Grenze International Journal of Engineering and Technology, Jan Issue

An Empirical Assessment of Botnets and Detection Methods

Shivanna D[1] and Pooja M R[2]
[1]Research scholar, Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India
Email: shivu.d@gmail.com
[2]Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India
Email: pooja.mr@vvce.ac.in

***Abstract*— a botnet is a group of devices which are controlled by the Attacker and used for fraud, scam and cyber-attacks. The attacker is also known as Bot-herder. The Bots are used to conduct many activities varies from stealing user credentials, sending spam, stealing data, ransom ware, fraudulently clicking on ads or distributed denial-of-service (DDoS) attacks. The traffic made up by bots is increasing day by day, now nearly a quarter of the whole traffic. Detecting them becomes an uphill task as the time passes. Bot developers are adopting new technologies at a rapid rate in an effort to avoid bot detection.**

While some bot problems are industry-specific, researchers found out that bot problems run across all industries. The top five industries which are targeted with bad bot traffic are the financial, education, IT & services industries, marketplaces, and government. This assessment paper briefs the different botnet detection methods that exist at present. The main aim of this paper is to give the insight knowledge of various bots and their implications on the present cyber world and also shedding some light on the current methods using to detect them...

***Index Terms*— bot, botnet, DDoS, DNS hijacking, Domain fluxing, Black-hole.**

I. INTRODUCTION

As the internet user base grows the network traffic is also increasing in the same manner, the problem they are facing is also increasing and diverse in nature, i.e. the security issues. The malware or malicious applications have acquired the vital position producing the security issues [1]. Even in the earlier days many security issues were present but nowadays as the application and usage of the internet grows day by day in different aspects such as email, web applications and VoIP and so on this provides the big platform for the extensive malware attacks which affect a huge number of systems or devices around the nook and corner of the world [19]. The Botnet term coined in 1993 for the first time through the introduction of first Botnet called Eggdrop [2]. The botnet is nothing but a compromised system or device which is controlled for malicious activities. Bots use distinct techniques to infect systems. They find easily accessible and unprotected machines to compromise and control them; and then forward details to the botmaster. The Bots keep hidden until they are aware by their botmaster to execute an attack or task [3].

A. Botnets Categories

Botnets are analysed using architectural design constructed by BotMaster. Based on their design, they have been

Grenze ID: 01.GIJET.10.1.306

classified into Centralized, Peer to Peer (Decentralized) [4] [5].

Centralized Botnets

In this structure, a central command and control oversees all bots. Each Bot is directly connected to a single of central command and control server. It requires a high bandwidth to operate this central point. A botmaster will communicate through sending and receiving messages to other bots in the network from a single C&C server [6].
HTTP based BOTS: Hypertext transfer protocol is a popular protocol in creating bots. It can easily conceal traffic of the Botnet with regular traffic. With the HTTP protocol, Botmaster is capable of circumvent the firewall and also hide malware traffic with original HTTP traffic. HTTP-based bots sometimes download instructions from the web-based C & C server that is more precise than the IRC-based botnet [6]. Http based botnet model has been used by some of the famous botnets such as Click Bot, Bobox, Rustock.
IRC based BOTS: The IRC (Internet Relay Chat) utilizes real-time internet text messages. IRC channels are used by botmasters to design and communicate with botnets. IRC botnets are normally used by botnet owners because of their straightforward design and great flexibility. This type of botnet is exceptionally good at designing and recycling bots. One drawback of IRC-based networks is their widespread use of the IRC (Internet Relay Chat) protocol for communication, which would be easily traced by analysing IRC traffic to be part of conventional or normal traffic.

Decentralized Botnets

The major flaw of a centralized (single unified) model is that having only one centralized server makes it easy to find or detect the botnet. So the bot developer created the decentralized botnet. In this technique the botnets aren't controlled by a single C&C server. Here the botmaster makes any bot behave or act like a botmaster. A one single botnet can contain a lot of bots because of the decentralized model, and also they are exceedingly hard to track down. P2P protocols are necessary for the decentralized model [6].
Peer to Peer botnets: P2P botnets are defined as having P2P-based C&C models, meaning they do not use a centralized server. Every bot member serves as both a client that accepts commands and a server that distributes commands. Because of this, P2P botnets are typically more resistant to defenses than conventional centralized botnets.

