THINK INDIA (Quarterly Journal)

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019

Detection And Prevention Of Botnet Attacks In Iot Devices Using Reverse Proxy

[1]Mrs.R.Deepika, [2]Mr.Muniyaraj, [3]K.Karan, [4]J.Dev Anand, [5]S.Ajith Kumar
[1]Assistant professor, [2]Trainer, [3][4][5]Dept of Computer science and Engineering
[2]Pyroferus technologies, [1][3][4][5]Sri SaiRam engineering college
[1]deepika.cse@sairam.edu.in, [2]muniyaraj517@gmail.com

***Abstract*--More than 20 billion devices are connected to internet today which also increases vulnerability,due to the inevitable growth of internet. An IOT BOTNET is a group of hacked computers, smart appliances and Internet-connected devices that have been co-opted for illicit purpose.Botnets use MALWARES and DDOS methods to attack and reduce the performance of the vulnerable IoT devices.Our proposed system includes the detection of the attacks and preventing them from further damage. We proposedreverse proxy algorithm and anomaly detection to detect a DDOS attack. An threshold value is fixed.If flooding of IP packets exceeds the threshold, then the system is set to showcase an anomaly(unusual deviation) and block the particular IP using reverse proxy algorithm,from the botnets trying to take the control of unaffectedIOT devices. This would help the users to protect their IOT devices from the unknown botnet attackers and also protect the information being hacked.**

Keywords :Iot devices, Botnet, DDOS Attack, Vulnerable ,Anomaly, Ip address,Reverse proxy.

I. INTRODUCTION

The term botnet is derived from the words robot and network. A bot, sometimes referred to as a zombie, is an individual device connected to an Internet Protocol (IP) network, typically the internet. Historically, this meant desktop computers, laptops, printers, home router, etc. were vulnerable to becoming a bot.Today however, as the Internet of Things (IoT) evolves our household devices are increasingly more often connected to the Internet. This means that the candidate list of potential botnet devices has greatly expanded. Included now are web cams, baby monitoring controls, and even toasters. After a device becomes infected with botnet malware, it can be leveraged via its network connectivity to conduct a slew of unauthorized and malicious activities.Botnet herders are actors who control bots remotely. They setup and deploy command and control (C&C) servers, and these serve as the interface to the bots. Coded within the botnet malware are C&C check-in IP addresses, schedules, and instructions. Their purpose is to establish communications channels from the herders to the bots. For example, IRC channels are frequently employed for this purpose. After communications are setup, the compromised hosts are often times further organized and issued updated instructions. They have now become an organized group of hosts under centralized control. Assuming that botnet attacks are unlikely to disappear,the fundamental question we address is as follows. Givena large number of heterogeneous IoT devices connectedto an organizational network, can we devise a centralized,automated method that is highly effective and accurate indetecting compromised IoT devices which have been added to a botnet and have been used to launch attacks? For detecting attacks launched from IoT bots we propose a network-based approach, which uses deep learningtechniques to perform anomaly detection. Specifically, weextract statistical features which capture behavioural snapshotsof benign IoT traffic, and store in database(one for each device) to learn the IoT's normal behaviours. Thedeep reverse proxy attempts to compress snapshots. Whenan algorithm fails to reconstruct a snapshot, then it is astrong indication that the observed behaviour is anomalous(i.e., the IoT device has been compromised and is exhibitingan unknown behaviours). An advantage of using deepautoencoders, is their ability to learn complex patterns, e.g.,of various device functionalities. This results in an anomalydetector with hardly any false alarms. We empirically showthat the autoencoders' false alarm rate is considerably lowerthan three other algorithms commonly used for anomalydetection.

A. Distributed Denial of Service (DDoS) Attack :

DDoS is a coordinated attack, generated by using many compromised hosts. An attacker initially identifies the vulner-abilities in a network to install malware programs on multiple machines to bring them under his control. Then the attacker uses these compromised hosts to send attack packets to the victim without the knowledge of the compromised hosts.

