IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 6, June 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|

Detecting and Mitigating Botnet Attacksin Software-Defined Networks using Deep Learning Techniques

C Sai Vanaja, G S Uday Kiran Babu, D William Albert

Department of Computer Science and Engineering, Bheema Institute of Technology & Science, Adoni, Kurnool, Andhra Pradesh, India

ABSTRACT: Software-Defined Networking (SDN) is an emerging architecture that enables flexible and easy management and communication of large-scale networks. It offers programmable and centralized interfaces for making complex network decisions dynamically and seamlessly. However, SDN gives opportunities to both businesses and individuals to build network applications based on their demands and improve their services. In contrast, it started to face a new array of security and privacy challenges and simultaneously introduced the threats of a single point of failure. Usually, the attackers launch malicious attacks such as botnets and Distributed Denial of Service (DDoS) to the controller through OpenFlow switches. Deep learning (DL)-based security applications are trending, effectively detecting and mitigating potential threats with fast response. In this article, we analyze and show the performance of the DL methods to detect botnet-based DDoS attacks in an SDN-supported environment. A newly self-generated dataset is used for the evaluation. We also used feature weighting and tuning methods to select the best subset of features. We verify the measurements and simulation outcomes over a self-generated dataset and real testbed settings. The main aim of this study is to find a lightweight DL method with baseline hyper-parameters to detect botnet-based DDoS attacks with features and data that can be easily acquired. We observed that the best subset of features influences the performance of the DL method, and the prediction accuracy of the same method could be variated with a different set of features. Finally, based on empirical results, we found that the CNN method outperforms the dataset and real testbed settings. The detection rate of CNN reaches 99% for normal flows, and 97% for attack flows.

KEYWORDS: Botnet attack, convolutional neural network, deep learning, distributed denial-of-service attack, network security, software-defined networking

I. INTRODUCTION

The development of the internet is rapidly growing; the limitations of traditional networks have been explored. The emerging issues of the conventional networks can be solved by patching the network, which makes the network more bloated and the control ability of the network becomes weaker. The invention of Software-Defined Networking (SDN) [1], [2] has resolved these problems by decoupling the data and control planes. SDN became famous among the network community due to its novel architecture and can fulfill the demands of fast-growing networks. SDN has a centralized control architecture, so the SDN controllers can access all the OpenFlow switches in their range and control the entire network through the open south API interfaces [3], [4]. It is also known as the three-layer network architecture, application, control, and data layers. The application layer runs all the policies and rules the network administrator defines, and the SDN controller can adopt these rules dynamically. Any modification in the application layer may change the behavior of the whole network. The application layer is an excellent development by the open-source platform, which does not force the administrator to entirely relies on vendors [5]. Positively, the SDN allows administrators to eliminate license constraints and cloud-develop customized network applications over general-purpose hardware. The controllayer is known as the brain of the architecture, and SDN controllers run in this layer. The controllers receive the rules from the application layer, decode them into readable messages, and forward them to the underlying data layer; after that, they collect the feedback from the data layer and pass it back to the application layer. Moreover, a decision is made on the control layer, and the rules are implemented in the data layer. The data layer is non-intelligent, and different hardware devices, such as routers, OpenFlow switches, etc., exist in this layer, and instructions are passed by the control layer [6].

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|

Furthermore, it simplifies and eases the network's development, deployment, and maintenance. The network functions and features can be easily enhanced by updating and introducing new applications. Besides being cost- effective, SDN-based networks require simple hardware devices and have almost no compatibility issues [7]. Network access is allowed without revealing the details of the different underlying layers.Although SDN is an inordinate invention that can improve the network's flexibility and controllability and is considereda double-edged sword due to central control, a single controller can easily manage the network. SDN also helps toimprove traditional networks' security measures, but SDN has yet to be extensively trustworthy security measures to support the vision of next-generation networking [8]. Due to its centralized control nature and innovative architecture, it may introduce new security threats and can lead toward a single point of failure. Furthermore, the centralized nature ofSDN architecture unlocks the way for attackers to launch different attacks such as botnets, DDoS [6], saturation, and other types.

