A Review Of Recent Detection Methods For HTTP Ddos Attacks

Priteshkumar Prajapati,. Nidhi Patel, Dr. Parth Shah

Abstract: There has been a tremendous increase in dependency on web technologies alongside which its disruption attempts have also increased. Web servers that run on Hypertext Transfer Protocol (HTTP) are exposed to Denial of Service (DoS) attacks. Distributed denial of service attack is among the most dangerous internet attacks with the ability to overwhelm a web server. The content provides review of recent detection methods in recognizing detection attacks at the application layer. Trojan, Botnets, Advanced Persistent threat, Distributed Denial of Service attacks are increasing the security risks and are posing a great threat to critical information. Also, applications of Machine learning and Deep learning helps in the exclusion of false positives.

Index Terms: Artificial Intelligence, AnonymousDoser tool, Bots, Botnets, Detection metrics, DDoS, DDoS attack, Deep Learning, flooding attacks, HTTP, IDPS, LOIC tool, Machine Learning, OpenStack, Propaganda Bots, SDN, smart controller, Slowloris tool.

1. INTRODUCTION

DDoS attacks at the application layer is complex to detect, because such attacks may be able to mimic a legal request with the purpose of using the system resources. A web server uses HTTP and HTTPS protocols to process the request from the users. These protocols are widely used in commercial to operate business routines among banks, credit card payment gateways, government web servers, online shopping servers, social media servers and broadcasting servers to name a few. The consequence of the DDoS attack against a web server leads to monetary loss and loss of trust among people. HTTP protocol is designed to have request and response so as to allow communication to take place between client and web server. This paper presents recent detection methods of DDoS attack at the application layer and highlights several recommendations for future research. It explains types of DDoS attacks at the application layer, strategies performed by the attack, presents techniques of prevention against DDoS.

2 LITERATURE SURVEY

Paper [1] is about DDoS attack at application layer and various detection methods for HTTP DDoS attack. HTTP and HTTPS are the protocols used to process a request from the user. HTTP is responsible for request and response, allows communication between client and server. Detection of DDoS attacks at application layer is very difficult as they are able to mimic a legal request with the purpose of using system resources. During these attacks it is difficult to differentiate normal packets from the request packets as both appear same. Consequences of this attack may exhaust resources, such as network bandwidth, CPU processing and memory. Lower bandwidth is the need of DDoS attacks at application layer as it prevents legitimate users from surfing web server. DDoS detection at application layer is difficult due to three factors : 1)obscurity, as HTTP uses Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) which makes separation of illegitimate users from legitimate users complex; 2)efficiency, fewer connections are enough to initiate DDoS attack; 3) lethality, web servers are immediately overwhelmed due to attack. DDoS attacks mainly send in large amount of GET requests into the server, detection becomes even more difficult when flash crowds come into play. Increase in number of legitimate HTTP GET requests by a web server due to some events such as result announcements, elections is termed as flash crowd. Botnets are used to generate plenty of traffic against a server. It is created by using the HTTP protocol that dismisses command-and-control server. These botnets are of two types: Botnets that are controlled and configured by PHP script and web-based botnets that operate to report website statistics. Attacks at the application are based on following categories: 1) Session Flooding Attack: High session request rates than valid users exhaust the resources of server. 2)Request Flooding Attack: Attacker initiates vast number of request in one session. Usually these requests are larger than the one's from the valid user. 3)Asymmetric Attack: Intruders seek for HTTP session with high workload of requests such as by downloading huge files or plenty of running queries from database server. 4) Slow Request/Response Attack: High workload of requests is sent by the attacker to initiate the attack in form of session. DDoS attack defence includes four phases: Prevention, Monitoring, Detection and Mitigation. Various detection techniques are proposed and applied in past five years. 1) HADEC: It is a framework to detect high live-rate of attacks including TCP-SYN. HTTP GET, UDP and ICMP [10]. 2)D-FACE: It is used to detect four traffic types: Legitimate user, Low-rate, High-rate and flash event traffic [11]. 3)Machine-learning based prevention method to distinguish botnets from legitimate users in various traffic detections [12]. 4) Machine-learning matrix with a bio-inspired bat algorithm to allow early and fast detection of attack [13]. 5)Statistical approach with covariance matrix for cloud based detection [14]. 6) A Multilayer perceptron with genetic algorithm [15]. 7) FPGA based real-time DDoS attack detection [16]. 8) DDoS attack detection based on Entropy using Artificial neural network [17] . Paper [2] is about Improving capacity of Cyber Security safeguarding. Details, Studies and solutions are all based on China and presented by Chinese Academy of Cyberspace Studies. Cyberspace is accounted as the fifth space. Leading countries see cybersecurity as a measure of competition between them. China aims to ensure development through security and promote security through development. Cybersecurity issues

