DUAL MACHINE LEARNING STRATEGY FOR PREVENTING AND DETECTING IOT BOTNET ATTACKS

¹D.Saikrishna, ²Burla Supriya ¹Assistant Professor, ²MCA Student Department Of MCA Student Sree Chaitanya College of Engineering, Karimnagar

ABSTRACT

In the context of the Internet of Things (IoT), the botnet assault is the most common cyberattack. It starts with scanning activity and culminates in a distributed denial of service (DDoS) attack. The majority of current research is on identifying botnet assaults that occur after hacked IoT devices begin executing DDoS attacks. In a similar vein, the majority of machine learning-based botnet detection algorithms now in use are only effective when used with a certain dataset. Because of the variety of attack patterns, these solutions therefore perform poorly on different datasets. Therefore, in this study, we generate 60 types of DDoS attacks and 33 types of scans in order to first create a general scanning and DDoS attack dataset. Furthermore, in order to effectively train the machine learning algorithms, we partially combined the scan and DDoS attack samples from three publicly available datasets for maximum attack coverage. We then suggest a dual machine learning strategy to stop and identify IoT botnet assaults. In order to stop IoT botnet assaults, we first developed a cutting-edge deep learning model, ResNet-18, to identify scanning activity in the early stages of an attack. In order to identify DDoS assaults and identify IoT botnet attacks, we trained a second ResNet-18 model in the second fold. Overall, the suggested two-pronged method

for preventing and detecting IoT botnet assaults shows 98.89% accuracy, 99.01% precision, 98.74% recall, and 98.87% f1score. In order to illustrate the efficacy of the suggested two-fold method, we trained three more ResNet-18 models for the detection of scan and DDoS assaults on three distinct datasets and contrasted their results with the suggested two-fold method. In comparison to previous trained models, the experimental findings demonstrate that the suggested two-pronged method may effectively prevent and detect botnet assaults.

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 hardcoded 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 anomalybased detection method is further classified into two sub types V

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

<u>Machine learning-based detection</u> 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 botinfected 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 further steps to 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 publiclyavailable 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 twofold 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

"Systematic literature review on IoTbased botnet attack,"

The adoption of the Internet of Things (IoT) technology is expanding exponentially because of its capability to provide a better service. This technology has been successfully implemented on various devices. The growth of IoT devices is massive at present. However, security is becoming a major challenge with this growth. Attacks, such as IoT-based botnet attacks, are becoming frequent and have become popular amongst attackers.IoT has a resource constraint and heterogeneous environments, such as low computational power and memory. Hence, these constraints create problems in implementing a security solution in IoT devices. Therefore, various kind of attacks are possible due to this vulnerability, with IoT-based botnet attack being one of the most popular. In this study, we conducted a comprehensive systematic literature review on IoT-based botnet attacks. Existing state of the art in the area of study was presented and discussed in detail. A systematic methodology was adopted to ensure the coverage of all important studies. This methodology was detailed and repeatable. The review outlined the existing proposed contributions, datasets utilised, network forensic methods utilised and research focus of the primary selected studies. The demographic characteristics of primary studies were also outlined. The result of this review revealed that research in this domain is gaining momentum. particularly in the last 3 years (2018-2020). Nine key contributions were also identified, with Evaluation, System, and Model being the most conducted.

"IoT-Flock: An open-source framework for IoT traffic generation,"

Network traffic generation is one of the primary techniques that is used to design and analyze the performance of network security systems. However, due to the diversity of IoT networks in terms of devices, applications and protocols, the traditional network traffic generator tools are unable to generate the IoT specific protocols traffic. Hence, the traditional traffic generator tools cannot be used for designing and testing the performance of IoT-specific security solutions. In order to design an IoT-based traffic generation framework, two main challenges include IoT device modelling and generating the IoT normal and attack traffic simultaneously. Therefore, in this work, we propose an open-source framework for IoT traffic generation which supports the two widely used IoT application layer protocols, i.e., MQTT and CoAP. The proposed framework allows a user to create an IoT use case, add customized IoT devices into it and generate normal and malicious IoT traffic over a realtime network. Furthermore, we set up a realtime IoT smart home use case to manifest the applicability of the proposed framework for developing the security solutions for IoT smart home by emulating the real world IoT experimental devices. The results demonstrate that the proposed framework can be effectively used to develop better security solutions for IoT networks without physically deploying the real-time use case.

"On data-driven curation, learning, and analysis for inferring evolving Internetof-Things (IoT) botnets in the wild,"

The insecurity of the Internet-of-Things (IoT) paradigm continues to wreak havoc in consumer and critical infrastructures. The highly heterogeneous nature of IoT devices and their widespread deployments has led to the rise of several key security and measurement-based challenges, significantly crippling the process of collecting, analyzing and correlating IoT-centric data. To this end, this paper explores macroscopic, passive empirical data to shed light on this evolving threat phenomena. The proposed work aims infer to classify and Internet-scale compromised IoT devices by solely observing one-way network traffic, while also uncovering, reporting and thoroughly analyzing ``in the wild" IoT botnets. To prepare a relevant dataset, a novel probabilistic model is developed to cleanse unrelated traffic by removing noise samples misconfigured network traffic). (i.e., Subsequently, several shallow and deep learning models are evaluated in an effort to train an effective multi-window convolutional neural network. By leveraging active and passing measurements when generating the training dataset, the neural network aims to accurately identify compromised IoT devices. Consequently, to infer orchestrated and unsolicited activities have generated by wellthat been coordinated IoT botnets. hierarchical agglomerative clustering is employed by scrutinizing a set of innovative and efficient network feature sets. Analyzing 3.6 TB of recently captured darknet traffic revealed a momentous 440,000 compromised IoT devices and generated evidence-based artifacts related to 350 IoT botnets. Moreover, by conducting thorough analysis of such inferred campaigns, we reveal their scanning behaviors, packet inter-arrival times, employed rates and geo-distributions. Although several campaigns exhibit significant differences in these aspects, some are more distinguishable; by being limited to specific geo-locations or by executing scans on random ports besides their core targets. While many of the inferred botnets belong to previously documented campaigns such as Hide and Seek, Hajime and Fbot, newly

