A Survey of Botnet Technology and Detection

Fatima Naseem, Mariam shafqat, Umbreen Sabir, Asim Shahzad
fatima@uettaxila.edu.pk, ,mariam@uettaxila.edu.pk, ,umbreen@uettaxila.edu.pk, asimshahzad@uettaxila.edu.pk

Department of Computer Engineering
University of Engineering and Technology,
Taxila, Pakistan 47040

Abstract

Apart from viruses, worms, Trojan horses, and network intrusions; there is a less familiar and exponentially growing threat that tends to be more disastrous: Botnets. The target of the botnet attacks on the integrity and resources of users might be multifarious; including the teenagers evidencing their hacking skills to organized criminal syndicates, disabling the infrastructure and causing financial damage to organizations and governments. In this context, it is crucial to know in what ways the system could be targeted. The major advantage of this classification is to identify the problem and find the specific ways of defense and recovery. This paper aims to provide a concise overview of major existing types of botnets on the basis of attacking techniques.

1. Introduction

The technological advancements are pushing the human life towards ease and trouble simultaneously. Emerging information technologies have made access to information so easy that was never before. But on the other hand, it has worsened the security level.

BOTNETS are proving to be the most recent and disastrous threat to the field of information technology.

The understanding of a layman about Botnets is that it is a network facilitating the malicious attacks on the user machines but technically speaking “Botnets are a collection of computers

on which ,a software, ‘bot’, is automatically installed without user intervention and are remotely controlled via command and control server” [8]. Despite of the fact that this network can be implied both for nefarious and beneficial purposes, its extensive deployment in the criminal and destructive purposes has made the title ‘botnets’ tantamount to malware.

An active Botnet initializes its attack by first exploiting vulnerabilities in the user computers. It then downloads the malicious binary and executes it locally. This program logs on to the Command and Control Server (C & C) and notifies its Host, commonly known as ‘Botmaster’ or ‘Botherder’, that the computer is now converted to a ‘Bot’. It can now be used to forward its affect to other computers by repeating the same procedure [9]. The major difference between botnets and other security threats is that a botmaster communicates regularly with the bots either via centralized communication channel or decentralized network. These bots perform any type of destruction on receiving the commands from the botmaster [13]. These botmasters send the commands, control all the bots, and then can attack a victim as a unit.

Botnets are developing at a very fast rate making it difficult to detect and recover from their side effects. However, some of their types extensively deployed can be classified to

provide for their remedy. This paper mainly deals with three major types of botnets: IRC botnets, peer-to-peer and HTTP botnets and suggests some techniques to identify and detect them.

Section 1 gives an introduction of botnets. Section 2 reviews their history. Botnet Command and Control is considered in Section 3. Section 4 analyzes Botnet life cycle. Three major types of botnets and their detection scenarios are considered in Section 5, 6 and 7 respectively. Section 8 proposes some of expected advances in this particular field as future work. Section 9 is dedicated to the overall conclusion of our study.

2. History of Malicious Bots

Before the evolution of Botnets; the major sources of malware were viruses, worms, Trojan Horses that used to affect only a single machine. With the evolution of Botnets; the concept of destruction was enhanced from a single machine to a network as a whole. The history of undertaking botnets for destruction roughly dates back to 1990. Prior to this, botnets were the major sources of maintaining control of the IRC channels. Their mischievous applications mainly took advantage of the centralized control of IRC for command and control. But centralized control structure was relatively easy to discover and track. Due to insecure nature of IRC botnets; they completely changed their structure form centralized to a peer-to-peer nature, which is a decentralized control structure. This ultimately makes it much harder to spy the communication among the bots and to track their origin. The most recent improvement is again the implementation of centralized C&C in HTTP botnets; but here the distinguishing feature is that the Botnets periodically connect and disconnect with the bot master. This further aggravates the problem of detection [6].

3. Command and Control Channel

The backbone of botnet is command and control channel; which is responsible for setting up the botnet, controlling the activities of the bots, issuing commands, and ultimately reaching the goals [13].

The command and control channel is stable during the operation of botnets i.e. once a botnet is established; the command and control channel remain the same throughout its operation. But on the other hand, once a C&C channel is detected; then the whole botnet is exposed.

4. Botnet Life Cycle

The success of any process mainly lies in how well the sequence of steps is organized. The major reason of dramatic success and spread of botnets is their well organized and planned formation, generation and propagation. The lifecycle of a botnet from its birth to disastrous spread undergoes the following phases [11], [12]:

1. Bot-herder configures initial bot parameters.
2. Registers a DDNS.
3. Register a static IP.
4. Bot-herder starts infecting victim machines either directly through network or indirectly through user interaction.
5. Bots spread.
6. Bot joins the Botnet through C&C server.
7. Bots are used for some activity (DDoS, Identity Theft etc.)
8. Bots are updated through their Bot-operator which issues update commands.