- *Mr. Priteshkumar Prajapati, K.D. Patel Dept. of IT, CSPIT, CHARUSAT, Changa, Gujarat, India. E-mail: pritesh.pnp.007@gmail.com*
- *Ms. Nidhi Patel, K.D. Patel Dept. of IT, CSPIT, CHARUSAT, Changa, Gujarat, India. E-mail: nidhip0520@gmail.com*
- *Dr. Parth Shah from K.D. Patel Dept. of IT, CSPIT, CHARUSAT, Changa, Gujarat, India. E-mail: parthshah.ce@charusat.ac.in*

are prominent in many areas. Website security problems occur frequently which includes website distortion, phishing and website backdoor installation in which intruder's upload remote control page into website's directory and control its server. Security risks for critical information infrastructure have been increasing with an alarming rate. Data breach is frequent which leaks sensitive information of millions of users available. China used to face a weak aspect that used to affect the operation of Internet in China which was caused by Domain Name System(DNS). Country faced a large-scale DNS breakdown during 2014 which disabled all web portals. By the 2016 they were able to tackle and upgrade China's DNS security. Mobile Internet Security, IoT smart devices and cloud-based services are most vulnerable to cyberattacks. House appliances, smart wearables and routers are in common use by many people in their daily life. Attackers get control of these devices and make them work as large-scale botnets and steal user's data. Alibaba cloud suffered a DDoS attack which was biggest of its kind back in 2015 which lasted for 14h traffic's peak value reached 453.8 Gbps. There are different means of Cyberattacks. China became a major victim of APT attack. In these kind of attacks invader mainly aims in stealing core data of the user. Hacker hides himself and steals data regularly in an organized way. After facing such problems, The Cyber Security System is being built and improved fast. The regulation system concerning Cyber Security is being improved. Laws show improvement, policies have been launched, review system is being established, multi-level protection of information security is carried out on a continuous basis. Cyber Security Standardization is speeding up, even the educational system has improved and helped in improvement. Paper [3] is about The Benefits of Artificial Intelligence in Cybersecurity. The ways of attacks have improved a lot and current defense methodologies are incompetent to solve them. The application of Artificial Intelligence(AI) increases the detection rate of Intrusion Detection and Prevention Systems(IDPS) and Machine Learning techniques help us find botnets and their sources. AI has multiple branches such as Machine Learning(ML) and Deep Learning(DL). ML helped researchers in enhancing current security techniques to detect phishing emails and malware detection. DL has a great capability of handling humongous datasets. It's high processing power helps in botnet detection and intrusion detection which it learns from experience by examining data. Botnets which are used to launch Distributed Denial of Service(DDoS) attacks and Intrusion Detection & Prevention Systems(IDPS) that generate huge number of false alarms are the constituents of existing problems which work as distraction for experts from finding real threats. Bots work in a network of computers known as botnets. Connection to these computers are responsible for malware infection. C&C server is used by the botmaster to provide commands to bots. AI has a capability of detecting these bots inside the network which in turn stops the infection from spreading to other devices, data leakage and DDoS attacks can be prevented. Network and system administrators use IDPS to detect intrusions. This technology prevents as well as detects intrusions. Proper configuration of tools can provide a high level of security. The only problem with IDPS systems is that they generate large number of false alarms, AI helps