Botnet attacks are considered malicious attacks that become a critical threat in the next generation of networking [9]. Thebotnet is a network-based attack that breaches multiple computers into “bots” to launch malicious activities such as DDoS, identifying theft, domain name system spoofing,spamming, phishing, etc. In a botnet attack, a malicious actor, “Bot master,” tries to get unauthorized access to a single device and then implements botnet malware to take control of the device without alienating its legitimate users. After that, establish a connection of bots with a Command and Control (C&C) center owned by the attacker, and the bots remain ready to launch malicious activities under the instructions of C&C.Currently, botnet technology is typically used to launch the most sophisticated DDoS attacks in SDN and Internet of Things (IoT) based networks. The power and flexibility of the technology enable the botnet technology to generate several types of DDoS attacks. There are four main reasons to use advanced technology by the attackers:
(i) a powerful flooding attack can be quickly generated with a large numberof bots,
(ii) there is difficulty in finding the actual attacker,
(iii) they can use different protocols to dodge the security mechanism
(iv) the attack and normal traffic have similarity,

so, the real-time detection is difficult. With the advancementin technology, the attacker becomes more competent, and they know the structure of the SDN-based networks. They know they can control the whole network if they access the controller. The attackers can easily breach multiple computers into bots and make a member of the botnet force perform malicious actions on the SDN controller. So, the botnets have resulted in the progress of severe and massive DDoS attacks against the SDN and may cause a single failure point.Positively, deep learning-based applications are widely used for the detection of intrusion in various fields of networking. Examining the behavior in SDN becomes a new research area because using deep learning to study the behavior could be a first attempt for early detection [10], [11], as machines can respond faster than humans. Deep learning techniques allow machines to make decisions through training, features,trial, and error methods. DL approaches use historical data torealize the network behavior and predict upcoming traffic flows. The DL techniques proved excellent performance in classifying normal and attack traffic flows. These techniques do not rely on packet payload and take a set of a particular set of features to make predictions. DL techniques could help the SDN controllers to realize the network's current status. In existing solutions [12], [13], most researchers try to detect the simple DDoS attack in the SDN using machine learning and deep learning approaches. Moreover, the botnet-based DDoS attack potentially affects the SDN due to its centralized nature, so we need to focus on botnet-based DDoS. In most studies [8], [14], the researchers used many traffic features to train ML/DL methods. Extracting and collecting significant features from a real traffic flow is challenging and time-consuming due to authorization oraccessibility. Furthermore, most existing methods are verified by simulations or experimental datasets only; real- time testbed validation is intermittent.

A. MOTIVATION AND CONTRIBUTION

This work aims to conduct a comparative study on deep learning methods with optimized feature selection from network traffic flow to detect the botnet-based DDoS attackin an SDN environment. The selection of optimal features can improve DL methods' detection performance and accuracy and can help to reduce the SDN controller’s overhead. As we know, botnet attacks can hijack the controller and launch malicious activities. So, DL-based techniques and dedicated bot management are optimal solutions for SDN to deal with botnet traffic. The progressive development in security measures of SDN has motivated us to develop a DL method named “DepBot” to detect and mitigate botnet-based DDoS. In short, our contributions are as follows:

- Optimal Features: Compressing multiple features and identifying the most optimal features is our onecontribution. In this, a comparison of DL methods is performed to detect the botnet-based DDoS. The methods have

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|

experimented on a custom-developeddataset.

- Factor Analysis: We select and retrieve the optimal features into different subsets. We use feature weighting and threshold tuning-based evaluation techniques to choose the best feature subset. Based on this, we have identified optimal features and converted them into five subsets.
- DL Analysis: We select and use the five DL methods, such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Multilayer Perception (MLP), Long Short Term Memory (LSTM), and Deep Neural Network (DNN), that mostly appears in network intrusion detection-based applications. We use time-series- based techniques in RNN and LSTM to develop thedetection methods.
- Performance: The performance of the DL methods is evaluated using accuracy, precision, detection rate, and F1-score metrics. These methods build a relationship with raw data through multiple levels of abstraction. We observed that the DL methods produce good results on subset 3. This study wouldhelp to select the optimal features and best DL method for botnet-based DDoS attack detection in SDN.
- Mitigation: We used graph theory and dynamic flow deletion concepts to adopt a more targeted flow-dropping strategy.

