International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

Two-Fold Machine Learning Approach to Prevent and Detect IOT BOTNET Attacks

Ms.M.Sandhya Vani [1] and Dr.C.Ramesh Kumar [2]
[1]Assistant Professor, Department of Information Technology, MREC (A), Hyderabad-500100
[1] Professor, Department of Information Technology, MREC (A), Hyderabad-500100

ABSTRACT

The botnet attack is a multi-stage and the most prevalent cyber-attack in the Internet of Things (IoT) environment that initiates with scanning activity and ends at the distributed denial of service (DDoS) attack. The existing studies mostly focus on detecting botnet attacks after the IoT devices get compromised, and start performing the DDoS attack. Similarly, the performance of most of the existing machine learning based botnet detection models is limited to a specific dataset on which they are trained. As a consequence, these solutions do not perform well on other datasets due to the diversity of attack patterns. Therefore, in this work, we first produce a generic scanning and DDoS attack dataset by generating 33 types of scan and 60 types of DDoS attacks. In addition, we partially integrated the scan and DDoS attack samples from three publicly-available datasets for maximum attack coverage to better train the machine learning algorithms. Afterwards, we propose a two-fold machine learning approach to prevent and detect IoT botnet attacks. In the first fold, we trained a state-of-the-art deep learning model, i.e., ResNet-18 to detect the scanning activity in the premature attack stage to prevent IOT BOTNET attacks. While, in the second fold, we trained another ResNet-18 model for DDoS attack identification to detect IoT botnet attacks. Overall, the proposed two-fold approach manifests 98.89% accuracy, 99.01% precision, 98.74% recall, and 98.87% f1-score to prevent and detect IoT botnet attacks. To demonstrate the effectiveness of the proposed two-fold approach, we trained three other ResNet-18 models over three different datasets for detecting scan and DDoS attacks and compared their performance with the proposed two-fold approach. The experimental results prove that the proposed two-fold approach can efficiently prevent and detect botnet attacks as compared to other trained models.

1. INTRODUCTION

Internet of Things (IOT) revolutionized the technology by enabling real-world objects/things to connect and communicate with each other over the internet to luxuriate human life [1], [2]. Over the past few years, the adoption of smart IOT devices like smart cameras, smart TV, smart wear ables, smart toys, smart bulbs, etc., is exponentially increasing in our daily life [3], [4]. Therefore, this new emerging trend in the field of computing has empowered our everyday life objects to connect and communicate with each other without human intervention. Despite the IOT devices are helping us in a lot of areas, these devices have negligible or very limited security features [3]. Furthermore, many IOT devices come with a fixed key or hard-coded default username and

password, which a user cannot change [5]. These security pitfalls makeit easy for hackers to exploit these insecure IOT devices and get control over them [4].

The recent trends reveal that the cyber-attacks increasing day by day with the rapid increase of insecure IOT devices [6]. Among the recently reported cyber-attacks, botnet and distributed denial of service (DDOS) attacks are the most prevalent attacks, which are increased both in frequency and magnitude over the last decade [4], [6]. A botnet attack is a cyber-attack in which an attacker first scans a network to look for weakly secured or vulnerable (IOT) devices. After analysing the scanning information, the attacker targets vulnerable (IOT) devices to install a bot program into them through malware [7].

The installed bot program connects the infected devices with a central server or a peer network from where the further commands are sent to them to perform different malicious activities like sending spams, _ooding DDOS [6], [8], etc., from plenty of infected IOT devices over the target server, website, etc. Once an IOT device gets infected and becomes part of a botnet, then the attacker uses the infected device to perform DDOS attacks.

The botnet attack is not only a serious threat to insecure IOT devices but also a crucial threat to the whole internet [6]. With the advent of the Mirai botnet attack in 2016, the IOT botnet attacks are continuously escalating [9]. After the public disclosure of the Mirai botnet source code, many variants and imitators of Mirai botnet have been evolved [9]. These new variants and imitators have infected millions of IOT devices [3], [9] and wreaked ever large and catastrophic DDOS attacks like GitHub [10], AWS [11], etc., over the past few years.

Nowadays, attackers can easily locate insecure IOT devices via online services such as Shodan [12], Censys [13], etc. These online search engine services provide a huge amount of information to attack insecure IOT devices [9]. By compromising the insecure IOT devices, an attacker can perform several cyber-attacks such as spamming, phishing, DDOS [6], [8], [9], etc., to wreak havoc against the other resources on the Internet. Some recent studies exposed that IOT devices are much prone to botnet and DDOS attacks, as a wide range of DDOS attacks are performed by compromised IOT devices [14], [15]. Likewise, Gartner recently predicted that 25% of the cyber-attacks are posed due to the insecure IOT devices [16].

