Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2024.054610

ARTICLE

Machine Learning Enabled Novel Real-Time IoT Targeted DoS/DDoS Cyber Attack Detection System

Abdullah Alabdulatif[1], Navod Neranjan Thilakarathne[2,*] and Mohamed Aashiq[3,*]

[1]Department of Cybersecurity, College of Computer, Qassim University, Buraydah, 51452, Saudi Arabia

[2]Department of ICT, Faculty of Technology, University of Colombo, Colombo, 00700, Sri Lanka

[3]Department of Computer Science and Engineering, Faculty of Engineering, South Eastern University of Sri Lanka, Oluvil, 32360, Sri Lanka

*Corresponding Authors: Navod Neranjan Thilakarathne. Email: navod.neranjan@ict.cmb.ac.lk; Mohamed Aashiq. Email: aashiqmnm@gmail.com

Received: 03 June 2024 Accepted: 16 July 2024 Published: 12 September 2024

ABSTRACT

The increasing prevalence of Internet of Things (IoT) devices has introduced a new phase of connectivity in recent years and, concurrently, has opened the floodgates for growing cyber threats. Among the myriad of potential attacks, Denial of Service (DoS) attacks and Distributed Denial of Service (DDoS) attacks remain a dominant concern due to their capability to render services inoperable by overwhelming systems with an influx of traffic. As IoT devices often lack the inherent security measures found in more mature computing platforms, the need for robust DoS/DDoS detection systems tailored to IoT is paramount for the sustainable development of every domain that IoT serves. In this study, we investigate the effectiveness of three machine learning (ML) algorithms: extreme gradient boosting (XGB), multilayer perceptron (MLP) and random forest (RF), for the detection of IoT-targeted DoS/DDoS attacks and three feature engineering methods that have not been used in the existing state-of-the-art, and then employed the best performing algorithm to design a prototype of a novel real-time system towards detection of such DoS/DDoS attacks. The CICIoT2023 dataset was derived from the latest real-world IoT traffic, incorporates both benign and malicious network traffic patterns and after data preprocessing and feature engineering, the data was fed into our models for both training and validation, where findings suggest that while all three models exhibit commendable accuracy in detecting DoS/DDoS attacks, the use of particle swarm optimization (PSO) for feature selection has made great improvements in the performance (accuracy, precsion recall and F1-score of 99.93% for XGB) of the ML models and their execution time (491.023 sceonds for XGB) compared to recursive feature elimination (RFE) and random forest feature importance (RFI) methods. The proposed real-time system for DoS/DDoS attack detection entails the implementation of an platform capable of effectively processing and analyzing network traffic in real-time. This involves employing the best-performing ML algorithm for detection and the integration of warning mechanisms. We believe this approach will significantly enhance the field of security research and continue to refine it based on future insights and developments.

KEYWORDS

Machine learning; Internet of Things (IoT), DoS, DDoS; cybersecurity; intrusion prevention; network security, feature optimization, sustainability

Copyright © 2024 The Authors. Published by Tech Science Press.
This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

The proliferation of IoT devices has transformed many domains by enabling seamless communication and automation [1]. Such domains include manufacturing, agriculture, transportation, the military, medical care, and so on [2]. The IoT is also increasingly becoming integrated into our daily lives and transforming into an integral part of us. Recent statistics indicate that by 2025, the number of IoT devices in use worldwide is projected to reach 30.9 billion units [1–4]. Despite the fact that the number of IoT devices is growing by the day, the majority of IoT devices lack sufficient computing capacity and acceptable security measures when connected to the Internet, threatening the long-term sustainability of the ubiquitous IoT ecosystem.

As such, the extensive incorporation of IoT technology is paving the way for cybercriminals to intrude into digital ecosystems if the security measures are not sufficient. DoS and DDoS represent two of the most formidable security threats encountered in the realm of IoT [1–3]. DoS and DDoS attacks occur when a malicious actor orchestrates a coordinated assault on a target system or network with the intention of rendering it unavailable to legitimate users by overwhelming it with a flood of network traffic [3–6]. In general, DDoS attacks have been the subject of in-depth research and analysis in conventional areas of cybersecurity, where mitigation measures have also been designed [7–10]. However, when these traditional methods are applied, they frequently prove insufficient owing to the sophisticated nature of modern DoS/DDoS attacks [10–13]. As a result, IoT networks and devices are left vulnerable to potentially catastrophic security breaches. The main issue in detecting such attacks is these attacks are often distinguished by an elevated volume of packets originating from a single IP address (in the case of DoS) or several IP addresses (in the case of DDoS) [4–7]. As IoT network complexity increases, so does the area susceptible to malicious intrusion. In most cases, Dos/DDoS attacks targeting the IoT can result in substantial disruptions, compromising end-user safety, causing financial losses, and undermining the very promise of IoT technology [10–12]. Consequently, innovative and adaptable security solutions that can protect IoT ecosystems from the ever-changing landscape of DoS/DDoS threats are required immediately [13–15]. Fig. 1 shows the total number of DoS/DDoS attacks encountered since 2018 and predicted numbers according to Cisco [14]. From the figure, it is evident that the number of DoS/DDoS attacks is growing in millions every year.

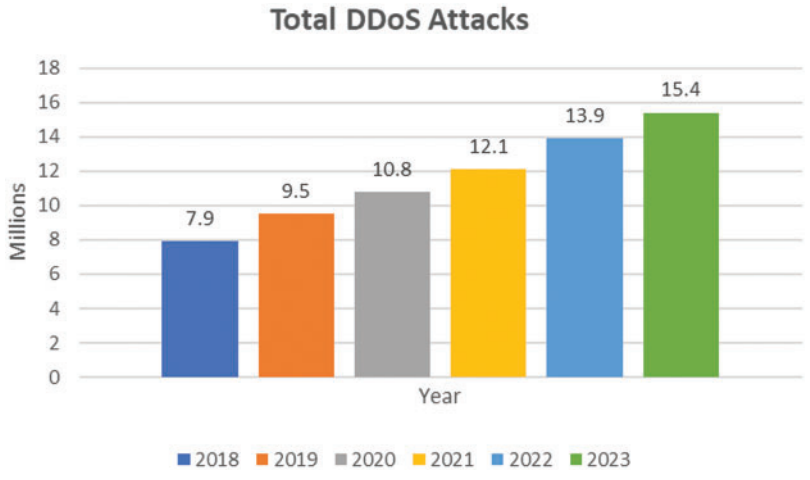

Figure 1: DoS/DDoS total attack history and predictions according to Cisco

In the battle against IoT-related DoS/DDoS attacks, Machine Learning (ML), a subset of Artificial Intelligence (AI), has emerged as a valuable ally. Its ability to process vast amounts of data, lightweight nature, cost-effectiveness, adaptability to evolving attack patterns, and anomaly detection make it a promising tool [4–6].

Thus motivated by the fact that employing ML for IoT-related DoS/DDoS attack detection, this research addresses the specific challenges posed by IoT-based DoS/DDoS attacks and explores the implementation of ML techniques for detecting such attacks by designing a novel ML-enabled real-time monitoring system. On the other hand, our study also presents the importance of combining ML with feature engineering to enhance DoS/DDoS detection in IoT environments. We employ three ML algorithms-XGB, RF, and MLP-along with three feature engineering approaches to improve accuracy and resource efficiency. Ultimately, our goal is to highlight the potential synergy between ML and IoT security, contributing to the development of robust, adaptable, and faster DoS/DDoS defenses to ensure the sustainable growth of the IoT ecosystem. The following section outlines the key contributions of this study:

- Propose three efficient ML models (RF, XGB, and MLP) for the classification of fifteen types of DoS/DDOS attacks using the CICIoT2023 dataset, which is the latest real-time dataset and benchmark for large-scale attacks in the IoT environment.
- Analyzing the effectiveness of three feature engineering methods (PSO, RFE, and RFI) with the aim of improving overall accuracy, reducing resource usage, saving costs, and enabling faster response, which many researchers have overlooked.
- The effect of training settings on classification accuracy (%), precision (%), recall (%), and F1-score (%) was investigated.
- To determine the effectiveness of employed feature engineering methods, we contrasted the employed ML models in order to assess their performance. The comparative analysis reveals that PSO made significant improvements in the performance of the ML models and their execution time compared to other employed methods.
- By employing the best-performing algorithm, design a prototype of a novel real-time DoS/D-DoS attack detection system that can effectively identify attacks before they are onset.

The structure of this study is as follows: After the introduction, Section 2 provides a concise overview of DoS/DDoS attacks and their detection using AI. Section 3 delineates the research approach. In Section 4, the research findings are presented, along with the implementation of the real-time detection system, and the study concludes by discussing the results and suggesting future directions.

2 Related Work

DoS and DDoS attacks would hinder the sustainable growth of many IoT-served domains [16,17]. These attacks are a major challenge for IoT owing to the growth of connected devices and their vulnerabilities. This section presents a concise synopsis of DoS and DDoS attacks pertaining to the IoT and highlights the application of AI in detecting such attacks.

2.1 T DoS/DDoS Attacks in the Context of IoT

DoS attacks are among the oldest types of cyberattacks. At their core, they aim to render a digital service (like a website or an online application) unavailable by overwhelming it with network traffic or exploiting specific vulnerabilities [18–20]. In the early days of the Internet, it would often take just one

computer with a decent Internet connection to launch a successful DoS attack. On the contrary, DDoS attacks are more sophisticated and more potent than DoS attacks. Instead of a single source, DDoS attacks use a network of compromised computers (often referred to as a botnet) to flood their target with a substantial quantity of network traffic. Such attacks can be catastrophic, leading to significant downtimes, financial losses, and a tarnished reputation for the targeted entity [21–23].

The convergence of the IoT and the increasing sophistication of cyber threats has created a perfect storm. Given their always-on nature, inadequate security provisions, and sheer numbers, IoT devices/networks/infrastructure are becoming attractive targets for cybercriminals [21–23]. Hence, the devices/networks/infrastructure themselves can be victims of such DoS/DDoS attacks, rendering them useless, or they can be harnessed as tools in a broader DDoS campaign against other targets [24–27].

The infamous malware known as the Mirai botnet exploits feeble security measures, including default usernames and passwords, to target IoT devices that are vulnerable [3–5]. When a device becomes infected, it is automatically incorporated into a vast botnet, which is a collection of compromised IoT devices. The Mirai DDoS attack on Dyn occurred in 2016 and caused significant disruption in its services, a prominent DNS provider, affecting a multitude of prominent online platforms [3–5]. This danger served as a stark reminder of the critical nature of IoT device security and prompted enhancements to IoT security awareness and practices. As illustrated in Fig. 2, DDoS attacks employ a network of compromised computers to inundate their target with an insurmountable volume of network traffic, in contrast to DoS attacks.