II. RELATED WORK

This section discusses the recently developed network intrusion methods for SDN. Researchers have recently proposed various ML and DL-based methods for detecting botnet attacks. Some appropriate research works have been done in DDoS and botnet attack detection. [15] proposed “high-level PSI-rooted subgraph-based features” to detect the botnets in IoT, and they also used a hybrid combination of machine learning and deep learning methods. Their mainaim was to detect the attack with a minimum number of features that can help to reduce the space and faster the detection speed. [16] developed an IoT botnet sensor system based on unsupervised learning. They do the hyper parameterization of SVM using the “Grey Wolf Optimization (GWO) algorithm” to determine the critical features for a botnet attack. The suggested technique helps todetect IoT botnet attacks launched from intelligent, vulnerable nodes. We observed another network-based anomaly detection approach called N-BaIoT in [17], which takes snapshots of the network behavior and employs deep autoencoders to detect the attack traffic. [18] proposed a hybrid method of PSO algorithms with a voting mechanismto detect botnet attacks in IoT. In this method, the PSO is used to pick the critical and remarkable features, and the election process is performed using DNN, Decision Tree (DT) C4.5, and SVM to detect the botnet attacks. In another research [19], the authors proposed a CNN-based method named BoTIDS. They observed that this method produced good results on a complete dataset and a predefined subset of 10 features compared to LSTM and RNN. However, the evaluation results on a subset are comparatively better than the full dataset. We observed another CNN-based Anomaly Intrusion Detection System (AIDS) [20]. This system is tested on BoT-IoT and Network Intrusion Detection (NID) datasets. The proposed system produced effective results onthe NID dataset compared to BoT-IoT. In [8], we observed a two-level DDoS attack detection method based on DL and Information Entropy (IE). First, they used information entropy to detect the suspicious port and components in coarse granularity. Then, they executed a CNN-based fine-grained detection mechanism to classify the attack and normal traffic. [21] developed a DNN-based network intrusion detection system to show the potential of DL methods for malicious traffic detection. A combined method of Autoencoder and RNN named DDoSNet has been proposed by [22] to detect DDoS attacks in SDN environments. [23] proposed an LSTM-based DDoS detection method and compared the performance of this method with Random-Forest (RF) based approach, and it reduced the error rate from 7.517% to 2.103%. [24] present a deep learning-based lightweight and practical DDoS detection system called LUCID, which shows the capabilities of CNN to classify traffic flows as normal or malicious. [25] proposed a method based on SVM, assisted by Genetic Algorithm (GA) and Kernel Principle Component Analysis (KPCA). In this method, the authors used the KPCA to reduce the dimensions of the features, and GA was used to tune the hyperparameters of the SVM. They introduce an improved Kernel Function (N-RBF) to reduce the noise caused by feature differences. [26] used Particle Swarm Optimization(PSO)-BP neural networks and normal entropy metrics in their research to detect the DDoS attack.[27] improved the DDoS attack detection performance usinga trigger mechanism for the first time. This method helps reduce the overhead of the switches and controller. Another hybrid method based on Artificial Neural Networks (ANNs) and DNN was proposed in [28]. To classify the major bot attacks as Emotet, Zbot, Dridex, and Sality, the authors used a combined method of RNN and LSTM [29]. Another hybrid method based on LSTM and CNN was proposed in [30]. They deployed this method in a real testbed which used data gathered from nine commercial IoT devices to detect the attack. In [31], the authors proposed a distributed method based on CNN and LSTM with an additional cloud-based component for detecting DDoS and phishing attacks. [32] introduced a method combining Fuzzy and Taylor-Elephant Herd Optimization (F-TEHO) inspired by Deep Belief Network (DBN) to detect DDoS attacks. This method

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|

compromises three modules: feature extraction, selection, and classification. The feature extraction module extracts feature from raw packet data, and then the Holo-entropy method is used to select essential features. Finally, the classification process is performed using the FT-EHO method. Another method based on cognitive-inspired computing with dual address entropy is proposed in [33]. In this method, the dual address entropy is first used to extract the features from the flow table, and then SVM classifies the traffic as normal or attack. This method helps to detect the attack at the preliminary stage and restore the usual communication in time.

III. OVERVIEW OF THE PROPOSED METHOD

To successfully protect the SDN against botnet-based DDoS attacks, detecting and blocking attack flows from the attackers is significant. A flow comprises several packets with the same information (source and destination IP address, source and destination port number, protocol, and other features are listed in TABLE 3). In an attack flow, the source IP address belongs to the attacker. Assume we have N number of flow samples and y classes. Let the flow samples are $\boldsymbol{X} = \{\boldsymbol{F_1}, \boldsymbol{F_2}, \boldsymbol{F_3}, \dots\dots\dots, \boldsymbol{F_N}\}$ s $\mathbb{R}^{d \times N}$, where $\boldsymbol{F_i}$ represents the ith flow, $\boldsymbol{d}$ is the number of original features in a flow, and $\boldsymbol{N}$ represents the total number of flows. In a $\boldsymbol{F_i}$ flow, the actual tables are defined as $\boldsymbol{y_i} = \{\boldsymbol{0}, \boldsymbol{1}\}$. We aim to develop an end-to-end method to predict a label as an actual label ($\boldsymbol{y_{pr}(i)} = \boldsymbol{y_i}$).

