Array 19 (2023) 100316

Contents lists available at ScienceDirect

Array

journal homepage: www.sciencedirect.com/journal/array

Capturing low-rate DDoS attack based on MQTT protocol in software Defined-IoT environment

Mustafa Al-Fayoumi, Qasem Abu Al-Haija *

Department of Cybersecurity, Princess Sumaya University for Technology (PSUT), Amman, 11941, Jordan

ARTICLE INFO

Keywords:
Cybersecurity
Denial-of-service attack (DoS)
Software defined network (SDN)
Internet of things (IoT)
MQTT protocol

ABSTRACT

The MQTT (Message Queue Telemetry Transport) protocol has recently been standardized to provide a lightweight open messaging service over low-bandwidth and resource-constrained communication environments. Hence, it is the primary messaging protocol used by Internet of Things (IoT) devices to disseminate telemetry data in a machine-to-machine approach. Despite its advantages in providing reliable, scalable, and timely delivery, the MQTT protocol is widely vulnerable to flooding and denial of service attacks, specifically, the low-rate distributed denial of services (LR-DDoS). Unlike conventional DDoS, the LR-DDoS attack tends to appear as normal traffic at a very slow rate, which makes it difficult to differentiate from legitimate packets, allowing the packets to move undetected by traditional detection policies. This paper presents an intelligent lightweight detection scheme that can capture LR-DDoS attacks based on MQTT protocol in a software-defined IoT environment. The proposed scheme examines the performance of four machine learning models on a modern dataset (LRDDoS-MQTT-2022) with a minimum feature set (i.e., two features only) and a balanced dataset, namely: decision tree classifier (DTC), multilayer perceptron (MLP), artificial neural networks (ANN), and naïve Bayes classifier (NBC). Our exploratory assessment demonstrates the arrogance of the DTC detection scheme achieving an accuracy of 99.5% with peak detection speed. Eventually, our best outcomes outdo existing models with higher prediction rates.

1. Introduction

IoT technology is developing quickly, making it possible to use in many fields, such as healthcare, agriculture, manufacturing, and city applications. IoT devices, on the other hand, have limited computing power, storage space, and user interfaces. This makes them vulnerable to security threats. A report from Statistica says that the number of Internet of Things (IoT) devices will triple from 9.7 billion in 2020 to over 29 billion in 2030 [1]. This shows that IoT networks need strong security measures to keep them safe. Fig. 1 shows how the number of IoT devices is expected to grow from 2019 to 2030.

IoT technology can help solve problems without human intervention and can be used to develop smart systems that monitor real-time IoT applications. The IoT architecture comprises various layers, and ensuring interoperability poses a significant challenge. Interoperability can be viewed from different perspectives, such as device, network, syntactic, semantic, and platform interoperability. However, IoT devices can be vulnerable to security attacks, particularly DDoS attacks that disrupt communication between IoT servers and users' devices [2,3]. The rise of the Internet of Things has opened the door for DDoS attackers, who can exploit many insecure devices to create botnets. With new techniques, attackers can launch attacks with minimal bandwidth by taking advantage of vulnerabilities in network devices to amplify their impact on the target system [4,5]. Fig. 2 illustrates a DDoS attack scenario occurring within IoT networks.

The development of interoperability solutions has enabled the wide deployment of IoT. One of these solutions is Software-Defined Network (SDN), which is referred to as SD-IoT [6,7]. The SDN control layer serves as both a traffic management hub and an Intrusion Detection System (IDS) component to patch security vulnerabilities in IoT networks [8,9]. Furthermore, SDN ensures network interoperability by flexibly managing and configuring all heterogeneous equipment in a network. SDN separates the control and data planes. The control plane consists of an SDN controller with network orchestration capabilities, and the data plane comprises network devices responsible for packet forwarding [10]. This approach enhances network agility and scalability while enabling centralized network management. The SDN paradigm has been successful in both wired and wireless networks. In essence, SDN acts as a

* Corresponding author.
E-mail addresses: m.alfayoumi@psut.edu.jo (M. Al-Fayoumi), q.abualhaija@psut.edu.jo (Q. Abu Al-Haija).

https://doi.org/10.1016/j.array.2023.100316
Received 14 March 2023; Received in revised form 27 June 2023; Accepted 2 August 2023
Available online 3 August 2023

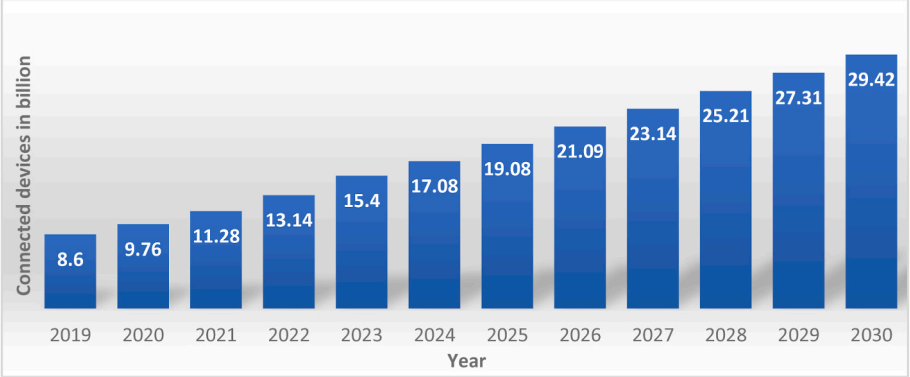

Fig. 1. The predicted increase in IoT devices (2019–2030).

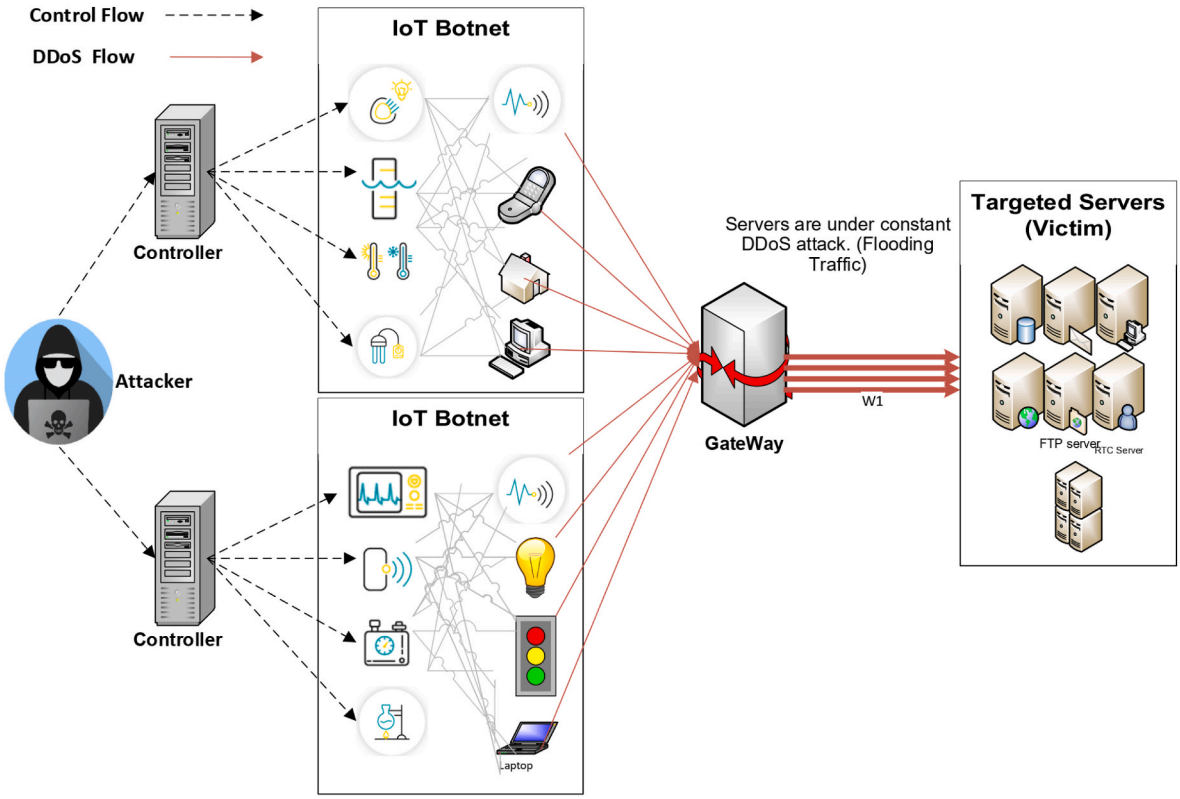

Fig. 2. DDoS attack scenario within IoT networks.

traffic facilitator for the network, managing resources and maintaining IoT network security [11,12].

