A survey on Malware, Botnets and their detection

Harvinder Singh, Anchit Bijalwan

Department of CSE, Uttaranchal University, Dehradun, Uttarakhand, India

Abstract— *The use of Internet and its related services is increasing day by day. Many million people everyday surf net and use it for various reasons. With so much use of internet, the threats related to security are the major concern of today. There are many security concerns or threats faced by the net surfers and that is because of malwares which have many forms such as viruses, worms, trojans horses, rootkits, botnets and various other forms of data attacks. Among all the threats mentioned above, botnet seems to be quite prevalent now days. It has already spread its roots in Wide Area Network (WAN) such as Internet and continuously spreading at very high pace. Botnet is a network of computers where the computers are infected by installing in them a harmful program. Each computer as a part of Botnet is called a bot or zombie. A Botnet is remotely controlled by a person who commands and controls the bots through a server called command and control sever(C&C). Such person who commands the bots is called a botmaster or bot herder. This paper is written to serve the objective to perform an extensive study of core problem that is the study and detection of Botnets.This paper focuses on the study of malwares where special emphasis is put on botnets and their detection.*

Keywords— Botnets, HTTP, IRC, Malware, P2P, Spam.

I. INTRODUCTION

Over the past few years the internet malwares attacks have grown to an extent that it appears next to impossible to get rid of them. The word malware is derived from malicious software. It is a type of file that contains harmful malcode. Malcode is a malicious code . The malicious codes are distributed to different computers through internet by the use of untrusted websites at an alarming rate. As soon as a malware enters into one's computer system, it starts performing the malware activity and corrupt the entire system. All this activity takes place without the knowledge of the owner of the computer.

Some of the malwares are easily detected and defended through antivirus scanners. But, now a days , the packers pack the malware in such a way that it plays hide and seek with antivirus scanners and malware wins the game. So, it has become a tough task for the antivirus softwares to detect the malwares [2].

Some forms of malwares are viruses, worms, tojans, rootkits, spywares, keyloggers etc. Now a days, botnet is adopted as a medium to launch the malware attacks.

This paper is a study based on malwares and botnets. The paper is organized in the following manner:

Section II explains the different forms of malwares. Section III explains the botnet, its historical overview and botnet phenomenon. Section III explains about different types of botnets or botnet categories. Section IV explains about finding the presence of botnet or detecting the botnets. Section V gives brief conclusion about the paper.

II. BACKGROUND STUDIES

Malware means malicious software, a software with some malicious intent. It enters into the computers without the owner's knowledge. There are different forms of malwares that appear as threat for the internet users.

2.1 Different forms of Malwares

The different forms of malwares that appear as threat for internet users are as follows :

a) Virus : Virus is a type of malware that enters into a computer system without knowledge of the computer user and attaches it to some executable file. It is capable of duplicating itself and can cause harm to other computers also. Its symptoms are , low system performance, data corruption etc.
According to Dr Cohen "A virus is a program that can infect other programs by modifying them to include a possibly evolved version of itself." A virus is by definition a computer program that spreads or duplicates by copying itself. The viruses have tendency to cause infection by performing modification in other programs by including their copies and then further infecting other programs[1] .

b) Worm: A Worm is a standalone malicious software that can operate independently and don't hook itself to propagate. The worms breach the weak security system of computer or network and spread themselves through the storage devices , e-mails etc. The symptoms of worms may be low performance of network, consumption of large amount of memory [2].

A computer worm may be considered similar to computer virus in many ways except it is a self contained program. The fundamental purpose of a worm is to gain access to another computer system so that it can replicate itself on the new machine and reproduce further [3].

c) Trojan: It is a form of malware which appears to be a useful software. It may enter into computers as a part of downloading file from the internet. Trojan horse keeps track of user activity, steals passwords, login details, deletion of files etc.
A Trojan horse is an executable file in the Winows Operating System. These executable files have certain peculiar characteristics. Multiple Windows system process will be called whenever a Trojan horse tries to execute any operation on the system[4].

d) Rootkit: It is a kind of malware disguised as a useful program. Its actual identity is concealed from the virus removal programs. It gets installed through Trojan and is involved in password stealing, recording keystrokes on keyboards. Rootkits hide the malicious program from the system's process list and try to avoid detection by antivirus program [5].

e) Spyware: A spyware is a form of malware that keeps track of user's activity without his consent and sends back the sensitive data to its creator. It may enter into a computer system as a part of freeware installation. It is a class of malicious code that is surreptitiously installed on victim's machine. Once active, it silently monitors the behavior of users, records their web surfing habits and steals their password [6].