in detection of those false alerts and increase the threat detection rates. IDPS system comes into two further bifurcations. Intrusion Detection System(IDS) and Intrusion Prevention System(IPS). This technology of IDPS relies on two systems, Signature-based and Anomaly-based. In signature-based systems the previous intrusions or the known threats are stored in the database, the incoming packets are examined and the signatures are extracted from those packets and matched with the available database if there is a match then system considers it as an intrusion. The major drawback for this problem is that it only works for known threats. On the other hand, the anomaly-based systems analyses the patterns and behaviors. In ML based approach we have various types for implementation in IDPS systems which can be used to reduce false positives out of which we will consider two methods; Artificial Neural Network(ANN) and Genetic Algorithm(GA). Artificial Neural Networks(ANN) is one such method which uses processing nodes or neurons. They are capable of recognizing unprecise patterns. Usually nowadays attackers deploy attacks in such a way that they do not trigger alerts in system which makes detection tough. ANN has a capability of recognizing whether the connection is legitimate or not and provides alert for the same. Genetic Algorithm learns from previous experience of anomaly Behaviour. It is useful for detecting common threats. DL based approach consists of various layers. Data is provided at the input layer of the neural network which is further passed to hidden layers for implementation of algorithms. A Deep Belief Network(DBN) was proposed by researchers from CRRC Qingdao Sifang Co. for Intrusion detection [18]. Restricted Boltzmann Machine(RBM) is trained in the first step. It acts as a foundation for DBN and is a hidden layer. RBM uses large number of hidden and visual layers. It allows DBN to process more data. Second step is the backpropagation, it adds backpropagation neural network which receives data from previous step and monitors the network. This layer decides and attempts to identify false positives and discards them accordingly to provide accurate results. Initially DBN receives the data produced by IDS. It processes the data and estimates the number of layers needed. Training dataset is prepared to start the process and perform calculations needed to determine whether it is a false positive or a real threat. Paper [4] is about propaganda bots which pose to be fellow humans and attack the affected platform without being detected. This is certainly a new thing evolving the involvement of bots to spread propaganda. So the techniques to determine the same are also at its formative stages. This new field is trying to improve its efficiency but still is inconsistent as bots continue to become more sophisticated and mitigating them keeps on posing challenges in the path of enhancement. According to some research there are three distinct generations of social bots available: simple bots, convincing ‘sock puppets’ and advanced ‘sock puppets’. Also it was discovered that twitter was the most vulnerable social media when propaganda bots came into play which resulted in various fake tweets and new feeds during USA elections of 2016-17. Paper [5] is about monitoring DDoS attack by placing smart controller in Software Defined Network(SDN). SDN can be a great threat for distributed denial of service as it can directly send huge amount of traffic to the controller. Here they have presented a hypothetical situation of using a smart controller to monitor health status of controllers and provide proper services during DDoS attacks. SDN is preferred by many scientists and researchers because it is financially savvy, programmable, centralized. SDN architecture has three layers application layer, control layer and the infrastructure layer. In OpenFlow

