Open Access Article

BOTNETS: WORKING MECHANISM AND SURVEY OF DETECTION TECHNIQUES

Rajesh Yadav
BML Munjal University, Gurgaon, India

Abstract:
As internet usage is growing day by day, cyber-crimes have also increased at a very high rate and cyber criminals are performing these crimes as profitable criminal activities The world of cyber security is facing botnet as an emerging threat and the use of Command-and-Control Server(C&C Server) makes this threat a more dangerous one in comparison to all other cyber-attacks. Multiple number of machines are compromised and the collection of such machines as a network creates a botnet, such a network is controlled from a remote location by a bot herder, and he performs different types of malicious activities with all the compromised machines working as bots or zombies. Botnet has the objectives like performing denial of service attack, identity theft, phishing as well as other malicious activities. Detection of these botnets is a very important and main issue; it has motivated me to perform a detailed survey of botnet detection techniques. This paper throws some light on the working principle of botnet and highlights the research work done by various researchers for detection of botnets using different techniques.
Keywords: Botnet, IoT, Denial of Service, Malware, Cyber-Security.

抽象的：
随着互联网使用量的日益增长，网络犯罪也以非常高的速度增加，网络犯罪分子将这些犯罪作为有利可图的犯罪活动进行网络安全世界正面临僵尸网络作为一种新兴威胁和命令和- **与所有其他网**络攻击相比，控制服务器（C&C **服**务器）使这种威胁更加危险。多台机器被攻陷，这些机器的集合形成一个僵尸网络，这种网络由僵尸牧羊人从远程位置控制，他执行不同类型的恶意活动，所有受感染的机器都作为机器人或僵尸。僵尸网络的目标是执行拒绝服务攻击、身份盗用、网络钓鱼以及其他恶意活动。检测这些僵尸网络是一个非常重要和主要的问题；它促使我对僵尸网络检测技术进行详细调查。本文重点介绍了僵尸网络的工作原理，并重点介绍了不同研究人员为使用不同技术检测僵尸网络所做的研究工作。
关键词：僵尸网络、物联网、拒绝服务、恶意软件、网络安全。

1. Introduction

Collection of interconnected devices when infected by malware forms a botnet environment, this allows attackers to control these devices from remote locations. The main purpose of creating botnet is to launch attacks through those

Received: July 02, 2021 / Revised: July 30, 2021 / Accepted: August 31, 2021 / Published: September 13, 2021
About the authors: Rajesh Yadav
Corresponding author- Rajesh Yadav

devices. Different protocols are there on which botnets rely ,these are hypertext transfer ,internet relay chat and peer to peer transmission. Many host machines which we call as zombies or bots are a part of such network and botmaster control these machines to perform various activities[1,2].Different attacks are a part of the main objective of botmaster, like spam, phishing, distributed denial of service attack and information theft, which is a crucial type of threat faced by internet[3]. Different security issues are raised by the botnet's development, identification as well as detection of botnets is now a big challenge in in the world of Internet. Distributed denial of service attack is a very crucial attack with the help of which botnets can cause calamitous interference in the network and this situation turns out to be very costly for the organizations. The entire network of botnet can be programmed to collect sensitive information from various sources like government, corporate, or any individual and then it can be sold at a very high price in the organized crime market since this type of information can be very much useful for some people or organizations. One of the very important features of botnets is that they are reliable and can be reused, this makes them special among the attackers, these attackers operate from remote locations, and it is a challenging task to detect the original source of botnet.

Various technologies are used by botnets, some of them are fast flux, bitcoin networks,p2p networks and zero-day vulnerabilities as their platforms for spreading and utilization[4-6]. On comparing botnet with network following conventional approach, we see that botnets spread faster and in addition to this, they have more technical content ,more channels of infection and they are more concealed with a high volume of destruction. Bot machines have a unique way of operating which is very dangerous and that is they maintain their operation silently and the connection is maintained only through the command-and-control server, it becomes difficult to detect them as they operate with conventional attacks methods. First, we discuss the working strategy of botnet in section two, and thereafter botnet detection techniques proposed by various researchers is discussed.

2. Botnet working strategy

2.1 Spreading Infection

A variety of techniques are used by the bot herder to spread infection among the victim machines and other devices which are connected to the network, then they are being transformed into bots. A script which we call as shell code is executed by the targeted host machine, an image of real bot binary is drawn by this code from locations through file transfer protocol, hypertext transfer protocol or even peer to peer mechanism. After the self-installation of bot binary on the compromised machine, the victim machine transforms into a bot and performs malicious code execution. The bot applications execute automatically once the bot is rebooted[7].

