# Survey on Botnet Detection Techniques

**Rahul Mishra and Sudhanshu Kumar Jha**

**Abstract** Due to unpleasant market competition, IT companies are releasing the software or applications without much considering the unintended security breaches presented inside it. The malware programs moving around the Internet is looking for such kind of the security breaches to attain the malicious intention. Botnet is a kind of malware program(s) looking for the vulnerable system and has now become the worldwide epidemic due to mostly applied for a malicious purpose. Many malware detection techniques have been discussed so far for botnet detection in the literature, however typically considering it on host device or on the traffic of a network. IoT devices are much vulnerable to botnet as the manufacturer has the main concern to releases with new feature in order to compete the market without much attention to weak point(s) inside it. This paper discusses the basic botnet detection techniques. For this purpose, the paper covers both static and dynamic detection techniques along with its advantages and shortcomings. On the basis of characteristics for botnet detection, this paper also deliberates the basic procedures to create botnet detectors, defining some parameters for botnet detection and categorizing the detection methodologies. Further, the paper reveals an implementable position in the system with advantages and drawbacks on the detection performance.

**Keywords** Botnet · Botnet detection technique · Anomaly-based detection · Signature-based detection · Specification-based detection · IoT Vulnerability

## 1 Introduction

Due to the present coronavirus COVID-19 pandemic situation, a huge demand is looking for various ICT tools to meet the immediate business need without much concern to the security breaches inside it. Self-propagating botnet programs are moving around the Internet in search of vulnerable system. As per report by World Health Organization (WHO), more than 450 e-mails addresses and passwords of WHO's were compromised during April 2020 [1] and a forecasted report by P&S

R. Mishra (✉) · S. K. Jha
Department of Electronics and Communication, University of Allahabad, Prayagraj 211002, India

K. Dahal et al. (eds.), *Internet of Things and Its Applications*, Lecture Notes in Electrical Engineering 825, https://doi.org/10.1007/978-981-16-7637-6_39

intelligence, Market Research Future (MRFR) shows a compound annual growth rate (CAGR) of 12.6% with total $119.9 billion financial loss by the end of 2030 due to cyber-attacks [2, 3], and thus, this area needs immediate attention among researcher.

The journey of malwares was first appeared in 1988 and continues their impact till date. McGraw and Morrisett [4] define malicious code or malware as "any code added, changed, or removed from a software system in order to intentionally cause harm or subvert the intended function of the system." Initially, Internet users were attacked with malwares, propagating through e-mails, freeware, lucrative software, and games or with some other medium [5]. However, by the time, this kind of approaches for infecting devices become not much effective, as the infection processes were easily detectable due to the actions initiated by the malware during infection or changes caused on the infected devices. With the advancement in anti-malware protection and elimination programs, the malware presence is being detected and immediately removed from the infected devices, and thus, the infected devices are no longer remain asset for the attacker. With the advancement in technology, the intention of trend for infection was changed, and now, attacker wish to hold the control of infected device as long as possible. Therefore, the presence of malware in the infected device needs to be shield. Various kinds of malwares are being used; botnet is one of them.

Botnets are utilizing the feature of self-propagation to target more and more devices by exploiting the vulnerabilities present in the devices such as open ports and default credentials. IoT devices are most suitable for performing such activities as the devices are implemented with least security mechanisms, continuous network connectivity, and limited computing resources in order to retain it simple to use. IoT devices are dynamic, heterogeneous, and interoperable, and due to these features, a uniform solution to prevent or mitigate the botnet in IoT devices is not feasible [6]. Most of the IoT devices are connected to Internet without firewall and available round the clock and are mostly configured with factory-enabled default username and password with open ports for various protocol, customer-care service, and thus, to contaminating an IoT device becomes quite easy. The device infected with malicious code called bot and group of infected devices together with Command and Control Centre (C&C) termed botnet [7]. In a botnet, infected devices (bots), take command(s) from the C&C to perform predefined action(s) on the basis of the received commands. These commands are given by the attacker who controls the botnet. Typically, the commands are being used for performing DDoS attack, sending spam mails, click fraud, or stealing financial and sensitive information from the infected device [8]. Many researchers believe that more than 25% of the IoT devices which are connected to the Internet without any proxy server are member of botnets [8–10].