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019

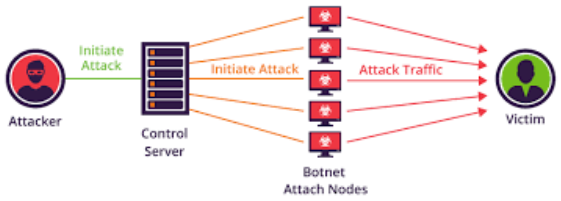


Fig.1.1 DDOS ATTACK

DDOS ATTACK IN IOT DEVICES

De-pending on the attack packet intensity and the number of hosts used for attacking, commensurate damage occurs in the victim network. If the number of compromised hosts is very large, it may disrupt a network or a Web server in a very short period of time. Some examples of DDoS attacks include smurf, fraggle and SYN flooding. The aim of a DDoS attacker is to disrupt a network so that it cannot provide any services to legitimate users (Fig. 1). To launch an attack, an attacker generally follows four basic steps. information gathering to scan a network to find vulnerable hosts to use them later to launch an attack, (ii) compromising the hosts to install malware or malicious programs in the com-promised hosts (called zombies) so that they can be controlled only by the attacker, (iii) launching the attack to command the zombies to send attack packets with specified intensities to the victim, and (iv) cleaning up to remove all records or history files from memory.A DDoS attacker often aims to attack one or more of the following targets (i) routers, (ii) links, (iii) firewalls and other defense systems, (iv) victim's computer and network infras-tructure, (v) victim's OS, (vi) current communications and victim's applications. Some prominent factors of DDoS attacks are (i) existence of high interdependencies in Internet security, (ii) limited availability of Internet resources, (iii) many conspiring against a few, (iv) not collecting intelligence and resources, (v) use of straightforward and simple routing prin-ciples, (vi) mismatch between core and edge network design issues and speed, (vii) laxity in network management, and (viii) the common practice of sharing of resources.

B. DDoS Attack Using Botnet

A botnet is a collection of many malware infected machines called zombies that are controlled by a malicious entity called the bot master. A bot master remotely controls the zombies and instructs them to perform malicious activities through com-mands. The way the bots are controlled depends on the architecture of botnet command and control mechanisms, which may be IRC, HTTP, DNS or P2P-based. These botnets are used to commit cyber crimes such as sending spam mail, launching denial-of-service attacks or stealing personal data such as mail accounts or bank credentials. It is common knowledge that approximately 80% of all email traffic is spam and most such messages are sent through botnets.

C.Types of DDoS Attacks:

Direct attack:In a direct attack, a large number of attack packets are sent to the victim machine directly. In this attack, the attacker spoofs the source IP address so that the response is misdirected and goes elsewhere

Reflector attack:In case of an reflector attack, many innocent intermediate nodes known as reflectors (Botnets or Zombies) are used to generate an attack. An attacker sends packets that require responses to the reflectors with the packets' inscribed source address set to the victim's address. The attack packets can be constructed using TCP, UDP, ICMP or IGMP protocols.

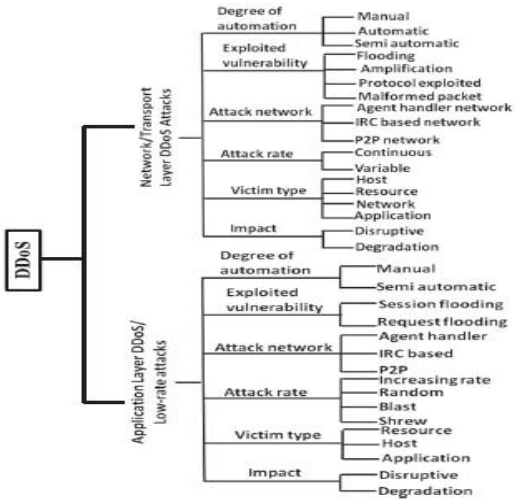


Fig 2.1 Taxonomy of DDoS attacks

Ddos attack on application layer:

An application layer distributed denial of service attack is usually initiated by hiring machines, bots, or taking control of remote systems. These components

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019

are used to ping multiple fake requests to server making the services of an application or server unavailable to its intended users. Such an attack targets everything that can eat huge chunks of the bandwidth, processing speed, and memory to slow down or disrupt services.