The SDN structure separates the IoT network's control and data delivery functions across distinct abstraction levels. Nevertheless, the centralized control of SDN is still susceptible to DDoS attacks. DDoS attacks are intended to overload the SDN's centralized management system by continuously sending bogus packets, which can exhaust the computing resources of the controller [13].

Another solution for ensuring interoperability across devices in the IoT perception layer is to utilize the MQTT protocol to handle the heterogeneous data produced by various smart objects. MQTT is a messaging protocol that allows devices to connect through a central server known as a broker, which is responsible for relaying data from machine to machine. MQTT is based on a publish/subscribe paradigm of operation. However, the widespread adoption of the MQTT protocol has also attracted the attention of cyber attackers who can launch different types of attacks, including LR-DDoS attacks [14]. The LR-DDoS attack is a stealthy and sophisticated attack that targets the MQTT protocol, taking advantage of the protocol's inherent characteristics. Unlike traditional DDoS attacks, which involve sending a high volume of traffic to a target, LR-DDoS attacks occur at a much lower rate, typically below the normal traffic threshold.

LR-DDoS attacks are a major threat to IoT systems, particularly those relying on the MQTT protocol. These attacks can congest network traffic, overload devices, and exhaust system resources, leading to service interruptions and data breaches. LR-DDoS attacks are difficult to detect and mitigate, making them attractive to attackers. Traditional detection mechanisms have difficulty differentiating between normal and malicious traffic, which allows attackers to move undetected. In addition to being susceptible to LR-DDoS attacks, IoT devices are also often resource-constrained, which makes them more vulnerable. Due to their limited memory, processing power, and battery life, IoT devices are unable to support complex security mechanisms. They are susceptible to attacks that exploit vulnerabilities in their communication protocols.

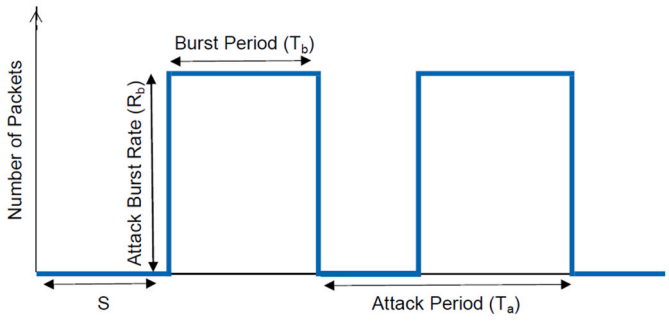

Fig. 3. A generic model of LR-DDoS attack.

With the increasing number of interconnected IoT devices, LR-DDoS attacks pose a significant challenge to the security and resilience of IoT systems. Fig. 3 illustrates an LR-DDoS attack, including the attack period T_a, burst period T_b, attack burst rate R_b, and starting time S [15, 16]. The attack can be identified by analyzing the source, destination IP, port, and protocol. In low-rate TCP DoS attacks, burst packets are sent periodically with a burst rate exceeding the bottleneck capacity, exploiting TCP's Minimum Retransmission Timeout property. LR-DDoS attacks often use multiple sources with their flow. Zhang et al. [17] categorized these attacks into four groups: Attack Burst Rate Intensification, Attack Frequency Intensification, Mixed Intensification, and Attack Burst Width Intensification.

Researchers have proposed various detection and mitigation mechanisms to address the challenge of LR-DDoS attacks on IoT systems. However, only some of these mechanisms could be more efficient or impractical due to the resource constraints of IoT devices [18]. Therefore, there is a need for an effective and efficient mechanism that can detect LR-DDoS attacks based on the MQTT protocol in software-defined IoT environments.

The proposed intelligent lightweight detection scheme can provide a practical defense mechanism against LR-DDoS attacks based on the MQTT protocol. The scheme leverages machine learning techniques to detect LR-DDoS attacks based on minimal features, making it practical for use in resource-constrained IoT devices. By enhancing the security of IoT systems, this research has significant implications for the IoT ecosystem's security and resilience. The main contribution of this paper is an intelligent lightweight detection scheme that can detect LR-DDoS attacks based on MQTT protocol in software-defined IoT environments. This research proposes an efficient and effective approach to tackle the challenges of LR-DDoS attacks that can enhance the security and resilience of IoT systems. The proposed scheme's simplicity and performance make it suitable for resource-constrained IoT devices and can be integrated into existing security mechanisms to enhance their detection capabilities.

1.1. Summary of contribution

This paper employs an AI-based approach to detect LR-DDoS attacks based on MQTT protocol in software-defined IoT environments. This paper's significant contributions are as follows:

- We propose an intelligent lightweight detection scheme that can provide a practical defense mechanism against LR-DDoS attacks based on the MQTT protocol in software-defined IoT environments.
- We implement and assess several machine learning techniques to detect LR-DDoS attacks based on minimal features, making it practical for use in resource-constrained IoT devices. By enhancing the security of IoT systems, this research has significant implications for the IoT ecosystem's security and resilience.

1.2. Paper organization

The remaining parts of this article are structured in the following manner: In Section 2, the literary works are presented and summarized. Section 3 provides the specifics of the proposed scheme, the methodology for evaluating its efficacy in detecting and preventing Low-Rate DDoS attacks, and details on experiments conducted. Section 4 analyzes the obtained results, compares the implemented schemes, and benchmarks the optical scheme with other state-of-the-art models. Finally, Section 5 concludes the article and suggests potential future directions.

2. Literature review

LR-DDoS attacks have been on the rise in IoT networks, leading to the development of various techniques to detect and mitigate them. The use of Machine Learning (ML) algorithms has gained significant attention in detecting and mitigating DDoS attacks due to the increasing number of IoT devices and their susceptibility to cyber-attacks. The literature review has been divided into three categories: (1) LR-DDoS Attacks based on IoT networks, (2) DDoS attacks based on MQTT Protocol in IoT Environment, and (3) DDoS attacks in a Software-Defined IoT Environment.

2.1. Low-rate DDoS attack based on IoT networks

Several studies have employed machine learning techniques to detect and prevent LR-DDoS attacks in IoT networks. Mugunthan et al. [19] proposed a method named HMM-RF, which combines the Hidden Markov Model and random forest classifier to identify LR-DDoS attacks in cloud data centers. The Hidden Markov Model analyzes traffic flow features, while the random forest classifier differentiates between normal and attacked flow based on the predicted attack level. Renyi entropy is used to calculate attack probabilities.