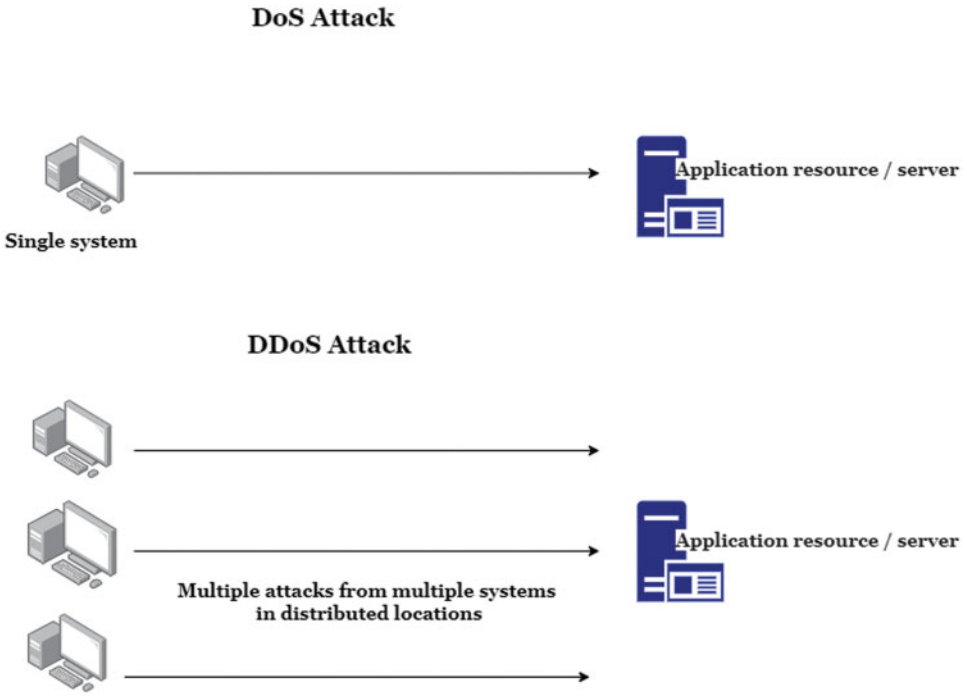

Figure 2: DoS *vs.* DDoS attacks

2.2 Use of AI for DoS\|DDoS Attack Detection in the Context of IoT

The growth of IoT devices has created a slew of security issues, including the potential for DoS and DDoS assaults. Traditional firewalls, intrusion detection systems (IDS), and intrusion prevention

systems (IPS) cannot guard against sophisticated DDoS assaults because they filter regular and suspicious traffic using static, predefined criteria. To counter these threats, the integration of AI has emerged as a powerful solution for early detection and mitigation [28–31]. As of now, IDS and IPS that use AI approaches to filter invasive efforts are more dependable and effective than static, predefined rules in earlier times [30].

AI-driven approaches leverage ML algorithms to analyze the network traffic patterns within an IoT network. These algorithms are trained on historical data to identify normal network behavior, which enables them to detect anomalies that may indicate a DoS or DDoS attack [31–35]. For instance, AI models can scrutinize network traffic for sudden spikes in data volume, unusual packet patterns, or a high number of connection requests, all of which are indicative of attack attempts [28–31]. The ability to swiftly recognize these anomalies allows for rapid reaction and alleviation, reducing the potential impact of the attack. Nonetheless, AI can also facilitate real-time analysis of IoT device behavior, enabling the identification of compromised devices that may be participating in a botnet-driven DDoS attack [36–40]. When identifying such attacks, many factors have to be considered owing to the dynamic nature of IoT and the DoS/DDoS attacks, necessitating feature engineering approaches for reducing computational resources and improving accuracy, which has been overlooked by many researchers.

While AI offers promising capabilities for enhancing DoS/DDoS attack detection in IoT, it is important to consider the dynamic nature of IoT networks and the potential challenges associated with scaling AI solutions for large-scale deployments. Additionally, ongoing research and development in this field are essential to stay ahead of evolving attack techniques and to ensure the robustness and reliability of AI-based security systems [40–43].

For better understanding, Table 1 summarizes and provides a brief comparison of recent related work in the area, which involves the use of AI for IoT-targeted DoS/DDoS attack detection. It offers insights into the underlying algorithm(s) used, the domain applied, and the scope of the study.

Table 1: Summary of recent related work in the area

Reference	Domain applied	Algorithm(s) used	Scope of the study
[1]	Banking industry	Support Vector Machine (SVM), RF, and K-Nearest Neighbors (KNN)	The authors employ the Banking dataset to detect DDoS assaults on financial organizations, and they apply SVM, KNN, and RF algorithms for prediction.
[3]	Smart agriculture	Convolutional Neural Network (CNN), deep neural network, and recurrent neural network	In the context of smart agriculture, the authors propose a Deep Learning (DL) based IDS for DDoS attacks based on three DL models. For the experimental purpose, they have used the CIC-DDoS2019 dataset and the TON_IoT dataset, which contain different DDoS attacks.
[4]	For the entire IoT domain	RF	The authors investigate DoS/DDoS attack detection for IoT using ML methods. The experiment was evaluated based on the Bot-IoT dataset, where the RF classifier shows an accuracy of 99.81%.
[6]	For the entire IoT domain	Autoencoder network model and an improved version of genetic algorithm	The researchers used an IGA-BP network to address the rising problem of Internet security, detecting intrusions with a 98.98% detection rate and 99.29% accuracy.

(Continued)

Table 1 (continued)

Reference	Domain applied	Algorithm(s) used	Scope of the study
[7]	For entire IoT domain	ResNet	The authors presented a method for transforming network traffic data into images and trained a cutting-edge CNN model (ResNet) with a 99.99% accuracy for recognizing DoS and DDoS in binary categorization. Furthermore, the proposed approach has an average accuracy of 87% for identifying eleven various types of DoS and DDoS attack patterns.
[9]	For the entire IoT domain	XGB	Through eleven ML algorithms, the researchers investigate distinct vulnerabilities in the NSL-KDD dataset that potentially affect sensor nodes and networks in IoT contexts. According to the data, XGB is the dominating algorithm among others, with a 97% accuracy and 99.6% Area Under the Curve (AUC) performance.
[11]	For the entire IoT domain	RF	The researchers demonstrate the utilization of ML approaches for four types of DDoS attacks using the CICDDoS-2019 dataset. The proposed ML model demonstrated an accuracy of 99.92%.
[12]	For the entire IoT domain	Multiple Linear Regression	The authors have used regression analysis techniques to build an ML model for DDoS attack detection utilizing the CICIDS 2017 dataset. It has been observed that their proposed model yields a prediction accuracy of 97.86%.
[13]	SDN-enabled IoT network infrastructure	SVM, Naive Bayes (NB), RF, KNN, and Logistic Regression	The researchers suggested an Adaptive Machine Learning-based framework for successful detection and mitigation of DDoS attacks in the context of SDN-enabled IoT infrastructure.
[15]	SDN-enabled IoT network infrastructure	advanced version of SVM algorithms	The authors present an SDN-enabled Distributed Denial-of-Services attack Detection and Mitigation System (SDN-DMS) that uses ML to develop a DDoS detection and mitigation system for IoT devices.
[16]	Smart home/IoT network infrastructure	Decision Tree (DT), RF, Neural Network (NN)	In this research, the authors have employed a variety of ML algorithms to prove that IoT-enabled automated home appliances can detect DDoS attacks using simple, cost-effective ML algorithms.
[17]	For the entire IoT domain	SVM, DT, Naïve Bayes (NB) and MLP	The authors have presented an Information Gain-Based Intrusion Detection System (IGIDS), which is a merger of a filter-based selection approach with an ML algorithm.
[19]	For the entire IoT domain	MLP, Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), KNN, SVM, DT, and RF	This research employs eight ML algorithms to diagnose DDoS attacks in the context of IoT, where they have utilized the NSLKDD dataset for the experiment. It is indisputable that MLP exhibits a higher level of performance, as nearly 99.9% of attacks are successfully identified, as indicated by the results obtained.

(Continued)

Table 1 (continued)

Reference	Domain applied	Algorithm(s) used	Scope of the study
[20]	For the entire IoT domain	NN	The researchers present a detection strategy that is based on a deep neural network that employs feed-forward back-propagation in order to successfully uncover several application layer DDoS attacks. The suggested model attained an accuracy rate of 98%.
[22]	For the entire IoT domain	KNN, DT, RF, and Artificial Neural Network (ANN)	The researchers categorized regular and DDoS attack traffic utilizing the CICDDoS2019 dataset amassed by the Canadian Institute of Cyber Security and employing a number of ML algorithms. ANN, among the implemented ML algorithms, has produced the most favorable outcomes.
[23]	For IoT network infrastructure	RF, XGB, Gradient Boosting (GB), and DT	The researchers utilized a variety of ML algorithms in order to prevent Man in the Middle (MTM) attacks and DoS attacks. In order to do so, they obtained related datasets from the Kaggle website, which was dedicated to MTM and DoS attacks. The findings that were collected demonstrated that these algorithms were extremely capable of detecting MTM and DoS assaults, which demonstrated the usefulness of safeguarding IoT networking devices from these types of attacks.
[24]	SDN-enabled IoT network infrastructure	KNN, ANN, SVM, and DT	In order to Classify SDN traffic as either malicious or benign, the researchers implemented a number of ML algorithms that were outfitted with Neighbourhood Component Analysis (NCA). The researchers utilized the DDoS-attack SDN dataset for this purpose. The experimental outcomes demonstrated that DT outperforms remaining algorithms in terms of accuracy.
[25]	SDN-enabled IoT network infrastructure	Advanced Support Vector Machine (ASVM)	The authors propose a design of an SDN SDN-based DDoS detection system where they have used an ASVM algorithm to detect DDoS attacks.
[27]	Cyber-physical systems	DT	The researchers adopted an ML approach for network anomaly detection and constructed various models based on data to detect distributed (DDoS) attacks on Industry 4.0 cyber-physical systems. In this regard, they have used network traffic statistics obtained from an actual semiconductor production factory and 11 different ML algorithms, where DT has proven to be more accurate, with an accuracy of 99.9%.

(Continued)

Table 1 (continued)

Reference	Domain applied	Algorithm(s) used	Scope of the study
[29]	For IoT network infrastructure	DT and RF	Using ML techniques, the researchers developed a system to detect DDoS attacks employing the Neighbor Discovery protocol, given the intensity of the incursion and the significance of the Neighbor Discovery protocol in Internet Protocol version 6. Overall, the accuracy of the results produced by the DT algorithm and RF algorithm is superior to that of the other algorithms.
[31]	For the entire IoT domain	DT and MLP	By exploiting the Bot-IoT dataset's class imbalance issue, the researchers constructed an innovative IDS utilizing ML and DL models. It is evident from their experiment that DT and MLP performed the best in detecting DDoS and DoS attacks across IoT networks.
[32]	For the entire IoT domain	SVM, RF, DT, Logistic regression, KNN and NB	The authors describe an ML-based attack detection technique with the intention of identifying attack traffic in consumer IoT.
[33]	For the entire IoT domain	RF and XGB	The researchers employed ML techniques to classify and forecast DDoS attack types, employing RF and XGB classification algorithms. In this regard, they have used the UNWS-np-15 dataset extracted from GitHub.
[43]	For the entire IoT domain	ANN	The researchers conducted research to detect and mitigate known and unknown DDoS attacks in real-time environments using the ANN algorithm.
Our work	For the entire IoT domain	RF, XGB, and MLP	Our research employs three ML models with three feature engineering methods for the classification of fifteen types of DoS/DDOS attacks using the CICIoT2023 dataset and, with the best-performing algorithm, design a novel real-time DoS/DDoS attack detection system. On the other hand, apart from evaluating the performance of ML models, we also investigated the efficacy of employed feature engineering methods, as the reduction of insignificant features would lead to an effective classification of such attacks when there is a large volume of data that has been overlooked by the state of the art.

According to the state of the art evaluation, it is clear that AI has already been applied to the majority of IoT domains, including SDN-enabled IoT. Besides this study, research [5] delved into current DDoS detection systems employing both singular and combined machine learning techniques in today's network contexts. It also reviews various ML-based DDoS protection mechanisms that use virtual environments, including cloud computing, software-defined networking, and network functions virtualization. In [33], the authors undertook a comprehensive literature analysis to investigate the present state of DDoS detection approaches and to find the most competent and effective DDoS

detection system utilizing AI. Reference [36] provided a method based on smart contracts and ML as a countermeasure against DoS/DDoS assaults in the 5G background by concealing a secured server within a blockchain network and dynamically limiting the size of DoS/DDoS via transaction fees. Authors in [37] focused their study on detecting DDoS assaults by designing an IDS customized to the Internet of Vehicle (IoV) systems using AI approaches. Researchers did a study in [38] of the literature on the application of DDoS to identify DDoS assaults. Research in [41] used RF and MLP models for predicting application layer DDoS attacks. Meanwhile, reference [40] presented an explainable artificial intelligence (XAI)-based innovative technique to detect DDoS assaults that identify irregular network traffic flows by analyzing the traffic at the network layer.

