Volume No:1, Issue No:1 (June-2015)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

An Innovative Method That Distinguishes Between Botnet Traffic and Legitimate Traffic in Internet Chats

E. likith
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

K. Latha
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

A.Pramod Reddy
Associate Professor,
Depatment of CSE,
TKR College of Engineering & Technology.

Abstract:

A botnet is a collection of Internet-connected programs communicating with other similar programs in order to perform tasks. This can be as mundane as keeping control of an Internet Relay Chat (IRC) channel, or it could be used to send spam email or participate in distributed denial-of-service attacks. The word botnet is a combination of the words robot and network. The term is usually used with a negative or malicious connotation.Peer-to-Peer botnets are legally taken by botmasters for the quick recovery against taking down effortsof the system. But it's a harder one for the botmasters, because modern botnets are hidden and performing maliciousactivities it makes the process inefficient. Additionally because of sudden growth of the network traffic there was anability to enlarge the malicious activities of the system. In this paper, the hidden P2P botnets are identified usingbotmasters. Our system first identifies the system which is all engaged in p2p communications. Then it analysis thebehavioral characteristics of identifying P2P and it finds the difference between P2P botnet traffic and legal p2p traffic. By doing this our scalability of our system increases. Alternatively it also increases the detection accuracy as well asscalability of our system.

Keywords: Botnets, Traffic, Legitimate traffic, P2P, Botmaster.

Introduction:

A botnet (also known as a zombie army) is a number of Internet computers that, although their owners are unaware of it, have been set up to forward transmissions (including spam or viruses) to other computers on the Internet.

Any such computer is referred to as a zombie - in effect, a computer "robot" or "bot" that serves the wishes of some master spam or virus originator. Most computers compromised in this way are home-based. According to a report from Russian-based Kaspersky Labs, botnets -- not spam, viruses, or worms -- currently pose the biggest threat to the Internet. A report from Symantec came to a similar conclusion.

Types of botnets:

Legal botnets:

The term botnet is widely used when several IRC bots have been linked and may possibly set channel modes on other bots and users while keeping IRC channels free from unwanted users. This is where the term is originally from, since the first illegal botnets were similar to legal botnets. A common bot used to set up botnets on IRC is eggdrop.

Illegal botnets:

Botnets sometimes compromise computers whose security defenses have been breached and control conceded to a third party. Each such compromised device, known as a "bot", is created when a computer is penetrated by software from a malware (malicious software) distribution. The controller of a botnet is able to direct the activities of these compromised computers through communication channels formed by standards-based network protocols such as IRC and Hypertext Transfer Protocol (HTTP). A botnet's originator (known as a "bot herder" or "bot master") can control the group remotely, usually through IRC, and often for criminal purposes. This server is known as the command-and-control (C&C) server. Though rare, more experienced botnet operators program command

Volume No:1, Issue No:1 (June-2015)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

protocols from scratch. These protocols include a server program, a client program for operation, and the program that embeds the client on the victim's machine. These communicate over a network, using a unique encryption scheme for stealth and protection against detection or intrusion into the botnet.[citation needed]A bot typically runs hidden and uses a covert channel (e.g. the RFC 1459 (IRC) standard, Twitter, or IM) to communicate with its C&C server. Generally, the perpetrator has compromised multiple systems using various tools (exploits, buffer overflows, as well as others; see also RPC). Newer bots can automatically scan their environment and propagate themselves using vulnerabilities and weak passwords. Generally, the more vulnerabilities a bot can scan and propagate through, the more valuable it becomes to a botnet controller community.

The process of stealing computing resources as a result of a system being joined to a "botnet" is sometimes referred to as "scrumping."Botnet servers are typically redundant, linked for greater redundancy so as to reduce the threat of a takedown. Actual botnet communities usually consist of one or several controllers that rarely have highly developed command hierarchies; they rely on individual peer-to-peer relationships.

Botnet architecture evolved over time, and not all botnets exhibit the same topology for command and control. Advanced topology is more resilient to shutdown, enumeration or discovery. However, some topologies limit the marketability of the botnet to third parties.[6] Typical botnet topologies are Star, Multi-server, Hierarchical and Random.This example illustrates how a botnet is created and used to send email spam.

- A botnet operator sends out viruses or worms, infecting ordinary users' computers, whose payload is a malicious application—the bot.
- The bot on the infected PC logs into a particular C&C server.

- A spammer purchases the services of the botnet from the operator.