A mechanism has been proposed by Singh et al. [20] for preventing and detecting IP spoofing, which is a common form of cyber-attack. The proposed mechanism utilizes a neural network as a classifier to prevent spoofed packets from entering the network. The neural network is optimized using an artificial neural network to improve performance. Several parameters have been considered to evaluate the mechanism's effectiveness, including throughput, packet delivery ratio (PDR), and energy consumption. These parameters measure the mechanism's performance and determine how well it can prevent and detect IP spoofing attacks. The proposed mechanism provides a reliable and effective defense against network IP spoofing by optimizing the neural network and defense ring key performance indicators. Zhijun et al. [21] presented a dynamic flow rule deletion-based mechanism to detect LR-DDoS attacks in SDN environments. The proposed approach utilized a multi-feature LR-DDoS attack detection method based on the Factorization Machine algorithm, using four features to detect attacks. The proposed approach was evaluated using three datasets, including NSL-KDD, DARPA98, and CAIDA, in a simulated environment, and it was found to achieve a high detection rate of 95.8%. Verma et al. [22] proposed an adaptive hybrid approach that selects the threshold value from different available thresholding techniques based on the incoming network traffic conditions to distinguish between DDoS and benign requests. The approach can distinguish between DDoS attacks, including TCP, UDP, and ICMP-based attacks [23]. Pérez et al. [24] proposed a flexible and modular security architecture that allows easy replacement of modules without affecting the others, using pre-trained machine learning models in the IDS module to detect flows. The architecture is evaluated using two different topologies and successfully mitigated all previously identified attacks by the IDS.

Cheng et al. [25] proposed a learning-based mechanism that detects LR-DDoS attacks on SDN control and switches nodes in IoT networks. The mechanism uses stateless and stateful features from OpenFlow

packages and learning algorithms to develop classifiers that distinguish normal traffic from LR-DDoS attacks. The proposed mechanism demonstrated high accuracy and outperformed traditional solutions.

Nugraha et al. [26] provide a deep-learning architecture for slow DDoS detection. The detection module analyzes SDN switch traffic flow statistics from the SDN controller's REST API to detect slow DDoS attacks. In the detection module, they used a hybrid CNN-LSTM model. This approach was evaluated on custom datasets. The best hyperparameters optimized the hybrid CNN-LSTM model. This framework used 12 features.

SlowITe is a new low-rate denial-of-service attack proposed by Vaccari et al. [27]. This attack targets MQTT and makes use of low-rate techniques. This study focuses on the vulnerabilities in IoT environments using the MQTT protocol. The authors also explain how the client can modify the Keep-Alive parameter of the server, which gives the attacking node control over the connection closure timeouts on the server. The same authors, Vaccari et al. [28], also introduced SlowTT as an SL-DoS attack. To prevent legitimate clients from setting up MQTT sessions and sending or receiving messages owing to a lack of open connection sockets, the attack aims to consume and block as many of the broker's open TCP connections as possible. Once the attacker has started a conversation with the broker, they use the MQTT protocol's network configuration settings, particularly the KeepAlive parameter, to maintain connections for a long time. By mimicking actual behavior using PINGREQ and PINGRESP packets, SlowTT may also sustain connections for a long time, even with lower KeepAlive levels.

A method for detecting LR-DoS attacks using hybrid deep neural networks that are comprised of a 1-D convolutional neural network and the recurrent gated unit is proposed by Xu et al. [29]. This method uses hybrid deep neural networks. The approach requires temporal statistics of network traffic to detect LR-DoS attacks. Real, legitimate traffic from a website was recorded in a data center, and several low-resource denial-of-service attacks were carried out in a laboratory environment on a copy of the website to record attack traffic. This was done so that the effectiveness of the proposed method could be evaluated. Nada et al. [30]used a novel dataset approach in the emulation procedure. Out of the original 21 features, the experiment also used feature selection using the logistic regression coefficient, producing eight features. For LR-DDoS prediction, the Random Forest (RF) approach was used. For packets sent at 200 packets per second, the accuracy was 98.7%, while the forecast loss value was 99.1%. (PPS).

AlMasri et al. [31] suggest a new method for identifying and preventing network attacks using ML techniques. The strategy involves using the NSL-KDD dataset to create a machine-learning system to recognize DoS and port scanning attacks. The method also integrates the results of the ML algorithm with a proven SDN architecture, which should enhance performance by leveraging several previously tried and true procedures. The categorical data were converted into non-categorical data using one-hot encoding as part of the data pre-processing, and the features were chosen using the ANOVA test. Various classifiers were examined using the original dataset and the chosen features from the ANOVA test. The Naive Bayes model delivered the most accurate outcomes. They are using the packet drop method. REPD is a Renyi entropy DDoS attack detection method that Ahalawat et al. [32] suggested. Using the packet-drop technique for prevention, evaluating the probability distribution of flow fluctuations, and achieving better results than the Shannon entropy is possible.

2.2. DDoS attack based on MQTT protocol in -IoT environment

In IoT networks, the MQTT protocol is used for machine-to-machine communication, making it vulnerable to DDoS attacks. In the study by Haripriya and Kulothungan [33], they found that spoofing attacks can be used against MQTT. This means an attacker can send a malicious packet that looks like a legitimate message packet. This vulnerability happens because the MQTT broker can't tell the difference between normal message packets and spoofed messages. An attacker could steal identities by getting access to information about MQTT clients. The authors devised a fuzzy logic-based Secure MQTT to solve this problem. This method uses a fuzzy rule interpolation mechanism to find malicious activity in MQTT communication between IoT devices. With fuzzy logic, the authors made an intrusion detection system (IDS) that only gave false positives 0.6% of the time. This IDS was made to find DDoS attacks in IoT setups. Still, it is unclear how the input parameters, Connection Message Ratio (CMR) and Connection Acknowledgement Message Ratio (CAMR) were set or how the crisp value for anomaly output was found. Also, it needed to be clarified what high and low values were. Kumar et al. [34] proposed a distributed intrusion detection method based on fog computing. To detect attacks early on, this approach employs AI in fog nodes. IPFS-based distributed file storage is recommended for off-chain IoT data balancing. Mutual information-based feature selection improves model detection accuracy (AC), processing time, and false alarm rate (FAR). The proposed distributed architecture was tested on the BoT-IoT dataset for accuracy, precision, F1-score, and detection rate (DR) and compared to state-of-the-art approaches [35]. The distributed framework had the highest DR and outperformed RF with eleven features. The blockchain-IoT network attacks it detects are incredibly effective.

In [36], Ghannadrad aimed to detect and classify DoS attacks for MQTT sensor networks by simulating a smart home scenario and collecting reliable data. Ghannadrad proposed a realistic MQTT-based dataset that includes legitimate and malicious flow-level traffic. Online and offline machine learning (ML) solutions were presented for detecting and classifying DoS attacks. Through the evaluation of the study, the author successfully differentiated between malicious and legitimate traffic as a binary classification, and the classifiers were able to detect and classify different categories of DoS attacks in a multi-value classification. Ghannadrad simulated a smart house scenario and collected credible data to detect and classify MQTT sensor network DoS attacks [36]. Ghannadrad presented a realistic MQTT-based dataset, including valid and malicious flow-level communication. Online and offline machine learning (ML) systems detected and classified DoS assaults. The study's evaluation showed that the author could distinguish malicious from valid traffic as a binary classification. The classifiers could detect and categorize multiple DoS attack categories as multi-value classifications. Aldhyani and AlKahtani [37] proposed an ML-based IDS for MQTT protocol in IoT networks using four algorithms, with the CNN-LSTM model having the highest precision of 98.94%. The proposed frameworks and solutions effectively detected attacks in blockchain IoT and MQTT networks.

2.3. DDoS attack in software Defined-IoT environment

SDN is becoming a popular way to manage and secure IoT networks due to its dynamic traffic management capabilities. Bhayo et al. [38] proposed a secure and fast Software-Defined Networking framework to prevent DDoS attacks in IoT systems (SDN). The framework improved detection accuracy and reduced false positive rates by assessing multiple metrics from a huge volume of SDN communication data. DDoS attacks in IoT systems can be managed securely and quickly using the proposed framework. SDN-based entropy-based DDoS detection and mitigation in IoT networks is proposed by Galeano-Brajones et al. [39]. OpenStack, an extension of Openflow, allows SDN switches to store information about previously processed flows, lowering controller load. The authors filter harmful traffic using entropy. The switch doesn't respond to the controller about states when its window size is too large.