3 Materials and Methods

Having provided a brief overview of related work on DoS/DDoS attacks in the context of IoT and the use of AI for such attack detection, this section provides an in-depth overview of the methodology followed, the dataset used, and performance indicators. According to the reviewed literature, the majority of IoT attack detection systems described in the literature are incapable of identifying the most recent DoS and DDoS attacks. The main reason for this is that the majority of these models were trained using obsolete datasets or datasets that fail to encompass a diverse array of DDoS attack types. These situations are both problematic. Hence, we utilized CICIoT2023, a real-time dataset and benchmark for large-scale attacks in an IoT environment, to conduct our experiment [44]. Overall it includes data from 105 IoT devices representing a variety of types and brands. The dataset captures both benign and malicious traffic, with 33 types of attacks categorized into seven classes: DDoS, DoS, Recon, Web-based, brute force, spoofing, and Mirai. The attacks are executed by malicious IoT devices targeting other IoT devices. This comprehensive dataset facilitates the development of security analytics for IoT environments. Following the filtration of the dataset to retrieve only DoS and DDoS, data preprocessing techniques were applied to the dataset.

The dataset underwent data preprocessing, which involved deleting null and empty values, as well as applying scaling and balancing algorithms. Subsequently, we utilized three feature engineering techniques to ascertain the most significant characteristics. Following this, the dataset was divided into 80% training data and 20% testing data. The training data was used to train the ML models, while the testing data was used to evaluate the models' performance after training. Afterwards, based on the outlined performance criteria, the best-performing model was used to deploy to the simulation environment for real-time detection of IoT-targeted DoS/DDoS cyber-attacks. The methodology of the research is depicted in Fig. 3.

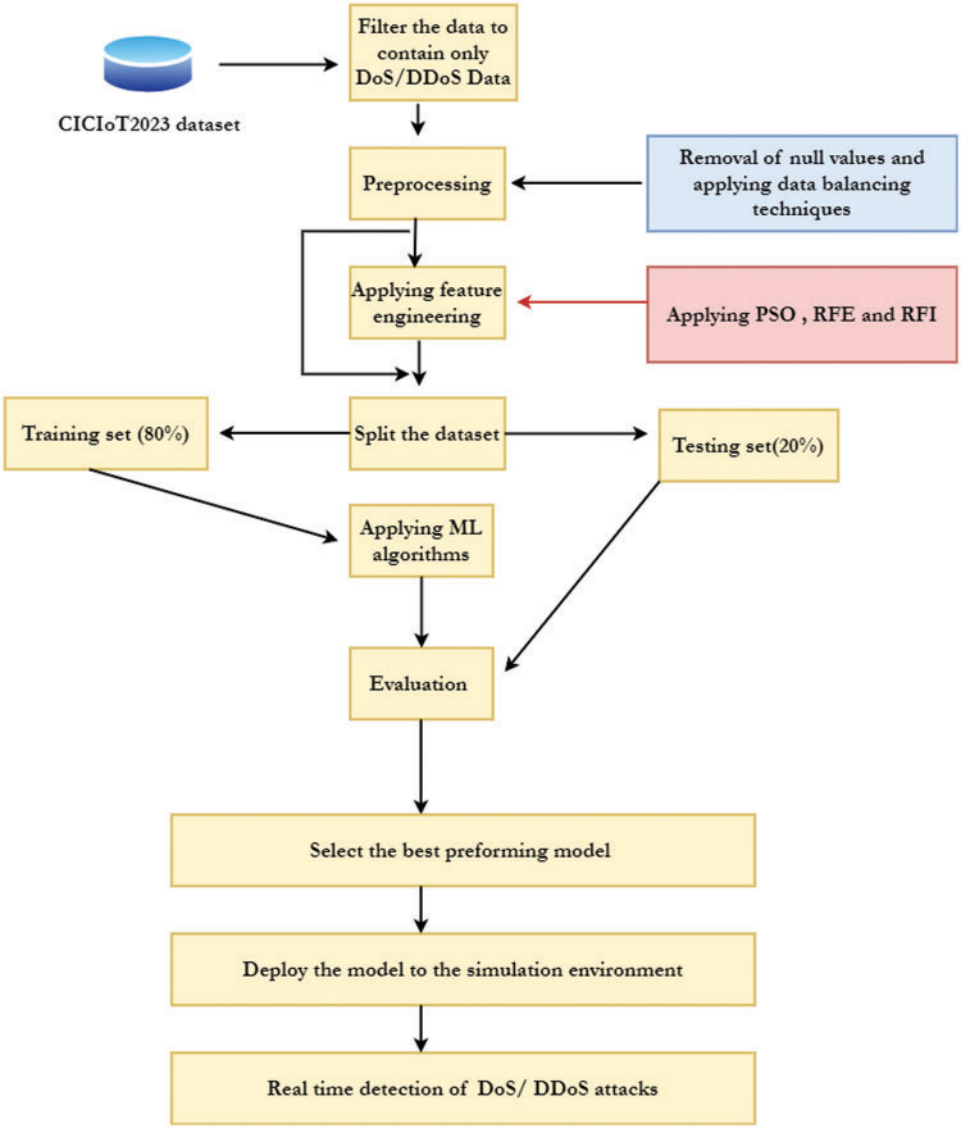

Figure 3: The methodology followed in the research

3.1 Dataset Description, Preprocessing, and Feature Engineering

Overall, once the dataset is filtered for DoS/DDoS attack types and following the preprocessing stage, it contains 956,382 records. The filtered-out dataset contains 15 classes of DoS/DDoS attack types that target IoT ecosystems, which are further described in the following for better understanding:

1. DDoS RST-FIN flood attack
   In this type of DDoS attack, the assailant inundates the target server with a large number of TCP packets containing both the RST (Reset) and FIN (Finish) flags [4,5]. These packets can disrupt ongoing TCP connections by closing them abruptly, causing service disruptions.
2. DoS TCP flood attack

An attack in which the attacker inundates a target with a large volume of TCP (Transmission Control Protocol) packets, overwhelming its resources and making it unresponsive to legitimate traffic [5–7].

3. DDoS ICMP flood
   In a DDoS ICMP (Internet Control Message Protocol) flood attack, the attacker sends a massive number of ICMP packets to the target, overloading its network and causing a denial of service [4–7].
4. DoS UDP flood attack
   A DoS attack wherein the intended recipient or network is inundated with an excessive quantity of UDP (User Datagram Protocol) packets, resulting in the depletion of its resources and the subsequent inaccessibility to authorized users [5–8].
5. DoS SYN flood attack
   In this form of DoS attack, the assailant inundates a large number of TCP SYN (Synchronize) packets to the target but does not complete the handshake process. This exhausts the target's resources, preventing it from handling legitimate requests [7–10].
6. DDoS synonymous IP flood attack
   A variant of DDoS SYN flood attack, where the attacker uses multiple IP addresses to send SYN packets, which makes it harder to mitigate the attack through IP-based blocking [11–13].
7. DDoS PSHACK flood attack
   In this form of DDoS attack, the attacker sends packets with the PSH (Push) and ACK (Acknowledgment) flags set. These packets can be used to saturate network bandwidth or overwhelm a server, causing a denial of service [14–18].
8. DDoS TCP flood attack
   This attack, which is coordinated by multiple compromised systems in a DDoS fashion [14–18], works similarly to the DoS TCP flood in that it floods a target with a large number of TCP packets.
9. DDoS UDP flood attack
   This is a DDoS attack that floods the target with a massive number of UDP packets, causing network congestion and rendering the target unresponsive [5–9].
10. DDoS ACK fragmentation attack
    In this form of DDoS attack, the attacker sends TCP ACK (Acknowledgment) packets that are fragmented, making it harder for network defenses to detect and mitigate the attack [4–7].
11. DoS HTTP flood attack
    In a DoS attack using HTTP flooding, the attacker overwhelms a web server with a high volume of HTTP requests, potentially causing the server to become unresponsive [5–8].
12. DDoS ICMP fragmentation attack
    A DDoS attack that involves sending fragmented ICMP packets to a target, potentially overwhelming its network infrastructure [6–9].
13. DDoS UDP fragmentation attack
    In this DDoS attack, the attacker sends fragmented UDP packets to the target [7–10], making it challenging to detect and mitigate the attack.
14. DDoS HTTP flood attack
    A form of DDoS attack where the target is flooded with a high volume of HTTP requests, overwhelming its capacity to serve legitimate users [14–16].
15. DDoS SlowLoris attack

SlowLoris is a type of DDoS attack that targets web servers by keeping many connections open but not completing the HTTP request, effectively tying up server resources and causing it to become unresponsive [14–18].

The dataset underwent preprocessing to prepare it for ML classifiers subsequent to the data filtration phase. Firstly, the null, NaN, empty values, and outliers were dropped from the dataset, and then the dataset was normalized using ***MinMax*** scaling method. Overall, ***MinMax*** scaling scales the features of the dataset to a specific range, typically between 0 and 1. This process standardizes the feature values, making them more consistent and comparable. The scaling formula, denoted as x, is transformed into the converted data, denoted as *x*'. The labels "*min*" and "*max*" refer to the lowest and highest values, respectively, in the column where x is located.

$$x' = ((x - min))/((max - min)) \tag{1}$$

Finally, a label encoding was done to convert the textual labels to numerical format to enable the underlying machine learning system to improve its decision-making capabilities in utilizing these labels. The dataset was partitioned into 80% training data and 20% testing data following the conclusion of the preprocessing and feature engineering phases. On the other hand, to provide the comparison of the employed feature engineering methods, directly after the preprocessing data set was split as training and test data without applying feature engineering, as highlighted in Fig. 3.

3.1.1 Particle Swarm Optimization (PSO) for Feature Selection

PSO offers a dynamic and efficient way to select the essential features for underlying ML models employing nature-inspired approaches [45–48]. The central concept of PSO is to simulate the social behavior of birds flocking or fish schooling [45–48]. Within the framework of feature selection, each particle symbolizes a prospective solution, specifically, a subset of characteristics derived from the dataset.

Suppose the dataset has no features. Each particle i in the swarm can be represented as a vector $X_i = [x_{i1}, x_{i2}, ..., x_{in}]$, where each x_{ij} can be a binary value indicating the presence (1) or absence (0) of the j-th feature in the i-th particle's selected feature subset.

Each particle has a velocity $V_i = [v_{i1}, v_{i2}, ..., v_{in}]$, which guides its movement through the search space (the set of all possible feature combinations). The velocity is updated based on the particle's own best-known position (personal best) and the best-known position among all particles in the swarm (global best). This update can be mathematically represented as:

$$V_i(t+1) = w \cdot V_i(t) + c_1 \cdot rand() \cdot (P_{best}, i - X_i(t)) + c_2 \cdot rand() \cdot (G_{best} - X_i(t)) \tag{2}$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{3}$$

Here, w is the inertia weight that controls the impact of the previous velocity, c_1, and c_2 are the acceleration coefficients that determine the influence of the personal best position and the global best position on the movement of the particle. The function *rand*() generates a random number between 0 and 1. P_{best}, i refers to the personal best position of particle i, whereas G_{best} represents the best location found by any particle in the swarm. The position $X_i(t+1)$ after the update indicates the new set of selected features.

The iterative process persists until a specified stopping threshold is achieved, like the highest number of iterations or a satisfactory error rate. PSO for feature selection is particularly effective in exploring and exploiting the search space, leading to an optimized subset of features in the ML model.

