IRAQI Academic Scientific Journals
العراقية المجلات الأكاديمية العلمية
University of Kut Journal
ISSN (E): 2616 - 7808 II ISSN (P): 2414 - 7419
js.uokut.edu.iq
k.u.c.j.sci@alkutcollege.edu.iq
Vol. 11, Issue 1, June 2026

Enhancing the Efficiency of DDoS Attack Detection in IoT Networks Using Machine Learning

Beagard Salih Hassen [1] , **Ahmad Khalil Ibrahim** [2]

Abstract

Modern connectivity has been revolutionized by the Internet of Things (IoT). The limited computational capabilities of IoT devices make them vulnerable to various attacks including, Distributed Denial of Service (DDoS) attacks that compromise the availability of these devices and their services. This research introduces a machine learning-based approach to enhance DDoS attack detection in IoT environments. Using the CICIoT2023 dataset, we evaluated several machine learning models: Random Forest, XGBoost, Decision Tree, and K-Nearest Neighbors. The results demonstrated high classification performance across all models. XGBoost achieved the highest accuracy of 99.97% with a prediction time of 0.4735 seconds, while Decision Tree delivered the best prediction time of 0.1879 seconds, maintaining a high accuracy of 99.94%.

The results of the article confirm that the suggested machine learning models for DDoS attack detection in IoT networks are effective and improve the security of such networks.

Keywords: IoT security, DDoS, Machine Learning, CICIoT 2023 dataset

Affiliations of Authors

[1] Department of Medical Device Engineering, Kut University, Iraq, Wasit, 52001

[2] Department of Medical Device Engineering, National University of Science and Technology, Iraq, Nasiriyah, 64001

[1] bykrdalbrznjy@gmial.com
[2] ahmed.k.ibrahim@nust.edu.iq

[1] **Corresponding Author**

Paper Info.
Published: Jun. 2026

تحسين كفاءة اكتشاف هجمات DDoS في شبكات إنترنت الأشياء باستخدام التعلم الآلي

م. م. بيكرد صالح حسن [1]، م. د. احمد خليل ابراهيم [2]

المستخلص

لقد أحدثت تقنيات الاتصال الحديثة ثورة بفضل إنترنت الأشياء (IoT). ومع ذلك، فإن القدرات الحسابية المحدودة لأجهزة إنترنت الأشياء تجعلها عرضة لهجمات متعددة، بما في ذلك هجمات حجب الخدمة الموزعة (DDoS) التي تؤثر على توفر هذه الأجهزة والخدمات التي تقدمها. تقدم هذه الدراسة نهجًا يعتمد على التعلم الآلي لتعزيز الكشف عن هجمات DDoS في بيئات إنترنت الأشياء. باستخدام مجموعة بيانات CICIoT2023، قمنا بتقييم عدة نماذج للتعلم الآلي، بما في ذلك الغابة العشوائية (Random Forest)، وXGBoost، وشجرة القرار (Decision Tree)، والجار الأقرب (K-Nearest Neighbors). أظهرت النتائج أداءً عالياً في التصنيف عبر جميع النماذج. حقق نموذج XGBoost أعلى دقة بنسبة 99.97% مع وقت تنبؤ بلغ 0.4735 ثانية، بينما حقق نموذج شجرة القرار (Decision Tree) أفضل وقت تنبؤ بلغ 0.1879 ثانية مع دقة عالية بلغت 99.94%، تؤكد هذه النتائج فعالية نماذج التعلم الآلي المقترحة في الكشف عن هجمات DDoS في شبكات إنترنت الأشياء وتعزيز أمن هذه الشبكات.

الكلمات المفتاحية: أمن انترنت الأشياء، هجوم حجب الخدمة الموزع، مجموعة البيانات CICIoT 2023.