An entropy-based technique for identifying and thwarting DDoS assaults in IoT networks utilizing SDN is proposed by Galeano-Brajones et al. [39]. The approach is based on the OpenState protocol, an addition to the OpenFlow protocol that enables SDN switches to save details about previously handled flows, lessening the burden on the controller. The writers use entropy to spot malicious traffic and create filtering rules

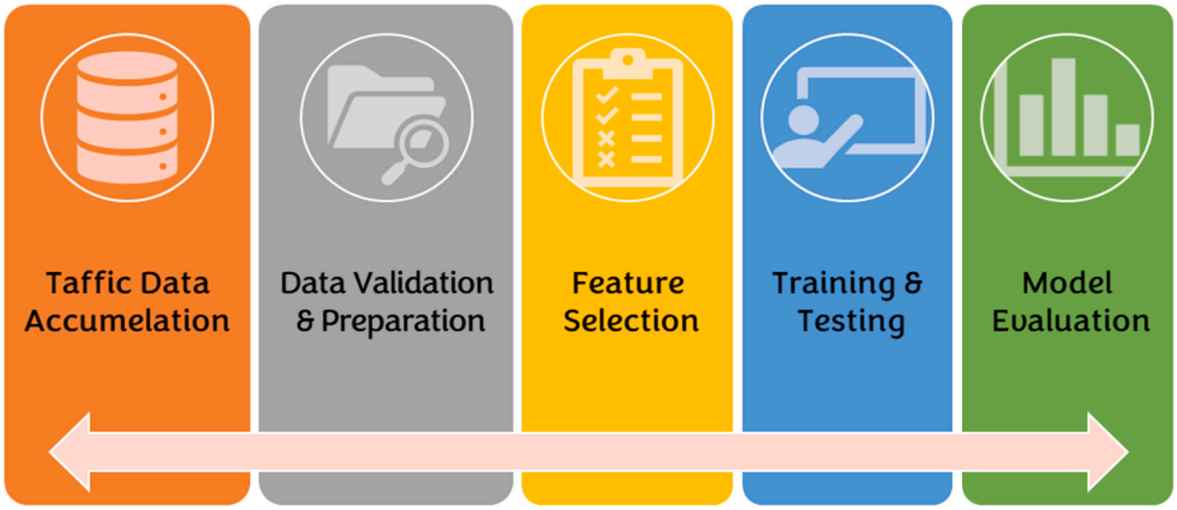

Fig. 4. Overall workflow diagram for developing and evaluating the proposed LR-DDoS detection system.

to lessen the threat. The proposed technique rapidly identifies and mitigates the DDoS attack by removing the flow entry from the switch using monitoring data acquired from flows and state tables at the data plan. The drawback of this strategy is that the switch needs to inform the controller of the states when the window size is too large. Anbarsu et al. [40] proposed a fuzzy logic-based IDS with deep neural networks (DNNs) in an SDN network to detect DDoS attacks. The combination of fuzzy logic and DNNs is motivated by their high classification accuracy and low false positive rate. The proposed IDS was evaluated using the "KDD CUP99" dataset. However, implementing deep learning algorithms requires expensive GPUs and specialized knowledge. Ivanova et al. [41] developed a feed-forward neural network model to recognize DoS and DDoS attacks that use various activities. When the model was put through its paces using the BotIoT dataset, it effectively detected DDoS attacks using 8 or 10 features, attaining an accuracy rate of 99.99% overall. The Adam optimization technique and a hyperbolic tangent activation function are used for each neuron in the neural network's single hidden layer.

The literature review highlights the knowledge gap in detecting LR-DDoS attacks based on MQTT protocol in SD-IoT environments using lightweight and efficient detection mechanisms. While previous studies proposed detection mechanisms, they require significant resources or could perform better on resource-constrained IoT devices. Additionally, few studies have explored the feasibility and effectiveness of ML-based detection mechanisms for LR-DDoS attacks based on the MQTT protocol. Therefore, the article proposes an intelligent lightweight detection scheme that uses machine learning models and minimal features to achieve high detection accuracy with peak detection speed. To the best of our knowledge, this study is the first ML-based study to use only two features and achieve the best accuracy with maximum detection speed compared to previous works.

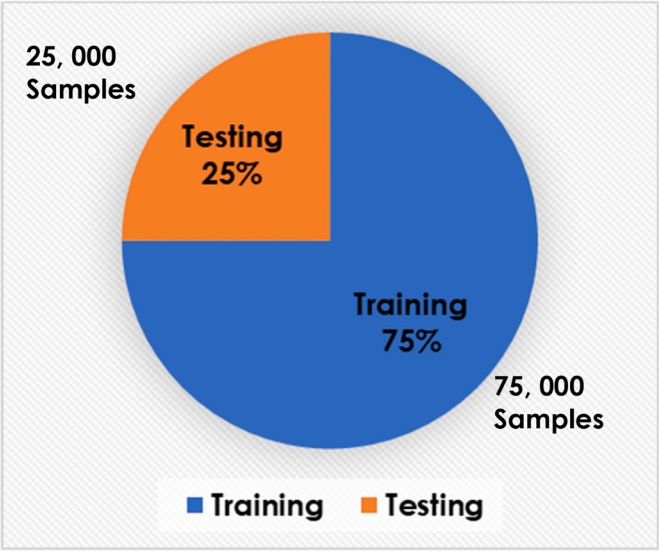

Fig. 6. Dataset distribution statistics.

3. LR DDoS recognition methodology

This section uses various learning and evaluation techniques to investigate the development and assessment approaches of the proposed low-rate DDoS attack detection model based on the MQTT protocol in the software-defined Internet of Things environment (SDN-IoT). Fig. 4 illustrates the overall workflow diagram for developing and evaluating the proposed LR-DDoS detection system.

Initially, the traffic data was accumulated from the Mendeley data

F1	DATAPATH_ID	F8	TTL	F15	PORT_NO
F2	VERSION	F9	PROTO	F16	RX_BYTES_AVE
F3	HEADER_LENGTH	F10	CSUM	F17	RX_ERROR_AVE
F4	TOP	F11	SRC_IP	F18	RX_DROPPED_AVE
F5	TOTAL_LENGTH	F12	DST_IP	F19	TX_BYTES_AVE
F6	FLAGS	F13	SRC_PORT	F20	TX_ERROR_AVE
F7	OFFSET	F14	DST_PORT	F21	TX_DROPPED_AVE

Fig. 5. Feature set in the original dataset (LR-DDoS-MQQT-2022 dataset [30]).