B. Attacks of Botnets

Email-Spamming attack

Today's major issue with the Internet has been spam emails which are delivered via botnets. Since it is simple and inexpensive to send a lot of emails, spam is expanding quickly. Combating spam is an interesting topic in computer security because it wastes user time and consumes resources like space and network bandwidth.

DDoS Attack

Distributed Denial-of-Service (DDoS) attack targets websites, web applications, networks, and infrastructure and make them inaccessible to authorized users by overloading the services and resulting in outages or crashes. DDoS attacks are categorized according to their attacks in network.

Application layer attack

Attacks at the application layer that overload the victim with a huge number of ostensibly (looks genuine) legitimate requests, measured in Requests Per Second (RPS).

Network-layer attack

Network-layer attacks are measured in Gbps (Gigabytes per second) or PPS (Packets per second) and aim to exhaust the target's upstream bandwidth until the network is saturated.

Brute-force approach

Brute-force attacks describe a hacker's trial-and-error approach to find encryption keys, login credentials, or the location of hidden web pages. Brute Force attacks refers using several attempts to 'force' their way into your private accounts [7].

IoT Botnets

An internet of things botnet, often referred as IoT Botnet is a group of IoT-connected devices, typically routers, which have been compromised by malware (particularly IoT botnet malware) and are now under the control of

malicious or hostile actors. IoT botnets are well-known for being employed in distributed denial-of-service (DDoS) assaults on target entities to obstruct their operations and services. The crucial role that routers play in networks creates more opportunity for criminal actors to exploit IoT botnets to launch more destructive attacks [20].

Phishing or social engineering

Phishing attacks use social engineering techniques to convince a victim to take a specific action that will give the perpetrator access to their personal information. A hacker sends emails to users that seem to be from a reliable source in an effort to convince them into disclosing private information like passwords and financial details [7].

Targeted intrusions

Small botnets made to infiltrate specific high-value organizational systems from which attackers can break in and spread farther throughout the network. Assailants particularly target organizations' most precious assets, including financial data, research and development, intellectual property, and consumer information, making these incursions very harmful for businesses [8].

Ad Fraud

Attackers may replicate genuine user activity maliciously through botnet compromised devices. A botnet, for instance, can be used to continue "click fraud," where systems or devices which are in connection with the botnet continually click on links or buttons in an advertisement. This attack has the capacity to seriously harm competitor's budgets, because the advertisers get paid when a user clicks on an ad (a payment model known as pay-per-click, or PPC). [9]

Crypto currency Mining

Some attackers establish bitcoin mining operations using botnets to increase their personal financial advantage. The "mining" method used to produce new coins in crypto currencies like Bitcoin requires a lot of processing power. Attackers can utilize a botnet to harness the processing power of the devices they control, creating new coins for themselves while the owners of the devices pay the price in higher electricity usage. [9]

C. Domain and fast fluxing techniques

Many botnets use the technology called domain fluxing, a technique used for keeping the botnet's in operation by constantly changing the botnet C&C server's domain name using Domain name generation algorithm(DGA). To further improve the concealment, the bots, communicate through a mechanism called Message Transmission Mechanism where an elliptic curve encryption method is used for communication messages or commands. [10]. Even DNS hijacking is also used for concealment from detection. DNS hijacking (DNS redirection) is an attack where DNS queries are incorrectly resolved, so that the users can be redirected to malicious websites. [10]
Fast Fluxing is a technology which constantly changes the association between domain names and IP addresses, where the IP addresses are tied to malicious sites. This technology is also used to break the domain names of some phishing websites and malicious websites. [10]

D. Is Botnet a biggest Security Threat? [10]

There are numbers of factors why Bot is becoming a biggest security threat.
Botnet development history is divided into 2 stages, like worm or virus at the first stage and later they transform themselves into Bots. It adds the advantage of rapid infection of virus and rapid transmission, so that it will be dangerous to control them onwards.
The C&C server and the hosts will communicate through the HTTP packets to make requests and connections. SO it needs only a few commands to the C&C servers to launch the different forms of attacks, which makes the Botnet more dangerous and stealthy. The Botnet controllers will make use of vulnerabilities like weak passwords to snatch the control of the host. Example for this is the Mirai botnet, which launched attacks on the IoT devices. The Mirai botnet will use the weak password vulnerability to hack the server's remote login portal telnet to take control of the host.