انتساب الباحثين

[1] قسم هندسة الأجهزة الطبية، جامعة الكوت، العراق، واسط، 52001

[2] قسم هندسة الأجهزة الطبية، الجامعة الوطنية للعلوم والتكنولوجيا، العراق، الناصرية، 64001

[1] bykrdalbrznjy@gmial.com
[2] ahmed.k.ibrahim@nust.edu.iq

[1] **المؤلف المراسل**

معلومات البحث
تأريخ النشر: حزيران 2026

1. Introduction

The Internet of Things plays an important role in life, transforming the physical world into a digital world. The use of the IoT has been increasing in the recent years. The size of the Internet of Things may reach 25 billion in 2030 [1]. Due to the limited computing and storage resources, complex defense systems cannot be hosted by the Internet of Things devices.

Networking limitations like mobility, scalability, and low data rates from low-power radios further increase their vulnerability, making IoT devices prime targets for attacks and potential weak points

in network security [2]. DDoS is a critical threat to IoT devices, attempting to disrupt their availability. Generally speaking, large botnets are employed by attackers to flood devices, which eventually make services unavailable to critical users. Mirai botnet was just such an example that exploited weak or default credentials in IoT devices to launch huge DDoS attacks [3]. Once infected, those devices become part of this botnet, amplifying the power of the attack. Therefore, securing IoT devices is very important to minimize such intermediate attacks and maintain network stability [4]. In the third quarter of 2022, DDoS attacks have surged dramatically, rising by 90% compared to the same period in the previous year [5].

DDoS can be divided into two groups: reflected-based and exploitation-based. The major difference between them is in the ways of their carrying out. Reflected-based DDoS consists of a huge amount of requests sent to servers and utilizes big-sized responses for flooding the victim, so-called reflection amplification. In contrast, exploitation-based DDoS attacks directly against vulnerabilities in the victim's device. Both classes may include several types of attacks depending on the protocols used in such an attack. The exploitation-based attacks include SYN flood, UDP flood, and UDP lag, while the reflected-based attacks involve protocols such as DNS, MSSQL, SSDP, LDAP, NTP, TFTP, and SNMP [6]. It is necessary to detect abnormal traffic from these variants of DDoS for prevention. Machine - learning has recently become popular in IoT networks due to its adaptability and scalability [7]. Machine learning algorithms can learn from past attacks and adapt to new patterns, enabling effective detection and mitigation [6]. This is because rapid identification of attack types allows devices to implement tailored countermeasures that are more effective than generalized solutions. This study applied four algorithms (XGBoost, Decision Tree, Random Forest, and KNN) for multiclass classification to detect different DDoS attack types and compared their performance.

2. Related Work

A large number of surveys have been conducted to analyze the quantity and aspects of Internet of Things (IoT) devices. To tackle the problem of class imbalance, the study in [8] developed a new Intrusion Detection System (IDS) for Bot-IoT using line Learning and Deep Learning techniques. To see how the prediction relies on the use of record timestamps, three different feature sets were used for both the binary and multiclass classifications. This approach reduced the feature dependencies that may have been brought in by the Argus flow data generator and it reached an average accuracy of 99%. Exhaustive experiments were conducted and time performance analysis was done and it was able to identify all DoS attacks and provided better state of the art for identifying denial of service (DoS) attacks. Decision Tree and Multi-Layer Perceptron (MLP) methods were found to be most effective in identifying DDoS and DoS attacks in IoT networks.

The researchers in study [9] utilized the CICDDoS2019 dataset to perform an enhanced study and features a fresh classification of DDoS attacks with the deployment of a novel classification scheme based on network flow. The authors have come up with DNN and LSTM method to detect DDoS attacks using a deep neural network. The researchers successfully detected over 99.90% of all three classes of DDoS attacks. Future disruptions can be predicted through the use

of deep learning, for predicting potential targeted attacks.

This research [10] proposes a machine learning based approach to identify DDoS attacks in an SDN-WISE IoT controller. A machine learning detection module was used. The detection module uses Naive Bayes (NB), Decision Tree (DT), and Support Vector Machine (SVM) algorithms for classification of the SDN-IoT network packets. In this paper, the accuracy of the framework was evaluated under various traffic simulation scenarios and the result shows 97.4% for NB, 96.1% for SVM, and 98.1% for DT.