Primarily botnets are configured for testing various features of network and however, later on, intruders started implementing bots with intention to perform malicious activities such as purloin financial information, security credential, sending spam mails, or performing DdoS attack to slow down or sometimes stop the services of the targeted system [11]. These botnets are capable of utilizing exact vulnerabilities available on specific devices from a manufacturer, in order to the keep the device easy to use, companies provide details on their website such as default credentials, open telnet ports for remote access of device. This publicly available information is utilized by the attackers to target any IoT device or a model of a manufacturer. Mostly, the botnet exploits the codes of exiting botnets or append new features to exploit the new vulnerabilities of the IoT devices [12]. The malware detector is a kind of program which basically looks for the description and the basic functionalities of malicious program. Unlike an anti-virus software, a malware detector is not necessarily supposed to reside on the device under observation (DUO) and senses the presence of malware on the basis of set of rules of the detection techniques. The performance of a malware detectors depends on the set of detection techniques it uses [7]. Many researches have been carried out to predict or mitigate the botnet attack. These detection techniques utilize various parameters of botnet or network such as botnet signature, network traffic, very-long connection time between client–server and so on [13–15].

The motivation and main contribution of this paper are as follows: Sect. 2 of the paper discusses the taxonomy of a botnet with common IoT vulnerabilities. Section 3 deliberates the basic detection techniques and a discussion about categories of botnet, based on common botnet attack along with a detailed review and comparison analysis of botnet detection techniques and tools followed by conclusion in Sect. 4.

## 2 Botnet Taxonomy

As described, a botnet is a network of connected bots that spread over network to perform various malicious activities such as spam mail generation, distributed denial of service (DDoS) attack, stealing sensitive financial (credit/debit card data) and security information, and tricking personal information for identity theft [13]. An autonomous program performing the above actions without taking instructions from any intruder is called bot, whereas the network of bots connected to Command and Control Server (C&C) taking commands to perform actions based on the commands are called botnet [8–10]. Botnets utilize the vulnerabilities present in the devices to infect them, and once a device is infected, it starts working as a bot and further search for the new device on the network for further infection. List of common vulnerabilities in Table 1 [13–15]. These botnets generally utilize default credentials of IoT devices, open ports or sometimes the vulnerabilities present in the software. Mirai botnet is the famous botnet that performs dictionary-based default credential attack on the devices.

The connection between bots and C&C defines the architecture of the botnet. The architecture of the botnet can be categorized as centralized, peer to peer (P2P), and hybrid [7]. The centralized architecture of botnet is easy to implement, generate quick response to bots' requests, quick and direct update to bots but the dependability on a single source make is less reliable.

Whereas the P2P architecture does not directly communicate to the bots, rather command is sent via another bot in the network so the detection of the C&C becomes

**Table 1** Common IoT vulnerabilities

| Vulnerability | Description |
|---|---|
| Default credentials | Many IoT devices have no mechanism to change the default credentials. The credentials of the devices are available on the manufacturer’s website for open access |
| Un-authorized access | IoT devices do not ensure who can access the data available with them |
| Insecure software/firmware | These devices also lack the update mechanism or if the patch is available, then it is not verified |
| Physical security | Sometimes these devices are installed on open places or their physical security is not considered and open to disassemble or can be accessed via USB or some removal storage medium |
| Open ports and network services | These devices are vulnerable to DDoS attack, and network ports are open for remote login purposes |

next to impossible. The hybrid architecture utilizes the features of both architectures to control the bots. Figure 1 shows the architecture of the botnet.

Authors [11, 12] classified the IoT attacks on the basis of how the attacker utilizes these devices after successful infection. Table 2 shows the categories of attacks.