These sequential steps make botnets more unavoidable and difficult to capture. Subsequently they become more successful and devastating.

5. Types Of Botnets

There is a variety of botnets causing the mass destruction. As already discussed in section II, the three major categories that we have considered in our study depend on the type of command and control they are based on [6], [13]. They are as follows:

- IRC botnets
- P2P botnets
- HTTP botnets

Now we will consider each one of them to briefly view their operation and detection mechanism.

5.1 IRC

The IRC (Internet Relay Chat) protocol was initially designed for real-time Internet text messaging. The building ground of IRC is TCP/IP protocol. It works by making a central location and then all the required users (clients) connect to that central location; and that central location is called ***server;*** while anything except server is called ***client.*** Clients are distinguished from each other by their nickname; which is a string composed of 9 characters. Any server must know the real name of the host the client is running on, the username of the host the client is running on, the user name of client on that host, and the corresponding server [13].

As IRC came into extensive use several variations in the protocol and structure were adopted. Automated clients called bots emerged as a new concept and the success was obvious. They served as a permanent point of contact for information exchange. With their popularity, their deployment in several unexpected tasks increased manifold. One of these was the emergence of botnets for nefarious purposes.

This emergence grew into a massive network that allow its operators to use it for running games, file distribution, or use it for user misbehavior.

[1] the most vulnerable feature of an IRC is its server. The IRC channel operator is connected to this server. If the server is crashed due to some reason, then the connection of this operator would automatically die and another member from the same channel would automatically be assigned the server status. This behavior proved to be disastrous, and allowed any user to snatch the server's honor, and therefore use the channel according to its own will.

The IRC bot is an assembly of programmed codes that behave as a client in an IRC channel. But unlike the traditional clients providing interactive access, it performs self-propelled functions.

The key feature of pioneer legitimate IRC bots called botnets; was to allow secure assignment of privileges between bots, sharing of user/ban lists and to control floods. This allowed the IRC operators to utilize the congregated power of many modules of bots together.

5.2 IRC Detection Techniques

A lot of techniques have been proposed for IRC Botnet detection. The basis of all these techniques is hounding of packets either at network layer or application layer.

In [3] the mechanism of detection is suggested on the network layer level. Here the hierarchy between routers and the IRC server is explored in bottom-up manner i.e. the tracking initiates from the victim and follows the path of infecting routers till the origin (bot-herder).

The author has proposed a frame work in [2], which sniffs the network traffic, filters it on the basis of application layer protocol, and then segregates them into either righteous or saboteur IRC traffic just by contemplating the IRC chat contents. The separating foundation between a normal human and botnet conversation is that the human language is alternating while the Botnet conversation is repeating.

[4] presents a pipelined approach which accomplishes the detection procedure in a number of steps. First it separates the black and white list traffic based on the DNS queries; this separated traffic is classified according to applications i.e. extract chat-like traffic. Next pair wise correlation of the traffic flows is done to identify similar traffic considering it to be originating from same botnet.

The study of these IRC detection techniques reveals that choice of the suitable detection technique depends on the required scenario. If the solution has to be managed at the network layer, [3] serves as the best option; while on application layer [4] and [2] serve the purpose. Regardless of their applications, each technique has its respective shortcomings which leave a large room for further suggestions and research.

6. P2P Botnets

Preliminary botnet architecture was based upon centralized architecture but that was much prone to detection; as the entire botnet can be apprehended just by tracking down a single central command.

To overcome this drawback, a rather new technology in the field of Botnets is peer-to-peer Botnets; where a peer (host) can act as both client and server alternatively. To enter the network a peer can connect to any other peer of the network using its IP address that was already present in its database. Finally when this peer is part of the network; it continually updates its database by interacting with other peers. Using this approach when any peer tries to send commands to the botnet, it sends a library call to its database to get the addresses of other bots; thus acting as commander and controller of the P2P botnet. This Commander and Controller now sends orders that are to be followed by the remaining peers of the network [5].

To track down a peer-to-peer network, initially the simplest possible solution was for the hacker to enter the botnet by pretending to be a new bot [5]. This newly entered bot will now be able to connect to any other peer of the network and thus be able to track down its activities. The biggest disadvantage of this approach is that the intruder can monitor the activity and thus track down only a single peer; the entire botnet activity can neither be monitored nor can be tracked down immediately. The entire Botnet tracking is obviously a time consuming operation.