Some of the botnet infection mechanisms are listed below:

a. Vulnerability in Software*:* The software used by victim machines is remotely exploited, after such successful exploitation, attacker is able to control the victim machine completely. The attacker can then install malicious programs on the victim machine for creation of user account with full administrative privileges. At last, the bot launches attack through the infected victim machines[8][9].

b. Instant Messaging: Machines which are remotely controlled are used to build large botnet network by using computer worms that can sent to other machine by instant messaging. In year 2006, a US company identified a worm known as W32.pipeline which spreads through AOL's instant messaging program. Either in the form of jpeg or in a website link, these types of worms are embedded[10].

c. P2P file sharing network: Another method of spreading infection among the machines is using P2P. A copy of itself is created by the malware binary in the P2P programs shared folder and legitimate names are used to fool a victim for opening the malicious binary. For further spreading, these binaries make use of social engineering concept[11].

2.2 Command and Control (C&C) Mechanism

The bot uses a C&C(command and control)channel to connect C& C server to bots by using variety of topologies, models and application like P2P,HTTP and IRC. This zombie then becomes a part of the botnet attack environment after C&C channel creation. This channel is used by bot herder or bot master

for sending commands to his army of bots[12]. Centralized, decentralized and hybrid are three types of botnet C&C architectures.

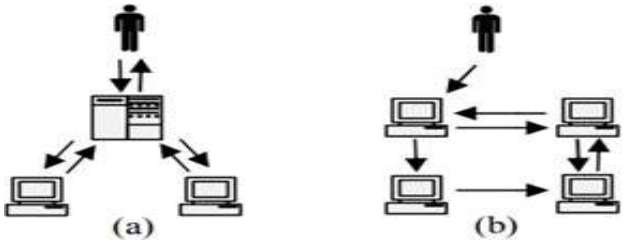

Figure 1: (a) Centralized and (b) Decentralized Mechanisms [13]

a. Centralized C&C: Figure 1(a) shows the centralized command and control architecture. Here the zombies are attached with the server for getting commands as well as updates and using these commands and updates they perform their task. Registration services to the available bots are provided by the C&C server for activity tracking. While performing all the activities in the botnet network, the bot herder is always connected to the C&C server so that it can assign tasks and commands to the zombies. This approach of centralized C&C server is common, and it applies simple steps for creation as well as giving direction to the bots, their response rate is very fast. It is of two types i.e., they are based on IRC and HTTP protocols on which they rely for connection establishment[13].

b. Decentralized C&C:

Figure 1(b) shows the decentralized C&C architecture, it is based on P2P model. In this approach, those machines which are infected act as both C&C as well as bots simultaneously. For sending commands to its nearby bots, each bot act as a server thereby replacing the need of centralized C&C server. In the botnet network based on this approach, one or more bots receive the command and then they pass on this command to other bots, this is how this network operates. For creation as well as management of this type of architecture, it requires a good level of expertise since it is a complex approach[13].

c.Hybrid C&C: Each of the previously mentioned approached of C&C have their advantages and disadvantages. Hybrid approach is formed to take advantages from both models. Just as an example, HTTP2P botnet network communicates with HTTP for avoiding firewalls over a P2P model in order to remove

shortcomings of central C&C server. This approach is not confined to usage of some architectures of services, bot herders always have a flexibility to make use of any protocols for implementation of this model[13].

3. Survey on Botnet detection techniques

In recent years, a lot of research work has been done by various researchers on different methods for detecting the botnet, we have analyzed the same in this section.

A wide-area, large scale system for botnet detection is proposed which includes combining new techniques to overcome the issues and challenges imposed by net flow data. Specifically, many characteristics have been identified which are responsible for disclosure to command-and-control channels from benign network traffic using records of net flow such as access patterns of clients, temporal behavior and flow size. At last, the authors evaluated this disclosure over two real world large networks thereby demonstrating that the disclosure can detect botnet as well as command and control channels in real time scenarios over the datasets with respect to billions of flows every day[14].

Two significant issues about the LSTM(Large Short-Term Memory) applications of detection of network behavior were analyzed by another group of researchers. For imbalanced traffic, two of the most popular sampling techniques i.e., oversampling and under sampling were evaluated on dataset in which case there is more botnet as compared to normal traffic. This evaluation showed that FAR i.e., false alarm rate will increase in absence of use of any of the sampling techniques. Because of being computationally efficient, under sampling technique was preferred. A deep analysis showed that those botnet behaviors which were different from normal were correctly detected by the LSTM[15].