f) Keyloggers: It is another form of malware which is a type of spyware. It secretly records the keystrokes as tapped by the user. It reads cookies and gathers the personal information. Keyloggers steal the usernames and passwords, credit card numbers, online banking details etc.
The keyloggers can be installed by gaining physical access to the computer or by downloaded programs. Their small footprint in terms of memory and processor utilization makes them practically untraceable. Keyloggers can email the file containing keystrokes back to a spying person [7].

g) Botnet: A network of compromised hosts that are remotely controlled by a master is called a botnet [8]. Botnet is a collection of infected computers that receive instructions from the botmaster, who is a corrupt hacker and uses the botnet for causing destruction or getting financial gains. Any computer can be compromised and taken as part of botnet if it has a weak security system.

2.2 Botnets

Botnets are emerging as the most serious threats against cyber security. A botnet is a group of infected end hosts under the command of a botmaster [9].

Botnet stands for Robot Network. It is a network of compromised machines that are infected with malicious programs that can be remotely controlled by an attacker through a command and control (C&C) architecture on IRC(Internet Relay Chat) channel or peer to peer network. Botnets most often consist of thousands of compromised machines which enable the attacker to cause a serious damage. Some terms related to Botnet are :

1. Bot: Bot is a malicious software program that can be installed on victim machine without the knowledge of owner. It is a self propagating application.
2. Command and Control (C&C): It is the channel used to manage a botnet. It may be thought of as a private infrastructure which can be used for malicious purpose. Bots are updated and directed through C&C.
3. Botmaster: Botmaster or botherder is the person or hacker behind the botnet. The group of compromised computers are controlled by one or group of attacker known as Botmaster [10]. He commands and controls the botnet for causing damage to the data and for financial gains.

Botnets are used for all DDOS(Distributed Denial Of Service) attacks, Spam, click fraud, information theft, phising attack, and distribution of other malware.

2.3 Historical overview

A botnet is a network of infected machines also called bots, which aims to distribute the malicious code over the internet without user intervention. The purpose of entire botnet is to increase the bot army for intentional destructive tasks. The difference between botnet and other types of network attacks is the existence of Command and Control(C&C) [12]. A botnet causes a number of serious offences on the internet; as it allows intruders to hijack several computers simultaneously (Paxton, Ahn et al 2011) [13].

The concept of botnet was evolved in 1993 by introducing the first botnet by the name Eggdrop(X wang 2003). Then , GTBOT and NetBus in 1998, SdBot and AgoBot in 2002, SpyBot and Sinit in 2003, Bobax and Bagle in 2004, Rustock in 2006, Cutwail and Srizbi in 2007, (conficker, mariposa, sality, Asprox, waledac, krakren) in 2008, (Maazben, Grum, Festi, Wopla, Zeus) in 2009, (Kelihos, TDL4, lowsec, Gheg) in 2010, Flashback in 2011, Chameleon in 2012, Boatnet in 2013 and many more botnets appeared quickly.

The sizes of botnets are varying from 10000 bots to 30,000,000 bots [12].

1.4 Life cycle of Botnet

The life cycle of a botnet is planned and well organized. The life cycle of a botnet from its inception to propagation is divided into series of steps that are as follows :-

1. The botmaster configures starting bot binaries.
2. The botmaster registers DNS space.
3. The static IP Address is being registered.
4. The botmaster starts victimizing or compromising the machines by different means.
5. The propagation of bots take place.
6. The bots start becoming the part of botnet using C&C server.
7. Bots are used for malicious activity.
8. Bots are continuously upgraded and updated by the botmaster by running specialized programs.

A typical advanced botnet is formed in five stages : Initial infection, secondary infection, connection, malicious C&C and finally update and maintenance [14].

In initial infection, the weakness or vulnerabilities of victim machines are exploited and machine gets infected. In secondary infection , the malcode or shellcode is executed on the victim machine which fetches the image of bot binary to get installed on the machine. In connection, the bot binary establishes command and control channel. Im malicious C&C stage, the C&C channel is used by the botmaster to send the commands and directions to bots or victim machines. In the final and last stage that is update and maintenance the botmaster requires to upgrade or update the bots for different types of purposes.

The defining characteristic of botnet is that each bot is controlled through the commands sent by the botmaster. The communication channel used to issue commands can be implemented using a variety of protocols eg.(HTTP,P2P etc). But the majority of botnets now a days use the IRC(Internet Relay Chat) protocol [15]. Upon initialization , each bot tries to communicate with the IRC server through the address given in the shell code. In many cases the DNS name resolving is done for the IRC server. As soon as the IP address of IRC server is obtained , the bot establishes an IRC session with the IRC server and joins the C&C channel as specified in the bot binary.