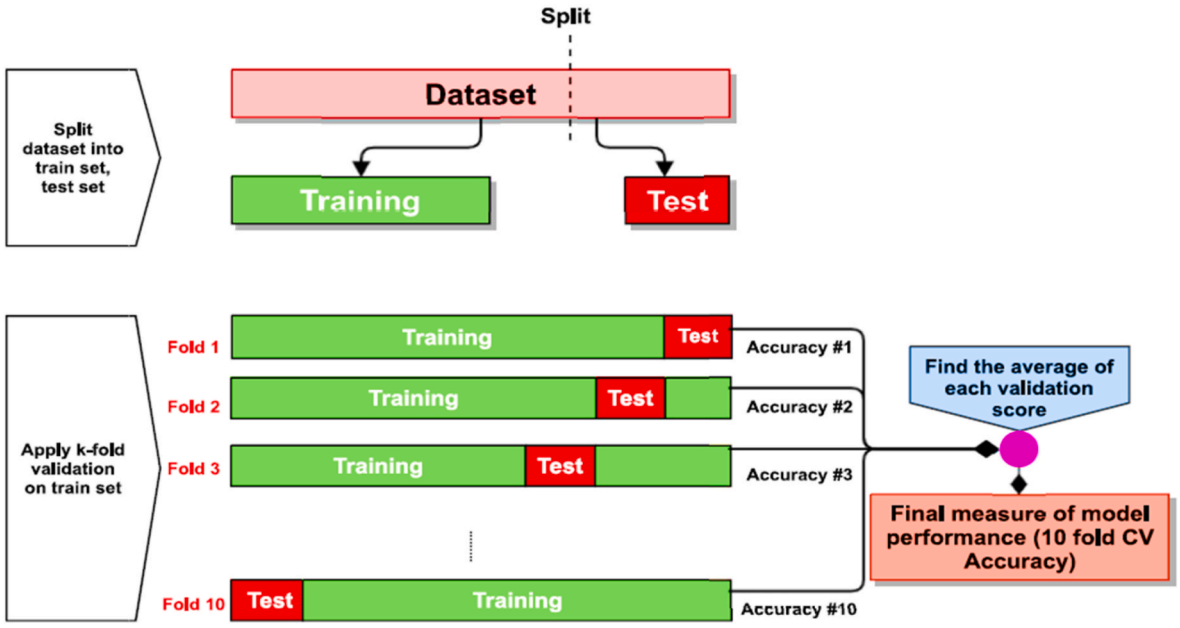

Fig. 7. Ten-fold cross-validation diagram [42].

repository, a global data warehouse for different applications and disciplines. LR-DDoS-MQQT-2022 dataset [30] has been collected for the purpose of model evaluation in capturing LR-DDoS attacks based on MQQT protocol in a software-defined IoT environment. LR-DDoS-MQQT-2022 dataset is an up-to-date and balanced dataset that has been extracted using the OpenFlow software tool for generating real-time DDoS attacks in SD-IoT. It comprises 200,00 traffic samples distributed equally into two target classes: 100,00 samples for the normal traffic and 100,00 samples for the LR-DDoS traffic. Besides, its feature set comprises 21 input features (shown below in Fig. 5) and one label feature used to classify the traffic as normal or LR-DDoS.

After that, the accumulated dataset underwent a validation and preparation phase to ensure the readiness of the traffic samples for the training procedure using the machine learning modules. This includes the samples checking against missing records/values, duplicated samples, errors in data entries, categorical data encoding or excluding samples randomization, and the dataset division into training and testing datasets. Fig. 6 shows the distribution of the LR-DDoS-MQQT-2022 dataset into training and testing datasets.

Once the dataset was ready for training/testing phases, we examined all features to extract those essential features in which the system performed with the highest performance metrics and the lowest prediction delay. The feature set has undergone a series of feature extraction experiments using principal component analysis (PCA) to pick up the minimum number of essential features that enable the prediction at the highest accuracy and maximum prediction speed (i.e., lightweight). As a result, only two main features (CSUM and SRC_PORT) have been finally used to develop the proposed lightweight detection system to capture the LR-DDoS attack based on MQTT protocol in software defined-IoT ecosystems.

The next stage is the learning stage which comprises the training and testing phases for the system using the prescribed training and testing datasets. In this module, we have examined the performance ability of four machine-learning models in discovering anomalous traffic, viz. decision tree classifier (DTC), multilayer perceptron (MLP), artificial neural networks (ANN), and naïve Bayes classifier (NBC). At each experiment, we employed the Ten-fold cross-validation in order to ensure effective validation for the learning process using the LR-DDoS-MQQT-2022 dataset. Fig. 7 shows the validation policy used in this research. The dataset was divided into ten parts; nine were taken to train the model, and one was used to test the model. The mean value E of the ten-fold test results is calculated to approximate the model accuracy for the current Ten-fold cross-validation model, where E_i represents the cross-validation error of the ith group.

4. LR DDoS recognition assessment

In order to select the best ML mode that provides the best

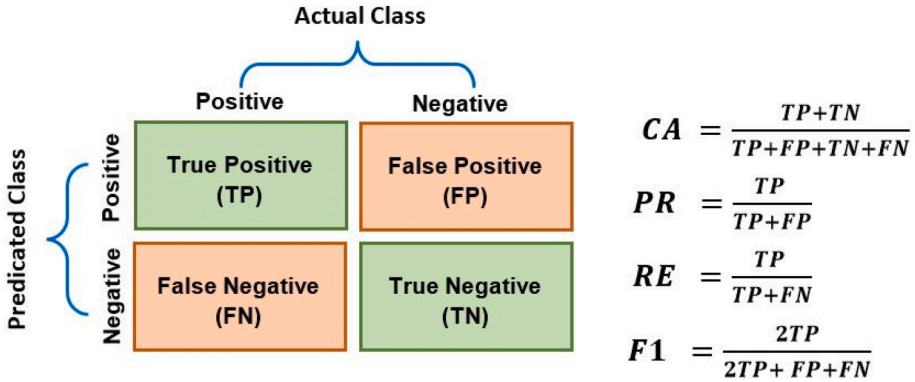

Fig. 8. Evaluation factors [43].

Table 1
Summary of review-related research.

Model	Method/Architecture	No. of Features	Advantages	Limitation
Mugunthan et al. [19]	RFC model	41	• Can detect and prevent low-rate DDoS attacks. • Uses multiple soft computing techniques to increase the accuracy and efficiency	• Could be a concern for real-world implementation. • Lacks extensive experimental results to prove the system's efficacy.
Singh et al. [20]	ABC-ANN	21	• A novel approach for preventing IP spoofing attacks.	• Have limitations in terms of scalability and generalizability to diverse types of networks and systems. • Not extremely fast
Zhijun et al. [21]	FM/CNN/RFC	64	• Several datasets were evaluated. • Good detection rate with low FP • Able to identify low-rate DDOS.	• Does not discuss the computational and storage requirements of the proposed approach. • Not extremely fast
Verma et al. [22]	MAD-RF	43	• Possibility to choose a threshold value based on network traffic conditions. • High accuracy in detecting DDoS attacks while maintaining low false positive rates	• Selecting a large number of features can lead to an increase in both detection and training time.
Pérez et al. [24]	MLP/SVM	44	• High-precision and recall LR-DDoS detection and mitigation. • Flexible and scalable architecture allows seamless integration with existing network infrastructures.	• Selecting a large number of features can lead to an increase in both detection and training time.
Cheng et al. [25]	SVM	19	• High accuracy in detecting LR-DDoS attacks	• Limited dataset (Small dataset) • The proposed approach may require a significant amount of computation, which could pose a challenge in resource-constrained IoT environments
Nugraha et al. [26]	CNN-LSTM	12	• The hyperparameters optimized the hybrid CNN-LSTM model. • High accuracy and low false-positive rates	• There is a small dataset available to evaluate the effectiveness of the deep learning model.
Vaccari et al. [27]	SlowITe DDoS attack. Exploit MQTT client-server vulnerability (connection-closure timeout)	N\A	• Behaves as a legitimate traffic	• Can be detected if investigated properly
Vaccari et al. [28]	SlowTT DDoS attack. Exploiting the broker's "Keep-Alive" parameter.	N\A	• Efficient way of launching a DDoS attack. • Can target a large number of IoT networks	• Can be detected if investigated properly
Xu et al. [29]	1-CNN-GRU	NA	• Low-rate detection with high results (accuracy, precision, recall, and F1-score)	• Not extremely fast
Nada et al. [30]		8	• Detection of LR-DDoS attacks with high results (accuracy, precision, recall, and F1-score) • Decreased false negative rate and reduced missed detection rate of LR-DDoS attacks.	• Not extremely fast (Training Time (s): 0.422 s)
AlMasri et al. [31]	NBC Model	13	• The system employs both unsupervised and supervised machine learning algorithms. • Using Anova for feature selection achieved high accuracy in detecting Probe attacks.	• Low accuracy in detecting DoS attacks. • Selecting a large number of features can lead to an increase in both detection and training time.
Ahalawat et al. [32]	Renyi Entropy with Packet Drop "REPD" DDoS attack detection technique	NA	• Used packet drop method for mitigation. • Can detect low-rate DDoS	• Not extremely fast.
Haripriya et al. [33]	Fuzzy logic-based approach	2	• Fast approach	• It is unclear how the input parameters Connection Message Ratio (CMR) and Connection Acknowledgement Message Ratio (CAMR) were determined.
Kumar et al. [34]	RF/Fog computing-based distributed intrusion detection framework	10	• Uses a few parameters and can be updated in real-time deployment.	• Using a common information-based feature selection method
Ghannadrad [36]	RF, KNN, & SVM with ANOVA as FS/Realistic MQTT dataset with benign and malicious flow-level communications.	10	• binary and multi-level classification was used for differentiating between malicious and legitimate traffic	• High computational Complexity: It requires a long time (>1 min) of data-gathering to perform better. • The proposed models are evaluated only on a specific scenario (smart home) and may not generalize well to other scenarios or applications.
Aldhyani et al. [37]	CNN-LTSM	29	• High intrusion detection accuracy. Accuracy is 98.9%, Precision is 99.2%, The recall is 99%, F1-score is 99.1%	• Complex model that requires resources
Bhayo et al. [38]	A framework was developed to improve the DDoS detection accuracy.	6	• High detection rate from 98% to 100% with low False-Positive. • Several IoT nodes and packet sizes were used.	• Complex and heavy model
Galeano-Brajones et al. [39]	entropy-based method for detecting and mitigating DDoS attacks based on OpenState protocol	4	• IoT-based stateful SDN DoS and DDoS detection and mitigation. • Quickly detects and prevents real-time DDoS attacks. • The detection rate is between 80% and 100% based on the window size.	• It may require significant computational resources to implement, which could limit its practicality for resource-constrained IoT devices. • It stops responding to the controller regarding the states when the window size is too large. • The experimental setup used in the study is small, limiting the results' generalizability to larger-scale IoT networks.
Anbarsu et al. [40]	Fuzzy logic-based IDS with Deep Neural Networks	6	• Good classification accuracy and low False Positive Rate	• Requires expensive GPUs and specialized knowledge.

