IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 25, Issue 3, Ser. II (May. – June. 2023), PP 53-58
www.iosrjournals.org

Identification of Botnet in Network Traffic: Taxonomy and Survey

[1]Ms. Priya, [2]Dr.Arvind Kumar

[1]Ms. Priya, Assistant Professor, priya.fca@mriu.edu.in, School of Computer and Applications,Manav Rachna International Institute of Research and Studies, Faridabad
[2]Dr. Arvind Kumar, Associate Professor,akdangi@gmail.com, School of Computer andApplications, Manav Rachna International Institute of Research and Studies, Faridabad

Abstract— *Now the most of our daily activities are automatized and available for the use on internet. In which security is rarely a top consideration. Which results The most prevalent and frequent type of botnet in use today cyberattack by which our personal information can be misused by others? Botnet is collection of the system over the internet which are controlled remotely. This is done by the malware. malware is any software which intentionally designed to cause damage to a computer and perform such tasks which is not interest of the user. The botnet master will communicate with the bots via existing IRC c&c channels.And the techniques of the detection which can be used to detect the botnet. Those techniques are signature detection technique, honeynets detection technique, mining detection technique and many more. This Paper contains an Introduction to Botnet, what they are and how are they created, Protocols used by Botnet for functioning, Taxonomy of Botnet handling techniques and the conclusion asto what the current and future state of botnetmight be.*

Keywords— Botnet, Intrusion, Malware, P2P,Mining, Detection-Techniques.

Date of Submission: 03-06-2023 Date of Acceptance: 17-06-2023

I. INTRODUCTION

One of the major threats to internet security is posed by botnets. The network of botnet in which the systems are in the influence of the malware. Botnet is controlled by the Bot Master and use those systems as resources and the platform to perform malicious activities. By using the botnet Distributed denial-of-service (DDoS) assaults, phishing, keystrokes and identity theft, virus distribution, and information exfiltration are the main types of attacks. The bot master will use established IRC c&c channels to communicate and controls the bots by which bots receive commands and coordinate attacks. Bot master uses c&c channels to communicate and controls the robots. First, PC's executes script which connect that system to the channel without the acknowledgement of machine owner. Then the infected PC joins the channel, it will automatically assign a randomized username (bot id). This id is generated by the malicious script. Now the victim's system is bot and that system will run all commands which will be given by the Bot Master. Some of the botnet protocols in which HTTP, P2P, TOR based botnets have been getting more extensive growth in fast few year. But IRC has been most popular method because of its built-in ability to control a channel and their actions. HTTP communications also allow cyber criminals to establish an endless set of communication domain. In P2P the communication take place within the botnet itself. Botnet detection techniques are signature detection technique, honeynets detectiontechnique, mining detection technique and many more.

The need to perform decentralized denial of service (DDOS), spread spam, perform click fraud, and steal personal user information (credit card numbers, social security numbers, etc.).Perform powerful distributed computing tasks by leveraging powerful computing resources provided by bots.Gives an attacker access to the connection and the device.Decentralized denial of service (DDOS), spam distribution, click fraud, and the theft of private user data (credit card details, social security numbers, etc.) all need the usage of botnets.Perform powerful distributed computing tasks by leveraging powerful computing re-sources provided by bots. Allows anattacker access to the device and its connection.

II. BOTNET

BOT word is combination of ROBOT and NETWORK.

A botnet is a group of computers connected to the internet that are managed remotely by a bot master using bots, or malicious software (reference no. 1).Bots are programs. These programs can execute

automatically in zombie computers. Zombie computers are those computers which are connected to the internet and controlled remotely by hacker via computer virus or malware. These viruses are used to execute malicious actions without the interest of the computer owner. Botnet is a network composed by bots. Bot herder and bot master are the person who controlling the botnet. They use command and control (c&c) software to manage the botnet. It is based on flow intervals, traffic monitoring and traffic analysis. Bot produced when malware from a malware distribution infiltrates a device. In botnet devices whose security & control is given to the third party. for example, computer, smart phones or IOT devices. Botnet attack can be devasting. Recent botnets increasingly frequently interact over peer-to-peer networks that already exist.. The botnet took advantage of unsecured IOT devices. For the taking advantage of these IOT devices, the hacker will install malware which will attack the DYN servers that route internet traffic. Botnet terms are commonly used in a negative or malicious sense. Botnet is mostly used by the cyber criminals for the different purposes.
Example of some known bonnets: Mirai,Reapa(a.k.a. IoTroop),Echobot, Emotet,Gamut and Necurs.