For Bot Cloud detection, a convolution neural network i.e., CNN based technique was proposed. It first starts with the basic network flow features extraction from data packets and then maps these characteristics into gray image. After that, CNN algorithm is used to extract and learn those characteristics which can express structural relationship as well as hidden model in network data flow. The results have shown that the method proposed by these researchers not only increases the detection accuracy, but it also minimized the detection time[16].

A two-level technique for detection of botnet as well as identification of those nodes(bots) who have been compromised even before the activation of botnet has been proposed. During the first level, by leveraging large deviations of an empirical distribution, this method detects anomalies. And at the second level, botnet detection is performed by applying the technique of graph based social network community detection which captures correlations of interactions between nodes. By raising modularity measure to maximum point in that graph, community detection is being done. This method is applied in a real time botnet traffic environment and its performance is compared with other techniques[17].

In order to evade their detection, botnets use domain name system which is popular for locating the command-and-control server and it enhances their chance of going undetected. A method was proposed for evasion as well as detection of domain name system-based botnets, in addition to this future research directions for detection as well as mitigation of such botnets is also being pointed out[18].

As we know that peer to peer botnets is one of the crucial security threats since they serve as a platform for different malicious activities. A method named peer hunter was presented by researchers which was based on community behavior analysis and it can perform botnet detection through a peer-to-peer structure. This method initiates from a P2P host detection part and then it utilizes contacts which are mutual as the main characteristic for clustering the bots into communities. At last, for the detection of potential botnet communities and also for identification of bot candidates, it uses community behavior analysis. It has been shown that this method of peer hunter can help in achieving high rate of botnet detection as well as low false positives[19].

A new detection technique which was based on nodes topological feature within a graph like clustering coefficient, in degree,in degree weight, out degree, out degree weight as well as eign vector centrality was proposed. In order to establish clusters of nodes, self-organizing map-based clustering method was applied based on those topological features in network. The researchers explained that their technique can isolate bots in small cluster sizes while keeping the most of the normal nodes in the same big cluster environment. This makes detection easy since the searching is being done on a limited number of nodes. In addition to this method, a filtration technique is also developed which increase the efficiency of algorithm by removing those nodes which are not active[20].

In IoT environment, botnets are considered as a major threat w.r.t. to security of IoT devices. Distributed denial of service attack is being performed by those machines(bots) who have been compromised. A solution has been proposed to detect activity of bots in IoT devices and environment. Deep learning technique is applied for development of detection model based on BLSTM-RNN i.e., Bidirectional Long Short-Term Memory based Recurrent Neural Network (BLSTM-RNN. This technique is being compared with LSTM-RNN method for the detection of those attack vectors which are used by very famous mirai botnet, and it is also evaluated for loss and accuracy. The researchers demonstrated that it increased processing time, but it turns out to be a progressive model over time[21].

Machine Learning field can also play an important role in detection of botnets. Due to the changes in botnet command and control methods as well as forms, manual selection of features becomes a difficult task. As a solution to this problem, researchers proposed deep learning-based system i.e., Bot Catcher for detection of botnets. This system can extract features from time and space dimension automatically and establishes classifier by multiple neural network constructions. It has no dependency on prior knowledge about topology and protocol and it performs its task without selecting features manually. It was shown through experimental results that this model is capable of performing botnet detection as well as identification of botnet traffic[22].

A survey was presented for new domain name system-based botnet detection methods classification, and it also provided an analysis of each of the methods with this category[23].

Researchers proposed an unsupervised evolutionary technique for detection of IoT botnet. This method detected those botnet attacks in IoT environment who have been launched from compromised machines or devices by means of exploiting Grey Wolf Optimization

algorithm (GWO) for the purpose of OCSVM parameters optimization as well as simultaneously identifying those characteristics which describe IoT botnet issue. Performance of this technique was evaluated through measures of anomaly detection over a new dataset. The technique was shown performing better as compared to other techniques w.r.t. false and true positive rate as well as G-mean for all IoT devices. Lowest botnet detection time was being achieved in addition to minimizing the number of selected features in this research work[24].

A comprehensive survey was presented for detection of malware in android environment using machine learning. A detailed background based on android applications involving the system architecture as well as security mechanisms and android malware classification was introduced. By using machine learning approach, the authors analyzed and summarized status of research by focusing on data preprocessing, selection of feature, sample acquisition, models of machine learning as well as effectiveness of evaluation of detection[25].

A technique was proposed for detection of new botnets, and it can be implemented on network side as well as on host side. An algorithm i.e., HANABot is used for extraction and preprocessing of features for differentiating the legitimate behavior and botnet behavior. The proposed solution is evaluated by using real data sets for both environments i.e., legitimate and malicious. High accuracy and low false positive rate are being shown in results,also existing methods are compared by focusing on their performance as well as some key features[26].