E. Botnets detection is a challenging exercise

Bot detection is difficult since they commonly use strategies like using proxy servers or frequently modify their IP address, making it difficult to determine their identity or location. Additionally, a lot of bot developers create many identities to disperse their bot's activities across various platforms in an attempt to avoid being discovered. Furthermore, it would be impossible to manually detect all bots. [11]

Another challenge in identify bots is that bots often imitate human behaviour. In an effort to imitate a real person, a Twitter bot may follow numerous individuals and retweet their content. Many bot developers even go so far as to construct fake accounts with images and biographies that give the impressions of the feeling that they are actual individuals [12].
The fact there is an ample amount of ambiguity surrounding what constitutes a bot and what constitutes a legal application of automation technologies makes the task of detecting bots much more challenging. For instance, a lot of software enables users to automate their social media activities.
It is getting more and more difficult to distinguish between bots and actual users. Bots rely on browsers that resemble human browsers in some ways. For organizations, bot detection is extremely difficult due to their resemblance to human behaviour. [21]

II. Methodologies For Botnet Detection

There are several methods which are applied and tested to detect the Botnet.

A. Signature based Detection

In this method it uses the signature of the currently existing botnet signatures to find them. This method provides number of benefits, which includes a very low rate of false alarms, quick detection, ease of implementation, and improved information related to the sort of detected assault. Only well-known botnets can be found using signature-based methods of detection. But if the system behaves differently, then this method doesn't work, to overcome this anomaly based detection method came into existence.

B. Anomaly Based Detection

The goal of the anomaly-based detection approach is to identify botnets by considering different network traffic anomalies, such as high traffic volume, high network latency, traffic on unusual ports and unusual system behavior that would be an indication that malicious bots are present in the network.
Anomaly-based detection methods are again further divided into host-based detection and network-based detection

C. Host-based detection

This detection method observes and examines a computer system's internal operations rather than the network activity on its ports. In this method each host is monitored carefully to find any suspicious activity like processing overhead, and access to suspicious files. The Host Intrusion Detection Systems (IDS) will notify the user or administrator if any suspicious activity is found. It takes a snapshot of the present state of the system files and compares it to the last snapshot. An alter signal is issued to the administrator to look into whether the crucial system files were altered or destroyed.

D. Network-based detection

This detection method detects Botnets by monitoring network traffic. Network-based detection methods can be categorized into two categories:
Active network monitoring: Capability of injecting sample test packets into the network, servers, or application to test how the network responds. This intern results in the increase of network traffic. Whether a human or bot is controlling that session can be determined by the injected packets. This method works in a cause-effect correlation because for a large portion of Botnet command-and-control channels, a command-and-control interaction has a deterministic command response pattern. This method shows effectiveness on IRC-based Botnet detection in real world.
Passive network monitoring: Monitor the data traffic in network and check for any suspect communications that may be given by bots or C&C servers. It doesn't increase the traffic for inspection at network.

III. Review Of Related Works

S. Haq and Y. Singh [13], has proposed a model which uses ML (machine learning) algorithms, have used clustering, classification and the hybrid approach that suggests the results and performance of classification & clustering on one side and hybrid approach at another side. They mainly proposed ideas based on unusual or unknown traffic and human interaction of the botnet.
They proposed a model called WEKA, because the botnet's traffic differs from regular traffic, it is necessary to first extract features from a large amount of network data before using a hybrid clustering and classification structure for precision and accuracy. The dataset in this paper is first focused on having a unique set of features.

Using the free and open-source WEKA tool, the features are pre-processed, trained, tested, and cross-validation is developed.