HTTP Flood:In HTTP flood DDoS attack the attacker exploits seemingly legitimate HTTP GET or POST requests to attack a web server or application. HTTP floods do not use malformed packets, spoofing or reflection techniques, and require less bandwidth than other attacks to bring down the targeted site or server. The attack is most effective when it forces the server or application to allocate the maximum resources possible in response to each single request

Slowloris:Slowloris is a highly-targeted attack, enabling one web server to take down another server, without affecting other services or ports on the target network. Slowloris does this by holding as many connections to the target web server open for as long as possible. It accomplishes this by creating connections to the target server, but sending only a partial request. Slowloris constantly sends more HTTP headers, but never completes a request. The targeted server keeps each of these false connections open. This eventually overflows the maximum concurrent connection pool, and leads to denial of additional connections from legitimate clients.

DDoS attack on network and transport layer:The main target of this type of attacks is to overwhelm the network infrastructure consisting of servers, routers and switches by sending a large volume of attack traffic. These attacks can be generated by exploiting protocol weaknesses. Network/Transport layer attacks can be further characterized according to degree of automation, exploited vulnerabilities, types of attack networks used, attacks rates generated, victim types and impacts of the attack.

3.1 SYN Flood:A SYN flood DDoS attack exploits a known weakness in the TCP connection sequence (the ―three-way handshake‖), wherein a SYN request to initiate a TCP connection with a host must be answered by a SYN-ACK response from that host, and then confirmed by an ACK response from the requester. In a SYN flood scenario, the requester sends multiple SYN requests, but either does not respond to the host's SYN-ACK response, or sends the SYN requests from a spoofed IP address. Either way, the host system continues to wait for acknowledgement for each of the requests, binding resources until no new connections can be made, and unlimitedly resulting in denial of services.

3.2 UDP Flood:This DDoS attack leverages the User Datagram Protocol (UDP), a sessionless networking protocol. This type of attack floods random ports on a remote host with numerous UDP packets, causing the host to repeatedly check for the application listening at that port, and (when no application is found) reply with an ICMP Destination Unreachable packet. This process saps host resources, and can ultimately lead to inaccessibility.

3.2ICMP (Ping) Flood:Similar in principle to the UDP flood attack, an ICMP flood overwhelms the target resource with ICMP Echo Request (ping) packets, generally sending packets as fast as possible without waiting for replies. This type of attack can consume both outgoing and incoming bandwidth, since the victim's servers will often attempt to respond with ICMP Echo Reply packets, resulting a significant overall system slowdown.

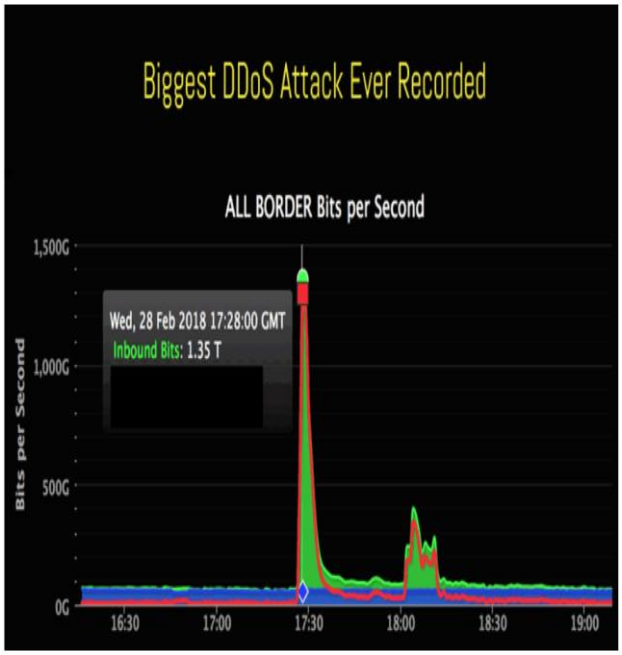


Fig. 3.1 Most Recorded DDOS ATTACK

GitHub knocked briefly offline by **biggest DDoS attack** ever. At its peak, inbound traffic reached a staggering 1.35 terabits per second (Tbps), outflanking the previously record-setting assault of