This section discusses a feature selection and deep learning-based comparative study for flow classification. It would help classify the normal and attack flows and improve efficiency and accuracy. Fig.1 shows the operation phases of the proposed research. First, the incoming packets are placed into a pretty table. Second, the features according to TABLE 3 are computed and extracted for each packet flow individually. Third, the optimal features are selected based on the optimal threshold values of the feature weights. Finally, the selected features are combined into a subset, passing this subset as input to the DL classifiers (e.g., RNN, CNN, MLP, LSTM, and DNN) for flow classification.

A. BRIEF DESCRIPTION OF DL METHODS AND HYPER-PARAMETERS SETTING

This section gives an overview of deep learning methods.

RECURRENT NEURAL NETWORK (RNN)

RNNs [34] are known as a dynamic type of feed-forward neural network. This network can learn the sequential data over time-steps. In conventional networks, each unit output depends on the current input with no dependency between the previous or output of the same unit. However, time- series-based applications rely on sequential data; each current sample depends on the last sample. Therefore, more than conventional networks is needed for these applications. RNNs use a time series mechanism to handle these issues. The sequential process for each time step is computed as follows:

$$= f\,(W_s\,x\,,x\,t\ + W_y\,s\,,s\,t-1\ + b\,s\,) \qquad (1)$$

$$Y\,t\ = (W\,y\,s\,,t-1\ + b\,y\,)$$

IV. EXPERIMENTAL ENVIRONMENT

The experiments are conducted in the Mininet virtual environment [41] with a POX controller [42]. We used Mininet version 2.3.2 in our experiments, which supports Open Virtual Switches (OVS) [43]. Usually, the SDN-based network emulations are performed in Mininet, and OVS is an open-source virtual machine that supports OpenFlow protocols. The POX is an interface-rich SDN controller that authorizes the network and research community to cultivate control and network applications using python. The operating system and hardware settings are intel Core i7, 8GB RAM, and Windows 10 operating system. The deep learning classifiers are implemented in python language with the Keras framework. A customized centralized network topology based on SDN is built in Mininet to conduct experiments. A complex tree network structure is adopted in this topology, and the SDN controller is installed in the control layer that centrally controls the hosts and switches at the data layer. The application layer consists of four modules: flow statistics collector, feature extractor, DL classifier, and mitigator. Furthermore, the network topology consists of a controller and seven OpenFlow switches. The switch S1 is connected to six hosts, and all the other switches with three hosts. The experimental network topology is shown in Fig. 2.

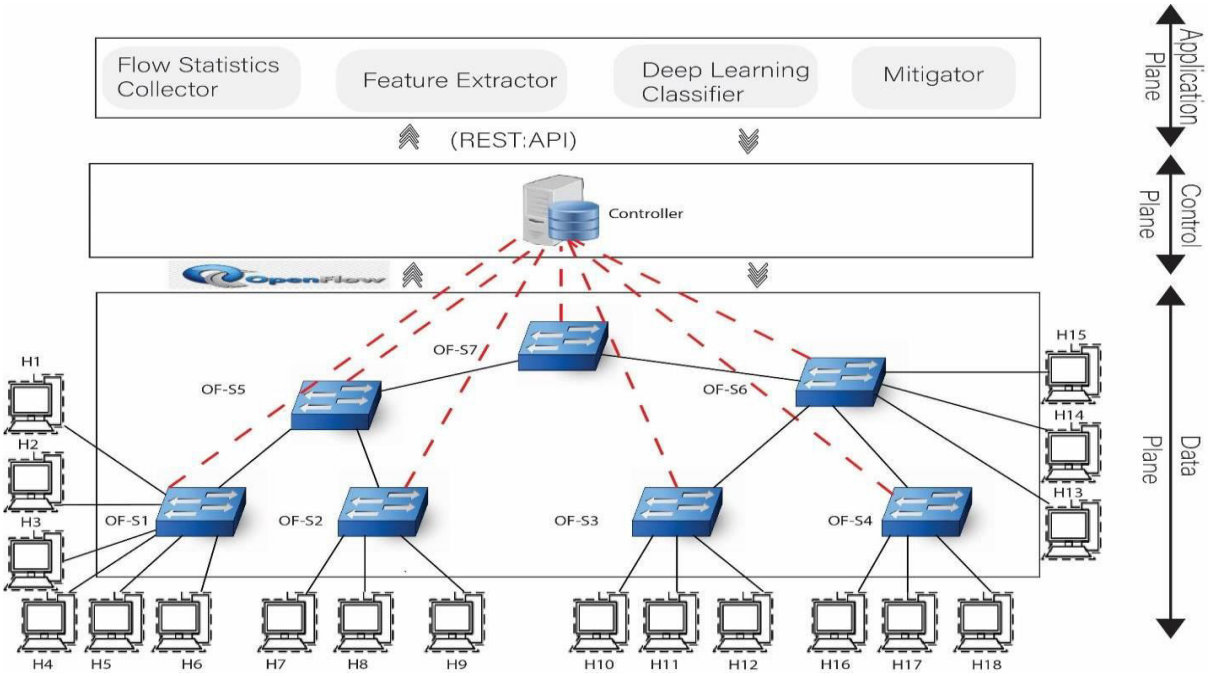
International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