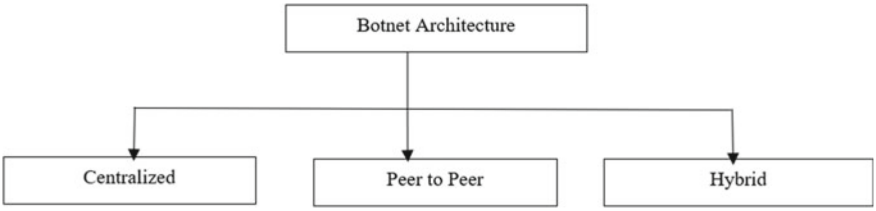

**Fig. 1** Botnet command and control topology

**Table 2** Common botnet attack categories

| Category | Description |
|---|---|
| Ignoring the functionality | In this type of attack, the intended works of the IoT devices are ignored and the IoT devices are considered only as the computing device that is connected to the Internet |
| Reducing the functionality | These attacks are designed to limit the functionalities of the IoT devices. But these kinds of attacks cost human life if the target is medical equipment |
| Misusing the functionality | Misusing the functionality of devices may sometime cause reverse impact of the intended functionality or doing something that is not expected from the device |
| Extending the functionality | These attacks are designed to extend the functionalities of the infected IoT devices; i.e., infected devices are performing works for which they were not designed |

These attacks were categorized on the basis of the impact caused by the malicious code after they have successfully infected the device or network.

## 3 Botnet Detection Techniques

Botnet detection techniques are the most discussed topic nowadays. Many works have been done to address this issue. Here in this paper, we are trying to find out the most relevant works and categories them how they are addressing detection of botnet. The botnet detector takes two inputs [12]. The first input is the knowledge of the malicious behavior of the botnet. Second input is the program that needs to be observed. Once the botnet detector has the knowledge of what is considered malicious behavior and the program under inspection, then it employs its detection techniques to decide that the program is malicious or benign. Sometimes IDS and malware detectors are used interchangeably but a malware detector is usually only a component of a complete IDS. Techniques used for detecting malware can be categorized into three categories: anomaly-based detection, specification-based detection, and signature-based detection [16]. Figure 2 provides the information about various botnet detection approaches. All the three categories have three subcategories, namely static, dynamic, and hybrid [13, 17–19].

An anomaly-based botnet detection technique uses its gathered information during the run-time of a program to constitute normal behavior of the program to decide the maliciousness of a program under observation. A sub-category of anomaly-based detection is referenced as specification-based detection.

Specification-based botnet detection techniques control some specification or set of rules of what is valid behavior in order to decide the maliciousness of a program under observation. Programs violating the set of rules of specification are considered anomalous and usually malicious [20].

Signature-based botnet detection technique uses its predefined set of rules to what is known to be malicious for the host under observation to decide the maliciousness of a program under inspection. It is clear that characterization of properties or signature

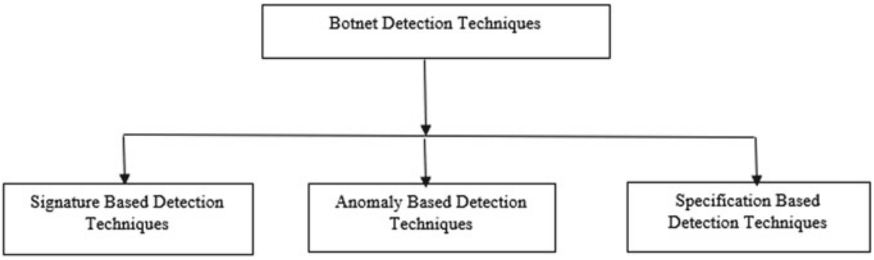
Botnet Detection Techniques
Signature Based Detection Techniques
Anomaly Based Detection Techniques
Specification Based Detection Techniques

**Fig. 2** Botnet detection techniques

of the malicious behavior is the key to a signature-based botnet detection method's effectiveness.