1 Tbps at French web hosting provider OVH in September 2016.

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019

D. Communication Mechanisms :Four types of communication mechanisms are typically used by most attackers [15]. (i) star topology, where the attacker gives more weight to a single centralized C&C resource com-ponent to communicate with all bot agents. This central com-ponent is responsible for issuing new instructions directly to each bot agent. (ii) ring-star topology, a logical extension of the star topology, uses multiple servers that are connected in a circular form to provide C&C instructions to bot agents. To manage the botnet, communications take place among multiple command systems, and if any of the individual severs fails or is detached permanently from the network, the remaining servers take up the responsibility of controlling the botnet. (iii) hierarchical topology follows the methods used in the compro-mise and subsequent propagation of the bot agents themselves. Bot agents have the ability to proxy new C&C instructions to previously propagated progeny agents. However, updated command instructions typically suffer latency issues, making it difficult for a botnet operator to use the botnet for real-time. The ensuing **DDoS attack** generated a flood of internet traffic that peaked at 1.35 Terabits per second, making it the **largest** on record. Fortunately, the software development site survived the disruption and was only down for few minutes

Detection Methods: When the Handshake Protocol is established with the target IoTdevices, attackers use Botnets to attack the victims device.Reverse proxy is used before the firewall to analyze the incoming IP requests.The request first passes through the reverse proxy, which send the IP's to the database.In the database the IP's are stored and updated depending on the frequency.When the particular IP exceeds the threshold, it is detected as anomaly.The particular IP's are stored separately in database and a separate text file is created, where list of IP's and execution time are stored.

(a) *Statistical*: Many statistical approaches have been used for detection of anomalies in a network. Such systems use statistical methods such as entropy, principal components anal-ysis, hidden Markov models, mutual information, correlation and covariance. Li *et. al.* [71] propose an entropy based DDoS attack detection method that calculates the distribution pattern of the attributes in network packet headers. Cumulative entropy is calculated to monitor network traffic behavior for a period of time instead of classifying the traffic as abnormal after initially detecting as abnormal in the first phase. In the second phase, an anomaly pattern is detected based on time instead of a threshold value set apriori. If abnormal behavior is continues for a certain period of time, only then the pattern is marked abnormal.

Feinstein *et. al.* [72] develop a method by computing entropy and frequency sorted distribution of packet attributes. In this method, entropy of the source addresses is computed for a packet window of size, say 1000, to determine the randomness or uniformity of the addresses. The amount of randomness is different for normal and attack conditions. In normal condi-tions, the entropy of the source addresses is less than in attack conditions. A low rate DDoS attack detection and traceback method using an information theoretic approach is proposed by Xiang *et al.* [49]. They calculate the difference between legitimate traffic and attack traffic using a generalized entropy metric and an information distance metric for detection of low rate DDoS attacks. Using experimental studies, they claim that the generalized entropy can detect an attack earlier than the traditional Shannon metric. The proposed information distance metric gives better result than Kullback-Leibler divergence approach.

Dynamic entropy can also be used to detect specific types of malicious traffic. A novel dynamic entropy-based model is pro-posed by Qi *et. al.* [73] to detect DoS attacks. This model uses netflow conversation correlation from different perspectives in a group of correlated events like request and reply. They compare dynamic and static entropy change rates in anomaly detection and find that the dynamic entropy method is more sensitive and more suitable for anomaly detection.