A. Can we Stop Botnet

Includes wide availability and ongoing purchaseof insecure devices.
It is almost impossible to easily lock an infectedmachine from the internet.
Difficulty tracking and protecting botnet creators

B. Steps to Create Botnet

There are the various stages for creating botnet.These stages are as follow:

1. Linking to a c&c server that is established forinstructions.
2. Creating IRC (Internet Relay Chat) traffic overa certain . set of ports.
3. creating multiple requests to the DNS that arethe same.
4. Producing email and traffic using the simplemail transfer . protocol (SMTP).
5. Decreasing workstation performance and Internet access . . to the point where end userscan clearly see it.

C. Illegal use of Botnet

Junk Email: Used by the messaging system to send junk email (spam) and repeatedly send messages on the same website. It is primarily used in advertising. Spam behavior is economically visible because advertisers have no operational costs. It can be used to spread computer viruses, Trojan horses and malicious software.Keylogging and identity theft: These malicious programmes collect sensitive and personal data, like bank and credit account numbers, after being covertly installed on a zombie system. Keyloggers are simple to download and can infect computers by visiting "legal" websites and "social networking" sites like Instagram.Software that is intended to harm a computer, server, or network is known as malware. These programs are called malware if they Bot system perform task against the interests of computer owner.Authentic sites that have been hewed: Subdomains are recurrently castoff by authentic portals to split a portal into coherent segments called subdomains. Malware scopes users concluded mistreatment of weaknesses. If there are achievements in outline of a site or in server uses, then it could hypothetically be a path of operators.The phenomenon known as "malvertising": In this the hackers uses the ads on the web. Hackers realize that they can reach millions of people if they can infiltrate an ad network.File sharing/P2P network and untrustworthy sites: In this there are reputation for being a hive of malware. In this they focus on the websites which is used for the file sharing.Misleading download intended to hoax you: Occasionally a malware is shaped by the writer and then is muddle with alternative prevalent sequencer. And then it waits until you run the program/software.
DDoS Attacks: In a distributed denial of service attack, numerous sources of incoming traffic overload the targets. Simply blocking one source will end the attack. Brute force attack: Consists of an attacker submitting a large number of passwords in the hope that they are correct. The hacker checks all key combinations until it finds the correct one. This allows the hacker to decrypt the encrypted data. Manipulation of online polls: polling is struggling with accuracy for many reasons. It produces sometimes junk data by which the outcomes are not accurate. And the alteration in the surveys is informal.Click fraud botnets: Click fraud is once a individual stopovers any page or website and then clicks on any ad or on any button. It occurs on the large scale and then the link is copied many times and some of hackers uses bots for this fraud over and overagain.

D. Botnet Topology

The method through which botnet relationships are organised is known as botnet topology. Topology comes in two flavours:
Centralized &control server (c&c server) centralized around a Decentralized using peer to peer(P2P) communication.

Using a star architecture, the bots are positioned around a single command and control server.The case one multi server, which includes numerous c&c servers for redundancy, is disabled.
a hierarchical structure where numerous C&C servers are grouped together for better dependability. The number of machines that can be found from a single group or bot detection can be decreased by the hierarchical structure.

III. RELATED WORK

[11] A botnet is a combination of systems over the web that is being controlled well. The bot pro compels the system, which performs dangerous activity against the client's interest. Web hand-off talk (IRC) is used to prevent correspondence plans. Right when the structure is polluted (zombie system) will find and connect with the IRC server. From there on out, the bot master will request and control (C&C) the zombie structure through IRC. In Botnet, we use a part of the topography where P2P is decentralized geology. In P2P individual bot can go as a server or a client [11]. There is a piece of the techniques used to recognize the Botnet. First is [5] Association Interference Acknowledgment Structure, or NIDS for short, is customizing that recognizes interferences. It finds the unapproved access by taking a gander at the association traffic. NIDS inspects an association's activity for bets while researching its activities. There are two fundamental methodologies for recognizing interferences:

Characteristic-based: It tends to standard traffic. That balances it with the traffic being thought about. Any deviations from the regular movement of traffic will be considered sporadic and may address malware or a gamble. Misuse-based: A model of intrusions, it is. It will do this by holding on for them to happen. These strategies use various frameworks. Consider data mining. Oppositely, the maltreatment-based approach uses data mining and model affirmation. [5] [6] Botnet revelation is a problematic issue. Anomaly-based botnet disclosure structure is free to the show and development of the botnets. An authentic association follows, Bot-Digger, which shows first-rate ID precision on various botnets. The different sort of botnets consolidates IRC-based, HTTP-based, and P2P-based botnets. This acknowledgment is the astoundingly low deceptive positive rate on common traffic. [6] Some botnet acknowledgment strategies are: [10] Imprint Based is Used for recognizable proof of seen Botnet. Anomaly-Based is a strategy used to recognize botnets considering remarkable instances of association traffic. BNS-Based is manufactured using DNS data made by a botnet. Moreover, Mining-Based is used to find Botnets. [10]IV. Logical classification OF BOTNET Acknowledgment Strategies It will generally be irksome, as we understand that bots are expected to work without the client's data. Regardless, thereare a couple of ordinary signs that a PC may be spoiled with contamination:

- Affiliation tries with known c&c servers. Various machines on the association make a comparative DNS request.
- High outbound SMTP traffic (on account of snap coercion).
- Astounding pop-ups.
- Slow enrolling high microchip use.
- The issue on the web is getting too.

These signs show bot pollution. These issues can, in like manner, occur due to malware or network issues. Distinguishing a botnet needs advanced taking-apart limits associated with the picked data for analysis. Then, at that point, the records and the characteristics of issues performed are made. In this, we use two kinds of assessment approaches these are:

Dynamic technique: It covers many assessment methodologies that make bot pro, clearly or by suggestion, informed about acknowledgmentactivity. It gets malware and subsequently deactivates its harmful parts.

Uninvolved procedure: It examines the Botnet's traffic without destroying or adjusting it. It revolves around the discretionary effects of botnet traffic.

For example, broken packages come about due toa distant DDoS attack.

Honeypots: They are just truly perfect for understanding what Botnet ascribes. They have the high capacity to distinguish security risks and to assemble malware marks. By which we fathom the motivation and technique behind the risk used by the executioner. In a wide scale association, different sizes of honeypots from honeynets.Honeynets are, by and large, used on Linux because of the abundance of tool compartments.

Interference Area structure:

Eccentricity-based botnet area: In this undertaking to recognize Botnet considering a couple of association traffic noticing is helpful. In this, the issue of settling dark botnets is tended to using the abnormality-based botnet area. It canfind the IRC channel, which the attacker doesn't require for the attack.

Signature-based area: It is used for the distinguishing proof of seen Botnet. In this, the data on the approach to the actions of any current botnet is important. Nevertheless, this system is simply important when the botnets are known.

DNS-based area: It relies upon DNS information made by a botnet. This technique resembles the anomaly-

based area methodology. Right when the bots start the relationship with the c&c server for orders, and then for access, they make DNS requests to find that c&c server that is routinely worked with by the DDNS provider by which it is possible to perceive botnet DNS traffic and idiosyncrasies. Mining-based acknowledgment: This is one of the most mind-blowing techniques to find the Botnet. When botnets use run-of-the- mill shows for c&c correspondence, the traffic resembles the common traffic. Along these lines, it becomes testing to distinguish Botnet c&c traffic. In this, the busy time gridlock isn't high volume and doesn't cause high association inactivity around, then the anomaly-based acknowledgment isn't helpful. Subsequently, for recognizing c&c traffic, a couple of mining-based strategies are used. These methodologies are artificial intelligence, gathering, and batching.

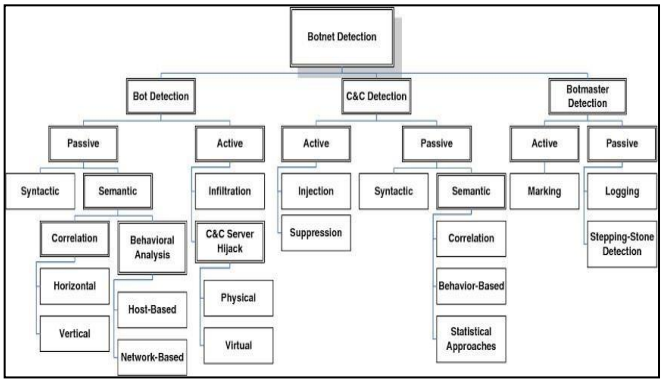

Fig. 1. Taxonomy of Botnet Detection (©gtastic.com)