3.1.2 Recursive Feature Elimination (RFE) for Feature Selection

RFE is another effective feature selection method used in ML. RFE starts with a set of all possible features and iteratively removes the least important ones [45–48]. Suppose the initial feature set is X $= \{x_1, x_2, ..., x_n\}$ where each x_i is a feature and n is the total number of features. The goal of RFE is to find a subset $X' \subseteq X$ that maximizes the performance of the ML model. The process begins with the full set X and iteratively removes features one at a time. At each step, the feature whose removal causes the least decrease in the performance metric *P* is eliminated.

This can be represented as:

$$X_k + 1 = X_k - \{argmin\, x \in X_k\, \Delta P(X_k, x)\} \tag{4}$$

Here, X_k is the set of features at the *k*-th step, and $\Delta P(X_k, x)$ is the change in performance when feature x is removed from the set X_k. The process continues until the desired number of features is reached or until there is a significant drop in model performance. This technique is particularly effective when working with data that has a large number of dimensions and needs to enhance the performance or interpretability of the ML model.

3.1.3 RF Feature Importance (RFI) for Feature Selection

RFI is a technique for figuring out how important each feature (input variable) is for generating predictions [45]. It measures how each variable affects or contributes to the overall predicted accuracy of the model [46]. Overall, it enables us to determine which features have the greatest impact on the model's predictions [45,46]. A popular method for determining RFI is to utilize the Mean Decrease in Accuracy (MDA) or Gini impurity, where, in our research, we employed the Gini impurity method for determining the feature relevance towards final classification.

1. Gini impurity
   A measurement of a group of data points disorder or impurity is called the Gini impurity. By calculating how frequently a feature is used to split data at decision tree nodes and how much it reduces the Gini impurity, one may determine the Gini significance of a feature in the context of an RF model [45,46].
2. MDA
   MDA quantifies the drop-in model accuracy that occurs when a certain feature is eliminated or has its values randomly rearranged. Features are deemed more significant if their removal results in a significant drop in accuracy. After determining the feature importance, this data could be used to identify the features that are most pertinent to the concerned problem, learn more about the connections between features and predictions, or even create a visual representation of the importance scores to help stakeholders understand the relative significance of various features [45,46]. Ultimately, this would be very helpful in reducing the computing resources needed to execute the ML algorithms while a huge number of variables are involved.

3.2 Employed Machine Learning Models

This section provides a brief overview of the ML algorithms we employed in the research.

3.2.1 XGBoost

XGB is a highly effective ML algorithm that has gained recognition for its remarkable predictive capabilities across a range of tasks, with a particular emphasis on structured and tabular data [2]. Its

architecture is comprised of an ensemble of decision trees, with each tree being constructed iteratively in order to rectify the mistakes made by its predecessors [2,12]. XGB utilizes a gradient-boosting architecture to construct a robust learner by aggregating the predictions of numerous weak learners (typically shallow decision trees). The algorithm optimizes both bias and variance in a principled manner [12] as it iteratively fits new trees to the residual errors from the previous iteration in order to minimize a loss function. This is accomplished by combining regularization techniques with an intelligent split-finding algorithm that efficiently evaluates potential divides in a dataset.

3.2.2 Random Forest

In both classification and regression tasks, the RF algorithm is a widely recognized ensemble learning method. It comprises a collection of decision trees, with each tree being built by employing bootstrapping (random sampling with replacement) on a unique subset of the dataset. Furthermore, to promote diversity among the individual trees during construction, at each split, a selection of features is made randomly. In order to arrive at a final prediction, the algorithm combines the outcomes of each tree. This is typically accomplished by aggregating the results for regression or majority voting, respectively, as depicted in Fig. 4 [1,13]. The mathematical expression denoting RF is derived from the aggregation method employed (e.g., mean for regression or mode for classification) and is composed of the contributions of individual decision tree predictions. To summarize, RF is an ensemble technique that generates numerous decision trees utilizing the collective intelligence of the trees to generate accurate and reliable predictions through random feature selection and data sampling [46–49].

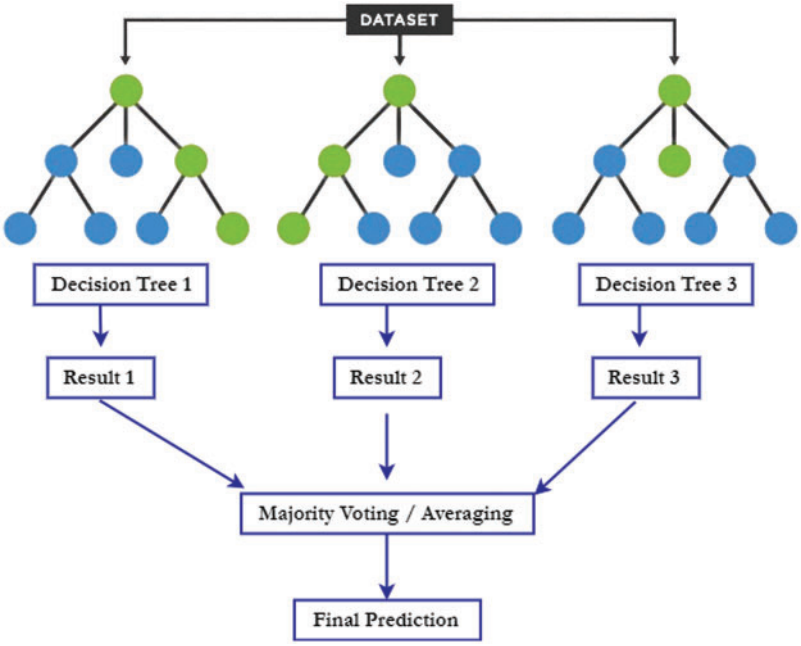

Figure 4: Random forest algorithm

3.2.3 MLP

The MLP algorithm is a fundamental artificial neural network architecture used for various machine learning tasks, including classification and regression. Its architecture consists of multiple layers of interconnected neurons, comprising an input layer, one or more hidden layers, and an output

layer [1,10,13]. Neurons inside each layer are linked to all neurons in the neighboring layers, with each connection being assigned a corresponding weight [50–52]. The mathematical formula for MLP involves a series of weighted sums and activation functions applied at each neuron. Sigmoid or rectified linear units (ReLU) activation functions are commonly used to incorporate non-linearity into the model, which facilitates the discovery of intricate patterns within the data.

Forward propagation and back-propagation are components of an MLP's training procedure [10]. During forward propagation, input data is transmitted through the network in order to generate predictions, while errors are calculated and utilized to iteratively update the weights with the objective of minimizing a loss function. In conclusion, the MLP is a flexible neural network architecture that models complex data relationships using interconnected layers of neurons with activation functions, rendering it a potent instrument for a vast array of ML tasks.

3.3 Performance Metrics

Four metrics were utilized in the study to assess the performance of the ML algorithms: F1-score, accuracy, precision, and recall [51–54]. Furthermore, the confusion matrix has been implemented to visually represent the rates of true positives and false positives, enabling a distinct differentiation between outcomes that were accurately classified and those that were misclassified. The metrics employed in this research include:

- TP (True Positives)
- TN (True Negatives)
- FP (False Positives)
- FN (False Negatives)

1. Accuracy
   This is responsible for assessing the performance of classification models by calculating the correct prediction percentage in a dataset using the following formula:

$$Accuracy = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}) \tag{5}$$

2. Precision
   The percentage of labels that were correctly detected in comparison to the total number of positive classifications:

$$Precision = TP/(TP + FP) \tag{6}$$

3. Recall
   Proportion of accurately recognized labels in relation to the total occurrences of a specific label within a dataset.

$$Recall = TP/(TP + FN) \tag{7}$$

4. F1-score
   F1-score denotes the harmonic average of precision and recall.

$$F1 = 2 \times (\mathrm{Pre} \times \mathrm{Rec})/(\mathrm{Pre} + \mathrm{Rec@.}) \tag{8}$$

4 Evaluation Results and Implementation

This section demonstrates how our underlying ML models perform with the DoS/DDoS data using our employed feature engineering methods and the implementation of the real-time detection system. Overall, there are 15 classes of DoS/DDoS attacks that would be used to test the generalization ability of proposed ML models. The experiment was executed on a personal computer with an Intel Core i5 2.70 GHz processor, 8 GB RAM, and 2 GB of video memory. Firstly, the ML models were trained without using feature optimization methods with all the features. Fig. 5 showcases the performance evaluation metrics of the XGB model. Overall, XGB showed an accuracy of 99.78%, along with the same value for other performance metrics. Fig. 6 showcases the confusion matrix based on the predictions made by the XGB model.

Figure 5: Performance evaluation metrics of the XGB model

The performance evaluation metrics pertaining to the RF model are depicted in Fig. 7. Accordingly, RF showed a 99.78% accuracy and 99.78% for precision, recall, and F1-score.

Fig. 8 showcases the confusion matrix based on the predictions made by the RF model.

The performance evaluation metrics pertaining to the MLP model are depicted in Fig. 9. Accordingly, MLP showed a 99.10% accuracy and 99.10% precision, recall, and F1-score.

Fig. 10 showcases the confusion matrix based on the predictions made by the MLP model.