(*continued on next page*)

Table 1 (*continued*)

Model	Method/Architecture	No. of Features	Advantages	Limitation
Ivanova et al. [41]	RNN, Adam optimization, hyperbolic tangent activation function	10 and 8	• Distinguish TCP, UDP, HTTP flood, keylogging, data exfiltration, OS fingerprinting, and service scan threats from typical network traffic. • Precision, Recall, F1, and CA is 99.9%	• The accuracy is 92% • Neural networks are not always able to accurately detect and identify complex patterns and can be prone to overfitting

Table 2
System evaluation using three machine learning techniques: DT, SVM, and NB, and in terms of classification accuracy, precision, recall, F1 score, MCR, FNR, FDR, and the prediction speed.

Model	CA	RE	PR	F1	MCR	FNR	FDR	PRS
DTC	99.5%	99.4%	99.4%	99.35%	0.50%	0.60%	0.60%	5600000 obs/sec
MLP	91.5%	91.5%	91.7%	91.70%	8.50%	8.50%	8.30%	2600000 obs/sec
ANN	83.8%	83.5%	85.5%	85.35%	16.20%	16.50%	14.50%	2100000 obs/sec
NBC	72.6%	72.5%	73.0%	72.95%	27.40%	27.50%	27.00%	2400000 obs/sec

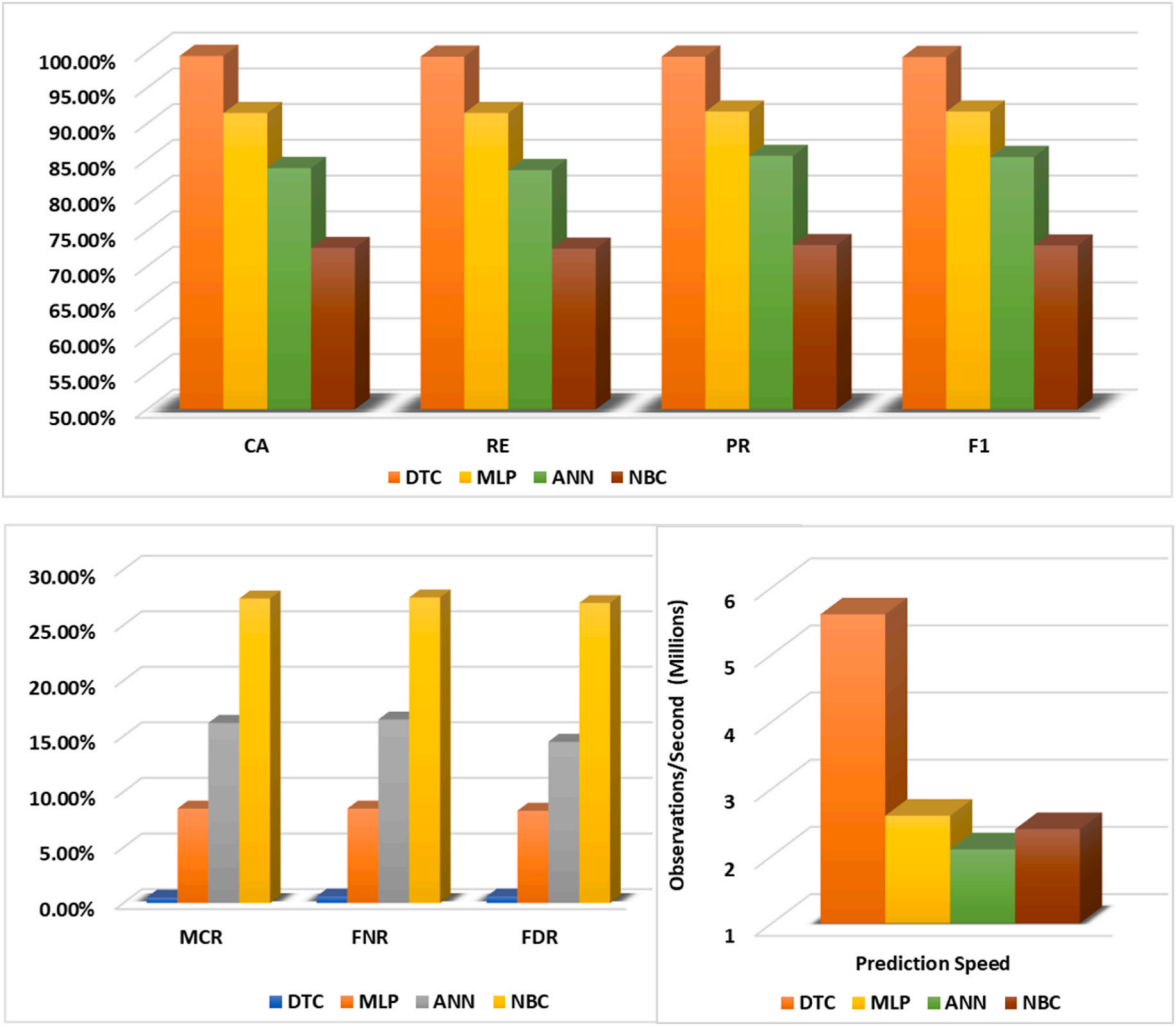

Fig. 9. System evaluation using three machine learning techniques: DT, SVM, and NB, and in terms of (a) Performance indicators (accuracy, precision, recall, F1 score), (b) False rates (MCR, FNR, FDR), and the prediction speed.

performance indicators and thus can be finally deployed to capture the anomalous LR-DDoS traffic packets, the following performance metrics have been used: classification accuracy rate (CA), true positive rate (TPR) A.K.A recall (RE), positive predictive value (PPV) A.K.A precision (PR), the harmonic score (F1), false negative rate (FNR) which can be calculated as (1-TPR)%, false discovery rate (FDR) which can be calculated as (1-PPV)%, misclassification rate (MCR) which can be calculated as (1-CA)%, and finally, the prediction speed in terms of the

Table 3
Comparing detection accuracy with existing models for LR-DDoS/DDoS attacks detection.