In order to secure the insecure IOT devices to become a bot and perform different DDOS attacks, there must be an efficient security system to detect IoT bots. The existing botnet and DDOS attack detection techniques are divided into two categories, i.e., host-based techniques and network based techniques [17]. Due to the resource constraint nature (i.e., limited memory, battery, and compute power) of IoT devices, the host-based

solutions are not feasible for IOT devices [1], [17]. However, the network-based solution is a better way to protect the IOT devices and network from these devastating cyber-attacks. The network-based techniques are subdivided into three main types [18]_[22]:

1) **Signature-based detection method:** relies on matching the network traffic with some specific rules defined in the rule database to detect and prevent potential attacks.

2) **Anomaly-based detection method:** analyses the normal behaviour of network traffic and builds a baseline profile of each device communicating in the network Any significant deviation from the baseline is considered as an anomaly. The anomaly-based detection method is further classified into two sub types V

_ **Statistics-based detection:** These methods detect anomalies based on a statistical distribution of intrusions.

_ **Machine learning-based detection method:** detects abnormalities based on packet and payload features. These methods mainly detect and prevent potential attacks using machine learning models.

_ **Knowledge-based detection method:** detects anomalies based on the profile or previous knowledge of a network. The profile or previous knowledge of the network is generated under different test cases to detect abnormalities in the network [22].

3) **Specification-based detection method:** performs intrusions detection based on the specifications or rules defined by a user [22].

The major drawback of the signature-based detection method is that it only detects the known threats for which the rules are available in its rules' database [20], [21]. On the other hand, the stateful protocol-based detection methods have limited ability to inspect the encrypted traffic. However, the traffic behaviour analysis, i.e., anomaly detection is very effective in both analysing the encrypted traffic and detecting the unknown attacks [19]. In case of anomaly detection methods, the machine learning approach has shown tremendous performance in recent years. The machine learning based detection methods are trained on datasets to learn and distinguish the behaviour and pattern of normal and attack traffic [20], [21]. Henceforth, by learning the normal and attack traffic patterns, the machine learning models are useful to detect new botnet and DDOS attacks that are derived variants or imitators of the existing botnet and DDOS attacks. The existing botnet attack detection methods detect the botnet after the IOT devices are compromised by some malware and start performing malicious activities as directed by the botmaster. Moreover, the performance of most of the existing machine learning based botnet detection models is limited to a specific dataset on which they are trained [6]. This is due to the fact that different datasets contain different types of botnet attacks. Further, the features used for detecting botnet attacks from one certain dataset, are not adequate to efficiently detect the botnet attacks from other datasets due to the diversity of botnet attacks [6]. As a consequence, these solutions do not perform well on other datasets due to the diversity of attack patterns [6]. However, in order to protect the IOT

devices from being compromised, there is a crucial need for providing a protection mechanism to safeguard the O devices from botnet and DDOS attacks during the premature stage (i.e., scanning) of the botnet attack. Therefore, in this work, we propose a novel two-fold approach to prevent a botnet attack during the premature stage (i.e., scanning attack) and to detect DDOS attack in IOT network in case an attacker compromises an IOT device and start performing a DDOS attack. As discussed earlier that an attacker can use the bot infected IoT devices to perform different malicious activities like sending spam emails, _ooding DDOS [6], [8], etc., however, in this work, we focus on detecting DDOS attacks performed by bot-infected IOT devices. The proposed twofold approach uses a state-of-the-art deep learning model, i.e., ResNet which is first trained for detecting the scanning activity and then trained for detecting the DDOS attack performed by the attacker or compromised IOT devices towards or outside the network. For preventing the IOT devices and network from IOT botnet attacks, in the first fold, we trained the ResNet-18 [23] model for scanning attack detection so that it can detect the premature attack stage and notify about the malicious attempt before an attacker goes to further steps for compromising the IOT devices. On the other hand, in the second fold, we trained the ResNet-18 [23] model for DDOS attack detection to detect and mitigate the botnet attack, in case an attacker invades the scanning attack detection model, install malware on IOT devices and starts performing DDOS attacks.

The key contributions of this work are as follows:

_ We analysed the frequently used scanning and DDOS attack techniques and produced a generic dataset by generating 33 types of scan and 60 types of DDOS attacks. In addition, we partially integrated the scan and DDOS attack samples from three publicly-available datasets for maximum attack coverage for better training of machine learning algorithms.