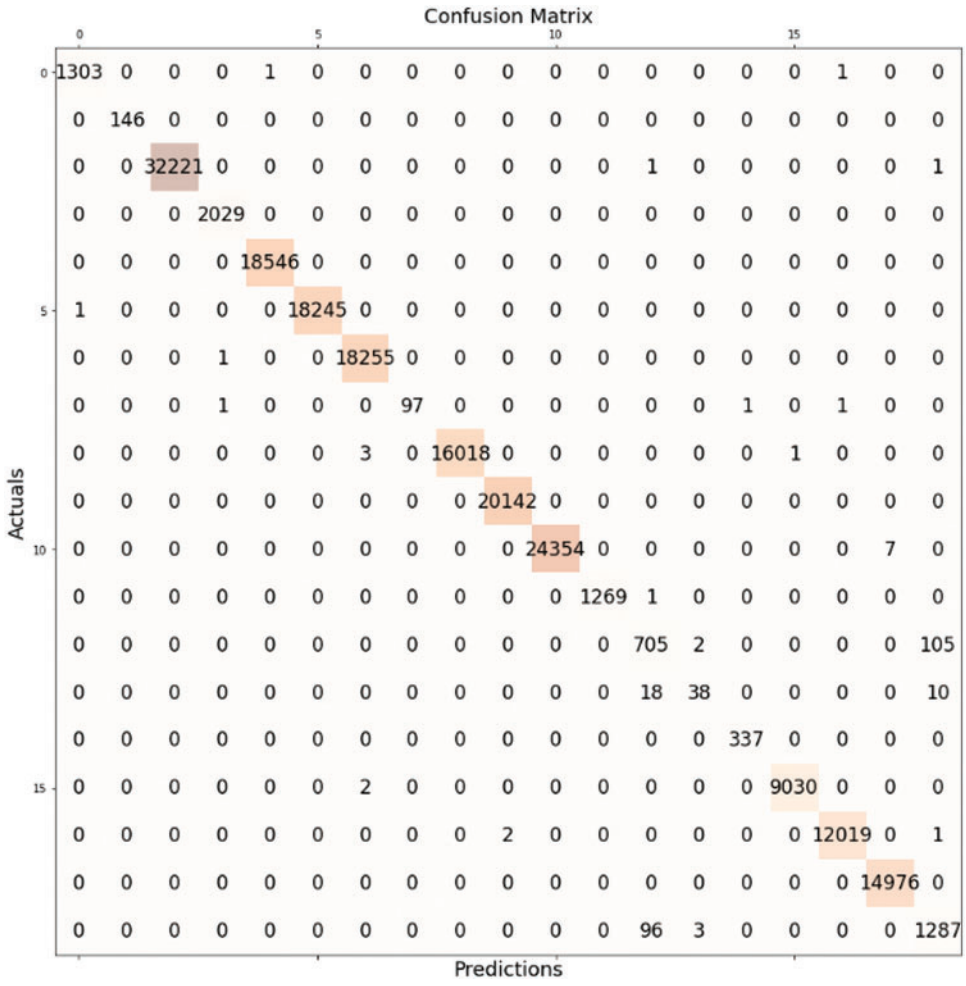

Figure 6: Confusion matrix pertaining to the XGB model

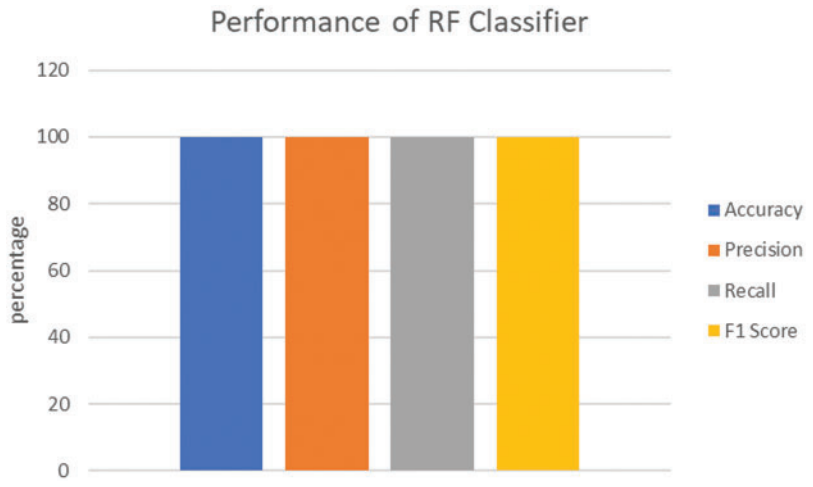

Figure 7: Performance evaluation metrics of the XGB model

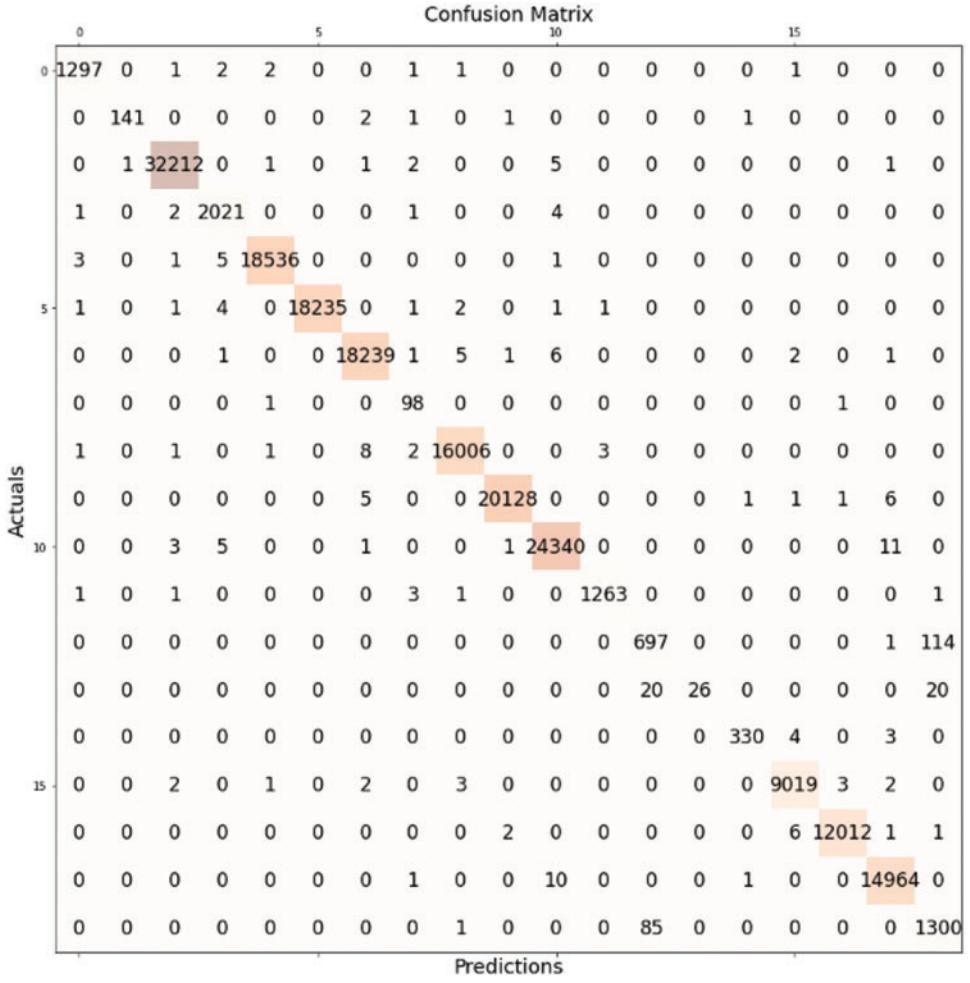

Figure 8: Confusion matrix pertaining to the XGB model

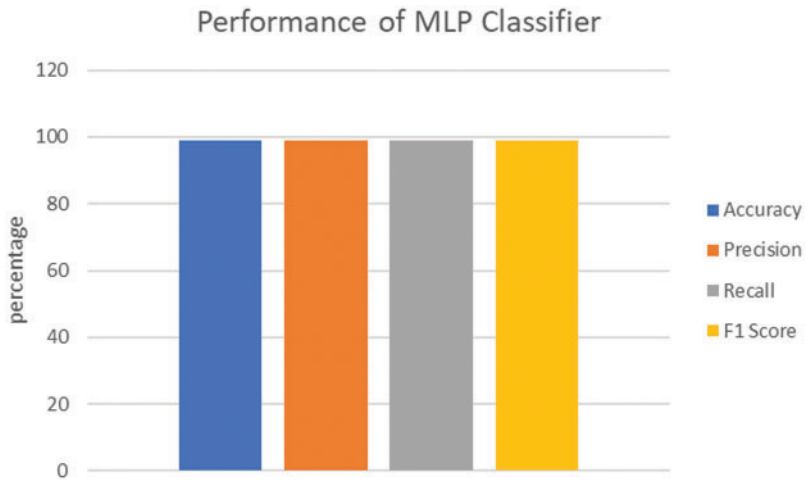

Figure 9: Performance evaluation metrics of the MLP model

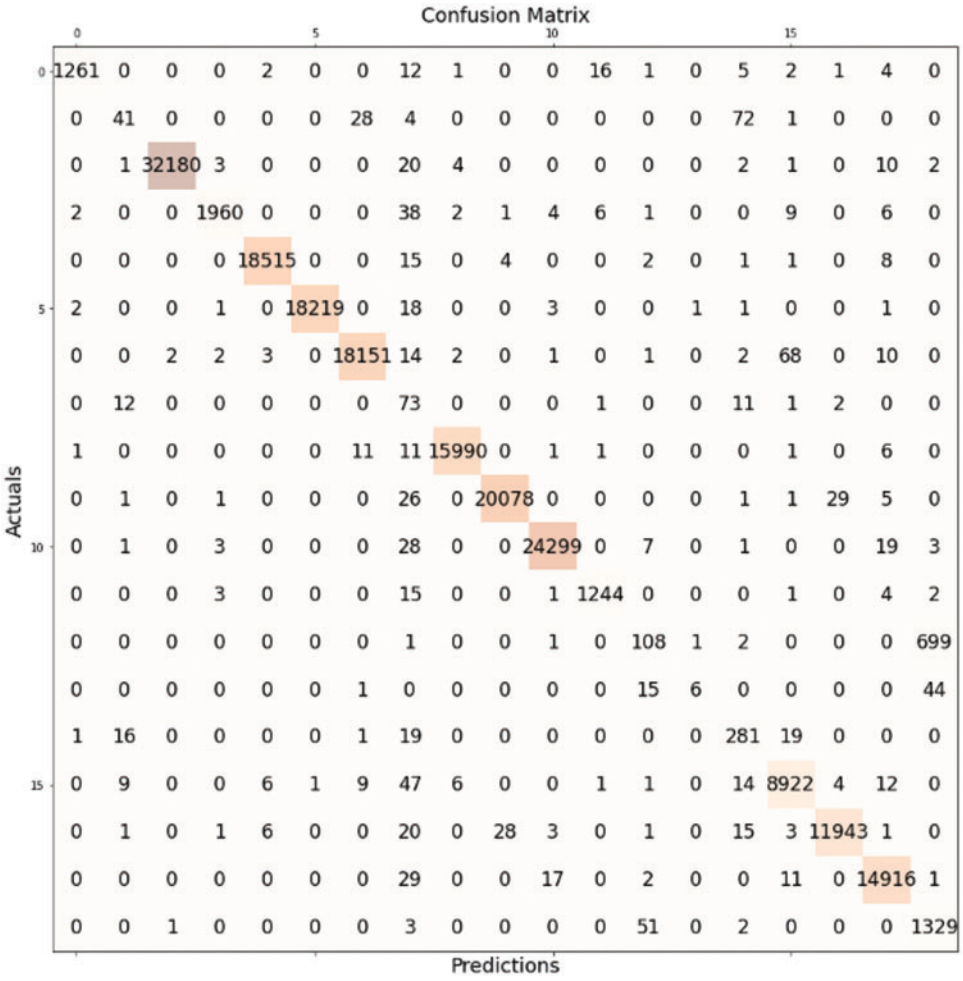

Figure 10: Confusion matrix pertaining to the MLP model

Table 2 presents the performance evaluation metrics of all ML models used, along with their execution time. Accordingly, it is evident that even though both XGB and RF offer a slightly higher accuracy compared to MLP, the execution time is considerably higher for XGB. On the other hand, RF showcased similar performance, with better execution time compared to XGB.

Table 2: Performance evaluation metrics of all models (without employing feature engineering)

Model	Number of features	Accuracy	Precision	Recall	F1-score	Execution time (seconds)
XGB	47	99.78	99.78	99.78	99.78	1050
RF	47	99.78	99.78	99.78	99.78	204
MLP	47	99.10	99.10	99.10	99.10	487

As the second step, all ML algorithms were trained employing RFI. Fig. 11 elaborates on the Gini importance values of all the features we have selected for classification. 21 variables out of 47 variables have Gini values of 0. Hence, those variables could be considered unnecessary for further analysis.