- The spammer provides the spam messages to the operator, who instructs the compromised machines via the control panel on the web server, causing them to send out spam messages.

How a Botnet works:

Types of attacks:

- In distributed denial-of-service attacks, multiple systems submit as many requests as possible to a single Internet computer or service, overloading it and preventing it from servicing legitimate requests. An example is an attack on a victim's phone number. The victim is bombarded with phone calls by the bots, attempting to connect to the Internet.

- Adware advertises a commercial offering actively and without the user's permission or awareness, for example by replacing banner ads on web pages with those of another advertiser.

- Spyware is software which sends information to its creators about a user's activities – typically passwords, credit card numbers and other information that can be sold on the black market. Compromised machines that are located within a corporate network can be worth more to the bot herder, as they can often gain access to confidential corporate information. Several targeted attacks on large corporations aimed to steal sensitive information, such as the Aurora botnet.

- E-mail spam are e-mail messages disguised as messages from people, but are either advertising, annoying, or malicious.

- Click fraud occurs when the user's computer visits websites without the user's awareness to create false web traffic for personal or commercial gain.

Volume No:1, Issue No:1 (June-2015)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

• Fast flux is a DNS technique used by botnets to hide phishing and malware delivery sites behind an ever-changing network of compromised hosts acting as proxies.

• Brute-forcing remote machines services such as FTP, SMTP and SSH.

• Worms. The botnet focuses on recruiting other hosts.

• Scareware is software that is marketed by creating fear in users. Once installed, it can install malware and recruit the host into a botnet. For example users can be induced to buy a rogue anti-virus to regain access to their computer.

• Exploiting systems by observing users playing online games such as poker and see the players' cards.

PROPOSED SYSTEM:

In this paper a novel scalable botnet detection system has been proposed. This detection system is capable of detectingstealthy P2P botnets whose malicious activities may not be observable in the network traffic.Our system aims todetect stealthy P2P botnets even if P2P botnet traffic is overlapped with the traffic generated by legal P2P applicationsrunning on the same compromised host.

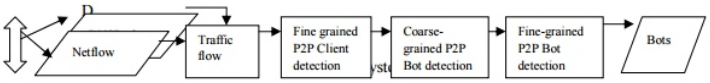
Our system identifies P2P bots within a monitored network by detecting theC&C communication patterns that characterize P2P botnets. The high scalability of the system can be achieved byusing the following techniques.1. P2P traffic profiling algorithm that is used to built the statistical fingerprints for various P2P applications.2. A flow-clustering based analysis approach to identify host that engaged in P2P communications.

3. A scalable design based on an efficient detection algorithms and parallelized computation.

4. A prototype system based on real world network traffic which demonstrated high detection accuracy.
The new model eradicates the necessity of keeping failed connections. Clustering based client detection algorithmenhances the efficiency of the model. The system is parallelized to boost scalability and efficiency.

The proposedsystem is effective over a large range of parameter values.

System Design:



A P2P botnet depends on P2P protocol to create transmission through C&C channel with botmasters. A P2P bots havea common network traffic patterns that is used to evolve P2P client applications as well as legal applications. It dividesin two phases (1) In first phase, Its aims is to detect all the network traffic which involved in peer-to-peer communications.

In figure we inspect the network flow at the edge and filter it to discard the flow which should becreated unexpectedly by peer-to-peer applications. From that we can analysis the network traffic and flow created bypeer-to-peer clients. (2) In second phase, our system will examine the network traffic generated by both legal P2Pclients and P2P bots. Then we explore the active time of the peer-to-peer client and recognize it as candidate P2P botand if there is a continuous change in host. We further analyze it by detecting 2 candidates P2P bots.

Finding out Peer-to-Peer Client: A Filter:

Filter component is used to filter the network traffic that is unrelated to P2P communications. This can be achieved byanalyzing DNS traffic. P2P clients contact their peers by looking up IPs from a routing table for the overlay networkrather than resolving a domain name. Most non P2P applications often connect to a destination address resulting fromdomain name resolution. This simple filter can eliminate a very large percentage of non P2P traffic and helps inretaining P2P communication.

B Peer-to-Peer Detector:

Client detector helps in detecting P2P clients by analyzing the remaining network traffic .For each host within themonitored network we identify two flow sets which contains flows related to sucessful outgoing TCP and UDPconnections.