[19] controller is like a brain for SDN. It manages and processes all the incoming packets and decides where to direct them [20]. DDoS attacker sends large amount of spoofed data packets to the controller and creates an unwanted fake traffic of new incoming packets and creates similar flow entries better harder stronger and faster. These spoofed entries take over the full control of flow tables in OpenFlow which can also lead to failure of SDN controller. DDoS attack propagation starts at infrastructure layer, attack packets are sent to data plane via switches, routers, wireless access points etc. Attack packets can spread to controller in control layer. During the final stage traffic propagates to the application layer. When SDN encounters DDoS attack, all switches start malfunctioning and can't serve well to the legitimate users. Whenever any controller malfunctions smart controller comes into play and manages the work of controller. During traffic when controller's get congested due to unwanted traffic they can make request to the smart controller to share with other controller or create rule. Smart controller's do not process the data; they just maintain the controller's. Upper bound part of the control plane of SDN is suitable location to locate the smart controller in architecture. Paper [6] is about Analysis of DDoS attacks using OpenStack Cloud Platform. Various tools are used to generate DDoS attacks. Here they have used three tools AnonymousDoser, LOIC and Slowloris. These tools are used to flood with HTTP and TCP/IP packets. AnonymousDoser sends large number of TCP SYN packets which overwhelm the server by making open connections due to which legitimate packets start dropping. Intensity of this attack is very large and it can bring down many servers in several seconds. Low Orbit Ion Cannon(LOIC) tool is a testing tool to attack web servers. It uses several flooding methods like TCP, UDP, ICMP to launch attacks against the targets. It can send large number of HTTP requests that can disrupt the services of the system. One of the major drawbacks of this tool is that it cannot spoof IP address of the operator or the handler. Slowloris tool is a tool which can exhaust device's web server even with minimum bandwidth. It creates connections by sending partial requests and tries to hold those connections open for longer time which in turn exhausts the web server's connection pool making it decline the legitimate requests to the server. DDoS analysis can be done using Wireshark, HTOP and Netstat. LOIC is the tool that can send large number of packets in just few seconds. Paper [7] is about practical approach to detect DDoS attack using a hybrid detection method. This paper is to test the previously available findings about entropy-based methods and custom-tailored methods. Custom tailored methods can only be used for the ones they are specially designed whereas anomaly-based methods can detect a wide range of anomalies and attack. The research done in this paper is only based on real network traffic. The proposed hybrid method is combination of both feature-based and volume-based detection along with high and low rate attacks which gets a comparison between proposed and previously known methods. Some of the most common DDoS attacks are ICMP flood, SYN flood, DNS amplification attacks and the earlier Smurf Attack and Fraggle Attack. A script file was developed to test the following attacks: ICMP flood attack with a high packet rate attack on the specified target. Using this approach, a real-life botnet is simulated. As in a real-life situation, the attack does not start with all attackers at the same time, but instead with attackers initiated after a random delay period. The experiment included two types of attacks: an ICMP flood attack, representing a high-rate attack and a TCP SYN flood attack, representing a low-rate attack. There are various methods a DoS detector can follow like based on simulation of traffic, real network dataset containing both attack traffic and baseline traffic or using Shannon entropy. Proposed method works with combination of packet rate and diversity, Exponential Moving Average (EMA). EMA is applied to both entities. One uses short hand period (FastEMA) and the other uses long period (SlowEMA). Paper [8] is about generating DDoS attack traffic using scapy framework. Here D-Scap bot is compared with previously available bots. These bots are made using scapy framework and coded in python language. Bots are responsible for infecting victim's device. They take in orders from C&C server which is in turn handled by botmaster. They are responsible for providing instructions to the bots. The source code for the bots is distributed into all devices via USB drives or flash drives. C&C server starts giving commands to the bots and is responsible for controlling all infected devices. Now the infected machines start different attacks on victim devices according to the instructions provided by botmaster through C&C server. C&C servers wait for the connections to establish and collect the IP and port addresses sent by the bots and sends the command to attack with type of attack, packet count and victim domain name as parameters. D-Scap bots have series of actions to perform. It starts by sending infected machine's IP and port address back to the botmaster. It follows every instruction provided and works for the same. Initially victim enters an infection stage where it encounters a bot for the first time, then after collecting the data bot sends data back to the botmaster via C&C server. This first time communication between the server and the bot is called the rallying stage. After this stage comes the attack performing stage where D-Scap bot generates ICMP or TCP-SYN or UDP flooding traffic. Paper [9] is about detection of Mirai-like bots in Large-scale networks through sub-sampled traffic analysis. Here they have presented a network-based network which can be used to detect IoT bots infected by Mirai-like malware. Bots scan the networks before getting into them. This scanning and propagation period last longer during which bots can be easily detected and stopped from participating in attacks. Traffic signatures and patterns of mirai malware are used to detect the presence of other such kind on IoT devices. Such devices are prone to attacks because they don't have firewalls installed in them. Even the network level firewalls are not meant to prohibit TELNET requests because they can be legitimate in most of the cases. Bot's delays and sampling frequencies can also be studied by studying simulations of previously known traces. IoT ecosystems lack full-fledged operating systems, has low power requirements therefore they are always network-based and never host-based. Machine learning is ineffective because of false positives. After a device gets infected from mirai it tries to initiate a connection with C&C server and open's socket connection. Bots scan the network by sending SYN packets to random IP addresses and waits for responses. As soon as it discovers open TELNET port it makes a socket connection and logs into device using default credentials. Scan are forwarded to loader and other bot binary is downloaded which enables other bots to connect to C&C server and scan the server. Even DDoS attacks are possible in Data Deduplication [21], [22], and in [23] Provable Data Possession Using Identity-Based Encryption.