In the paper [11], the authors proposed two techniques for identifying the reflective distributed denial of service (DrDoS) attacks on IoT networks. The first method is based on a hybrid intrusion detection system (HIDS) for detecting DoS attacks on IoT and the second method uses deep learning models like long short-term memory (LSTM) networks which are trained on the latest distributed denial of service (DDoS) attack set. The CICDDoS2019 dataset was used for training and evaluation. The model achieved an excellent accuracy of 99.19%.

The study [12] is developed to propose a Machine Learning (ML) based approach for anomaly detection in a network for identifying DDoS attacks in Industry 4.0 Cyber Physical Production Systems (CPPSs) using data driven models. This study is different from the previous approaches which are based on the use of artificially simulated data or information from IT networks and small-scale testbeds but instead utilized real-world Network traffic data from a semiconductor production factory. Extensive simulations were performed in order to compare 11 supervised, unsupervised and semi-supervised algorithms. The results indicate that the supervised algorithms are more accurate in the detection as opposed to their unsupervised and semi-supervised counterparts. Most importantly, the Decision Tree model was found to be very efficient in detecting anomalies in an industrial network environment with a high accuracy of 99.9%.

Research [13] compared machine learning classifiers for DDoS detection from the CICDDoS2019 dataset in two important ways. First, they implemented DDoS detection using four algorithms: K-Nearest Neighbors (KNN), eXtreme Gradient Boosting (XGBoost), Decision Tree, and Random Forest. Second, they applied three feature selection methods—Chi-square, ExtraTreeClassifier and Analysis of Variance (ANOVA)—with each classifier to select 20 features. The highest accuracy (98%) was achieved by XGBoost combined with ANOVA among the tested combinations.

Study [14] employed five different algorithms to detect DDoS traffic in IoT networks. A notable strength of their study was the use of lightweight features derived from self-collected traffic data, with a focus on memory efficiency. Their findings showed that four models achieved an accuracy of 99% .

IoT Device DDOS Detection for a multi-layer system was presented by research [15] that used protocols verification in an authentication mechanism at the first layer and a Decision Tree algorithm at the second layer. The third layer implemented traffic-blocking rules inside an SDN controller. It allowed binary classification accuracy as 97% with flooded sensor data, while the Network is experiencing Flooding at an Accuracy of 99%. A strong point in this paper was that for the detection, it showed a blocking mechanism, which made it stronger against the DDoS attack.

3. Methodology

The methodology shown in Figure 1 was followed, which consists of 7 stages. Each stage will be explained separately.