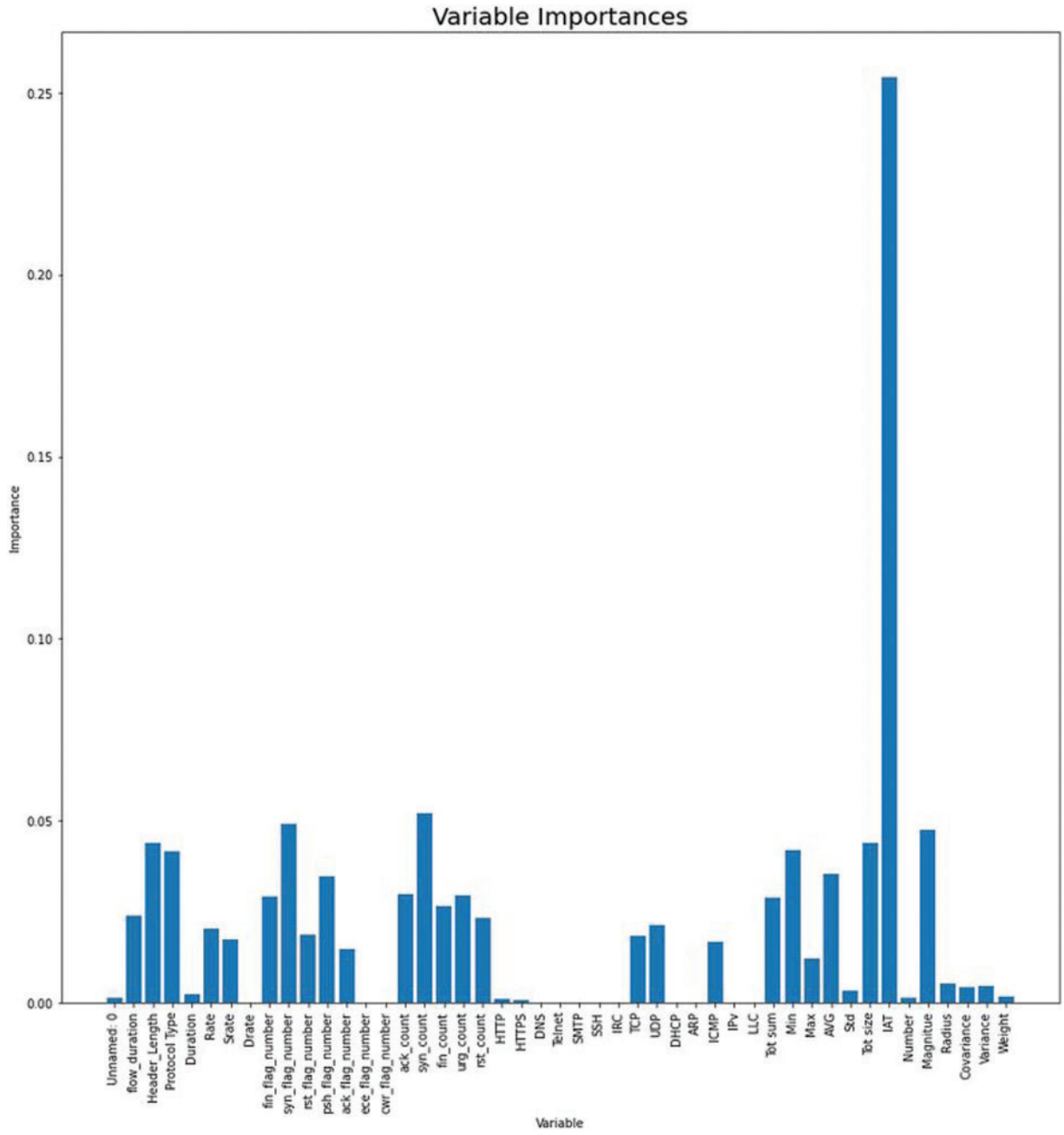

Figure 11: Normalized Gini importance values of each variable based on the RF algorithm

Following applying the RFI to all the features, we have measured the execution time for the classification of DoS/DDoS attacks with a reduced set of feature variables. Table 3 summarizes the performance metrics along with execution times each ML model has taken after employing the RFI method.

Table 4 summarizes the performance metrics along with execution times each ML model has taken, after employing the PSO method.

Table 3: Performance evaluation metrics of all models (after employing RFI)

ML model	Number of features	Accuracy	Precision	Recall	F1-score	Execution time (seconds)
XGB + RFI	27	99.80	99.80	99.80	99.80	530.45
RF + RFI	27	99.80	99.80	99.80	99.80	198.155
MLP + RFI	27	99.00	99.00	99.00	99.00	256.735

Table 4: Performance evaluation metrics of all models (after employing PSO)

ML model	Number of features	Accuracy	Precision	Recall	F1-score	Execution time (seconds)
XGB + PSO	26	99.93	99.93	99.93	99.93	491.023
RF + PSO	26	99.90	99.90	99.90	99.90	147.554
MLP + PSO	26	99.40	99.40	99.40	99.40	280.16

Table 5 summarizes the performance metrics along with execution times each ML model has taken after employing the RFE method.

Table 5: Performance evaluation metrics of all models (after employing RFE)

ML model	Number of features	Accuracy	Precision	Recall	F1-score	Execution time (seconds)
XGB + RFE	27	99.90	99.90	99.90	99.90	520.14
RF + RFE	27	99.81	99.81	99.81	99.81	198.1
MLP + RFE	27	99.20	99.20	99.20	99.20	248.780

Figs. 12 and 13 illustrate the accuracy variations and execution time variations of all three machine learning algorithms after using with PSO method. It clearly shows that PSO has significantly enhanced the performances of all three algorithms.

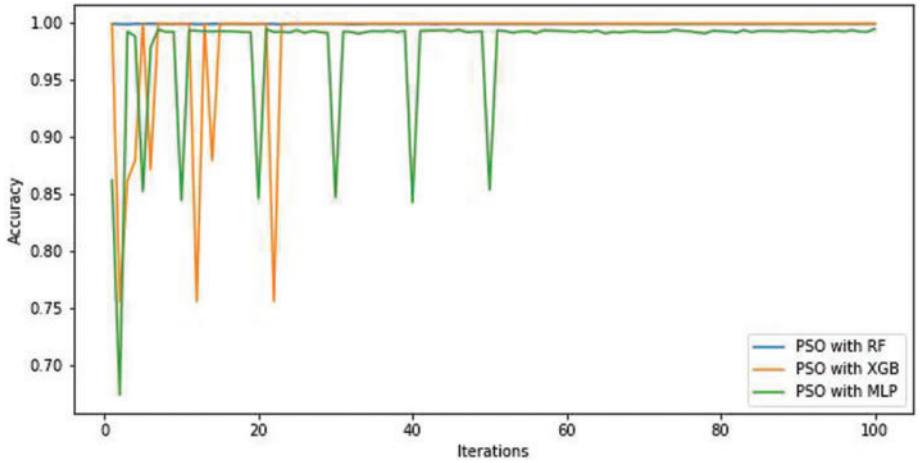

Figure 12: Accuracy metric variations of RF, XGB, and MLP algorithms with PSO technique

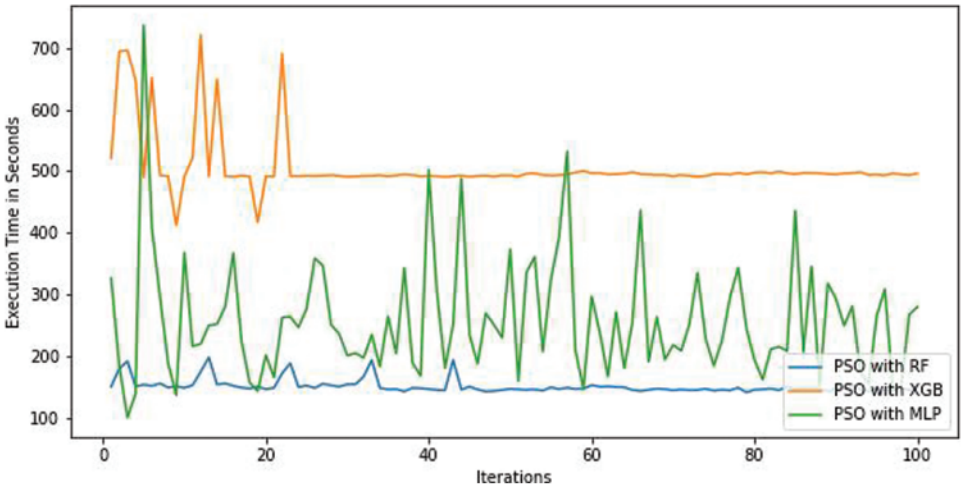

Figure 13: Execution time variations of RF, XGB, and MLP algorithms with PSO technique

From the results highlighted in Tables 4–6, it is evident that feature engineering has played a crucial role in improving the performance and reducing the execution time taken for the underlying ML algorithms. A significant improvement can be seen in the performance metrics, as well as the execution time when employing PSO for feature selection. As a result, this research would provide a vital contribution to research, dealing with a huge number of feature variables and restricted to limited computing resources. Overall, PSO excels in feature selection for ML due to its ability to balance exploration and exploitation effectively in high-dimensional spaces. Its versatility allows it to handle various types of optimization problems without the need for gradient information. On the other hand, it is simple to implement, requires fewer function evaluations to reach satisfactory solutions, and reduces the risk of overfitting by selecting the most relevant features. Overall, with the results obtained from the research, it is evident that feature engineering is an essential component of ML when it comes to detecting DoS/DDoS attacks, and XGB outperforms all other employed ML algorithms in terms of accuracy, precision, recall and F1-score where it also spends a less execution time.

Table 6: Specification of the Raspberry Pi device

Processor	Quad-core Cortex-A72 1.8 GHz
Memory (GB)	8
Connectivity	2.4 and 5.0 GHz IEEE 802.11ac wireless, Bluetooth 5.0, BLE Gigabit Ethernet

4.1 Implementation of the Real-Time Attack Detection System

With our experiment results, it is evident that XGB outperforms all other algorithms in terms of outlined performance criteria when used along with PSO for feature engineering. Thus, for our real-time DoS/DDoS attack detection system, we intend to use the XGB algorithm to detect DoS/DDoS attacks in real-time. Once the best-performing model has been selected, the next step involves serializing/saving the XGB model using the Python pickle module to design the detection system. The implementation steps are highlighted in Fig. 14.

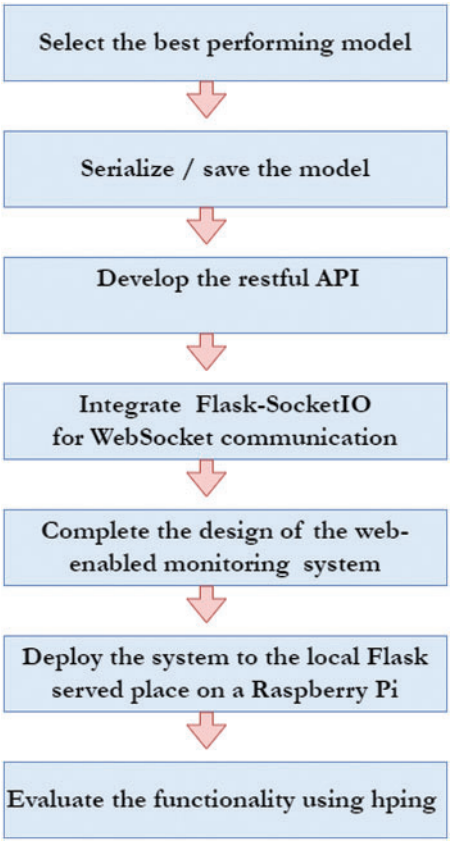

Figure 14: Implementation steps

For the design of the real-time monitoring system, a Python-based framework, Flask, was used to develop the front end of the monitoring system. Further, in order to create an interactive web application, HTML (Hypertext Markup Language), JavaScript and CSS programming languages were also used, employing Agile software methodology. The Flask-SocketIO extension was used to integrate WebSocket functionality with the system so it can communicate with the underlying server to which the application is deployed to retrieve the incoming network traffic. Once the development was done, the web-enabled system was deployed to a local Flask web server setup in a Raspberry Pi device (which acts as a vulnerable IoT device) with the following specifications (depicted in Table 6).

Upon the successful deployment of the system in the network, the real-time attack detection system was tested for its accuracy using hping3, which is designed for crafting and sending arbitrary IP packets and is often used for tasks like network stress testing, which was installed on a separate personal computer with Kali Linux operating system installed; attached to the same network that the Raspberry Pi is linked. Fig. 15 depicts the hping3 command used to simulate the DoS SYN flood attack against the IoT Raspberry Pi device.

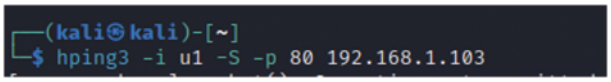

Figure 15: DoS SYN flood attack against the IoT Raspberry Pi device

Fig. 16 showcases the illustration of the developed real-time ML-enabled DoS/DDoS attack detection system, which primarily comprised of two modules to generate alerts and visualization of the historical data (number of occurrents DoS/DDoS attacks received on each day). Overall, with the evaluation results, it is evident that the developed system is working as expected without any bugs, and the system is able to generate alerts in real-time.