Botnet c&c servers essentially utilize 4 components for progressing tasks and communi- cations with the tainted bots or we can say botnet protocoals IRC (Web Transfer Visit) is a talking framework that trades the client's message (message information) through TCP/IP convention through server. IRC has been the most famous technique. IRC have underlying capacity to control a channel the permissible individuals and their activities. The bot expert will utilize laid out IRC c&c channels to impart and controls the bots. It makes conceivable to send directions to numerous bots by using the multicast conveyance system if IRC. These channels are set up as "welcome as it were" for actually forestalling passage. (Hypertext move convention) HTTP additionally have brought together engineering. By which the whole botnet can be disturbed by essentially coming up short of the IRC server. An online botnet in which c&c bots utilizes HTTP convention. This gives qualified and stable network between the client and waiter. P2P (peer 2 companion) Mid-2000's bots use peer-2-peer plans, individual bots go about as both client and server, makes an organization which isn't comprise of any unified point which might be debilitated. Pinnacle (Third-age Onion Switch) It is an organization of workers, which give large number of transfers used to course correspondence in the Peak organization. The encryption convention is intended to overlook the solicitation of giving the transfers admittance to the information. The information is about they are directing. All the order and control correspondence happen inside the Pinnacle organization. The bot ace maintains a strategic distance from the utilization of the leave hubs.The bot ace attempts to reduce the botnetdiscernibility.

A. ***Correlation of A few Botnet Discovery Methods Honeynets:*** It is an organization with purposeful weaknesses. The motivation behind the honeynets is to welcome goes after additional the exercises and techniques for assault are contemplated.

Interruption recognition framework: It screens a few organizations in desires to settle the obscure botnet divulgence that could conceivably be scrambled.

Signature based recognition: It is utilized to identify botnet with known conduct of any current botnet. DNS based location: It is utilized to screen botnets in view of DNS data produced by a botnet. This strategy is like the peculiarity basedrecognition procedure.

Mining based identification: When botnets utilize typical conventions for c&c correspondence then the traffic is like the ordinary traffic. This is where mining-based discovery helps in uncovering botnets.

Approach	Unknown Bot Disclosure	Protocol & Structure Self-Governing	Encrypted Bot Disclosure	Real Time Disclosure	Low False Positive
Honeypots	✓	✗	✗	✗	✗
Intrusion Detection System	✓	✗	✓	✗	✗
Signature Based Detection	✗	✗	✗	✗	✗
DNS Based Detection	✓	✗	✓	✓	✓
Mining Based Detection	✓	✓	✓	✗	✓

Table 1. Botnet Technique Comparison Table

IV. CONCLUSION

Botnet is one of the main danger in network security. Botnets are started by malware. There are various botnet identification methods like Honeypots, Interruption discovery framework, Mark based location, DNS Based Recognition, Mining Put together Location thus with respect to. A large portion of these methods can undoubtedly perform Obscure Bot Discovery aside from Mark Based Location, which requires extra data about the objective prior to going on with the Recognition Part. Mining Based Discovery is a very rare example of procedures that are Convention and Design Free, support Scrambled Bot Recognition and Have a Low Misleading Positive Rate yet can't work Progressively Recognition. Encoded Bot recognition then again are not Convention and Construction Free like the previous strategy, yet at the same time support Scrambled Bot Identification with a Low Misleading positive rate and every last bit of it Progressively Recognition. All things considered, Mining Based Recognition is viewed as an overall strategy on the off chance that not having Continuous Location isn't a big issue. In any case, assuming it is, DNS Based Recognition is the following best Method to work with.

REFERENCES

[1]. Honeynet project, know your Enemy trackingBotnets, march 2005. **www.honeynet.org/paper/bots**

[2]. R. Lemos, Bot software looksto improve peer-age. **www.securityfocus.com/news/11390,2006**.

[3]. Saad, S., Traore, I., Ghorbani, A., Sayed, B., Zhao, D., Lu, W., ... & Hakimian, P. (2011, July). Detecting P2P botnets through network behavior analysis and machine learning. In *2011 Ninth annual international conference onprivacy, security and trust* (pp. 174-180). IEEE.

[4]. Zeidanloo, H. R., Manaf, A. B., Vahdani, P.,Tabatabaei, F., & Zamani, M. (2010, June). Botnet detection based on traffic monitoring. In *2010International Conference on Networking and Information Technology* (pp. 97-101). IEEE.