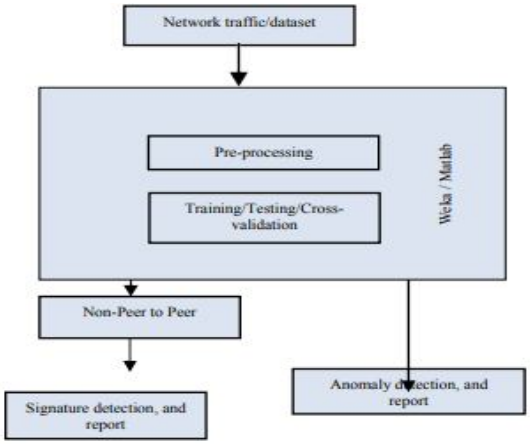

Figure 1. WEKA Tools working

They implemented the detection system based on random data-partitioning policy and also quoted the way of selecting data random data-partitioning policy, but did not produce favorable outcome and proposed an efficient data-partitioning policy as a future scope. Using efficient data-partitioning policy the data which will be used as input to the approach will produce the more accurate results. Hence the efficient data-partitioning policy can be considered as a future work.

Moorthy and Nathiya [16], proposed a method of detection of botnets using artificial intelligence. Algorithms used in this method are Support Vector Machine (SVM), decision tree and multi-layer perceptron, at first data set is trained and tested. After the training and testing, the decision tree model is used to find and the model has good accuracy and performance metrics around 92%. The active monitoring is done by injecting the packets into the network, with the support of the injected packets whether the system's session is managed by the botnet will be decided. Passive monitoring method without injecting any packets will be carried out for botnet detection.

Parameters of Dataset considered are

Start time: time when the transmission of packet transferred started.

Duration: time duration the source and destination machine were connected together through the network Protocol

i. *Source address*: IP address of the source machine
ii. *Scr Bytes*: No of bytes carried by source machine
iii. *Sport*: Source port
iv. *Dport*: Destination port and so on.

The data is processed using the Artificial intelligence. Machine learning and deep learning models are two components of artificial intelligence that cross disciplinary boundaries. Decision tree, SVM, KNN, XG boost, random forest, and extra tree classifier are among the ML and DL models that are used to build the detection system models for this study. The multi-layer perceptron model is employed in deep learning models. Decision tree has the highest accuracy of any of the classification models, at 0.9221. Therefore, it is believed that this model should be used in the detection algorithm.

But the problem with this is it takes a lot of time in collecting packets and classifying them through active monitoring and also lags in finding the packets with active monitoring.

Pradeepthi and kannan [17] proposed a model using neuro-fuzzy to get more accuracy and to bring down the false positives. The model is implemented using eucalyptus cloud having many number of virtual machines.

But the system was not tested or experimented for any other real cloud environments. The working model has been assessed and analyzed only in a Simulation environment and evaluated using the botnet simulation tools.

Javier, Victor, Eduardo and Enrique [18] Approached to detect the botnet using ML (machine learning) on large bandwidth networks. Analysis of network traffic is done using TCP and UDP based communications between machines with bandwidths up to 10Gbps. Time window of one second is used to examine the different features (Source IP address, Destination IP address, Source port & Destination port and so on) by different models. Detecting the botnet is based on knowledge database i.e. already existing botnets would be detected, if you want to detect new botnets the knowledge database would be detected frequently. In this model the botnet identification

is based on the packets dropped when the network experience saturation. In normal traffic it is unavoidable that certain packets are dropped by the routers or other network devices. Selected 10% of packets being dropped as threshold if the packet drop rate exceeds the administrators should be alerted. The objective is to identify botnets in real-time while accepting a reasonable accuracy loss as a trade-off. The methodology used here is F1 score on saturated networks score. The method is made up of a first module that separates the flows, or each independent communication between two devices, inside each time frame of one second that the incoming traffic is divided into. These flows can be identified by the two IP addresses and the two ports that are in use. Next, the module extracts four features for each 1s flow:

i. The source port number.
ii. The destination port number.
iii. The number of packets inside the time window.
iv. The total bytes transmitted in that second.

F1 score achieved by our proposal using period of time (I) one second or (ii) 180 seconds will be compared for different bots. Detection results were similar for most botnets considered, whether 1 or 180 seconds were used.

