Botnet: Survey and Case Study

Chao Li
National Computer network
Emergency Response technical
Team/Coordination Center of China
Beijing, 100029, China

Wei Jiang
Research Center of Computer
Network and Information Security
Technology
Harbin Institute of Technology
Harbin, 150001, China
E-mail: jiangwei @pact518.hit.edu.cn

Xin Zou
National Computer network
Emergency Response technical
Team/Coordination Center of China
Beijing, 100029, China

***Abstract*—Botnet is a hybrid of previous threats integrated with a command and control system and hundreds of millions of computers are infected. Although botnets are widespread development, the research and solutions for botnets are not mature. In this paper, we present an overview of research on botnets. We discuss in detail the botnet and related research including infection mechanism, botnet malicious behavior, command and control models, communication protocols, botnet detection, and botnet defense. We also present a simple case study of IRC-based SpyBot.**

Keywords-security; survey; botnets; bot; C&C mechanism

I. INTRODUCTION

Botnets have become the biggest threats on the Internet and been used for launching attacks and committing fraud. A study shows that, on a typical day, about 40% of the 800 million computers connected to the Internet in a botnet [1]. Those infected machines engage in many illegitimate activities, such as distributing spam, stealing sensitive information, launching denial-of-service attacks, and spreading new infections.

A better research of botnets will help us to develop new technologies to defeat the biggest security threat. In this paper, we provide an overview of current botnets technology research. The rest of the paper is organized as follows. Section II discusses background of botnets. An overview of botnets techniques is presented in Section III. In Section IV, present a simple case study of IRC-based SpyBot.

II. BACKGROUND OF BOTNET

In order to further discussion and better understand botnet, we first introduce some key terms. Then, we present a timeline of bot and botnet evolution which provides insight into their current and future trend.

A. *Related Definitions*

Bot - Compromised computer which waits for commands from the Botmaster.

Botnet - A network of bot under the control of the Botmaster and usually used for malicious activities.

C&C - Command & Control channel, a Botmaster usually uses an IRC channel to send commands.

IRC - Internet Relay Chat, a chat system that provides one-to-one and one-to-many instant messaging over the Internet. Each network has a "channels" for various topics.

B. *Botnet Evolution*

Historically, the bots can be traced their roots to the Eggdrop bot created by Jeff Fisher for assisting in IRC channel management in 1993 [2]. The Eggdrop was well-known and widely used of the time as non-malicious IRC bot. However, IRC bots with more malicious purposes emerged.

The first malicious bot was GT-Bot that we could find were in the April, 1998[3]. At present there are at least a hundred variants of GT-Bot which included an IRC client, mIRC.exe, as part of the bot [4]. PrettyPark worm [5] was the first worm which emerged to make use of IRC as a means of remote control in June 1999. In April 2002, Agobot's source code was published on many Web sites [6]. Slapper [7] was first worm with P2P communications protocol appeared in September, 2002. SDBot [8] began to appear in October, 2002. SDbot variants provided own IRC client for better efficiency. Compared with Agobot, SDBot is a simpler and terser code written in C. The SpyBot [9] was mentioned in the April, 2003. Sinit [10] was early malicious Peer-to-Peer bot using random scanning to find peers emerged in September, 2003. Phatbot [11] was another Peer-to-Peer bot based on WASTE. The appearance of Nugache [12] was documented widely in the anti-virus/ malware community in late April, 2006. In January 2007, the Trojan.Peacomm bot [13] came into our horizon. It is the most recently known peer-to-peer bot and the botnets uses the Overnet peer-to-peer protocol for controlling the bots. Figure 1 shows the timeline of botnet evolution. We can find peer-to-peer bots are now under widespread development.