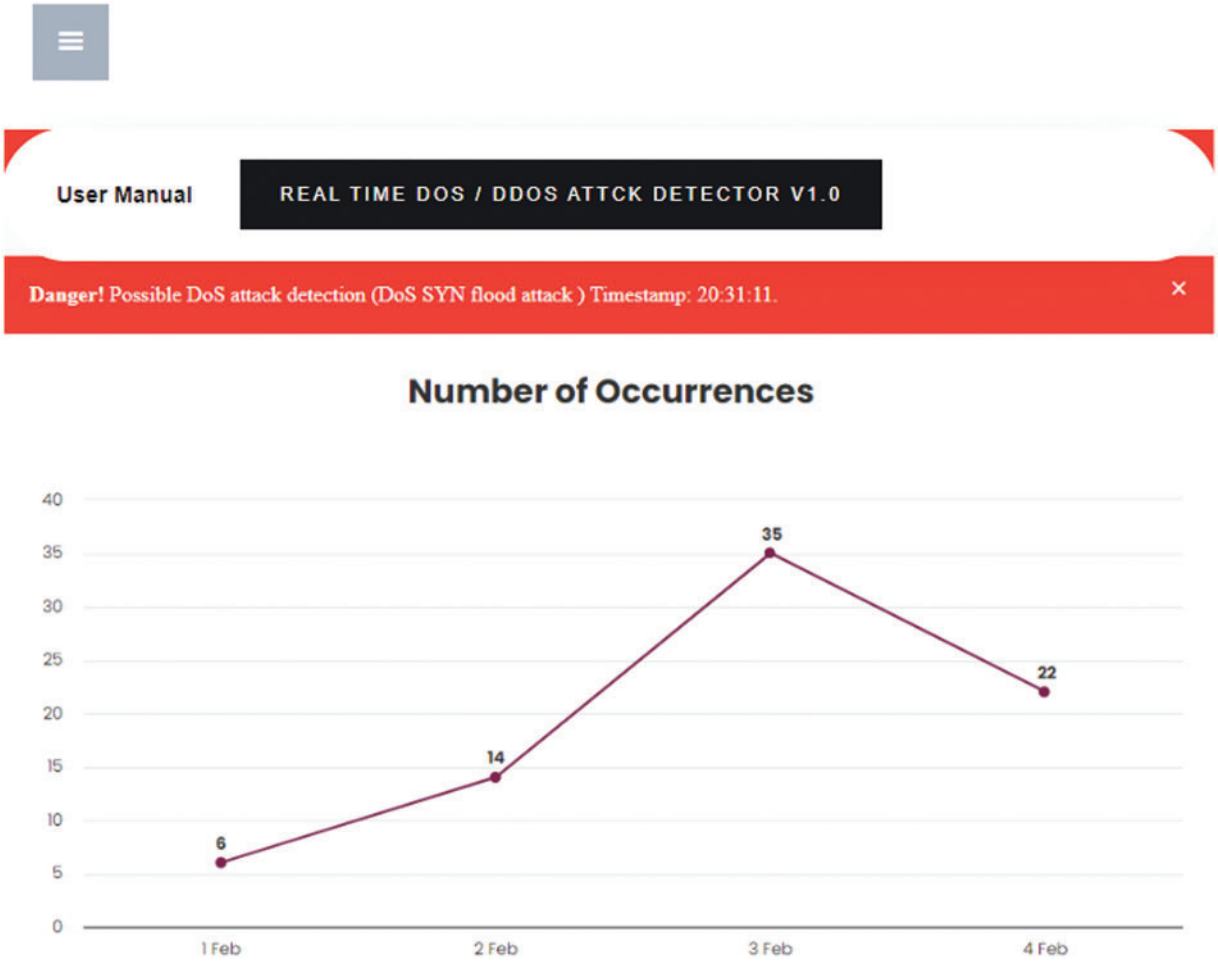

Figure 16: Developed ML-enabled real-time DoS/DDoS attack detection system

Table 7 presents a brief comparison of our research with similar studies that have been done in the area of the use of ML for DoS/DDoS attack detection in the context of IoT. Overall, it is evident that, apart from encompassing real-time monitoring capability, our research outweighs other research in terms of obtained accuracy metrics, employing novel feature engineering approaches and a variety of DoS/DDoS attacks considered.

Table 7: Comparison of similar research (✓-Yes, ×-No)

References	Algorithm used	Feature engineering was considered employing PSO, RFI and RFE	Variety of DoS/DDoS attack considered	Dataset used	Encompasses real-time monitoring capability	Accuracy metrics achieved
[1]	RF, SVM, and KNN	×	01	Banking sector dataset	×	SVM–99.5%, KNN–97.50% and RF–98.74%
[4]	RF	×	01	Bot-IoT dataset	×	99.81%
[11]	RF	×	04	CICDDoS-2019	×	99.92%
[12]	Multiple linear regression	×	01	CICIDS 201	×	97.86%
[33]	RF and XGB	×	09	UNWS-np-15 dataset	×	RF–89%, XGB–90%
Our work	RF, MLP, and XGB	✓	15	CICIoT2023	✓	XGB–99.93%, RF–99.90%, MLP–99.40%

5 Conclusion

In conclusion, this research underscores the urgent need for robust and efficient DoS and DDoS attack detection systems tailored to the unique characteristics of IoT devices. With the exponential growth of IoT and the inherent security vulnerabilities of many devices, the importance of effective security solutions cannot be overstated to guarantee the sustainable development of domains in which IoT is served. Our research focused on evaluating the effectiveness of three machine learning algorithms–XGB, RF, and MLP–and emphasized the role of feature engineering in identifying DoS/DDoS attacks targeting IoT devices. Nonetheless, by employing the best-performing ML algorithm with an accuracy of 99.93%, which is XGB, a prototype of a novel real-time DoS/DDoS attack detection system was developed, significantly enhancing the network's resilience against malicious threats and showcasing remarkable effectiveness in fortifying network security infrastructure, which also can be integrated with any IoT networks. Overall, our findings revealed that while all three algorithms exhibited commendable accuracy in detecting these attacks, employing PSO for feature engineering made a significant improvement in accuracy (high accuracy) and resource usage (less execution time) compared to the RFI and RFE.

Looking ahead, future research in this field should consider several key avenues. Firstly, expanding the dataset to incorporate a broader spectrum of IoT devices and network conditions will enhance the models' adaptability to diverse real-world scenarios, focusing on features that are commonly associated with DoS/DDoS attacks across various contexts, such as rate of requests, packet sizes, and traffic irregularities. Secondly, research should explore adaptive and self-learning models that can autonomously adapt to new attack vectors and tactics without requiring manual retraining. Additionally, implementing inherent security measures in IoT devices can significantly reduce their vulnerability to DoS and DDoS attacks, warranting further investigation. Lastly, integrating the

developed models into existing IoT security frameworks will provide a comprehensive approach to IoT security, combining intrusion detection with other protective measures. In summary, as the IoT continues its transformative impact across various industries, safeguarding the security and resilience of IoT devices against DoS and DDoS attacks remains a top priority. The insights gained from this research lay the groundwork for advanced and adaptable security solutions capable of safeguarding the expanding IoT landscape against evolving cyber threats.

Acknowledgement: The researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support.

Funding Statement: This research was funded by Qassim University (QU-APC-2024-9/1).

Author Contributions: The authors confirm their contribution to the paper as follows: study conception and design: Abdullah Alabdulatif, Navod Neranjan Thilakarathne, Mohamed Aashiq; data collection: Navod Neranjan Thilakarathne, Mohamed Aashiq; analysis and interpretation of results: Abdullah Alabdulatif, Navod Neranjan Thilakarathne, Mohamed Aashiq; draft manuscript preparation: Abdullah Alabdulatif, Navod Neranjan Thilakarathne, Mohamed Aashiq. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The required dataset was generated from a testing environment by the researchers. Dataset can be obtained from the corresponding authors upon the request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] U. Islam *et al.*, "Detection of Distributed Denial of Service (DDoS) attacks in IoT based monitoring system of banking sector using machine learning models," *Sustainability*, vol. 14, no. 14, pp. 8374, Jul. 2022. doi: 10.3390/su14148374.

[2] T. H. H. Aldhyani and H. Alkahtani, "Cyber security for detecting distributed denial of service attacks in Agriculture 4.0: Deep learning model," *Mathematics*, vol. 11, no. 1, pp. 233, Jan. 2023. doi: 10.3390/math11010233.

[3] M. A. Ferrag, L. Shu, H. Djallel, and K. -K. R. Choo, "Deep learning-based intrusion detection for distributed denial of service attack in Agriculture 4.0," *Electronics*, vol. 10, no. 11, pp. 1257, May 2021. doi: 10.3390/electronics10111257.

[4] A. Mihoub, O. B. Fredj, O. Cheikhrouhou, A. Derhab, and M. Krichen, "Denial of service attack detection and mitigation for Internet of Things using looking-back-enabled machine learning techniques," *Comput. Electr. Eng.*, vol. 98, no. 3, pp. 107716, Mar. 2022. doi: 10.1016/j.compeleceng.2022.107716.

[5] A. Aljuhani, "Machine learning approaches for combating distributed denial of service attacks in modern networking environments," *IEEE Access*, vol. 9, pp. 42236–42264, 2021. doi: 10.1109/ACCESS.2021.3062909.

[6] 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, pp. 494, Feb. 2022. doi: 10.3390/electronics11030494.

[7] F. Hussain, S. G. Abbas, M. Husnain, U. U. Fayyaz, F. Shahzad and G. A. Shah, "IoT DoS and DDoS attack detection using ResNet," in *2020 IEEE 23rd Int. Multitop. Conf. (INMIC)*, Bahawalpur, Pakistan, IEEE, Nov. 2020, pp. 1–6. doi: 10.1109/INMIC50486.2020.9318216.

[8] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated deep learning for cyber security in the Internet of Things: Concepts, applications, and experimental analysis," *IEEE Access*, vol. 9, pp. 138509–138542, 2021. doi: 10.1109/ACCESS.2021.3118642.

[9] J. Liu, B. Kantarci, and C. Adams, "Machine learning-driven intrusion detection for Contiki-NG-based IoT networks exposed to NSL-KDD dataset," in *Proc. 2nd ACM Workshop Wirel. Secur. Mach. Learn.*, Linz, Austria, ACM, Jul. 2020, pp. 25–30. doi: 10.1145/3395352.3402621.

[10] N. F. Syed, Z. Baig, A. Ibrahim, and C. Valli, "Denial of service attack detection through machine learning for the IoT," *J. Inf. Telecommun.*, vol. 4, no. 4, pp. 482–503, Oct. 2020. doi: 10.1080/24751839.2020.1767484.

[11] P. Singh Samom and A. Taggu, "Distributed Denial of Service (DDoS) attacks detection: A machine learning approach," in S. M. Thampi, J. Lloret Mauri, X. Fernando, R. Boppana, S. Geetha, A. Sikora (Eds.), *Applied Soft Computing and Communication Networks*, Singapore: Springer Singapore, 2021, vol. 187, pp. 75–87. doi: 10.1007/978-981-33-6173-7_6.

[12] S. Sambangi and L. Gondi, "A machine learning approach for DDoS (Distributed Denial of Service) attack detection using multiple linear regression," in *14th Int. Conf. Interdisciplinar. Eng.—INTER-ENG 2020*, Târgu Mureş, Romania, MDPI, Dec. 2020, pp. 51. doi: 10.3390/proceedings2020063051.