An information theory based DDoS attack detection algo-rithm is presented by Yu *et al.* [74] to classify attack traffic from the legitimate. During botnet attacks, the attacker uses con-trolled function(s), called zombies, to send malicious packets to the victim and hence the attack flow shows properties which are not followed by a legitimate flow in a short period of time. The method calculates distance between packet distribution behavior and suspicious network traffic flows and confirms DDoS attack flow if the distance is less than a predefined threshold; otherwise, the flow is marked legitimate.

COMPARISON OF EXISTING STATISTICAL METHODS

In the existing system,Dos and DDoS Attack Prevention Firewall Algorithm is developed for the

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019

purposes of protect on DoS and DDoS attack. Presently the main problem of computer networks are Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks which can block them. The methods mostly base on using firewall and IDS/IPS mechanisms to fight the attacks.Firewall System is developed for the purposes of protect on DoS and DDoS attack.In this project, a regular performance will help protect server and websites.DoS and DDoS attack able to block Web servers. Such attacks could be started from anywhere in the network.

CONSTRAINTS

They are susceptiable to certain type of TCP/IP protocol attacks.Packet filtering using firewall cannot authenticate information coming from a specific user**.**It cannot discriminate between good user and bad user.Therefore there may be a chance that it may not aware of Botnet attackers requests.For each firewall rule, fireman only examines all preceding rules but ignores all subsequent rules when performing anomaly analysis.

Jin*et al.* [75] propose a model using Multivariate Correlation Analysis (MCA) for detecting SYN flooding attacks. This method is very simple, but can effectively differentiate between normal and attack traffic in real-time. They use correlation anal-ysis on multiple features of normal network traffic and generate a normal profile. When testing network traffic, the method gen-erates a test profile using the same correlation analysis and if the test profile deviates from the normal profile beyond a predefined threshold value, the test profile is marked attack traffic. This method can also detect subtle attacks, which are difficult to differentiate from normal behavior. Experimental results show high detection accuracy and real-time effectiveness for DDoS attack detection.Yu*et al.* [76] propose a discrimination algorithm to distin-guish DDoS attacks from flash crowds. They use flow corre-lation coefficient as the similarity metric for suspicious flows. They observe that similarity among DDoS attack flows is higher than that in flash crowd flows in a community network. An ap-plication layer DDoS attack monitoring method is proposed by Xie*et al.* [77] using the concept of document popularity. They capture spatial-temporal patterns of a normal flash crowd using an access matrix, and apply principal components analysis, and use the independent components to extract a multidimensional access matrix. During attack detection, first they obtain the dynamics of the access matrix using a hidden semi-Markov model and then detect attacks. Xie*et al.* [78] introduce a new scheme that detects application-layer-based DDoS attacks in early stages. They use a hidden semi-Markov model to describe browser behavior during application layer attacks. The browser behavior of a Web user is related to two factors: the structure of a Website and the way the user accesses Web pages. A new on-line algorithm called the M-algorithm is proposed to detect anomalies, reducing memory requirement and improving computational efficiency.

II. PROPOSED METHOD:

Reverse proxy:

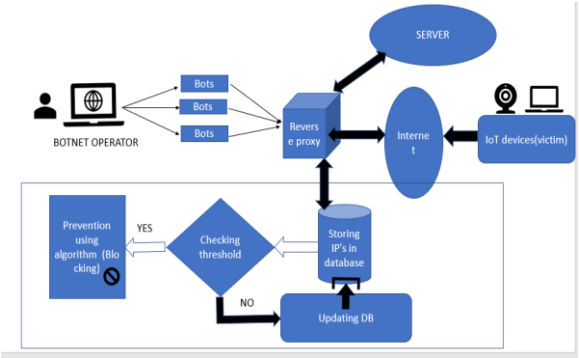
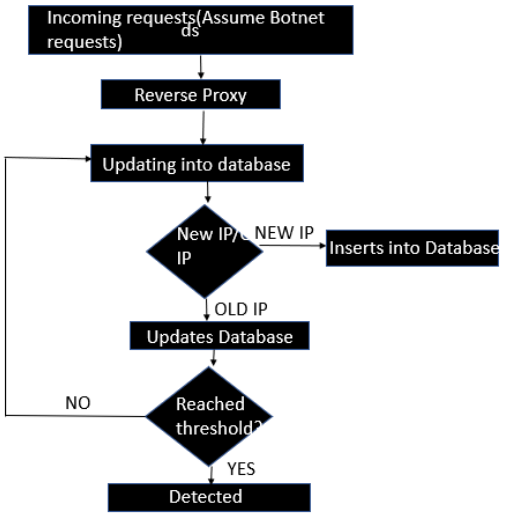