For detection of botnet in the IoT environment, CNN based deep learning model is proposed which has and 8-layer CNN as well as module for data processing. Before this model is applied, collected power consumption data is segmented and normalized so that their CNN model can reach high accuracy level. Processed data is classified into four class out of which botnet is one of the class and it is the main target. For performance demonstration, three IoT devices were taken and tests like cross-device evaluation, self-evaluation as well as leave-one-device out and leave-one-botnet-out are being conducted. The IoT devices were router, security cameras and voice assistant device. The classification accuracy achieved by these tests for detection of botnet was 90 % for cross-evaluation,96.5% for self-evaluation and 90% accuracy for leave-one-out[27].

Another method for detection of botnet in the IoT environment was proposed by researchers, this approach had three parts, in the first part a high-level PSI-rooted subgraph-based feature is presented and then some features are generated which require less space, have precise behavioral description, and also minimize the processing time. At last, in third part, strength as well as the usefulness of PSI-rooted subgraph-based characteristics, with classifiers involving decision tree, bagging, random forest as well as support vector, each one of these get more than 97% rate of detection and also turns out to consume less time for detection[28].

A survey was presented in which the life cycle of botnet, mechanism features of botnet architecture ,command and control channel were studied and presented classification of botnet detection methods. Main focus was on use of advanced technologies like complex network, swarm intelligence, deep learning and software defined networking. A botnet detection evaluation system(CBDES) is proposed by the authors based on four different points i.e., intelligence,

collaboration, service and assistant. This survey has proposed quantitative evaluation and presented a visual representation of the detection techniques[29].

4. Conclusion

Botnets continue to be a major threat in internet. A good amount of research has been done on techniques for botnet detection. This survey introduces the concept of botnet, its working strategy and summarized the latest research done in the field of botnet detection. One after another new botnets are emerging in the world of internet today, therefore not only understanding the botnet environment, but it is also important to put emphasis on various technologies with their comprehensive applications and it can be the focus of research in the future. The survey done in this paper is of great help for those security personals who wants to analyze and fight against botnets, and it can also help the research community to focus on better tools and techniques for remediation as well as mitigation of botnet threats.

References

[1] B.Fang,X.Cui,andW.Wang,"Surveyofbotnets,"Journal of Computer Research and Development, vol. 48, no. 8, pp. 1315–1331, 2011, (in Chinese).

[2] G.Vormayr,T.Zseby,andJ.Fabini,"Botnet communication patterns,"IEEECommunicationsSurveys&Tutorials,vol.19, no. 4, pp. 2768–2796, 2017.

[3] A. Karim, R. B. Salleh, M. Shiraz et al., "Botnet detection techniques: review, future trends, and issues," Journal of Zhejiang University Science, vol. 15, no. 11, pp. 943–983, 2014.

[4] M.CasenoveandA.Miraglia,"Botnetovertor:theillusionof hiding,"inProceedingsofthe6thinternationalconferenceon cyber conflict, CyCon 2014, tallinn,Estonia, pp. 273–282, Tallinn, Estoni, June 2014.

[5] T. Curran and D. Geist, "Using the bitcoin blockchain as a botnet resilience mechanism," 2016, https://www.os3.nl/ media/2016-2017/courses/ot/dana/tom.pdf.

[6] A. Kurt, E. Erdin, M. Cebe, K. Akkaya, and A. S. Uluagac, "LNBot:acoverthybridbotnetonbitcoinlightningnetwork for fun and profit," in Computer Security – ESORICS 2020. ESORICS 2020, L. Chen, N. Li, K. Liang, and S. Schneider, Eds., Springer, Berlin, Germany, 2020.

[7] Li, Chao, Wei Jiang, and Xin Zou. "Botnet: Survey and case study." In 2009 Fourth International Conference on Innovative Computing, Information and Control (ICICIC), pp. 1184-1187. IEEE, 2009.

[8] Tyagi, Amit Kumar, and G. Aghila. "A wide scale survey on botnet." International Journal of Computer Applications 34, no. 9 (2011): 10-23.

[9] Mailewa, Akalanka, Jayantha Herath, and Susantha Herath. "A Survey of Effective and Efficient Software Testing." In The Midwest Instruction and Computing Symposium. Retrieved from http://www.micsymposium.org/mics2015/ProceedingsMICS_2015/ Mailewa_2D1_41. pdf. 2015.