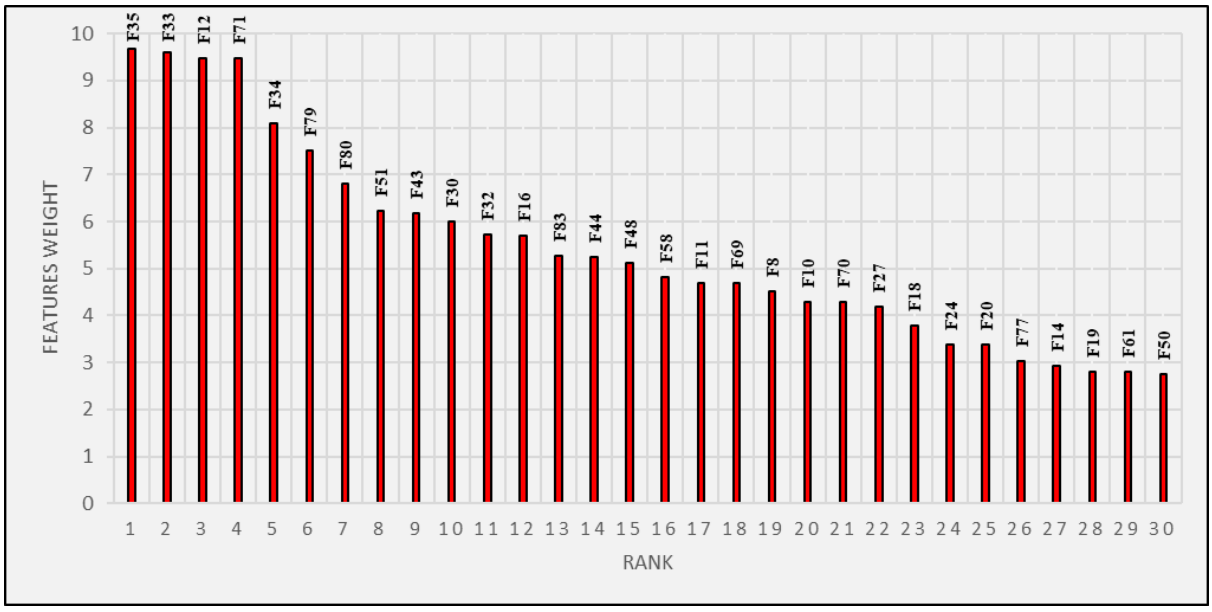
Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|





V. DISCUSSION AND FUTURE WORK

In this paper, we study and employ DL methods to contributeto the detection of botnet-based DDoS attacks in an SDN-supported environment. We explore the efficiency of the DL methods with baseline hyper-parameters and optimal features to detect the attack. We evaluate the performance ofthe DL methods using different evaluation metrics (e.g., accuracy, detection rate, training, detection time, etc.). A core contributing factor of this study is the generation and collection of the training dataset, which is purely developed in an SDN environment instead of relying on old or traditional datasets. In most studies [17], [20], the researchers used traditional datasets (e.g., Na-BaIoT, NSL- KDD, CIC-DDoS 2019, DARPA, etc.), which are notdatasets suffer imbalanced problems. In addition, we do notrely on a dataset with many features but divide the whole dataset into subsets (e.g., optimal features) based on the feature's importance; then, these subsets have been tested toobserve the impact of optimal features on the method performance. Simulation results show a variation in theperformance of each DL method on the same subset of features, so we can