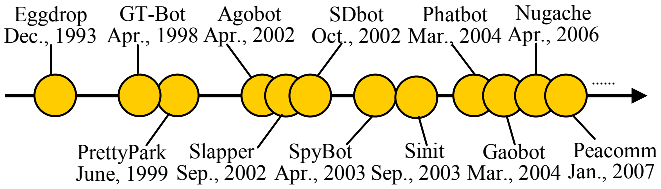

Figure 1. Timeline of botnet evolution

978-0-7695-3873-0/09 $29.00 © 2009 IEEE

III. RESRARCH TAXONOMY FOR BOTNETS

In this paper, we survey the botnet and related research including include: (i) infection mechanism, (ii) botnet malicious behavior, (iii) command and control models, (iv) communication protocols, (v) botnet detection, (vi) botnet defense. An overview of these techniques is shown in Figure 2.

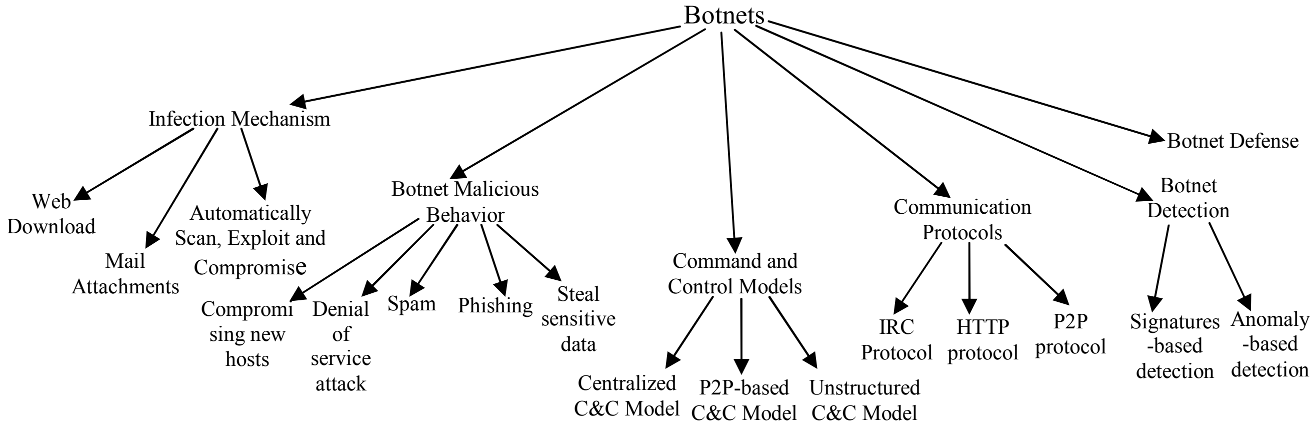

Figure 2. Overview of botnets.

A. *Infection Mechanism*

There are various types of methods for attacker to distribute a particular bot. Basudev et al. [14] discussed three methods of bots propagation.

1) Web Download: A recent google study showed that web-based infection vectors are now commonplace [15]. Web-based malware creates botnet-like structures in which compromised machines query web servers periodically for instructions and updates.

2) Mail Attachments: E-mail attachments with mass-mailing worms can contain bots. Spam techniques simplify and enable fast spreading of bots easily.

3) Automatically Scan, Exploit and Compromise: The bots automatically infect the host that have vulnerabilities.

Rajab et al. [16] summarized the various stages in a typical botnet life-cycle as shown in Figure 3. Botnet usually recruit new victims by remotely exploiting a vulnerability of the victim's machine via mentioned methods above. When the infection has achieved, the victim executes a script and downloads bot binary from some location. The bot binary installs itself to the victim and automatically run. The new bot contact a DNS server for getting the IP address of the IRC server. Then it will establish an IRC session with the server and join the C&C channel. Then the bot can automatically parse and executes the channel topic, which contains the default commands. Infected host executes commands from Botmaster via IRC server, for example, launching a distributed denial-of-service attack, sending mass spam mailings, or logging keystrokes.

B. *Botnet Malicious Behavior*

Botnet can be used for a wide variety of illegitimate activity [16, 17].