[10] Hachem, Nabil, Yosra Ben Mustapha, Gustavo Gonzalez Granadillo, and Herve Debar. "Botnets: lifecycle and taxonomy." In 2011 Conference on Network and Information Systems Security, pp. 1-8. IEEE, 2011.

[11]Saad, Sherif, Issa Traore, Ali Ghorbani, Bassam Sayed, David Zhao, Wei Lu, John Felix, and Payman Hakimian. "Detecting P2P botnets through network behavior analysis and machine learning." In 2011 Ninth Annual International Conference on Privacy, Security and Trust, pp. 174-180. IEEE, 2011.

[12] Zeidanloo, Hossein Rouhani, and Azizah Abdul Manaf. "Botnet command and control mechanisms." In 2009 Second International Conference on Computer and Electrical Engineering, vol. 1, pp. 564-568. IEEE, 2009.

[13] Rahimipour, Maryam, and Shahram Jamali. "A Survey on Botnets and Web-based Botnet Characteristics." International Journal of Science, Engineering and Computer Technology 4, no. 11 (2014): 282.

[14] Bilge, Balzarotti, Robertson, Kirda and Kruegel,"Disclosure: Detecting Botnet Command and Control Servers Through Large-Scale NetFlow Analysis." ACSAC '12: Proceedings of the 28th Annual Computer Security Applications Conference December 2012 Pages 129–138.

[15] P. Torres, C. Catania, S. Garcia, and C. G. Garino, "An analysis of recurrent neural networks for botnet detection behavior," in Biennial Congress of Argentina (ARGENCON), Springer, Berlin, Germany, 2016.

[16] K. Guang, G. Tang, S. Wang, H. Song, and Y. Bian, "Using deep learning for detecting Bot cloud," Journal of Communications, vol. 37, no. 11, pp. 114–128, 2016.

[17] J. Wang and I. C. Paschalidis, "Botnet detection based on anomaly and community detection," IEEE Transactions on Control of Network Systems, vol. 4, no. 2, pp. 392–404, 2017.

[18] X. Li, J. Wang, and X. Zhang, "Botnet detection technology based on DNS," Future Internet, vol. 9, no. 4, p. 55, 2017.

[19] D. Zhuang and J. M. Chang, "PeerHunter: detecting peer-to-peer botnets through community behavior analysis," in Proceedings of the 2017 IEEE Conference on Dependable and Secure Computing, pp. 493–500, Taipei, China, September 2017.

[20] Sudipta Chowdhury, Mojtaba Khanzadeh, Ravi Akula, Fangyan Zhang, Song Zhang, Hugh Medal, Mohammad Marufuzzaman & Linkan Bian." Botnet detection using graph-based feature clustering". Journal of Big Data volume 4,2017.

[21] C. D. McDermott, F. Majdani, and A. V. Petrovski, "Botnet detection in the internet of things using deep learning approaches," in Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, December 2018.

[22] Di WU,Binxing FANG,Xiang CUI,Qixu LIU. "BotCatcher:botnet detection system based on deep learning". Journal on Communications, 2018, 39(8): 18-28.

[23] M. Singh, M. Singh, and S. Kaur, "Issues and challenges in DNS based botnet detection: a survey," Computers & Security, vol. 86, pp. 28–52, 2019.

[24] A. Al Shorman, H. Faris, and I. Aljarah, "Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection," Journal of Ambient Intelligence and Humanized Computing, vol. 11, pp. 2809–2825, 2020.

[25] Kaijun liu1,, shengwei xu, guoai xu, miao zhang, dawei sun, and haifeng liu." A Review of Android Malware Detection Approaches based on Machine Learning". IEEE Access,2020.

[26] S. Almutairi, S. Mahfoudh, S. Almutairi, and J. S. Alowibdi, “Hybrid botnet detection based on host and network analysis,” Journal of Computer Networks and Communications, vol. 2020, Article ID 9024726, 16 pages, 2020.

[27] W. Jung, H. yang, M. Zhao, L. Sun, and G. Zhou, “IoT botnet detection via power consumption modeling,” Smart Health Smart Health, vol. 15, Article ID 100103, 2020.

[28] H.-T. Nguyen, Q.-D. Ngo, D.-H. Nguyen et al., “PSI-rooted subgraph: a novel feature for iot botnet detection using classifier algorithms,” ICT Express, vol. 42, 2020.

[29] Ying Xing & Hui Shu & Hao Zhao & Dannong Li & Li Guo, 2021. "Survey on Botnet Detection Techniques: Classification, Methods, and Evaluation," Mathematical Problems in Engineering, Hindawi, vol. 2021, pages 1-24, 2021.