_ We proposed a two-fold machine learning approach to prevent and detect both inbound and outbound botnet attacks in the IOT network environment. The proposed two-fold approach prevents IOT botnet attacks by detecting the scanning activity, while it detects the IOT botnet attack by identifying the DDOS attack.

_ Finally, to demonstrate that the performance of the proposed two-fold approach is not limited to a single dataset, we trained three ResNet-18 [23] models over three different datasets and compared their performance with the proposed two-fold approach for detecting and preventing IOT botnet attacks.

2. LITERATURE SURVEY

Nguyen *et al.* [16] proposed a graph-based approach to detect the IoT botnet via printing string information (PSI) graphs. The authors used PSI graphs to get high-level features from the function call graph and then trained a convolution neural network (CNN), a deep learning model, over the generated graphs for IoT botnet detection. Likewise, Wang *et al.* [24] proposed an

automated model named as BotMark. Their proposed model detects botnet attacks based on a hybrid analysis of flow-based and graph-based network traffic behaviors. The flow-based detection is performed by k-means, which calculates the similarity and stability scores between flows. While the graph-based detection uses the least-square technique and local outlier factor (LOF) which measures anomaly scores. Similarly, Yassin *et al.* [25] proposed a novel method that compromises a series of approaches such as the utilization of the frequency process against registry information, graph visualization and rules generation. The authors investigated the Mirai attacks using the graph-theoretical approach. In order to identify similar and dissimilar Mirai patterns, the authors used directed graphs. The proposed approach only focuses on the Mirai
attack.

Almutairi *et al.* [27] proposed a hybrid botnet detection technique that detects new botnets implemented on three levels, i.e., host level, network level and a combination of both. The authors focused on focused HTTP, P2P, IRC, and DNS botnet traffic. The proposed technique consists of three components: host analyser, network analyser, and detection report. The authors used two machine learning algorithms, i.e., Naïve Bayes and a decision tree for traffic classification. Similarly, Blaise *et al.* [28] proposed a bot detection technique named BotFP, for bot fingerprinting. The proposed BotFP framework has two variants, i.e., BotFP-Clus which groups similar traffic instances using clustering algorithms and BotFP-ML is designed to learn from the signatures and identify new bots using two supervised ML algorithms, i.e., SVM and MLP. Likewise, Soe *et al.* [30] developed a machine learning-based IoT botnet attack detection model. The proposed model consists of two stages: a model builder and an attack detector. In the model builder stage, data collection, data categorization, model training and feature selection are performed step by step. While in the attack detector stage, the packets are first decoded and then the features are extracted in the same way as in the model builder phase. Finally, the features are passed to the attack detector engine where artificial neural network (ANN), J48 decision tree, and Naïve Bayes machine learning models are used for botnet attack detection.

Sriram *et al.* [31] proposed a deep learning-based IoT botnet attack detection framework. The proposed solution specifically considered network traffic flows, which are further converted into feature records and then passed to the deep neural network (DNN) model for IoT botnet attack detection. Nugraha *et al.* [32] evaluated the performance of four deep learning models for botnet attack detection by performing a couple of experiments. The experimental results revealed that CNN-LSTM outperformed all deep learning models for botnet attacks detection.

Disadvantages

- An existing methodology prevents botnet attacks by detecting the scanning attack activity while it detects the botnet attack by identifying the

DDoS attack for both inbound and outbound traffic.

- IoT botnet attack doesn't initiates with the scanning activity and ends at the DDoS attack.

3. PROPOSED SYSTEM

The proposed system analyzed the frequently used scanning and DDoS attack techniques and produced a generic dataset by generating 33 types of scan and 60 types of DDoS attacks. In addition, we partially integrated the scan and DDoS attack samples from three publicly-available datasets for maximum attack coverage for better training of machine learning algorithms. The system proposed a two-fold machine learning approach to prevent and detect both inbound and outbound botnet attacks in the IoT network environment. The proposed two-fold approach prevents IoT botnet attacks by detecting the scanning activity, while it detects the IoT botnet attack by identifying the DDoS attack. Finally, to demonstrate that the performance of the proposed two-fold approach is not limited to a single dataset, we trained three ResNet-18 models over three different datasets and compared their performance with the proposed two-fold approach for detecting and preventing IoT botnet attacks.

The system proposed a novel two-fold machine learning approach to prevent and detect botnet attacks in IoT networks.

The proposed methodology stops an attacker during the scanning activity so that an attacker cannot proceed to further attack stages.