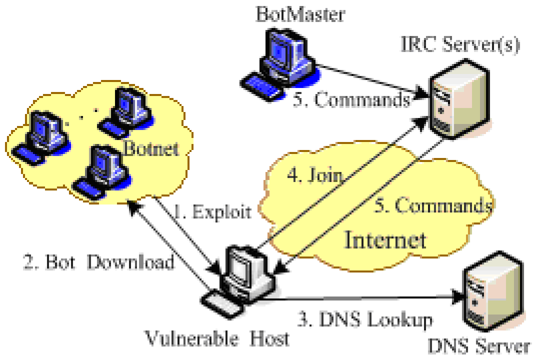

Figure 3. The process of a typical botnet infection[16]

1) Compromising new hosts: To make the botnets more stronger, the Botmaster often recruit new hosts using social engineering and distribution of malicious emails.

2) Denial of service attack: DDoS attack capability is a common feature of the botnet. The botnet always contains a set of flooding mechanisms, such as SYN flood, ICMP flood, and HTTP flood, for sending those packets to the targeted network, or just sending thousands of legitimate http, ftp requests to the site.

3) Spam: Spam bots can use the SMTP server to send spam on attacker's will. Most of today's e-mail spam is sent by botnet. Phatbot is one such bot widely being used for spamming.

4) Phishing: In most case, bots can be used for hosting phishing sites. Attackers can extract information from bots by turning them into web servers or DNS servers to conduct phishing.

5) Steal sensitive data: With screen capture, password theft, file upload and key-logging software, Botmaster can easily get enough victims' passwords and information. For example, the SDBot uses advanced key-logging software to collect personal information.

C. Command and Control Models

Understanding the C&C mechanism in botnet has great importance for us defending against botnet. Cooke et al. [18] identified three possible C&C communication topologies and investigated their associated benefits and weaknesses.

1) Centralized C&C Model[18]: A centralized model is characterized by a central point that forwards messages between clients.The centralized model has some advantages such as simple implementation and customization. However, the centralized C&C model will be detected and destroyed easier. Although this model hascertain drawbacks, most bots use the centralized C&C model such as AgoBot[6], SDBot[8], and Zotob.

2) P2P-based C&C Model[18]: For drawbacks of centralized model, the Botmaster shifts to P2P-based botnet. Compared with the centralized C&C model, the P2P based C&C model is much harder to discover and destroy. Botmaster can send commands from any peer. However, it is a more complex job for designing p2p systems. Some bots such as Phatbot[11]and Peacomm [13], have used P2P communication as a means to control botnet.

3) Unstructured C&C Model: A bot will not actively contact other bots or the Botmaster, and would listen to incoming connections from its Botmaster. The Botmaster randomly scan the Internet and pass along the encrypt message when it detected another bot.

D. Communication Protocols

Botnet usually use well defined communication protocols. Studying the communication protocols can help us determine the origins of a botnet attack and decode conversations between the bots and the Botmasters. In [19], the communication protocols was be classified in three different categories.

1) IRC Protocol: This is the most common protocol used by Botmasters to communicate with their Bots. IRC protocol mainly designed for one to many conversations but can also handle one to one, which is very useful for Botmasters control their botnet. However, secutiy devices can be easily configured to block IRC traffic.

2) HTTP protocol: The HTTP protocol is a popular communication method by botnet which is difficult to be detected. Using the HTTP protocol, botnet usually bypass security devices.

3) P2P protocol: Recently, more advanced botnet used P2P protocols for their communications[20]. Some recent variants of Phatbot and AgoBot [6], Nugache [12], Peacomm [13] used P2Pcommunication protocols.

E. Botnet Detection

Detecting and possibly defense botnet is an important research task. There are mainly two approaches of botnet detection and tracking methods. One is signatures based method and the other is based on anomaly.