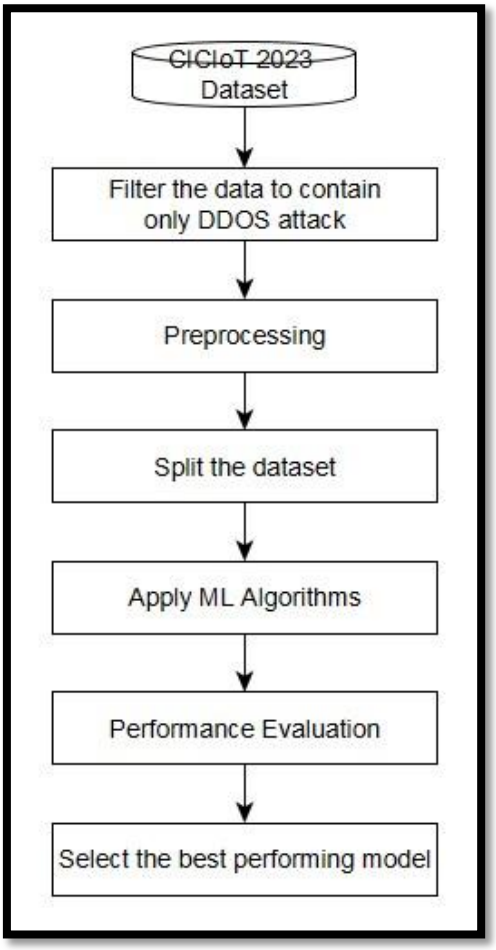

Figure (1): Methodology Diagram

3.1 CICIoT 2023 Dataset

CIC IoT Dataset 2023 was used. It is a real-time dataset designed for analyzing attacks in IoT environments. The dataset includes 33 different attacks in an IoT topology consisting of 105 devices. These attacks are divided into seven categories: DDoS, DoS, Reconnaissance, Web-based, Brute Force, Spoofing, and Mirai. All the attacks emanate from the compromised IoT devices to the target IoT devices [16]. The dataset contains 218,805 rows and 47 columns. The columns include features such as flow_duration, Header_Length, Protocol Type, and a label column representing activity categories (e.g., DDoS-SYN_Flood, BenignTraffic). They are intended to be used in this article to represent network traffic with attributes for anomaly or attack detection .

3.2 Preprocessing

This includes filtering the data to include only relevant data and then converting it to a NumPy matrix for further processing. Another step-in data preprocessing is ensuring that the dataset used for training the DoS attack detection model contains no missing values or duplicate records. This ensures the integrity and reliability of the data.

3.3 Split the dataset

The pre-processed dataset is split into training data (80%) and testing data (20%). The training data is

used to train machine learning models, and the testing data is used to evaluate the performance of those models.

3.4 Apply ML Algorithms

Four different machine learning algorithms are applied to the training dataset including: XGBoost, Decision Tree, Random Forest, and KNN.

1. Random Forest

It is an ensemble learning technique in which many decision trees participate collectively. It makes each tree to be trained on a different part of the data and some of the features, not on all of them, and all of the data. Each tree classifies the possible class of DDoS traffic and the final prediction is made by taking the majority vote of all the trees in the forest. This algorithm works well on large datasets with no problems of outliers and missing values that one may encounter [13]. The Random Forest was used in this study using the Random Forest Classifier function from the scikit-learn package.

2. XGBoost

XGBoost is a gradient-boosting ensemble machine learning algorithm that uses decision trees to classify data. The two major goals of any machine learning project are: high execution speed and superior model performance; these are also properties of XGBoost [17]. The Random Forest was also implemented in this study using the XGBClassifier function from the scikit-learn package.

3. Decision Tree

Decision trees are a family of supervised learning algorithms that use a tree-like model of decisions and possible outcomes. In the case of DDoS traffic, the result of a test is used to branch, and different attack types take the leaf nodes. Decision tree was implemented in this study using decision tree classifier function from the scikit-learn library.

4. K-Nearest Neighbors (KNN)

It is a simple, easily implemented supervised machine learning algorithm that can be used for classification and regression. To make any assumption, this algorithm takes new data points as covering the distribution of existing ones and will assign the category that is most similar to the available categories [18]. In this study, KNN was implemented using the K Neighbors Classifier function from the scikit-learn library.

3.5 Performance Evaluation:

After training the four models, their performance is evaluated using the test set. The metrics of accuracy, precision, recall, F1 score, and confusion matrix will be used to evaluate how well the models perform.

3.6 Select the best-performing model:

Based on the evaluation metrics, the model that performs best (highest accuracy and lowest prediction time) in detecting DDoS attacks is selected.

4. Results and Analysis

This section will explain the dataset that was used as well as analyze the results obtained from implementing different machine learning models.

4.1 Random Forest Model

Random Forest Classifier was trained and tested. Prediction time was 0.9105 seconds. Overall accuracy was 99.93%. Confusion matrix shown in Figure (2) reveals excellent performance across classes.