Table 3 presents a detailed analysis of variances among static and dynamic approach. Static analysis uses syntax or structural properties of the program under observation to predict its maliciousness. Whereas the dynamic approach works on various features on the host or the network under observation such as connection time between client–server, DNS and starts when the code starts executing [21]. The specific approach of an anomaly-based or signature-based technique is determined by how the techniques gather information to detect malware.

**Table 3** Review of botnet detection techniques and tools

| Techniques | Description | Tools |
|---|---|---|
| Dynamic anomaly-based detection technique | The information is gathered from the program under observation (POU) when starts executing on the host. This technique also | PAYL[22], Computer Forensic Method for Privacy invasive Software [23] |
| Static anomaly-based detection technique | Characteristics of file's structure of the PUO are used to detect. The malware can be detected even it is not being executed | Fileprint Analysis [24] |
| Hybrid anomaly-based detection technique | Uses the features of both dynamic anomaly-based detection and static-based detection | Strider GhostBuster [25] |
| Dynamic specification-based detection techniques | Tries to categories specification-based behavior at run-time to detect the malicious code | Monitoring Security-Critical Programs [26], Using Dynamic Information Flow to Protect Application [27] |
| Static specification-based detection techniques | Focuses on structural properties of files of PUO | Static Detection of Malicious Code in Executables [28], Detecting malware in Firmware [29] |
| Hybrid specification-based detection techniques | Specification-based behavior at run-time to detect the malicious code and structural properties of PUO | DOME [30] |
| Dynamic signature-based detection techniques | Uses the information gathered during execution of PUO | Rule-Based IDS Approach [31], Behavioral Approach to Worm Detection [32] |
| Static signature-based detection techniques | This technique uses the sequence of the code of PUO | Generic Virus Scanner [33] |
| Hybrid signature-based detection techniques | Uses the properties of static and dynamic detection techniques to detect the malwares | Analyzing and Detecting Malicious Mobile Code [30] |

These botnet detection techniques can be further classified as on host and on network-based detection techniques [9, 11]. On host-based detection technique, the detection of malware is done locally on the host itself. These techniques are local to the machines under observation. While in network-based detection techniques, the network traffic is monitored for the detection of malware. Network-based monitoring technique can be further divided into active monitoring techniques and passive monitoring techniques.

In active monitoring techniques especially, crafted packets are injected into the network traffic and their responses are apprehended for presence of malware in the network. Many legitimate DNS that expires are used by intruders for malicious activities. Such domains are specially used for sending the patches for many software, and when these patches are installed on machine, it starts performing many malicious activities such as keystroke recording, stealing valuable information from the device and sending it to the attacker's device. Passive detection technique monitors such DNS queries for malicious activities that can be a part of any botnet.

### *3.1 Limitations in Botnet Detection*

Botnet detection has many limitations [20, 25, 26] for designing single uniform solution such as heterogeneity, functionalities, and management policies for IoT devices. Sometimes the governing policies, goals of the Internet or network may also limit the applicability of botnet detection mechanisms. Lack of information related to connected devices to a network, sometimes it tough to decide that a particular device belongs to a network. The Command and Control channels also cause problem in detection of botnet many C&C channels use push methodology, whereas some C&C channels use pull methodology for communication. The protocols such as HTTP and IRC are used for communication between client and server. These protocols are also one of the limiting factors for botnet detection.

## 4 Conclusion

Due to the wide range of applications and easy deployment, IoT devices become a popular choice among people; however, due to unpleasant market demand, manufacture is not much concern about the security breaches in their products. Malware(s) is (are) one of the malicious programs always looking for such kind of vulnerable devices, and thus, IoT devices become one of the great choices. Due to the availability of open source-code of many malwares on various online forums and available for free download, intruders are using their knowledge to add new features to utilize the new vulnerabilities available in the system. In this paper, various methods have been compared based on the static analysis of malware detection. A detailed common IoT vulnerabilities have been presented based on common botnet attack categories.