1) Signatures-based detection: There is some very recent work on the signatures based detection of botnet[21-25]. Honeypots and honeynets are effective detection and analysis techniques at a reasonable cost and without false positives, and hence there has been much recent research in this area. The Honeynet project [21] deploys an architecture that consists of a Honeywall and the honeypot network. Chinese Honeynet Project [22] presented a malware collection tool based on the high interaction honeypot called HoneyBow. Tang and Chen [23] presented a novel "double-honeypot" detection system to effectively detect Internet worm attacks. Bothunter [24] modeled the bot infection phase as a set of ordered communication flows that are exchanged between an internal host and external entities and used this model to compare suspected infection events. Goebel et al. a regular expressions to represent sets of suspicious IRC nick names, and used n-gram analysis and scoring systems to evaluate the nick names to determine if a particular conversation belongs to a bot host [25].

Signatures-based detection cannot detect unknown attacks for which there is no signature available, and some botnet adopting some circumvention techniques such as polymorphic, metamorphic, obfuscation and packer.

2) Anomaly-based detection: As mentioned previously, some anomaly-based detection techniques are introduced to overcome this drawback. Binkley et al.[26]proposed an anomaly based system that combines IRC statistics and TCP work weight for detecting IRC-based botnet.Karasaridis et al. [27] developed an anomaly-based passive analysis algorithm that is able to detect IRC botnet controllers running on any random port without the need for known signatures or captured binaries.

However, anomaly-based detection techniques have high false alarm rate and the complexity involved in determining what features should be learned in the training phase. What's more, there are no anomalies of botnet until the botnet has been used.

F. Botnet Defenses

According the botnets structure, defender may consider the following method to prevent botnets. One possible way to defense botnets is to contact the owner of the compromised host to kill the bot malware and update the system. The large number of bot maybe make above approach cumbersome and infeasible. Defender should develop an automated notification system to do that job. Another approach to prevent botnet attacks aims at destroying the actual infrastructure, such as C&C server. The third approach to prevent botnet attacks aims at controlling the botnet and taking over the Botmaster, to shutting down the botnet

Alex Brodsky et al. [28] proposed a distributed content independent spam classification system to defend from botnet generated spams.

IV. Case Study: SpyBot

The earliest references to SpyBot [9] were in the April, 2003. And its codebase is relatively compact, written in under 3,000 lines of C. It has has strong spreading abilities and has over several hundred variants currently. Based on the investigation of Barford's research [4], we summarize the related techniques of SpyBot.

A. Infection Mechanisms

The scanning mechanisms included in SpyBot are quite simple and consists of horizontal and vertical scanning.

A typical command is:

- scan <start IP address> <port> <delay> <spreaders> <log filename>

B. Command and Control

SpyBot's commands are similar with SdBot's. Partial command language is listed as fellow:

ogin < password >;
disconnect < secs >;
server < new server addr >;
download <url> <filename>.

The set of host control capabilities provided in SpyBot is quite comprehensive. They include commands for local file manipulation, key-logging, process/system manipulation and remote command execution.

Key-logging functionality automatically launched as a new thread from the bot Main ():

- Threat_Handle =CreateThread(NULL,0,&keylogger, NULL, 0, &id);
- sprintf(buf, “Keys logging to %s\\%s”,sysdir,keylog filename);
- addthread(buf,0,Threat_Handle,2, “\0”);

C. Exploits and Attack Mechanisms

SpyBot's DDoS interface includes the capabilities for launching simple UDP, ICMP and TCP SYN floods.

The DDoS commands are:

- syn <host> <port> <delay> <number>
- spoofdsyn <host> <port> <delay> <number>
- ping <host> <port> <delay> <number>

Acknowledgment

This work is supported in part by the National High-Tech Research and Development Plan of China under Grant Nos. 2007CB311100, 2009AA012437, 2007AA01Z442.

References

[1] Botnet scams are exploding. http://www.usatoday.com/tech/news/computersecurity/2008-03-16-computer-botnets_N.htm.

[2] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Eggdrop: Open source IRC bot. http://www. eggheads.org/,1993.

[3] C. Associates. GTBot1. http://www3.ca.com/securityadvisor/pest/pest.aspx?id=453073312, 1998.

[4] P. Barford and V. Yegneswaran, “An Inside Look at Botnets”, ser. Advances in Information Security,Malware Detection. Springer, 2007.