A bot, in order to communicate with an IRC server is required to prove its authenticity and hence authenticates itself by following different techniques.

LIFE CYCLE OF BOTNET

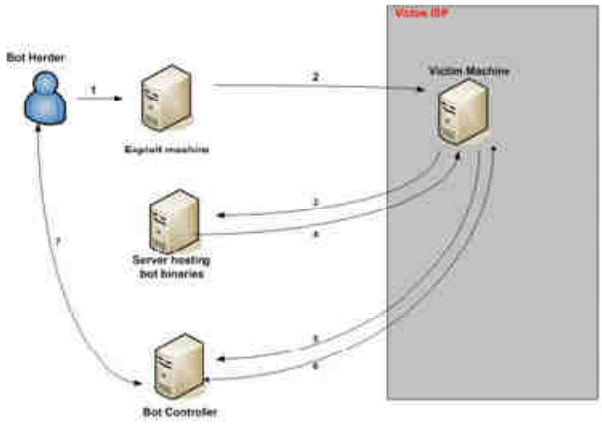

Fig. 1:(Life Cycle)

III. CATEGORIES OF BOTNETS

There are two main categories of botnets on the basis of command and control channel used, the Centralized model and the Decentralized model. The further categories of decentralized model are :

1. IRC(Internet Relay Chat) Botnet
2. P2P(Peer to Peer) Botnet
3. HTTP(Hyper Text Transfer Protocol) Botnet
4. Hybrid Botnet

Now a days botmasters also use SMS and Bluetooth as the command and control channel to perform malicious tasks in smart mobile phones . such types of botnets are called mobile botnets. A new technology that is cloud technology is also used in setting up botnets. These types of botnets are called cloud based botnets. The above mentioned botnet categories are classified under centralized and decentralized

3.1 Centralized Botnet

The centralized botnet is a type of botnet structure in which there is a centralized command and control structure. In this type of botnet there is a centralized server through which the commands are sent to the bots. Each bot machine is connected to the C&C server. In case the C&C server stops working the entire botnet is failed.

The botnets with centralized architecture provide a simple ,low latency,anonymous and efficient real time communication platform for the botnet controllers. Most of the latest detected large scale botnets are based on centralized structure with HTTP or customized protocols [16]. Example : IRC and HTTP botnets.

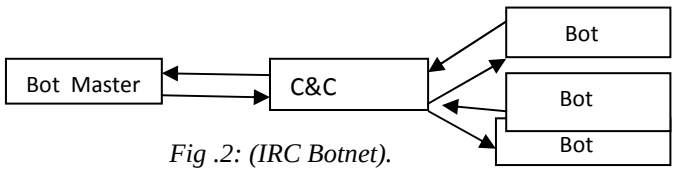

Fig .2: (IRC Botnet).

3.2 Decentralized Botnet

In decentralized botnet there is no central command and control server. Each bot is connected to another and further connected to botmaster. It is very difficult to shut down the decentralized botnet due to its structure. Each bot in this type of structure acts as a client as well as server. Example P2P botnet.

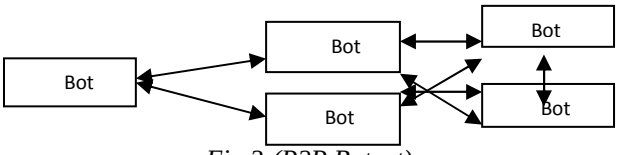

Fig.3:(P2P Botnet)

3.2.1 IRC Botnet

In this type of botnet, the botherder uses IRC as the C&C channel to command and control the bot machines. Once the bots receive commands from the botmaster through IRC server the individual bots start the malicious activity. The entire botnet can stop working if the IRC server is collapsed. In order to send the command to a particular bot, the botmaster first verifies the username and password. Once the verification is completed then only the commands are given to bots to perform the desired task. The IRC is a form of real time Internet text messaging or synchronous conferencing. The protocol is based on the client server model, which can be used on many computers in distributed networks [17].

3.2.2 P2P Botnet

In this type of botnet the P2P protocol is used. It is a decentralized combination of nodes. Each bot in this structure behaves both as the client as well as server. A special type of search key is used by the botmaster to send commands to different bots. If bots in this type ob botnet are taken offline, the botnet can still continue to operate under the control of Botmaster [18].