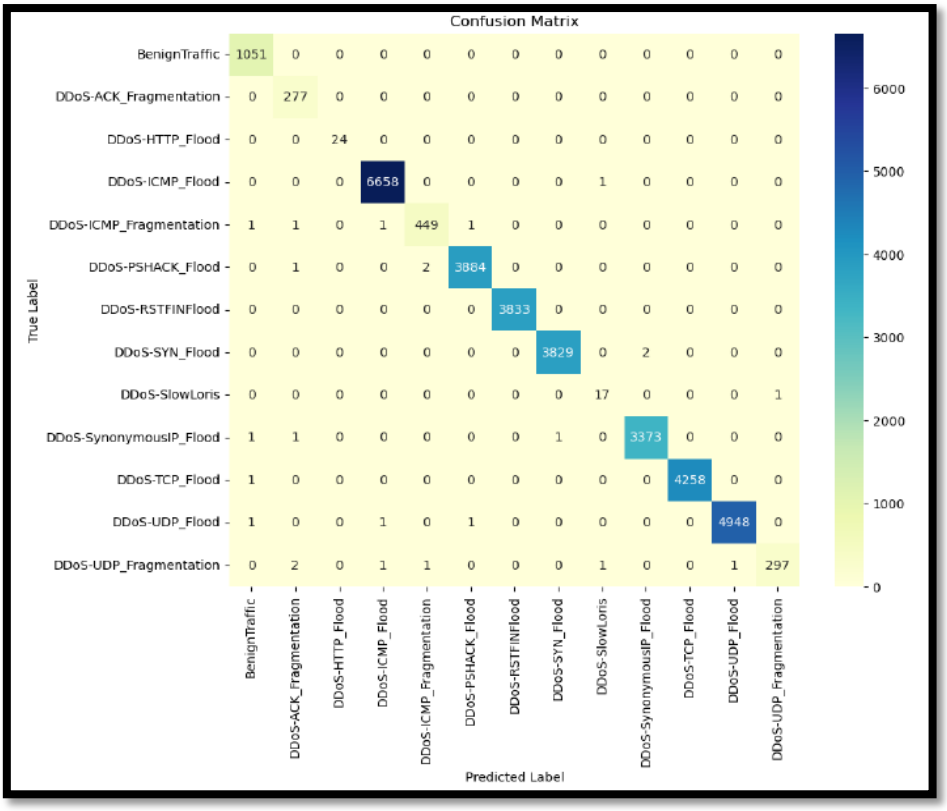

Figure (2): Confusion Matrix of Prediction for Random Forest Model

4.2 XGBoost Model

Another model, XGBoost, was implemented using label encoding for taxonomic tags. The prediction time was 0.4735 seconds; the overall accuracy was 99.97%. Improved accuracy and prediction time to random forest were observed. The confusion matrix of prediction for the XGBoost model is shown in Figure (3).