[5] The PrettyPark Worm/Trojan http://www.nwi.net/~pchelp/bo/pretty park. htm.

[6] Sophos. Troj/Agobot-A. http://www.sophos.com/virusinfo/analyses/trojagobota.html, 2002.

[7] Arce I, Levy E. An analysis of the slapper worm. IEEE Security & Privacy, 2003,1(1):82−87.

[8] Sophos. Troj/SDBot. http://www.sophos.com/virusinfo/analyses/trojsd bot.html, 2002.

[9] McAfee.W32-Spybot.worm. http//vil.nai.com/vil/content/v100282.htm,2003.

[10] Sinit P2P Trojan Analysis. http://www.secureworks.com/research/threats/sinit

[11] Phatbot Trojan Analysis. http://www.secureworks.com/research/threats/phatbot

[12] J. Nazario. Nugache: TCP port 8 Bot, May,2006. http://asert.arbornetworks.com/2006/05/nugache-tcp-port-8-bot/.

[13] M. Suenaga and M. Ciubotariu, “Symantec: Trojan.peacomm.” http://www.symantec.com/security response/writeup.jsp?docid=2007-011917-1403-99, February 2007.

[14] Basudev Saha and Ashish Gairola, “Botnet: An Overview”, CERT-In White Paper, CIWP-2005-05, Jun. 2005.

[15] Niels Provos, Dean McNamee et al.. “The Ghost In The Browser Analysis of Web-based Malware,” In: Proc. of the 1st Workshop on Hot Topics in Understanding Botnets (HotBots 2007). 2007.

[16] Rajab MA, Zarfoss J, Monrose F, Terzis A. “A multifaceted approach to understanding the botnet phenomenon”, In: Almeida JM, Almeida VAF, Barford P, eds. Proc. of the 6th ACM Internet Measurement Conf. (IMC 2006). Rio de Janeriro: ACM Press, 2006. 41-52.

[17] The Honeynet Project & Research Alliance, “Know your enemy: Tracking botnets,” http://www.honeynet.org, March 2005.

[18] E. Cooke, F. Jahanian, and D. McPherson, “The Zombie Roundup: Understanding, Detecting, and Disrupting Botnets,”Usenix Workshop on Steps to Reducing Unwanted Traffic on the Internet, July 2005.

[19] Taxonomy of Botnet Threats. Trend Micro Inc. White Paper, November, 2006.

[20] J. B. Grizzard, V. Sharma, C. Nunnery, B. B. Kang, and D. Dagon. “Peer-to-peer botnets: Overview and case study,” In Proc. of ot Topics in Understanding Botnets (HotBots'07), 2007:198~201.

[21] Honeynet Project and Research. http://www.honeynet.org/papers/bots.

[22] Chinese Honeynet Project. http://www.honeynet. org.cn/.

[23] Y.Tang and S.Chen, “Defending against internet worms: A signature-based approach”, In Proc. of the IEEE INFOCOM, May 2005.

[24] Gu G, Porras P, Yegneswaran V, Fong M, Lee W. “BotHunter: Detecting malware infection through IDS-driven dialog correlation,” In: Proc. of the 16th USENIX Security Symp. (Security 2007). 2007

[25] J. Goebel and T. Holz. Rishi: Identify bot contaminated hosts by irc nickname evaluation. In First Workshop on Hot Topics in Understanding Botnets (HotBots'07), Cambridge,MA, April 2007.

[26] J. R. Binkley and S. Singh. ‘An algorithm for anomaly-based botnet detection”, In Proceedings of USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI), July 2006, pp. 43–48.

[27] Karasaridis A, Rexroad B, Hoeflin D. “Wide-Scale botnet detection and characterization”, In: Proc. of the 1st Workshop on Hot Topics in Understanding Botnets (HotBots 2007). 2007.

[28] A. Brodsky and D. Brodsky. “A distributed content independent method for spam detection”, In: Proc. of the 1st Workshop on Hot Topics in Understanding Botnets (HotBots 2007). 2007.