[5]. Erquiaga, M. J., Catania, C., & García Garino, C. (2012). An analysis of network traffic characteristics for Botnet detection. In *XVIII Congreso Argentino de Ciencias de la Computación*.

[6]. Gu, G., Perdisci, R., Zhang, J., & Lee, W. (2008). Botminer: Clustering analysis of network traffic for protocol-and structure-independent botnet detection.

[7]. Banday, M. T., Qadri, J. A., & Shah, N. A. (2009). Study of Botnets and their threats to Internet Security. *Sprouts: Working Papers*

on Information Systems, 9(24).
[8]. Ianelli, N., & Hackworth, A. (2005). Botnetsas a vehicle for online crime. CERTCoordination Center, 1(1), 28.
[9]. Binkley, J. R., & Singh, S. (2006). Analgorithm for anomaly-based botnet detection. SRUTI, 6, 7-7.
[10]. Kannan, R., & Ramani, A. V. Bot Detection using Traffic Monitoring and Traffic Analysis.
[11]. Zhao, D., Traore, I., Sayed, B., Lu, W., Saad, S., Ghorbani, A., & Garant, D. (2013). Botnet detection based on traffic behavior analysis and flow intervals. Computers & Security, 39, 2-16.
[12]. Casenove, M., & Miraglia, A. (2014, June). Botnet over Tor: The illusion of hiding. In2014 6th International Conference On Cyber Conflict (CyCon 2014) (pp. 273-282). IEEE.
[13]. Feily, M., Shahrestani, A., & Ramadass, S. (2009, June). A survey of botnet and botnet detection. In 2009 Third International Conference on Emerging Security Information,Systems and Technologies (pp. 268273). IEEE.
[14]. Zeidanloo, H. R., Shooshtari, M. J. Z., Amoli, P. V., Safari, M., & Zamani, M. (2010, July). A taxonomy of botnet detection techniques. In 2010 3rd International Conference on Computer Science and Information Technology (Vol. 2, pp. 158-162). IEEE.
[15]. Kambourakis, G., Kolias, C., & Stavrou, A. (2017, October). The mirai botnet and the iot zombie armies. In MILCOM 2017-2017 IEEE Military Communications Conference(MILCOM) (pp. 267-272). IEEE.
[16]. P. Barford and V.Yagneswaran, "An Inside Look at Botnets ". In: Special Workshop on Malware Detection, Advances in Information Security, Springer, Heidelberg (2006).
[17]. B. Saha and A, Gairola, "Botnet: An overview," CERT-In White PaperCIWP-200S-0S,200S.
[18]. RVillamarin-Salomon and J.C. Brustoloni, "IdentifYing Botnets using Anomaly Detection Techniques Applied to DNS Traffic," in proceeding 5th IEEE ConsumerCommunications and Networking confernce.(CCNC 2008), 2008, pp. 476-481.
[19]. T. Karagiannis, K. Papagiannaki, and M. Faloutsos, "BLINC:multilevel traffic classification in the dark," In Proceedings of the 200S Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 229-240, Philadelphia, Pennsylvania, 2015
[20]. J. R Binkley and S. Singh. An algorithm for anomaly-based Botnet detection. In Proceedings of USENIX SRUTI'06, pages 43- 48, July 2006.
[21]. J. R. Binkley and S. Singh. An algorithm for anomaly-based botnet detection. In Proceedings of the 2nd conference on Steps to Reducing Unwanted Trac on the Internet - Volume 2, SRUTI'06, pages 77, Berkeley, CA,USA, 2006. USENIX Association.
[22]. C. A. Catania and C. García Garino.Automatic network intrusion detection:Current techniques and open issues.Computers and Electrical Engineering, 2012.Accepted. In Press. DOI:10.1016/j.compeleceng.2012.05.013.
[23]. P. Barford and V.Yagneswaran, "An Inside Look at Botnets ". In: Special Workshop on Malware Detection, Advances in Information Security, Springer, Heidelberg (2006).
[24]. N. Ianelli, A Hackworth, Botnets as a Vehicle for Online Crime, CERT, December 200S.
[25]. 1. Leonard, S. Xu, and R. Sandhu, "A Framework for Understanding Botnets," in Proceedings of the International Workshop on Advances in Information Security (WAIS at ARES), (Fukuoka, Japan), Fukuoka Instituteof Technology, March 2009.