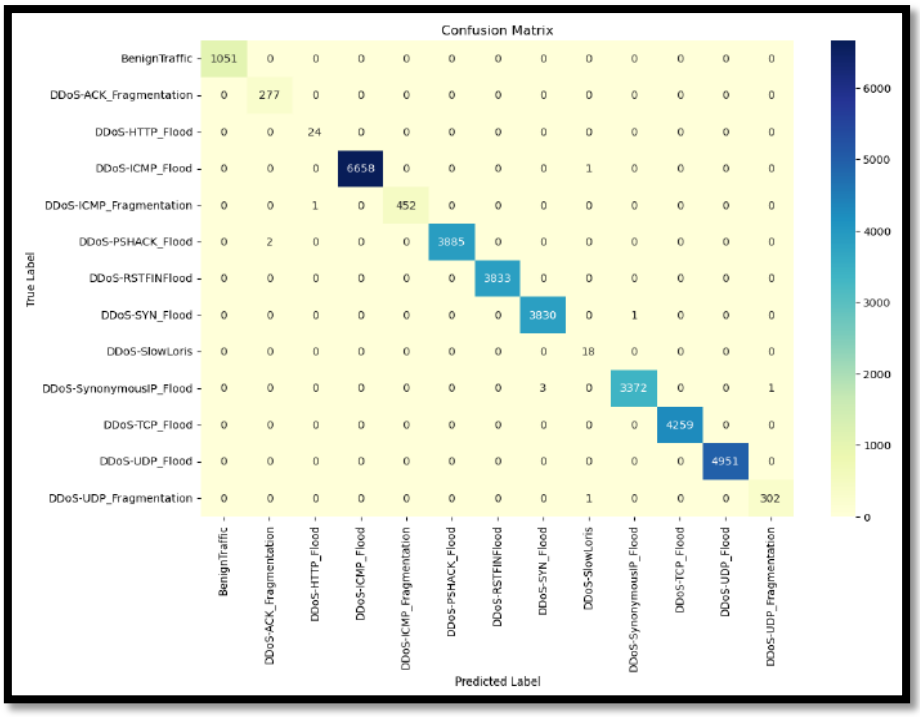

Figure (3): Confusion Matrix of Prediction for XGBoost Model

4.3 Decision Tree Model

The simplest decision tree model is used for classification and the model achieves an accuracy of 99.94% is slightly lower than XGBoost, but the prediction time is 0.1879 seconds which is better than XGBoost. The confusion matrix as shown in Figure (4) reveals excellent performance across classes .

```
Confusion Matrix:
[[1049    0    0    0    1    0    0    0    0    0    0    1    0]
 [   0  277    0    0    0    0    0    0    0    0    0    0    0]
 [   0    0   24    0    0    0    0    0    0    0    0    0    0]
 [   0    0    0 6657    1    0    0    1    0    0    0    0    0]
 [   0    2    0    0  450    0    0    0    0    0    0    0    1]
 [   0    0    0    0    0 3885    0    1    0    0    1    0    0]
 [   0    0    0    0    0    0 3833    0    0    0    0    0    0]
 [   0    0    1    0    0    0    0 3829    0    0    1    0    0]
 [   1    0    0    0    0    0    0    0   17    0    0    0    0]
 [   0    0    0    1    0    0    0    0    0 3374    1    0    0]
 [   2    0    0    0    0    0    0    0    0    0 4257    0    0]
 [   0    0    0    0    0    0    0    0    0    0    0 4949    2]
 [   2    0    0    0    0    0    0    0    0    0    0    0  301]]
```

Figure (4): Confusion Matrix of Prediction for Decision Tree Model

4.4 K-Nearest Neighbors (KNN) Model

A K-Nearest Neighbors (KNN) model was applied for classification. The accuracy was 99.58% (slightly lower than previously applied models). Prediction time is significantly higher (~111 seconds), reflecting the computational cost of KNN on large datasets. The confusion matrix shown in Figure 5 reveals good performance across different classes .

```
Confusion Matrix with KNN:
[[1051    0    0    0    0    0    0    0    0    0    0    0    0]
 [   0  274    0    0    0    0    1    0    0    1    0    0    1]
 [   0    0   24    0    0    0    0    0    0    0    0    0    0]
 [   0    1    3 6640    0    0    2    1    0    1    0   10    1]
 [   0    5    0    1  436    0    5    0    1    1    2    1    1]
 [   1    5    0    2    0 3867    8    3    0    1    0    0    0]
 [   0    3    0    2    0    1 3825    1    1    0    0    0    0]
 [   0    0    0    2    0    0    0 3817    0    0    8    4    0]
 [   0    1    0    1    1    0    5    0    7    2    0    0    1]
 [   0    2    0    0    3    1   11    1    2 3356    0    0    0]
 [   0    2    0    3    0    0    0    1    0    1 4250    2    0]
 [   0    0    2    6    0    0    1    4    0    0    0 4938    0]
 [   0    2    0    0    0    0    1    0    2    0    0    0  298]]
```

Figure (5): Confusion Matrix of Prediction for K-Nearest Neighbors (KNN) Model

Conclusions

This study aimed to implement multi-class classification to improve the efficiency of DDoS detection in IoT using machine learning using XGBoost, Decision Tree, Random Forest, and KNN models. The main goal was to train these

models to distinguish between different types of DDoS traffic and achieve high accuracy. The results show that the four models achieved high classification accuracy close to 1 in classifying various types of network traffic. The XGBoost model performed the best in terms of accuracy, reaching 99.97% and prediction time reaching 0.4735 seconds, while decision tree performed the best in terms of prediction time 0.1879 seconds with high accuracy 99.94%. As for the KNN model, despite its good accuracy, its prediction time was significantly higher. Overall, this study aims to build a model that can effectively identify and classify various kinds of DDoS traffic and distinguish between benign and malicious activities. Moreover, the study focused on improving the efficiency of DDoS detection in IoT using machine learning, and future research can explore defense mechanisms for different types of DDoS attacks and integrate them with machine learning detection systems.