3 CONCLUSION

DDoS attacks have been a really alarming issues since past many years and it continues to be so. There can be various ways of dealing with them or even cut down some possibilities of attacks. This can be done in many ways, such as by implementation of Machine Learning, Artificial Intelligence. Bots are the basis for most problems and spread of attacks to the systems. IoT devices are vulnerable to DDoS attacks and can easily be used to spy on any sort of data.

REFERENCES

[1] Jaafar, Ghafar A., Shahidan M. Abdullah, and Saifuladli Ismail. "Review of Recent Detection Methods for HTTP DDoS Attack." Journal of Computer Networks and Communications2019 (2019).

[2] Chinese Academy of Cyberspace Studies. "Improving Capacity of Cyber Security Safeguarding." China Internet Development Report 2017: Translated by Peng Ping (2019): 101-130.

[3] Calderon, Ricardo. "The Benefits of Artificial Intelligence in Cybersecurity." (2019).

[4] Williamson III, William, and James Scrofani. "Trends in Detection and Characterization of Propaganda Bots." In Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019.

[5] Haque, Muhammad Reazul, Saw C. Tan, Zulfadzli Yusoff, Ching K. Lee, and Rizaludin Kaspin. "DDoS Attack Monitoring using Smart Controller Placement in Software Defined Networking Architecture." In Computational Science and Technology, pp. 195-203. Springer, Singapore, 2019.

[6] Bhardwaj, Aanshi, Atul Sharma, Veenu Mangat, Krishan Kumar, and Renu Vig. "Experimental Analysis of DDoS Attacks on OpenStack Cloud Platform." In Proceedings of 2nd International Conference on Communication, Computing and Networking, pp. 3-13. Springer, Singapore, 2019.

[7] Bojović, P. D., I. Bašičević, S. Ocovaj, and M. Popović. "A practical approach to detection of distributed denial-of-service attacks using a hybrid detection method." Computers & Electrical Engineering 73 (2019): 84-96..

[8] Kumar, Guntupalli Manoj, and A. R. Vasudevan. "D-SCAP: DDoS Attack Traffic Generation Using Scapy Framework." In Advances in Big Data and Cloud Computing, pp. 207-213. Springer, Singapore, 2019.

[9] Kumar, Ayush, and Teng Joon Lim. "Early Detection of Mirai-Like IoT Bots in Large-Scale Networks through Sub-sampled Packet Traffic Analysis." In Future of Information and Communication Conference, pp. 847-867. Springer, Cham, 2019

[10] Hameed, Sufian, and Usman Ali. "HADEC: hadoop-based live DDoS detection framework." EURASIP Journal on Information Security 2018, no. 1 (2018): 11.

[11] Behal, Sunny, Krishan Kumar, and Monika Sachdeva. "D-FACE: An anomaly based distributed approach for early detection of DDoS attacks and flash events." Journal of Network and Computer Applications 111 (2018): 49-63.

[12] Singh, Karanpreet, Paramvir Singh, and Krishan Kumar. "User behavior analytics-based classification of application layer HTTP-GET flood attacks." Journal of Network and Computer Applications 112 (2018): 97-114.

[13] Sreeram, Indraneel, and Venkata Praveen Kumar Vuppala. "HTTP flood attack detection in application layer using machine learning metrics and bio inspired bat algorithm." Applied computing and informatics (2017).

[14] Aborujilah, Abdulaziz, and Shahrulniza Musa. "Cloud-based DDoS HTTP attack detection using covariance matrix approach." Journal of Computer Networks and Communications 2017 (2017).

[15] Singh, Khundrakpam Johnson, and Tanmay De. "MLP-GA based algorithm to detect application layer DDoS attack." Journal of information security and applications 36 (2017): 145-153.

[16] Hoque, Nazrul, Hirak Kashyap, and D. K. Bhattacharyya. "Real-time DDoS attack detection using FPGA." Computer Communications 110 (2017): 48-58.

[17] Johnson Singh, Khundrakpam, Khelchandra Thongam, and Tanmay De. "Entropy-based application layer DDoS attack detection using artificial neural networks." Entropy 18, no. 10 (2016): 350.

[18] Qu, Feng, Jitao Zhang, Zetian Shao, and Shuzhuang Qi. "An Intrusion Detection Model Based on Deep Belief Network." In Proceedings of the 2017 VI International Conference on Network, Communication and Computing, pp. 97-101. ACM, 2017.

[19] McKeown, Nick, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. "OpenFlow: enabling innovation in campus networks." ACM SIGCOMM Computer Communication Review 38, no. 2 (2008): 69-74.

[20] Mousavi, Seyed Mohammad, and Marc St-Hilaire. "Early detection of DDoS attacks against SDN controllers." In 2015 International Conference on Computing, Networking and Communications (ICNC), pp. 77-81. IEEE, 2015.

[21] Prajapati, Priteshkumar, and Parth Shah. "Efficient cross user data deduplication in remote data storage." In International Conference for Convergence for Technology-2014, pp. 1-5. IEEE, 2014.

[22] Prajapati, Priteshkumar, Parth Shah, Amit Ganatra, and Sandipkumar Patel. "Efficient Cross User Client Side Data Deduplication in Hadoop." JCP 12, no. 4 (2017): 362-370.

[23] Kadvani, Smit, Aditya Patel, Mansi Tilala, Priteshkumar Prajapati, and Parth Shah. "Provable Data Possession Using Identity-Based Encryption." In Information and Communication Technology for Intelligent Systems, pp. 87-94. Springer, Singapore, 2019.