Volume No:1, Issue No:1 (June-2015)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

TCP connection is considered successful if SYN,SYN/ACK,ACK I handshake is available.UDP connection is considered successful if there is at least one request and a consequent response packet is found. Inorder to detect P2P clients we first consider the fact that each P2P client frequently exchanges control messages with-other peers. Even though the characteristics of these messages such as size and the frequency of the exchanged packetsare same ,they vary depending upon the P2P protocol and network in use. If two network flows are generated by thesame P2P applications they carry the same control messages. In addition P2P client exchanges the control messageswith large number of peers that is distributed in different networks. The destination IP addresses of network flows that carry these control messages will spread across a large number of networks where each network can be represented by its BGP prefix.

Finding Peer-to-Peer Bots:
Coarse-Grained Detection of P2P Bots:

Since bots are malicious programs used to perform profitable malicious activities, they represent valuable assets for the botmasters, who will intuitively try to maximize utilization of bots. This is particularly true for P2P bots because in order to have a functional overlay network, a sufficient number of peers needs to be always online. In other words the active time of the bots should be comparable with the active time of the underlying compromised system. If this was not the case the botnet overlay network would risk degenerating into a number of disconnected subnetworks due to the short lifetime of each single node. In contrast the active time of the legitimate P2P applications determined by users, which is likely to be transient.

Fine-Grained Detection of P2P Bots:

The objective of this component is to identify P2P bots from all persistent P2P clients. We leverage one feature; the overlap of peers connected by two P2P bots belonging to the same P2P botnets is much larger than that contacted by two clients in the same legitimate P2P network. Assume two hosts in the monitored network are running the same legitimate P2P file sharing application (e.g., Emule). Users of these two P2P clients will most likely have uncorrelated usage patterns. It is reasonable to assume that in the general case the two users will search for and download different contents(e.g., different media files or documents)from the P2P network. This translates into a divergence between the set of IP addresses contacted by hosts.

Conclusion:

The computers that form a botnet can be programmed to redirect transmissions to a specific computer, such as a Web site that can be closed down by having to handle too much traffic - a distributed denial-of-service (DDoS) attack - or, in the case of spam distribution, to many computers. In this paper we proposed a novel scalable P2P botnet detection system that is able to identify stealthy P2P botnets.

To perform this task statistical fingerprints of P2P communications have been derived to detect P2P clients and furtherdistinguish between those that are part of legitimate P2P networks and P2P bots. The results shows that the proposedsystem accomplishes high accuracy on detecting stealthy P2P bots and great scalability.

References:

[1] Junjie Zhang, Roberto Perdisci, Wenke Lee, Xiapu Luo, and Unum Sarfraz, Building a Scalable System for Stealthy P2P-Botnet Detection, EEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 9, NO. 1, JANUARY 2014.

[2] A. Ramachandran, N. Feamster, and S. Vempala, " Filtering spam with behavioral blacklisting", In ACM CCS, 2007.

[3] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, " A multifaceted approach to understanding the botnet phenomenon",In IMC, 2006.

[4] D. Dittrich and K. E. Himma, "Active Response to Computer Intrusions,"in The Handbook of Information Security, edited by H. Bidgoli(Wiley,New York, 2005).

[5] Y. Zhao, Y.Xie, F.Yu and Y.Yu, "Botgraph : Large scale spamming botnet detection", in Proc. 6th USENIX NSDI, 2009, pp 1-14.

Volume No:1, Issue No:1 (June-2015)

International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

[6] G.Gu ,R.Perdisci, J.Zhang and W.Lee, "Botminer: Clustering analysis of network traffic for protocol and structure independent botnet detection",in Proc. UNI-SEX security, 2008, pp.139-154.

[7] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, " A multifaceted approach to understanding the botnet phenomenon"In IMC, 2006.

[8] D. S. Anderson, C. Fleizach, S. Savage, and G. M. Voelker. Spamscatter: Characterizing internet scam hosting infrastructure. In USENIX SecuritySymposium, 2007

[9] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, and Osipkov. Spamming botnets: Signatures and characteristics.In SIGCOMM, 2008.

[10] Y. Yu, M. Isard, D. Fetterly, M. Budiu, U. Erlingsson,P. K. Gunda, and J. Currey, " DryadLINQ: A system forgeneral-purpose distributed data-parallel computing using a high-level language", In OSDI, 2008.

[11] G.Bartlett, J.Heidemenn and J. Pepin, "Estimating P2P traffic volume at USC",USA,Tech. Rep. ISI-TR-2007

About Author's:

E. likith
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

K. Latha
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

A.Pramod Reddy
Associate Professor,
Depatment of CSE,
TKR College of Engineering & Technology.