7. HTTP Botnets

The most recent Botnet till date is HTTP botnet. It works by exchanging web requests using port 80. It sets up its communication with certain URL's using internet with an HTTP message. This HTTP message contains unique identifiers for the bots. The server under consideration will reply to these HTTP messages with further investigation commands (e.g. GET). This interrogating command ultimately becomes the reason of downloading the infecting malicious commands. Again it uses the centralized command and control channel as IRC botnet uses but a few advantages compared to IRC exists:

- Here the command and control server is web server as compared to IRC botnets where IRC serves as the C&C.
- In IRC bot once connected to C&C doesn't disconnect but here the bots regularly connects with the server after regular intervals of time; which is set by the web server.

The traffic of the HTTP botnets flows with the regular traffic. However, the bot packets are different from normal packets making the detection procedure easy [7]. [6] Discusses the most commonly deployed detection technique for HTTP botnets. Here a degree of periodic repeatability (DPR) is employed. This parameter represents the repeated reconnection of bots with botmaster after regular interval that is configured by the botmaster. The more number of times, same client connects to the same server after same interval of time, depicts greater probability of a client being a bot and server being a botmaster.

More work on several other techniques is underway to timely detect the modern HTTP botnet attacks.

8. Future work

Botnets is a center of inclination for both the attackers and the researchers. This concept evolved two decades ago and proved to be a blitz for internet fraternity in this short period. There seems to be a state of war going on between the botnet attackers and defenders or researchers. The researchers are implementing more advanced and organized strategies to detriment the internet users and researchers are consistently trying to cope with their advances. Being an emergent field there is an open room for research and future work.

Deep analysis of different classifications can lead to one generalized model of botnets. Furthermore, every technique mentioned has false positives and negatives which can be improved.

The most recent issue which has called for the consideration of researchers is that now the botnet herders try to track honeypots by injecting the binary into the network and examine who is spying their activities; thus banning the hackers when they find them out [10].

All this discussion reveals that botnets are still in evolutionary phase and provide a capacious field for research.

9. Conclusion

It is impossible to defy the significance of botnets in the current circumstances. The ravage they have caused to the finances and solidarity of several government and private organizations has devoted attention of the IT specialists to find the remedy.

In this paper we discussed briefly the emergence of botnets, their organization and architecture and botnet life cycle steps. Next the reputed botnet types; the architecture they use and their different possible detection techniques are presented. Although different in architectures, all types of botnets are of great threat to the internet community. They can be used both for good and bad purposes. This paper gives you a roller coaster ride of the botnet world, giving a concise but complete view of different flavors of botnets.

References

[1] John Canavan, "The Evolution of Malicious IRC Bots", *Proceedings of the VB2005 Conference*, April, 2005.

[2] Claudio Mazzariello, " IRC traffic analysis for botnet detection" *Proceedings of Fourth International Conference on Information Assurance and Security*, 2008.

[3] Zhenhua Chi, Zixiang Zhao, "Detecting and Blocking Malicious Traffic Caused by IRC Protocol Based Botnets", *IFIP International Conference on Network and Parallel Computing*, 2007.

[4] W. Timothy Strayer, Robert Walsh, Carl Livadas, and David Lapsley, "Detecting Botnets with Tight Command and Control", *IEEE*, 2006.

[5] Antti Nummipuro, "Detecting P2P Controlled Bots on the Host", *Seminar on Network Security*, Oct. 11, 2007.

[6] Jae-Seo Lee, HyunCheol Jeong, Jun-Hyung Park, Minsoo Kim, Bong-Nam Noh, "The Activity Analysis of Malicious HTTP-based Botnets using Degree of Periodic Repeatability", *IEEE International Conference on Security Technology*, 2008.

[7] Bogdan Botezatu, "Anatomy of a Botnet", *MalwareCity News*, Sep. 15, 2008.

[8] Zhaosheng Zhu, Guohan Lu, Yan Chen, "Botnet Research Survey", *Annual IEEE International Computer Software and Applications Conference*, 2008.

[9] David Barroso, "Botnets – The Silent Threat", *ENISA Position Paper No. 3*, Nov, 2007.

[10] Cliff C. Zou, Ryan Cunningham, "Honeypot – Aware Advanced Botnet Construction and Maintenance", *Proceedings of the 2006 International Conference on Dependable Systems and Networks (DSN'06)*, 2006.

[11] http:// www.wikipedia.com/ Botnets.

[12] Jivesh Govil, Jivika Govil, "Criminology of Botnets and their Detection and Defense Methods", *IEEE EIT 2007 Proceedings*, 2007.

[13] J. Oikarinen, D. Reed, "Internet Relay Chat Protocol", *Network Working Group; Request for Comment (RFC) 1459"*, May, 1993.

[14] Micheal Bailey, Evan Cooke, Farnam Jahanian, Yunjing Xu, Manish Karir, "A Survey of Botnet Technology and Defences", *Cybersecurity Applications & Technology Conference For Homeland Security*, 2009.