Fig 4.2.1 Reverse proxy method

A DDOS attack is carried out with various hacking tools such as metasploit framework, which results in attacking of target IoT device using Botnets(Reflectors).The IoT devices connected to a network(LAN,Wifi routers) is flooded with IP packets by Botnets, causing DDOS attack in turn reducing the system(IoTdevice) performance.We propose a meathod using Reverse proxy algorithm to detect anomaly and prevent the DDOS attack .MySQL is used to store the incoming IP addresses and timestamps. A threshold value is fixed, if the timestamps of a particular IP exceeds the threshold then the system is set to showcase an anomaly(unusual deviation).The detected IP is blocked and the ports to particular IP is closed and the attack is terminated.

ISSN: 0971-1260
Vol-22-Issue-2-April-June-2019



Using database to do packet filtering

ADVANTAGES:

- Avoid the expense of installing another web server. A reverse proxy server increases the capacity of existing servers.
- Serve more requests for static content and thus free up bandwidth to serve more dynamic content.
- Provide another layer of protection by hiding the internal IP address.
- Highly reliable over Dynamic IP requests.

III. CONCLUSION

As another extension to the current study, we also plan to evaluate transfer learning techniques by assessing the accuracy of models trained on specific devices when they are applied to identical devices, possibly when connected to other organizational networks. This can help (1) save time (e.g., organizations can deploy models previously learned elsewhere, without the need to collect data and train the models themselves), and (2) detect compromised IoT devices which have been contaminated prior to connecting to the organizational network, such that the organization has no benign data of them for model training.

REFERENCE:

[1] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[2] N. Hoque, D. Bhattacharyya, and J. Kalita, "MIFS-ND: A mutual information-based feature selection method," *Expert Syst. Appl.*, vol. 41, no. 14, pp. 6371–6385, Oct. 2014.

[3] D. K. Bhattacharyya and J. K. Kalita, *Network Anomaly Detection: A Machine Learning Perspective*. Boca Raton, FL, USA: CRC Press,2013.M. Feily, A. Shahrestani, and S. Ramadass, "A survey of botnet and botnet detection," in *Proc. IEEE 3rd Int. Conf. Emerging Security Inf.,Syst. Technol. SECURWARE*, 2009, pp. 268–273.

[4] M. Bailey, E. Cooke, F. Jahanian, Y. Xu, and M. Karir, "A survey of botnet technology and defenses," in *Proc. IEEE CATCH*, 2009, pp. 299–304

[5] R. A. Rodríguez-Gómez, G. Maciá-Fernóndez, and P. García-Teodoro, "Survey and taxonomy of botnet research through life-cycle," *ACMCSUR*, vol. 45, no. 4, p. 45, Aug. 2013.

[6] S. S. Silva, R. M. Silva, R. C. Pinto, and R. M. Salles, "Botnets: A survey," *Comput. Netw.*, vol. 57, no. 2, pp. 378–403, Feb. 2013.

[7] V. Igure and R. Williams, "Taxonomies of attacks and vulnerabilities in computer systems," *IEEE Commun. Surveys Tuts.*, vol. 10, no. 1, pp. 6–19, 1st Quart. 2008.

[8] J. Yuan and K. Mills, "Monitoring the macroscopic effect of DDoS flooding attacks," *IEEE Trans. Depend. Secure Comput.*, vol. 2,