Model	Method	No. of Features	CA	F1
Mugunthan et al. [19]	RFC model	41	97.34%	95.45%
Singh et al. [20]	ABC-ANN	21	78.50%	57.75%
Zhijun et al. [21]	FM model	64	95.80%	94.80%
Zhijun et al. [21]	CNN model	64	90.90%	90.30%
Zhijun et al. [21]	RFC model	64	90.30%	88.80%
Verma et al. [22]	MAD-RF	43	60.00%	76.00%
Pérez et al. [24]	MLP Model	80	95.00%	94.98%
Pérez et al. [24]	SVM Model	80	93.10%	93.00%
Xu et al. [29]	1-CNN-GRU	NA	98.68%	NA
AlMasri et al. [31]	NBC Model	43	86.90%	NA
This work	DTC model	2	99.50%	99.35%

number of observations (samples) per second; this metric particularly generated by MATLAB. The following figure, Fig. 8, summarizes the stated metrics with their equations.

- Classification Accuracy (CA): This performance indicator can be defined as the proportion of correctly classified samples (either positive or negative) among the overall number of samples.
- Classification Recall (Re): This performance indicator is the proportion of actual positive samples correctly (only positive) identified by the system out of all the positive samples in the dataset.
- Classification Precision (PR): This performance indicator can be defined as the proportion of correctly predicted positive samples out of all instances predicted as positive by the system.
- Classification F1 Score (F1): This performance indicator can be defined as the balanced proportion of precision and recall measuring the system's effectiveness by considering the ability to correctly identify positive samples (precision) and the ability to avoid false negative samples (recall).
- Misclassification Rate (MCR): This performance indicator can be defined as the proportion of incorrectly classified samples (either positive or negative) among the overall number of samples.
- False Negative Rate (FNR): This performance indicator is the proportion of actual positive samples incorrectly (only positive) identified by the system out of all the positive samples in the dataset.
- False Discovery Rate (FDR): This performance indicator can be defined as the proportion of incorrectly predicted positive samples out of all instances predicted as positive by the system.
- Prediction Speed (PRS): This performance indicator can be defined as the number of predictions (observations) that can be performed by the model in 1 s (see Table 1).

Furthermore, Table 2 and Fig. 9 present the system evaluation using three machine learning techniques: DT, SVM, and NB, regarding classification accuracy, precision, recall, F1 score, MCR, FNR, FDR, and prediction speed (PRS). Based on the table and figure results, we can plainly observe the advantage of the DTC model over other models in terms of all performance statistics. It is a high-performant model since it can predict individual communication traffic accurately with 99.5%. It's also lightweight since it can predict individual communication traffic rapidly with only 179 ns. Besides, it exhibits the least false alarm rates (MCR, FNR, FDR)among all models. Such outcomes indicate the robustness of the model being deployed to work in the real-time applications of the SDN-IoT ecosystem.

Lastly, Table 3 compares the detection accuracy of our DTC-based detection model with existing models for LR-DDoS/DDoS attacks detection in IoT ecosystems. Diverse models were developed and proposed to detect the LR-DDoS/DDoS attack detection in IoT ecosystems. To benchmark our results and show the advantages of our lightweight model, we have contrasted our model with several state-of-the-art models, including (a) Zhijun et al. model [21]: developed the LR-DDoS detection model using Factorization Machine (FM), convolutional neural network (CNN), and random forest classifier (RFC), (b) Mugunthan et al. model [19]: developed the LR-DDoS detection model using RFC Classifier, (c) Singh et al. model Singh et al. [20]: developed the LR-DDoS detection model using Artificial Bee Colony (ABC) with artificial neural network (ANN), (d) Verma et al. model [22]: developed the LR-DDoS detection model using the mean absolute deviation (MAD) technique with an RFC (MAD-RF), (e) Xu et al. model [29]: developed the LR-DDoS detection model using a one-dimensional convolutional neural network and recurrent gated unit (1-CNN-RGU), (f) AlMasri et al. model [31]: developed the LR-DDoS detection model using a naive Bayes Classifier (NBC), and (g) Pérez et al. model [19]: developed the LR-DDoS detection model using the multi-layer perceptron (MLP) model and support vector machine (SVM) model [24].

5. Conclusions and future work

This paper presents, discusses, and evaluates a new intelligent system to investigate the impact of LR-DDoS attacks based on the MQTT protocol in the SD-IoT ecosystem. Specifically, the proposed system employs several supervised learning schemes to distinguish LR-DDoS traffic from legitimate traffic. To evaluate the performance of the learning models, a new dataset, LRDDoS-MQTT-2022, extracted from real-time traffic of SD-IoT communication comprising both DDoS attacks and normal packets, has been used. The minimum number of input features has been used to train the learning models with a balanced number of samples per class (100,000 samples for normal traffic and 100,000 samples for LR-DDoS traffic) to obtain the swiftest traffic prediction rate. As a result, the detection-based DTC model has reported the highest accuracy proportion, with 99.5% at the lowest prediction lagging. In the future, we will seek to investigate other types of cyber-attacks based on the MQTT protocol in the SD-IoT ecosystem, such as brute force authentication (BFA) attacks. Another perceptible future work is the generation of a real-time or simulated dataset covering diverse IoT protocols and attacks. Furthermore, mitigation or prevention techniques can be investigated as an extension for the presented detection system to provide a completely autonomous security system. Finally, incorporating the analysis of 5G attacks and persistent threats using advanced learning techniques can investigate in a similar context to this research [44,45].

Credit author statement

Mustafa Al Fayoumi: Conceptualization, Methodology, Software, Formal analysis, Resources, Investigation, Data curation, Visualization, Software, Validation, Funding acquisition, Writing – original draft, Writing- Reviewing, and Editing. Qasem Abu Al-Haija: Conceptualization, Methodology, Software, Formal analysis, Resources, Investigation, Data curation, Visualization, Software, Validation, Writing – original draft, Writing- Reviewing, and Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] IoT connected devices worldwide 2019-2030 | Statista." [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. [Accessed: 7-March-2023].

[2] Dantas Silva FS, Silva E, Neto EP, Lemos M, Venancio Neto AJ, Esposito F. A taxonomy of DDoS attack mitigation approaches featured by SDN technologies in IoT scenarios. Sensors 2020;20(11). MDPI AG.

[3] Abu Al-Haija Q, Al Badawi A, Bojja GR. Boost-Defence for resilient IoT networks: a head-to-toe approach. Expet Syst 2022;39(10):e12934. https://doi.org/10.1111/exsy.12934.

[4] Liu G, Quan W, Cheng N, Zhang H, Yu S. Efficient DDoS attacks mitigation for stateful forwarding in Internet of Things. J Netw Comput Appl 2019;130(June 2018):1–13.

[5] Faek R, Al-Fawa'reh M, Al-Fayoumi M. Exposing bot attacks using machine learning and flow level analysis. In: International conference on data science, vol. 2021. E-learning and Information Systems; 2021. p. 99–106.

[6] Wang J, Liu Y, Su W, Feng H. A DDoS attack detection based on deep learning in software-defined Internet of things. In: IEEE vehicular technology conference, vol. 2020; 2020. p. 7–11.

[7] Tuan NN, Hung PH, Nghia ND, Van Tho N, Van Phan T, Thanh NH. A DDoS attack mitigation scheme in ISP networks using machine learning based on SDN. Electronics (Switzerland) 2020;9(3):1–19.

[8] Al-Fayoumi M, Ahmad Y, Tariq U. A heterogeneous framework to detect intruder attacks in wireless sensor networks. Int J Adv Comput Sci Appl 2017;7(12):52–8.