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.423

|www.ijirset.com | A Monthly Peer Reviewed & Referred Journal | e-ISSN: 2319-8753; p-ISSN: 2347-6710|

|Volume 13, Issue 6, June 2024|

|DOI: 10.15680/IJIRSET.2024.1306175|

conclude that the optimal features couldimprove the detection rate.

VI. CONCLUSION

SDN redefines the network's management and communications and leads to reforms. Although SDN has vast capabilities, it also introduced new emerging security challenges, which need the great attention of both the network and research community. The botnet attacks target the devices in the data plane or the controller in the control plane. Detecting the botnet and DDoS attacks in SDN becomes more challenging than in traditional networks due to the sophistication of traffic flow features. The SDN's decoupled architecture would help to build and deploy a method that flexibly detects the botnet-based DDoS attack. Nevertheless, the new network environments shall be adopted to provide a reliable definition and behavior of botnet-based DDoS attacks in SDN. Nowadays, deep learning-based network applications are trending and could be utilized to secure the SDN. The DL methods use the historical training sets in predicting real-time network state and inform the controller. In our study, first, we produced a dataset in a pure SDN-supported environment and then used DL techniques to predict the botnet-based DDoS attack to resolve these issues. The DL procedures, from dataset generation to attack detection, are done over a single controller and gratitude to the SDN's centralized control. Inreal-time, the controller sends a packet-In request to all the OpenFlow switches to collect the flow statistics, and then these flow statistics are used by the DL module inside the controller to detect the attacks. The real-time detection rate of CNN with 30 features reaches 99% for normal flows and 97% for attack flows, a remarkable contribution of this studyto secure the SDN from botnet-based DDoS attacks. We named the CNN-based method "DepBot," simplifying the data preparation in the DL training phase. Although the DL methods achieved reasonable detection rates and accuracies with 30 features in the adopted network scenario, it is suggested to use advanced feature selection methods to better the prediction accuracy further.

Data Availability: The source code to generate the botnet attack and the normal and attack traffic dataset utilized to conduct this study is available at: https://github.com/Waqas-

REFERENCES

1. L. Tan, Y. Pan, J. Wu, J. Zhou, H. Jiang, and Y. Deng, "A new framework for DDoS attack detection and defense in SDN environment," IEEE Access, vol. 8, pp. 161908–161919, 2020.
2. S. Wang et al., "Detecting flooding DDoS attacks in software defined networks using supervised learning techniques," Eng. Sci.Technol. an Int. J., vol. 35, p. 101176, 2022.
3. Y. Cui et al., "Towards DDoS detection mechanisms in software-defined networking," J. Netw. Comput. Appl., vol. 190, p. 103156,2021.
4. J. Ye, X. Cheng, J. Zhu, L. Feng, and L. Song, "A DDoS attack detection method based on SVM in software defined network," Secur. Commun. Networks, vol. 2018, 2018.
5. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," Futur. Gener. Comput. Syst., vol. 82, pp. 761–768, 2018.
6. J. A. Perez-Diaz, I. A. Valdovinos, K.-K. R. Choo, and D. Zhu, "A flexible SDN-based architecture for identifying and mitigatinglow-rate DDoS attacks using machine learning," IEEE Access, vol. 8, pp. 155859–155872, 2020.
7. R. K. Chouhan, M. Atulkar, and N. K. Nagwani, "A framework to detect DDoS attack in Ryu controller based software defined networks using feature extraction and classification," Appl. Intell.,pp. 1–21, 2022.
8. Y. Liu, T. Zhi, M. Shen, L. Wang, Y. Li, and M. Wan, "Software-defined DDoS detection with information entropy analysis and optimized deep learning," Futur. Gener. Comput. Syst., vol. 129,pp. 99–114, 2022.
9. O. Habibi, M. Chemmakha, and M. Lazaar, "Imbalanced tabular data methodization using CTGAN and machine learning to improve IoT Botnet attacks detection," Eng. Appl. Artif. Intell., vol. 118, p. 105669, 2023.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details