A brief review of botnet detection techniques and tools also has been discussed so far. Static botnet detection techniques have fixed set of rules to detect the botnet that makes them quick and easy to implement but this also becomes their limitation when it comes to detect any new botnet as their rules are not known to the static-based detection techniques. Whereas dynamic detection techniques detect malware by utilizing the features of botnet when they are executing and trying to detect the malicious action being performed by any device. If so then the alarm is triggered for botnet. This detection technique is complex to implement but has its own advantages as it works on the attributes of the traffic and behavior of the PUO. Discussion on detailed common IoT vulnerabilities and review of botnet detection techniques and tools are main contributions of this paper.

## References

1. WHO's report, April 2020. WHO reports fivefold increase in cyber-attacks, urges vigilance. https://www.who.int/news/item/23-04-2020-who-reports-fivefold-increase-in-cyber-attacks-urges-vigilance. Accessed on 22 Nov 2020
2. MRFR report on (Sept 2020). Botnet Detection Market Research Report, by Vertical (Government & Defense, IT & Telecommunications), by Organization Size (Large Enterprise, SMEs), by Application (Mobile-based, Web-based) by Deployment (On-Cloud, On-premise) — Global Forecast till 2023, https://www.marketresearchfuture.com/reports/botnet-detection-market-6477. Accessed on 22 Nov 2020
3. P&S Intelligence press journals. July, 2020. Cyber Security Market Research Report: By Component (Solutions, Services), Security Type (Application, Network, Endpoint, Cloud, Enterprise), Deployment (On-Premises, Cloud), Enterprise (Large Enterprises, SME), Use Case (Security Monitoring, Network Traffic Analysis, Threat Hunting, Incident Response, Data Exfiltration), Industry (Aerospace & Defense, Government, BFSI, Healthcare, Retail, IT & Telecom, Manufacturing) - Global Industry Analysis and Growth Forecast to 2030. https://www.psmarketresearch.com/market-analysis/cyber-security-market. Accessed on 22 Nov 2020
4. McGraw G, Morrisett G (2000) Attacking malicious code: a report to the InfoSec research council. IEEE Softw 17(5):33–41
5. Biozid B, Mohiuddin A (2020) Deep learning meets malware detection: an investigation. In: Combating security challenges in the age of big data, Springer, pp 137–155
6. Yerima SY, Alzaylaee MK (2020) Mobile Botnet detection: a deep learning approach using convolutional neural networks. In: 2020 International Conference Cyber Situational Awareness, Data Analysis Assessment, Cyber SA 2020
7. Karim A, Bin Salleh R, Shiraz M, Shah SAA, Awan I, Anuar NB (2014) Botnet detection techniques: review, future trends, and issues. J Zhejiang Univ Sci C 15(11):943–983
8. Alieyan K, Almomani A, Manasrah A, Kadhum MM (2017) A survey of botnet detection based on DNS. Neural Comput Appl 28(7):1541–1558
9. Tyagi AK, Aghila G (2011) A wide scale survey on botnet. Int J Comput Appl 34(9):9–22
10. Feily M, Shahrestani A, Ramadass S (2009) A survey of botnet and botnet detection. In: Proceeding—2009 3rd International conference emerging security information, system technology security 2009, pp 268–273
11. Daya AA, Salahuddin MA, Limam N, Boutaba R (2019) A graph-based machine learning approach for bot detection. In: 2019 IFIP/IEEE Symposium on Integrated Network and Service Management IM 2019, pp 144–152 April 2019

