes. ZDNet. [Online]. Available: https://www.zdnet.com/article/a-new-iot-botnet-is-infecting-androidbased-set-top-boxes.

[22] N. A. Alrajeh and J. Lloret, "Intrusion detection systems based on artificial intelligence techniques in wireless sensor networks", International Journal of Distributed Sensor Networks 9.10, p. 351047, 2013.

[23] J. Cannady, "Next generation intrusion detection: Autonomous reinforcement learning of network attacks." 23rd national information systems security conference, pp. 1-12, 2000.

[24] A. Shenfield, D. Day and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks", ICT Express, 4(2), pp. 95-99, 2018.

[25] S. Koutsoutos, I. T. Christou and S. Efremidis, "An Intrusion Detection System for Network-Initiated Attaclcs

Using a Hybrid Neural Network", Artificial Intelligence Applications and Innovations, Springer US, pp. 228-235, 2006.

[26] H. Liu and B. Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey", Applied Sciences, vol. 9, no. 20, p. 4396, 2019.

[27] K. M. Shayshab Azad, N. Hossain, M. J. Islam, A. Rahman and S. Kabir, "Preventive Determination and Avoidance of DDoS Attack with SDN over the IoT Networks," 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 *(*ACMI*)*, 2021, pp. 1-6, doi: 10.1109/ACMI53878.2021.9528133.

[28] H. -C. Chu and C. -Y. Yan, "DDoS Attack Detection with Packet Continuity Based on LSTM Model," *2021* IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE), 2021, pp. 44-47, doi: 10.1109/ECICE52819.2021.9645650.

[29] S. Rachmadi, S. Mandala and D. Oktaria, "Detection of DoS Attack using AdaBoost Algorithm on IoT System," 2021 International Conference on Data Science and Its Applications (ICoDSA), 2021, pp. 28-33, doi: 10.1109/ICoDSA53588.2021.9617545.

[30] A. Chopra, S. Behal and V. Sharma, "Evaluating Machine Learning Algorithms to Detect and Classify DDoS Attacks in IoT," 2021 8th International Conference on Computing for Sustainable Global Development ((IndiaCom*)*, 2021, pp. 517-521.