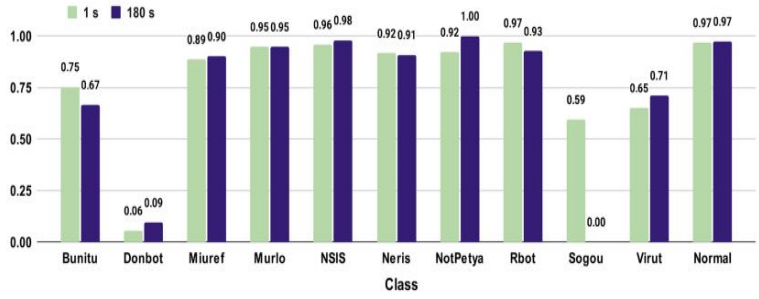

Figure 2. The F1 score achieved by different Botnet detection tools for window size 1sec and 180 seconds.

The features used are the source port and destination port, which do not change with the window size, the total number of packets and the total number of bytes transmitted inside the window will change. Since botnets are programs that try to be as small as possible to avoid detection, this suggests that the behavior of these botnets is consistent over time.

The experiment's findings revealed that using windows of one second did not significantly affect the score, but that using windows of one to two seconds did enable us to process traffic and identify botnets. This leads us to assert that, by examining traffic second by second, it is possible to identify botnet-like traffic in a real-time environment.

A. Black-hole or Sink holing technique

The Black-hole or Sink holing [14] method is another alternative technique to stop or evade the botnet traffic. When a large quantity of data is obtained from 3 major spamming bots, MegaD and sribzi, the data generated by these botnets will be considered malicious and redirected through Black-hole. The Black-hole is a virtually void that drops the incoming traffic without delivering the data traffic to the destination intended, but sometimes it could drop the original or legitimate traffic. When original traffic is dropped this would result in loss of legitimate data and this would make the data to be resent and in tern it would increase the network traffic. So this technique would not be a feasible.

B. The botnet's potential future

Botnet's Potential Future [15] includes Social botnets whether they exploit the social Medias like Facebook and Twitter, in such social media the messages can be very easily and quickly propagated and users trust the links which they receive and clicks and proceed further and now onwards it will be easy for infection. The Mobile botnets becoming a serious threat, the goal of the mobile bot’s attack is to get an access to the contents & the resources of the user's device and pass the controlling commands to the C&C server. The bots can even affect the clouds they named as dark clouds and they are controlled by cybercriminals.

IV. Conclusion And Future Direction

As internet usage increases the threat is also at an alarming rate in different forms like viruses, worms etc… Bot is a new form of threat but the effect and infection is huge and rapid. The effects of botnets should be controlled, otherwise it would damage the resources of the business and it may also result in financial loss. This research paper focuses on the implications or repercussions of botnets & different approaches and also the technologies used to identify or detect the botnets. The detection techniques are further categorized based on the various terminologies of botnets. This research paper even focuses on the different techniques adopted in finding botnets by some scholars. However, the implemented methods have not given any comprehensive approach to find botnets. There are still number of challenges persist in the field of botnet detection. Many researchers are facing a problem in validating their proposed techniques or methods in real time. The approaches which have been discussed shed some light on current methods used in detecting them, but still they lack in complete finding. S. Haq and Y. Singh [13] proposed to use efficient data partitioning policy to detect the botnets in feature in addition to the existing random data-partitioning policy where the results are satisfactory. Pradeepthi and kannan [17] have proposed to use the implemented system for any real cloud environment as a future enhancement. Javier, Victor, Eduardo and Enrique [18] will work in future on testing the proposed model on other botnets, creating a knowledge database of botnets that can be quickly detected with this model.

Some botnet detection tools work at system level and some work at network level, botnet identification should be done both at system and network level for a comprehensive detection approach. Botnet detection is still a potential research area as the new botnets hitting the world. Their new behavior makes the earlier detection techniques obsolete. Mobile botnets are emerging as a significant danger in addition to traditional botnets. The bots even affected some earlier versions of the iOS. Mobile botnet is becoming a new research domain and became a hot research topic. Hence research challenges and opportunities still present in detecting botnets.