12. Li H, Hu G, Yuan J, Lai H (2012) P2P botnet detection based on irregular phased similarity. In: International Conference on Instrumentation, Measurement, Computer, Communication and Control IMCCC 2012, pp 79–82
13. Meidan Y, Sachidananda V, Peng H, Sagron R, Elovici Y, Shabtai A (2020) A novel approach for detecting vulnerable IoT devices connected behind a home NAT. Comput Secur 97(101968)
14. Rahal BM, Santos A, Nogueira M (2020) A distributed architecture for DDoS prediction and bot detection. IEEE Access 8:159756–159772
15. Mohd Aman AH, Yadegaridehkordi E, Attarbashi ZS, Hassan R, Park YJ (2020) A survey on trend and classification of internet of things reviews. IEEE Access 8:111763–111782
16. Dong X, Hu J, Cui Y (2018) Overview of botnet detection based on machine learning. In: Proceeding—2018 3rd International Conference Mechanical Control Computer Engineering ICMCCE 2018, pp 476–479
17. Soe YN, Feng Y, Santosa PI, Hartanto R, Sakurai K (2020) Machine learning-based IoT-botnet attack detection with sequential architecture. Sensors (Switzerland) 20(16):1–15
18. Blaise et al. A (2020) Detection of zero-day attacks : an unsupervised port-based approach to cite this version : HAL Id : hal-02889708. Comput Netw, Elsevier 180:107391
19. Damodaran A, Di Troia F, Visaggio CA, Austin TH, Stamp M (2017) A comparison of static, dynamic, and hybrid analysis for malware detection. J Comput Virol Hacking Tech 13(1)
20. Ngo QD, Nguyen HT, Le VH, Nguyen DH (2020) A survey of IoT malware and detection methods based on static features. ICT Express, no. xxxx, 2020 (in press)
21. Thorat SA, Khandelwal AK, Bruhadeshwar B, Kishore K (2008) Payload content based network anomaly detection. In: 1st International Conference on the Applications of Digital Information and Web Technologies ICADIWT, pp 127–132
22. Oliveira LB, Pereira FMQ, Misoczki R, Aranha DF, Borges F, Nogueira M, Wangham M, Wu M, Liu J (2018) The computer for the 21st century: present security & privacy challenges. J Internet Serv Appl 9(1):24
23. Cui B, Jin H, Carullo G, Liu Z (2015) Service-oriented mobile malware detection system based on mining strategies. Pervasive Mob Comput 24:101–116
24. Richer TJ, Neale G, Osborne G (2015) On the effectiveness of virtualisation assisted view comparison for rootkit detection. In: Proceedings of the 13th Australasian information security conference (AISC 2015), vol 27, pp 30 Jan 2015
25. Chandramohan M (2018) Scalable analysis for malware and vulnerability detection in binaries (Doctoral dissertation)
26. Muppidi SR, Kodeswaran PA, Mukherjea S, Nandakumar V, Kapoor S International Business Machines Corp, 2016. Policy-based dynamic information flow control on mobile devices. U.S. Patent 9,253,209
27. Cui Z, Du L, Wang P, Cai X, Zhang W (2019) Malicious code detection based on CNNs and multi-objective algorithm. J Parallel Distrib Comput 129:50–58
28. Sallam AS, McAfee LLC (2016) System and method for firmware based anti-malware security. U.S. Patent 9,317,690
29. Khan WZ, Khan MK, Muhaya FTB, Aalsalem MY, Chao HC (2015) A comprehensive study of email spam botnet detection. IEEE Commun Surv Tutorials 17(4):2271–2295
30. Kumar V, Sinha D, Das AK, Pandey SC, Goswami RT (2020) An integrated rule-based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset. Clust Comput 23(2):1397–1418
31. Norouzi M, Souri A, Samad Zamini M (2016) A data mining classification approach for behavioral malware detection. J Comput Netw Commun
32. Wressnegger C, Freeman K, Yamaguchi F, Rieck K (2017) Automatically inferring malware signatures for anti-virus assisted attacks. In: Proceedings of the 2017 ACM on Asia conference on computer and communications security, pp 587–598, Apr 2017
33. Alme C, McAfee LLC (2016) System and method for detecting malicious mobile program code. U.S. Patent 9,246,938