References

[1] P. Malhotra, Y. Singh, P. Anand, D. K. Bangotra, P. K. Singh, and W.-C. Hong, "Internet of things: Evolution, concerns and security challenges," Sensors, vol. 21, no. 5, p. 1809, 2021.

[2] P. Williams, I. K. Dutta, H. Daoud, and M. Bayoumi, "A survey on security in internet of things with a focus on the impact of emerging technologies," Internet of Things, vol. 19, p. 100564, 2022.

[3] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," Computer (Long. Beach. Calif)., vol. 50, no. 7, pp. 80–84, 2017.

[4] M. H. Ali et al., "Threat analysis and distributed denial of service (DDoS) attack recognition in the internet of things (IoT)," Electronics, vol. 11, no. 3, p. 494, 2022.

[5] S. Görtz, S. Fischer, and R. Hackenberg, "Generation of Distributed Denial of Service Network Data with Phyton and Scapy," CLOUD Comput. 2023, p. 17, 2023.

[6] Y. Xie, "Machine learning-based DDoS detection for IoT networks," Appl. Comput. Eng., vol. 29, pp. 99–107, 2023.

[7] P. Kumari and A. K. Jain, "A comprehensive study of DDoS attacks over IoT network and their countermeasures," Comput. Secur., vol. 127, p. 103096, 2023.

[8] J. G. Almaraz-Rivera, J. A. Perez-Diaz, and J. A. Cantoral-Ceballos, "Transport and application layer DDoS attacks detection to IoT devices by using machine learning and deep learning models," Sensors, vol. 22, no. 9, p. 3367, 2022.

[9] T. Khempetch and P. Wuttidittachotti, "DDoS attack detection using deep learning," IAES Int. J. Artif. Intell., vol. 10, no. 2, p. 382, 2021.

[10] J. Bhayo, S. A. Shah, S. Hameed, A. Ahmed, J. Nasir, and D. Draheim, "Towards a machine learning-based framework for DDOS attack detection in software-defined IoT (SD-IoT) networks," Eng. Appl. Artif. Intell., vol. 123, p. 106432, 2023.

[11] M. Shurman, R. Khrais, and A. Yateem, "DoS and DDoS attack detection using deep learning and IDS," Int. Arab J. Inf. Technol, vol. 17, no. 4A, pp. 655–661, 2020.

[12] F. B. Saghezchi, G. Mantas, M. A. Violas, A. M. de Oliveira Duarte, and J. Rodriguez, "Machine learning for DDoS attack detection in industry 4.0 CPPSs," Electronics, vol. 11, no. 4, p. 602, 2022.

[13] V. Gaur and R. Kumar, "Analysis of machine learning classifiers for early detection

of DDoS attacks on IoT devices," Arab. J. Sci. Eng., vol. 47, no. 2, pp. 1353–1374, 2022.

[14] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning ddos detection for consumer internet of things devices," in 2018 IEEE Security and Privacy Workshops (SPW), IEEE, 2018, pp. 29–35.

[15] Y.-W. Chen, J.-P. Sheu, Y.-C. Kuo, and N. Van Cuong, "Design and implementation of IoT DDoS attacks detection system based on machine learning," in 2020 European Conference on Networks and Communications (EuCNC), IEEE, 2020, pp. 122–127.

[16] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," Sensors, vol. 23, no. 13, p. 5941, 2023.

[17] X. Wang and X. Lu, "A Host-Based Anomaly Detection Framework Using XGBoost and LSTM for IoT Devices," Wirel. Commun. Mob. Comput., vol. 2020, no. 1, p. 8838571, 2020.

[18] M. A. Mohsin and A. H. Hamad, "Performance evaluation of SDN DDoS attack detection and mitigation based random forest and K-nearest neighbors machine learning algorithms," Rev. d'Intelligence Artif., vol. 36, no. 2, p. 233, 2022.