3.2.3 HTTP Botnet

It is a type of centralized botnet which uses HTTP protocol as the command and control server. The malicious intent of the botherders is actually hidden along with the normal data traffic and are not caught by the antivirus , firewalls etc. A particular IP address is used by the botherder to make connection which also works as C&C server. The HTTP botnets are largely used by the hackers for phising acts and financial crimes. The HTTP bots frequently demand and download instructions from web servers under the attacker's control. As a result, detecting bots with web based controlling is complex than bots with IRC based controlling[19].

3.2.4 Hybrid Botnet

A botnet formed by combining the features of two or more known botnets is called hybrid botnet. The hybrid botnet is formed by combining the centralized as well as decentralized botnets. As per Anchit Bijalwan et.al in [26] a hybrid botnet is divided into servant and client bot. The servant bot receives the commands from the bot herder and forwards it to the clients.

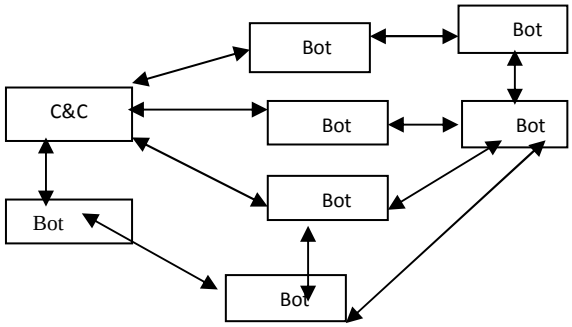

Fig. 4: (Hybrid botnet)

IV. BOTNET DETECTION

The detection of botnet has always been a big challenge to the organizations and individuals. It is very difficult to detect the presence of bot or botnet. To detect a botnet actually requires the use of advanced analyzing capabilities. The two approaches used for detection of botnets include:

1. Setting up honeynets
2. Passive traffic monitoring
   (a) Signature based detection
   (b) Anamoly based detection
   (c) DNS based detection
   (d) Mining based detection

4.1 Setting up Honeynet

A honeynet or honeypot can be thought of as a system in which the weaknesses or vulnerabilities are intentionally injected and then such systems are monitored for attracting the attacks and intrusions. It is a computer system that is used to trap to draw the attention to attach this computer. Such computers have a strong ability to detect security threats, to collect malware signatures and to understand the motive and method behind the threat used by the botmaster [21]. The honeypot method is not very successful or strong method as we have to wait until a bot infects the system.

4.2 Passive Traffic Monitoring

It means the data traffic movement is being monitored and the trails of intrusion are tied to be deleted. It has four categories:

4.2.1 Signature based detection

In this approach of detecting the botnets , the help of known malware is taken. The network traffic is thoroughly being monitored to detect yhe marks of intrusion. It is a rule based method , which detects the harmful traffic fitting into the rule. This detection technique can only be employed for detecting the Botnets

that are the known ones. The fundamental approach is to extract feature information on the packets from the data traffic and match the patterns registered in the knowledge base of existing bots. It has several disadvantages:

1. It can't identify the unidentified bots.
2. It should always update the knowledge base with new signatures.
3. The new bots may launch attacks before knowledge base are patched.

Examples are snort, Rishi and NEDRS etc. [22]

4.2.2 Anomaly based detection

This approach is used to detect the botnets that are unknown. In this technique the anomalies present in the network traffic are observed to predict about the presence of bot. The various anomalies could be high network latency, high volume of traffic , traffic on unusual ports and unexpected system working etc. The purpose of anomaly based detection is to find the signs that are different from the other available details. Bijalwan et.al in [23] identified UDP bot flooding through the lab experiments.

4.2.3 DNS based detection

The DNS based approach is a kind of passive technique. In such techniques there is full transparency but are unknown to botmasters. DNS based approach is based on the property that in order to access the C&C server, bots carry out DNS queries to locate the particular C&C server that is typically hosted by DDNS(Dynamic DNS) provider. So DNS monitoring will be easy approach to detect Botnet DNS traffic and detect DNS traffic anomalies. This is most famous and easy technique of botnet detection [24].

4.2.4 Mining based detection

The data mining based technique helps in recognizing the useful patterns to find out certain type of regularities and irregularities in available sets of data. Data mining techniques can be used for the purpose of optimization. In this method the sufficient amount of data is available from the network log file to work upon and analyse. The various data mining methods are correlation, classification, clustering, statistical analysis and aggregation for extracting the useful information from the available data[25].