[9] Al-Fawa'reh M, Al-Fayoumi M, Nashwan S, Fraihat S. Cyber threat intelligence using PCA-DNN model to detect abnormal network behavior. Egypt Inf J Jul. 2022;23(2):173–85.

[10] Abu Al-Haija Q, Al-Badawi A. Attack-aware IoT network traffic routing leveraging ensemble learning. Sensors 2022;22:241. https://doi.org/10.3390/s22010241.

[11] Karmakar KK, Varadharajan V, Nepal S, Tupakula U. SDN-enabled secure IoT architecture. IEEE Internet Things J 2021;8(8):6549–64.

[12] Li J, Cheng Y. Design and implementation of voice-controlled intelligent fan system based on machine learning. In: Proceedings of 2020 IEEE international conference on advances in electrical engineering and computer applications. AEECA 2020; 2020. p. 548–52.

[13] Hakiri A, Berthou P, Gokhale A, Abdellatif S. Publish/subscribe-enabled software defined networking for efficient and scalable IoT communications. IEEE Commun Mag Sep. 2015;53(9):48–54.

[14] Al-Haija QA, McCurry CD, Zein-Sabatto S. Intelligent self-reliant cyber-attacks detection and classification system for IoT communication using deep convolutional neural network. In: Ghita B, Shiaeles S, editors. Selected papers from the 12th international networking conference. INC 2020. Lecture notes in networks and systems, vol. 180. Cham: Springer; 2021. https://doi.org/10.1007/978-3-030-64758-2_8.

[15] Kandoi R, Antikainen M. Denial-of-service attacks in OpenFlow SDN networks. In: Proceedings of the 2015 IFIP/IEEE international symposium on integrated network management. IM 2015; 2015. p. 1322–6.

[16] Zhang C, Cai Z, Chen W, Luo X, Yin J. Flow level detection and filtering of low-rate DDoS. Comput Network 2012;56(15):3417–31.

[17] Tarasov Y, Pakulova E, Basov O. Modeling of low-rate DDoS-attacks. In: ACM international conference proceeding series; 2019. p. 10–3.

[18] Ibrahim RF, Abu Al-Haija Q, Ahmad A. DDoS attack prevention for Internet of thing devices using ethereum blockchain technology. Sensors 2022;22:6806. https://doi.org/10.3390/s22186806.

[19] M SR. Soft computing based autonomous low rate Ddos attack detection and security for cloud computing. J Soft Comput Paradigm 2019;2019(2):80–90.

[20] Singh R, Thakur K, Singh G, Gupta S. Prevention of IP spoofing attack in cyber using artificial bee colony and artificial neural network. In: ACM international conference proceeding series; 2019.

[21] Zhijun W, Qing X, Jingjie W, Meng Y, Liang L. Low-rate DDoS attack detection based on factorization machine in software defined network. IEEE Access 2020;8:17404–18.

[22] Verma P, Tapaswi S, Godfrey WW. An adaptive threshold-based attribute selection to classify requests under DDoS attack in cloud-based systems. Arabian J Sci Eng 2020;45(4):2813–34.

[23] Abu Al-Haija Q, Zein-Sabatto S. An efficient deep-learning-based detection and classification system for cyber-attacks in IoT communication networks. Electronics 2020;9:2152. https://doi.org/10.3390/electronics9122152.

[24] Perez-Diaz JA, Valdovinos IA, Choo KKR, Zhu D. A flexible SDN-based architecture for identifying and mitigating low-rate DDoS attacks using machine learning. IEEE Access 2020;8:155859–72.

[25] Cheng H, Liu J, Xu T, Ren B, Mao J, Zhang W. Machine learning based low-rate DDoS attack detection for SDN enabled IoT networks. Int J Sens Netw 2020;34(1):56–69.

[26] Nugraha B, Murthy RN. Deep learning-based slow DDoS attack detection in SDN-based networks. In: 2020 IEEE conference on network function virtualization and software defined networks. NFV-SDN 2020 - Proceedings; 2020. p. 51–6.

[27] Vaccari I, Aiello M, Cambiaso E. SlowITe, a novel denial of service attack affecting MQTT. Sensors 2020;20(10):1–16.

[28] Vaccari I, Aiello M, Cambiaso E. SlowTT: a slow denial of service against IoT networks. Information 2020;11(9).

[29] Xu C, Shen J, Du X. Low-rate DoS attack detection method based on hybrid deep neural networks. J Inf Secur Appl 2021;60(June):102879.

[30] Nanda WD, Sumadi FDS. LRDDoS attack detection on SD-IoT using random forest with logistic regression coefficient. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi) Apr. 2022;6(2):220–6.

[31] Almasri T, Snober MA, Al-Haija QA. IDPS-SDN-ML: an intrusion detection and prevention system using software-defined networks and machine learning. In: APICS 2022 - 2022 1st international conference on smart technology, applied informatics, and engineering, proceedings; 2022. p. 133–7. no. Ml.

[32] Ahalawat A, Babu KS, Turuk AK, Patel S. Corrigendum to 'A low-rate DDoS detection and mitigation for SDN using Renyi entropy with packet drop. J Inf Secur Appl 2022;70(October):103344.

[33] Haripriya AP, Kulothungan K. Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for Internet of things. EURASIP J Wirel Commun Netw 2019;2019(1).

[34] Kumar P, Kumar R, Gupta GP, Tripathi R. A Distributed framework for detecting DDoS attacks in smart contract-based Blockchain-IoT Systems by leveraging Fog computing. Trans Emerg Telecommun Technol 2021;32(6):1–31.

[35] Abu Al-Haija Q. Top-down machine learning-based architecture for cyberattacks identification and classification in IoT communication networks. Front. Big Data 2022;4:782902. https://doi.org/10.3389/fdata.2021.782902.

[36] Ghannadrad A. Machine leaning-based DoS attacks detection for MQTT sensor networks. Politecnico di Milano; 2021.

[37] Alzahrani A, Aldhyani THH. Artificial intelligence algorithms for detecting and classifying MQTT protocol Internet of things attacks. Electronics (Switzerland) 2022;11(22):1–16.

[38] Bhayo J, Jafaq R, Ahmed A, Hameed S, Shah SA. A time-efficient approach toward DDoS attack detection in IoT network using SDN. IEEE Internet Things J 2022;9(5):3612–30.

[39] Galeano-Brajones J, Carmona-Murillo J, Valenzuela-Valdés JF, Luna-Valero F. Detection and mitigation of DoS and DDoS attacks in iot-based stateful SDN: an experimental approach. Sensors 2020;20(3).

[40] Anbarsu S, Rayan AXA, Vetrian V. Software-defined networking for the Internet of things: securing home networks using SDN. In: Real-time data analytics for large scale sensor data. Elsevier; 2020. p. 215–70.

[41] Ivanova V, Tashev T, Draganov I. Detection of IoT-based DDoS attacks by network traffic analysis using feedforward neural networks. Int J Circ, Syst Signal Process 2022;16:653–62.

[42] Al Fayoumi M, Al Fawareh M, Nashwan S. VPN and non-VPN network traffic classification using time-related features. Comput Mater Continua (CMC) Mar. 2022;72(2):3091–111.

[43] Abu Al-Haija Q, Al Badawi A. High-performance intrusion detection system for networked UAVs via deep learning. Neural Comput Appl 2022;34:10885–900. https://doi.org/10.1007/s00521-022-07015-9.

[44] Ahad A, Ali Z, Mateen A, Tahir M, Hannan A, Garcia NM, Pires IM. A comprehensive review on 5G-based smart healthcare network security: taxonomy, issues, solutions, and future research directions. Array 2023:100290.

[45] Abdullayeva FJ. Advanced persistent threat attack detection method in cloud computing based on autoencoder and softmax regression algorithm. Array 2021;10:100067.