discovered events portray the evolving nature of such IoT threat phenomena by demonstrating growing cryptojacking capabilities or by targeting industrial control services. To motivate empirical (and operational) IoT cyber security initiatives as well as aid in reproducibility of the obtained results, we make the source codes of all the developed methods and techniques available to the research community at large.

"IoT DoS and DDoS attack detection using ResNet,"

The network attacks are increasing both in frequency and intensity with the rapid growth of internet of things (IoT) devices. Recently, denial of service (DoS) and distributed denial of service (DDoS) attacks are reported as the most frequent attacks in IoT networks. The traditional security solutions like firewalls, intrusion detection systems, etc., are unable to detect the complex DoS and DDoS attacks since most of them filter the normal and attack traffic based upon the static predefined rules. However, these solutions can become reliable and effective when integrated with artificial intelligence (AI) based techniques. During the last few years, deep learning models especially convolutional neural networks achieved high significance due to their outstanding performance in the image processing field. The potential of these convolutional neural network (CNN) models can be used to efficiently detect the complex DoS and DDoS by converting the network traffic dataset into images. Therefore, in this work, we proposed a methodology to convert the network traffic data into image form and trained a state-of-the-art CNN model, i.e., ResNet over the converted data.

The proposed methodology accomplished 99.99% accuracy for detecting the DoS and DDoS in case of binary classification. Furthermore, the proposed methodology achieved 87% average precision for recognizing eleven types of DoS and DDoS attack patterns which is 9% higher as compared to the state-of-the-art.

3. EXISTING SYSTEM

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 kmeans, 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 learningbased 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.

4. 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 [23] models over three different datasets and compared their performance with the proposed two-fold approach for detecting and preventing IoT botnet attacks.

Advantages

- The system proposed a novel twofold 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.

5. SYSTEM ARCHITECTURE Architecture Diagram



6. IMPLEMENTATION Modules 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.

7. RESULTS























8. CONCLUSION AND FUTURE ENHANCEMENT

In this study, we suggested a dual machine learning strategy to stop and identify IOT botnet assaults. We developed a cuttingedge deep learning model, ResNet-18, for scanning attack detection in the first fold and called it the ResNetScan-1 model. In the second fold, if the scanning detection model is unable to stop a botnet assault, we trained a second ResNet-18 model (called the ResNetDDoS-1 model) to identify the DDOS attack. We conducted a few experiments to verify the effectiveness of ResNetScan-1 the suggested and ResNetDDoS-1 models. We used scan and DDOS traffic samples from three publicly accessible datasets to train the ResNet-18 model, and we saved the resulting Res Net Scan and Res Net DDOS models. The resulting Res Net Scan and Res Net DDOS models were then evaluated using the test set of additional datasets that were not used for training. According to the experimental results, when evaluated on datasets that were not used for training, the performance of all Res Net Scan and Res Net DDOS modelsaside from the suggested ResNetScan-1 and ResNetDDoS-1 models-significantly decreased.

Additionally, the testing findings demonstrated that the suggested ResNetScan-1 and ResNetDDoS-1 models continued to function and performed better than any other models in detecting DDOS and scan assaults. As a result, the suggested dual strategy is effective and reliable for stopping and identifying IOT botnet assaults with a wide range of attack patterns.

Only thirty-three scanning kinds and sixty DDOS attack types are covered in the present effort. In order to properly train the suggested framework for more effective prevention and detection of IOT botnet and DDOS assaults, we want to cover more scanning and DDOS attack methodologies in the future. Additionally, we may use the suggested two-pronged strategy in an IDS to examine its efficacy on real-time network data.

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 AttackEver Recorded. Accessed:May 3, 2021.[Online]. Available:

https://github.blog/2018-03-01-

ddosincident-report/

[11] AWS Said it Mitigated a 2.3 Tbps DDoS Attack, Largest Ever. Accessed: May 3, 2021. [Online]. Available: https://www.zdnet.com/article/awssaid- itmitigated-a-2-3-tbps-ddos-attack-thelargest-ever/

[12] Shodan. Accessed: May 3, 2021.

[Online]. Available: https://www.shodan.io/

[13] Censys. Accessed: May 3, 2021.[Online]. Available: https://censys.io/

[14] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, ``DDoS in the IoT:Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80 84, 2017.

[15] R. Hallman, J. Bryan, G. Palavicini, J. Divita, and J. Romero-Mariona, 'IoDDoS_The internet of distributed denial of sevice attacks," in *Proc.* 2nd Int. Conf. Internet Things, Big Data Secur. Setúbal, Portugal: SciTePress, 2017, pp. 47 58.