4. IMPLEMENTATION

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as

Login, Browse and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Botnet Detection Status, View Botnet Detection Status Ratio,

Download Predicted Data Sets, View Botnet Detection Status Ratio Results,, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGINPREDICT BOTNET DETECTION TYPE, VIEW YOUR PROFILE.

5. CONCLUSION

In this work, we proposed a two-fold machine learning approach to prevent and detect IOT botnet attacks. In the first fold, we trained a state-of-the-art deep learning model, i.e., ResNet-18 for scanning attack detection, and named it ResNetScan-1 model. While in the

second fold, we trained another ResNet-18 model (named as ResNetDDoS-1 model) in order to detect the DDOS attack in case if the scanning detection model fails to prevent a botnet attack. In order to authenticate the performance of the proposed ResNetScan-1 model and ResNetDDoS-1 model, we performed a couple of experiments in which we take the scan and DDOS traffic samples from three publicly-available datasets, trained the ResNet-18 model over these datasets, and saved the resultant Res Net Scan and Res Net DDOS models. We then tested each resultant Res Net Scan and Res Net DDOS model over the test set of other datasets on which they were not trained. The experimental results revealed that the performance of all Res Net Scan and Res Net DDOS models except the proposed ResNetScan-1 and ResNetDDoS-1 model crucially reduced when tested over the datasets on which they were not trained. Furthermore, the experimental results proved that the proposed ResNetScan-1 and ResNetDDoS-1 models persisted in their performance and outperformed all other models for detecting the scan and DDOS attacks. Hence, the proposed two-fold approach is efficient and robust to prevent and detect IOT botnet attacks with a large attack patterns coverage.

The current work only covers 33 types of scanning and 60 types of DDOS attacks. In future, we aim to cover more scanning and DDOS attacks techniques in order to well train the proposed framework for more efficient prevention and detection of IOT botnet and DDOS attacks. Further, we can deploy the proposed two-fold approach in an IDS to investigate its effectiveness on live network traffic.

REFERENCES

[1] I. Ali, A. I. A. Ahmed, A. Almogren, M. A. Raza, S. A. Shah, A. Khan, and A. Gani, ``Systematic literature review on IoT-based botnet attack," *IEEE Access*, vol. 8, pp. 212220_212232, 2020.

[2] S. Ghazanfar, F. Hussain, A. U. Rehman, U. U. Fayyaz, F. Shahzad, and G. A. Shah, ``IoT-Flock: An open-source framework for IoT traf_c generation," in *Proc. Int. Conf. Emerg. Trends Smart Technol. (ICETST)*, Mar. 2020, pp. 1_6.

[3] M. Safaei Pour, A. Mangino, K. Friday, M. Rathbun, E. Bou-Harb, F. Iqbal, S. Samtani, J. Crichigno, and N. Ghani, ``On data-driven curation, learning, and analysis for inferring evolving Internet-of-Things (IoT) botnets in the wild," *Comput. Secur.*, vol. 91, Apr. 2020, Art. no. 101707.

[4] 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 *Proc. IEEE 23rd Int. Multitopic Conf. (INMIC)*, Nov. 2020, pp. 1_6.

[5] S. Dange and M. Chatterjee, ``IoT botnet: The largest threat to the IoT network," in *Data Communication and Networks*. Singapore: Springer, 2020, pp. 137_157.

[6] F. Hussain, S. G. Abbas, U. U. Fayyaz, G. A. Shah, A. Toqeer, and A. Ali, ``Towards a universal features set for IoT botnet attacks detection," in *Proc. IEEE 23rd Int. Multitopic Conf. (INMIC)*, Nov. 2020, pp. 1_6.

[7] A. O. Proko_ev, Y. S. Smirnova, and V. A. Surov, ``A method to detect

Internet of Things botnets," in *Proc. IEEE Conf. Russian Young Res. Electr. Electron. Eng. (EIConRus)*, Jan. 2018, pp. 105_108.

[8] B. K. Dedeturk and B. Akay, ``Spam _ltering using a logistic regression model trained by an arti_cial bee colony algorithm," *Appl. Soft Comput.*, vol. 91, Jun. 2020, Art. no. 106229.

[9] N. Vlajic and D. Zhou, ``IoT as a land of opportunity for DDoS hackers," *Computer*, vol. 51, no. 7, pp. 26_34, 2018.

[10] *GitHub Survived Biggest DDoS Attack Ever Recorded*. Accessed: May 3, 2021. [Online]. Available: https://github.blog/2018-03-01-ddosincident- report/