References

[1] https://www.enzoic.com/blog/bot-mitigation/

[2] Botnet detection techniques - review, future trends, and issues by Ahmad Karim, Rosli Bin Sallehi, Muhammad Shirazi, Syed Adeel Ali Shahi, Irfan Awan, nor badrul anuar. Karim et al. / j Zhejiang Univ-Sci C (Comput & Electron) 2014 15(11):943-983.

[3] H.R. Zeidanloo, M.J.Z. Shooshtari, P.V. Amoli, M. Safari, and M. Zamani, "A taxonomy of Botnet detection techniques," In Proceedings of the 3rd International Conference on Computer Science and Information Technology (ICCSIT), IEEE, vol. 2, pp. 158-162, 2010

[4] P. Amini, M. A. Araghizadeh and R. Azmi, "A survey on Botnet: Classification, detection and defense," 2015 International Electronics Symposium (IES), Surabaya, Indonesia, 2015, pp. 233-238, doi: 10.1109/ELECSYM.2015.7380847.

[5] S. Gaonkar, N. F. Dessai, J. Costa, A. Borkar, S. Aswale and P. Shetgaonkar, "A Survey on Botnet Detection Techniques," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-6, doi: 10.1109/ic-ETITE47903. 2020.Id-70.

[6] N. Kaur and M. Singh, "Botnet and botnet detection techniques in cyber realm," 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2016, pp. 1-7, doi: 10.1109/INVENTIVE.2016.7830080.

[7] https://securityscorecard.com/blog/what-is-a-botnet-attack/

[8] https://www.paloaltonetworks.com/cyberpedia/what-is-botnet

[9] https://www.eccouncil.org/cybersecurity-exchange/penetration-testing/botnet-attack-prevention/.

[10] https://www.intechopen.com/chapters/69332

[11] https://jpdias.me/botnet-lab//countermeasures/detection.html

[12] https://fingerprint.com/blog/what-are-bots-how-to-detect-bots/

[13] S. Haq and Y. Singh, "Botnet Detection using Machine Learning," 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), Solan, India, 2018, pp. 240-245, doi: 10.1109/PDGC.2018.8745912.

[14] @inproceedings {Shin2011CrossAnalysisOB, title= {Cross-Analysis of Botnet Victims: New Insights and Implications}, author= {Seungwon Shin and Raymond H. Lin and Guofei Gu}, booktitle= {International Symposium on Recent Advances in Intrusion Detection}, year= {2011}, url={https://api.semanticscholar.org/CorpusID:10322821}}

[15] Karim, A., Salleh, R.B., Shiraz, M. et al. Botnet detection techniques: review, future trends, and issues. J. Zhejiang Univ. - Sci. C 15, 943–983 (2014). https://doi.org/10.1631/jzus.C1300242

[16] R. Sri Skandha Moorthy, N. Nathiya, Botnet Detection Using Artificial Intelligence, Procedia Computer Science, Volume 218, 2023, Pages 1405-1413, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2023.01.119.

[17] K. V. Pradeepthi and A. Kannan, "Detection of Botnet traffic by using Neuro-fuzzy based Intrusion Detection," 2018 Tenth International Conference on Advanced Computing (ICoAC), Chennai, India, 2018, pp. 118-123, doi: 10.1109/ICoAC44903.2018.8939109.

[18] Real-time botnet detection on large network bandwidths using machine learning by Javier Velasco-Mata, Víctor González-Castro, Eduardo Fidalgo and Enrique Alegre.
[19] Om Prakash, Jena, et al., editors. Methods, Implementation, and Application of Cyber Security Intelligence and Analytics. IGI Global, 2022. https://doi.org/10.4018/978-1-6684-3991-3
[20] H. L., Gururaj, et al. "Machine Learning (ML) Methods to Identify Data Breaches." Methods, Implementation, and Application of Cyber Security Intelligence and Analytics, edited by Jena Om Prakash, et al., IGI Global, 2022, pp. 52-64. https://doi.org/10.4018/978-1-6684-3991-3.ch004
[21] M. R., Pooja, et al. "Emerging Challenges in Cyber Security." Methods, Implementation, and Application of Cyber Security Intelligence and Analytics, edited by Jena Om Prakash, et al., IGI Global, 2022, pp. 25-34. https://doi.org/10.4018/978-1-6684-3991-3.ch002