V. CONCLUSION

This paper is a thorough study and analysis of malware and their categories. In this research based exhaustive survey I have tried my level best to explain botnet, its formation and working. The purpose behind the formation is also covered to greater extent. I have also tried to throw some light upon the different types of botnets and their behavior. The different techniques used to detect botnets are also discussed. Even though some detection techniques are available but still the botnets are big challenge to the society. The field requires a lot of research so that a concrete solution should be found to fight with the challenge and mitigate its impact.

REFERENCES

[1] Manoj Kumar Dhruv,Yogita Dewangan, Purushottam Patel, "An introduction to Computer Virus, History and its evolution" in International Journal of Research, vol 03, issue 04, Feb 2016.

[2] Dolly Uppal, Vishakha Mehra and Vinod Verma,"Basic svrvey on malware analysis, tools and techniques", in IJCSA, vol4,no. 1, feb 2014.

[3] Munna Kumar et al., "Predator-Prey models on Interaction between Computer worms, Trojan horses and anti virus software inside a computer system", International Journal of Security and its Applications", vol 10, No 1 (2016) P. 173-190

[4] Prof.Abuzneid Abdelshakour et al., "Detection of rojan Horse by Analysis of System Behaviour and Data Packets", Dept of CS, University of Bridgeport, CT

[5] Ishita Basu et al.,"Malware detection based on some data using data mining : A survey", Ametican Journal of Advanced Computing, vol3910, p.18-37

[6] Mannel Egele et al.,"Dynamic Spyware Analysis", 2007, USENIX Annual Technical Conference, 2007.(p. 233-246)

[7] Kishore Subramanyam et al., " Keyloggers : The overlooked threat to Computer Security", Dept. of Mathematics & CSE, Northern Kentucly University, KY-41099

[8] Fariba Hadadi et al.," On the effectiveness of Different Botnet detection approaches" Springer International publishing, Switzerland 2015.

[9] Moheeb Abu Rajab et.al, "A multifaceted approach to understanding the botnet phenomenon",Rio de Janeiro, Brazil, ACM, IMC 06, 2006.

[10] Haritha S Nair,Vinedh Ewards, " A study on Botnet Detection Techniques", International Journal of Scientific and Research Publications, vol 2, issue 4, April 2012

[11] Parmar Riya H, Harshita Kanani, "A botnet detection techniques", in ISRJ vol 4, issue 4, May 2014.

[12] Karim et.al, "Botnet detection techniques,review, future trends and issues", journal of Jhejiang University", Jan 2014.

[13] Napolean C.Paxton et al, "Master Blaster : Identifying influential players in Botnet Transactions", in 35th IEEE conference, 2011

[14] Fariba Haddadi et al, "On Botnet behavior, Analysis using GP and C45", Faculty of CS, Dalhousie University, Canada.
[15] C.Calt Internet relay Chat : Client Protocol, RFC 2812,(Informational), April 2000.
[16] Wang, Tao, and Shun Zheng Yu, "Centralized botnet detection by traffic aggregation", Parallel and distributed processing with applications, 2009, IEEE, International Symposium on IEEE 2009.
[17] Vania, Jignesh, Arvind Meniya and H.B Jethra. "A review on botnet and detection technique", International Journal on Computer Trends Technology, vol 4, no.1(2013) pg 23-29
[18] Prabhu, S Nagendra and D Shanthi, "A survey an anomaly detection of Botnet in network", International Journal 2.1(2014)
[19] Sultan Mohd Shahid, " Monitoring HTTP based command and control Botnets in Traffic using Bot sniffer", Diss, Texas A7M University Corpus Christi 2015.
[20] D. Seerinivasan, K Shanthi, "Categories of Botnet : A Survey", in IJCEACIE, vol:8, No.9, 2014.
[21] Jignesh Vania et al, " A review on Botnet and Detection Techniques" in IJCTT, vol 4 issue 1, 2013
[22] Ghafir, Ibrahim, Jakob Svoboda, and Vaclav Prenosil, " A svrvey on Botnet command and control Traffic Detection", International Journal of Advances in Computer Networks and its security(ICJNS)5(1)(2015).
[23] Bijalwan A, Wazid M, Pilli ES, Joshi R C, ' Forensics of random – UDP flooding attacks" in Journal of Networks. 2015 May 27,10(5): 287-293
[24] Amit Dange, Prashant Gosavi, " Botnet Detection through DNS based approach", in IJAIEM, vol2, issue 6, June 2013.
[25] Alireja Shahrestani et al.,"Architecture for applying data mining and visualization on network flow for botnet traffic detection",IEEE, pages 33-37, 2009
[26] Anchit Bijalwan et al., " Survey and Research Challenges of Botnet Forensics", in IJCA, Vol 75-No 7, Aug 2013