[13] M. Aslam *et al.*, "Adaptive machine learning based distributed denial-of-services attacks detection and mitigation system for SDN-Enabled IoT," *Sensors*, vol. 22, no. 7, pp. 2697, Mar. 2022. doi: 10.3390/s22072697.

[14] "Five most famous DDoS attacks and then some|A10 networks," 2024. Accessed: Oct. 19, 2023. [Online]. Available: https://www.a10networks.com/blog/5-most-famous-ddos-attacks/

[15] M. Aslam, D. Ye, M. Hanif, and M. Asad, "Machine learning based SDN-enabled distributed denial-of-services attacks detection and mitigation system for Internet of Things," in X. Chen, H. Yan, Q. Yan, and X. Zhang (Eds.), *Machine Learning for Cyber Security*, Cham: Springer International Publishing, 2020, vol. 12486, pp. 180–194. doi: 10.1007/978-3-030-62223-7_16.

[16] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in *2018 IEEE Secur. Privacy Workshops (SPW)*, San Francisco, CA, USA, IEEE, May 2018, pp. 29–35. doi: 10.1109/SPW.2018.00013.

[17] S. Dwivedi, M. Vardhan, and S. Tripathi, "Distributed denial-of-service prediction on IoT framework by learning techniques," *Open Comput. Sci.*, vol. 10, no. 1, pp. 220–230, Aug. 2020. doi: 10.1515/comp-2020-0009.

[18] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Gener. Comput. Syst.*, vol. 82, no. 6, pp. 761–768, May 2018. doi: 10.1016/j.future.2017.08.043.

[19] M. Esmaeili, S. H. Goki, B. H. K. Masjidi, M. Sameh, H. Gharagozlou and A. S. Mohammed, "ML-DDoSnet: IoT intrusion detection based on denial-of-service attacks using machine learning methods and NSL-KDD," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–16, Aug. 2022. doi: 10.1155/2022/8481452.

[20] M. Asad, M. Asim, T. Javed, M. O. Beg, H. Mujtaba and S. Abbas, "DeepDetect: Detection of distributed denial of service attacks using deep learning," *Comput. J.*, vol. 63, no. 7, pp. 983–994, Jul. 2020. doi: 10.1093/comjnl/bxz064.

[21] Y. Al-Hadhrami and F. K. Hussain, "A machine learning architecture towards detecting denial of service attack in IoT," in L. Barolli, F. Hussain, and M. Ikeda (Eds.), *Complex, Intelligent, and Software Intensive Systems. CISIS 2019. Advances in Intelligent Systems and Computing*, Cham: Springer, vol. 993, 21 Jun. 2019. doi: 10.1007/978-3-030-22354-0_37.

[22] Z. He, T. Zhang, and R. B. Lee, "Machine learning based DDoS attack detection from source side in cloud," in *2017 IEEE 4th Int. Conf. Cyber Secur. Cloud Comput. (CSCloud)*, New York, NY, USA, IEEE, Jun. 2017, pp. 114–120. doi: 10.1109/CSCloud.2017.58.

[23] R. Amrish, K. Bavapriyan, V. Gopinaath, A. Jawahar, and C. Vinoth Kumar, "DDoS detection using machine learning techniques," *J. IoT Soc. Mob. Anal. Cloud*, vol. 4, no. 1, pp. 24–32, May 2022. doi: 10.36548/jismac.2022.1.003.

[24] S. A. M. Al-Juboori, F. Hazzaa, Z. S. Jabbar, S. Salih, and H. M. Gheni, "Man-in-the-middle and denial of service attacks detection using machine learning algorithms," *Bull. Elect. Eng. Inf.*, vol. 12, no. 1, pp. 418–426, Feb. 2023. doi: 10.11591/eei.v12i1.4555.

[25] Ö. Tonkal, H. Polat, E. Başaran, Z. Cömert, and R. Kocaoğlu, "Machine learning approach equipped with neighbourhood component analysis for DDoS attack detection in software-defined networking," *Electronics*, vol. 10, no. 11, pp. 1227, May 2021. doi: 10.3390/electronics10111227.

[26] M. M. Oo, S. Kamolphiwong, and T. Kamolphiwong, "The design of SDN based detection for distributed denial of service (DDoS) attack," in *2017 21st Int. Comput. Sci. Eng. Conf. (ICSEC)*, Bangkok, Thailand, IEEE, Nov. 2017, pp. 1–5. doi: 10.1109/ICSEC.2017.8443939.

[27] G. De La Torre Parra, P. Rad, K. -K. R. Choo, and N. Beebe, "Detecting Internet of Things attacks using distributed deep learning," *J. Netw. Comput. Appl.*, vol. 163, no. 4, pp. 102662, Aug. 2020. doi: 10.1016/j.jnca.2020.102662.

[28] 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, pp. 602, Feb. 2022. doi: 10.3390/electronics11040602.

[29] B. Alotaibi and M. Alotaibi, "A stacked deep learning approach for IoT cyberattack detection," *J. Sens.*, vol. 2020, no. 7, pp. 1–10, Sep. 2020. doi: 10.1155/2020/8828591.

[30] A. A. Alsadhan, A. Hussain, and M. M. Alani, "Detecting NDP distributed denial of service attacks using machine learning algorithm based on flow-based representation," in *2018 11th Int. Conf. Develop. eSyst. Eng. (DeSE)*, Cambridge, UK, IEEE, Sep. 2018, pp. 134–140. doi: 10.1109/DeSE.2018.00028.

[31] F. S. D. Lima Filho, F. A. F. Silveira, A. de Medeiros Brito Junior, G. Vargas-Solar, L. F. Silveira, "Smart detection: An online approach for DoS/DDoS attack detection using machine learning," *Secur. Commun. Netw.*, vol. 2019, no. december, pp. 1–15, Oct. 2019. doi: 10.1155/2019/1574749.

[32] 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, pp. 3367, Apr. 2022. doi: 10.3390/s22093367.

[33] B. B. Gupta, P. Chaudhary, X. Chang, and N. Nedjah, "Smart defense against distributed denial of service attack in IoT networks using supervised learning classifiers," *Comput. Electr. Eng.*, vol. 98, no. 7, pp. 107726, Mar. 2022. doi: 10.1016/j.compeleceng.2022.107726.

[34] Ismail *et al.*, "A machine learning-based classification and prediction technique for DDoS attacks," *IEEE Access*, vol. 10, no. 12, pp. 21443–21454, 2022. doi: 10.1109/ACCESS.2022.3152577.

[35] M. Al-Naeem, M. A. Rahman, A. A. Ibrahim, and M. M. H. Rahman, "AI-based techniques for DDoS attack detection in WSN: A systematic literature review," *J. Comput. Sci.*, vol. 16, no. 6, pp. 848–855, Jun. 2020. doi: 10.3844/jcssp.2020.848.855.

[36] A. Jaszcz and D. Połap, "AIMM: Artificial intelligence merged methods for flood DDoS attacks detection," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8090–8101, Nov. 2022. doi: 10.1016/j.jksuci.2022.07.021.

[37] L. Fang, B. Zhao, Y. Li, Z. Liu, C. Ge and W. Meng, "Countermeasure based on smart contracts and AI against DoS/DDoS attack in 5G circumstances," *IEEE Netw.*, vol. 34, no. 6, pp. 54–61, Nov. 2020. doi: 10.1109/MNET.021.1900614.

[38] H. H. R. Sherazi, R. Iqbal, F. Ahmad, Z. A. Khan, and M. H. Chaudary, "DDoS attack detection: A key enabler for sustainable communication in internet of vehicles," *Sust. Comput.: Inf. Syst.*, vol. 23, no. 3, pp. 13–20, Sep. 2019. doi: 10.1016/j.suscom.2019.05.002.

[39] M. Mittal, K. Kumar, and S. Behal, "Deep learning approaches for detecting DDoS attacks: A systematic review," *Soft Comput.*, vol. 27, no. 18, pp. 13039–13075, Sep. 2023. doi: 10.1007/s00500-021-06608-1.

[40] A. Singh and B. B. Gupta, "Distributed Denial-of-Service (DDoS) attacks and defense mechanisms in various web-enabled computing platforms: Issues, challenges, and future research directions," *Int. J. Semant. Web Inf. Syst.*, vol. 18, no. 1, pp. 1–43, Apr. 2022. doi: 10.4018/IJSWIS.297143.

[41] C. S. Kalutharage, X. Liu, C. Chrysoulas, N. Pitropakis, and P. Papadopoulos, "Explainable AI-based DDOS attack identification method for IoT networks," *Computers*, vol. 12, no. 2, pp. 32, Feb. 2023. doi: 10.3390/computers12020032.

[42] M. J. Awan *et al.*, "Real-time DDoS attack detection system using big data approach," *Sustainability*, vol. 13, no. 19, pp. 10743, Sep. 2021. doi: 10.3390/su131910743.

[43] M. Elhoseny *et al.*, "Security and privacy issues in medical Internet of Things: Overview, countermeasures, challenges and future directions," *Sustainability*, vol. 13, no. 21, pp. 11645, Oct. 2021. doi: 10.3390/su132111645.

[44] A. Saied, R. E. Overill, and T. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks," *Neurocomputing*, vol. 172, no. 7, pp. 385–393, Jan. 2016. doi: 10.1016/j.neucom.2015.04.101.

[45] 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, pp. 5941, Jun. 2023. doi: 10.3390/s23135941.

[46] Md. A. M. Hasan, M. Nasser, S. Ahmad, and K. I. Molla, "Feature selection for intrusion detection using random forest," *J. Inf. Secur.*, vol. 7, no. 3, pp. 129–140, 2016. doi: 10.4236/jis.2016.73009.

[47] U. Grömping, "Variable importance assessment in regression: Linear regression versus random forest," *Am. Stat.*, vol. 63, no. 4, pp. 308–319, Nov. 2009. doi: 10.1198/tast.2009.08199.

[48] A. Alabdulatif and N. N. Thilakarathne, "Bio-inspired Internet of Things: Current status, benefits, challenges, and future directions," *Biomimetics*, vol. 8, no. 4, pp. 373, Aug. 2023. doi: 10.3390/biomimetics8040373.

[49] M. Aamir and S. M. A. Zaidi, "DDoS attack detection with feature engineering and machine learning: The framework and performance evaluation," *Int. J. Inf. Secur.*, vol. 18, no. 6, pp. 761–785, Dec. 2019. doi: 10.1007/s10207-019-00434-1.

[50] D. K. Sharma *et al.*, "Anomaly detection framework to prevent DDoS attack in fog empowered IoT networks," *Ad Hoc Netw.*, vol. 121, no. 5, pp. 102603, Oct. 2021. doi: 10.1016/j.adhoc.2021.102603.

[51] G. Cosoli, G. Iadarola, A. Poli, and S. Spinsante, "Learning classifiers for analysis of Blood Volume Pulse signals in IoT-enabled systems," in *2021 IEEE Int. Workshop Metrol. Industry 4.0 & IoT (MetroInd4.0&IoT)*, Rome, Italy, IEEE, Jun. 2021, pp. 307–312. doi: 10.1109/MetroInd4.0IoT51437.2021.9488497.

[52] 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, no. 1, pp. 106432, Aug. 2023. doi: 10.1016/j.engappai.2023.106432.

[53] E. Altulaihan, M. A. Almaiah, and A. Aljughaiman, "Anomaly detection IDS for detecting DoS attacks in IoT networks based on machine learning algorithms," *Sensors*, vol. 24, no. 2, pp. 713, Jan. 2024. doi: 10.3390/s24020713.

[54] P. Kumari and A. K. Jain, "A comprehensive study of DDoS attacks over IoT network and their countermeasures," *Comput. Secur.*, vol. 127, no. 2, pp. 103096, Apr. 2023. doi: 10.1016/j.cose.2023.103096.