Techniques to Break the Botnet Attack

Gaurav R. Gabada[1], Mohammad Usman[2] and Jai Sharma[3]

[1]CSE, Department, Jawaharlal Darda Institute of Engineering & Technology, Yavatmal, India
gauravgabada@gmail.com)
[2]CSE, Department, Jawaharlal Darda Institute of Engineering & Technology, Yavatmal, India
uv.usman@gmail.com)
[3]CSE, Department, Jawaharlal Darda Institute of Engineering & Technology, Yavatmal, India
sharmajr49@gmail.com)

ABSTRACT

The term botnet is a group of computers, which are remotely controlled by hackers to launch various network attacks, and day by day botnets are emerging as the most serious threat against internet-security as they provide a dispersed platform for several illegal activities such as launching distributed denial of service(DoS) attacks against confidential targets, malware distribution, phishing, and click racket. The botnets makes the use of command and control servers through which they can be controlled. Recently, this topic has emerged out as one of the most intresting topic botnet related to cyber-threat and cyber-crime prevention. This paper is related to botnet and botnet detection. The paper describes the nature of botnet and discusses various botnet detection techniques. This paper classifies botnet detection techniques into different classes: honeypot, anomaly-based, signature-based, mining-based, and DNS-based and also explains other new techniques. It summarizes botnet detection techniques in each class and provides a latest report on increase or decrease in botnet attacks in recent years.

Keywords: Botnet, C&C Server, IRC Protocol, Fast Flux, Domain Flux

1. INTRODUCTION

The term bot is short for robot and refers to the clients in a botnet. Alternative names for bots are zombies. Any unprotected computers in any locations in the world may become bots as long as they are connecting with Internet and have been infected by botnet malware code. The botnet uses Internet Relay Chat (IRC) protocol, which appeared from 1989 onwards.

The first known IRC bot is Eggdrop, first published in 1993 and further developed since. Next ,following the release of Eggdrop, malicious IRC bots appeared, adopting the basic idea, but created primarily in order to attack other IRC users or even entire servers. Shortly after, Denial of Service (DoS) and then Distributed Denial of Service (DDoS) were implemented in these bots. A survey shows 90.4% of total emails were spam in June 2009. Among all spam, 83.2% was sent through botnets.

In addition, many spam emails are also used for injecting viruses, phishing attacks, and web-based malware. Therefore, sending spam through the use of botnet has influenced the cyber-crimes. Some IRC-based approaches were developed to overcome this problem. New techniques also have been determined to trace or detect more sophisticated botnets that use complex mechanism to prevent their detection. The two techniques used by the botmasters to prevent their detection are: Fast Flux and Domain Flux. This paper puts the main emphasis on the techniques to detect the sophisticated botnets.

2. WHAT IS BOTNET?

A bot is a program that runs on an end-system performing tasks automatically. A botnet is typically seen as a network of bots that use computing resources for a malicious end. The botnet is generally controlled by a single entity called as botmaster. Botnets infect new machines using techniques common to most classes of malware, they are distinguished by their use of command and control (C&C) server. The master computer sends instruction to its bots through a command and control (C&C) server, which passes commands from the botmaster to bots, and sends stolen information from bots to their master.The attacked bots can also infect other computers

enabling them to be botnet members. In recent years, fluxing techniques have been applied into many botnets to prevent detection. This brings additional challenge to botnet researchers.

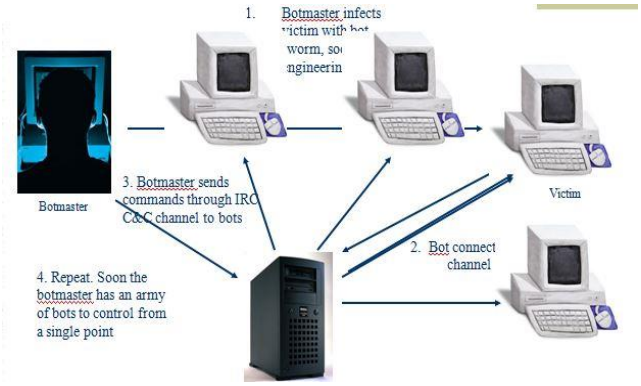


Fig1: Botnet Mechanism

The two advanced botnet mechanisms are:

- Fast flux (FF): A mechanism that a set of IP addresses change frequently corresponding to a unique domain name.
- Domain flux (DF): A mechanism that a set of domain names are generated automatically and periodically corresponding to a unique IP address.

3. FAST FLUX

Fast flux (FF) is a way to prevent botnet detections.By IP fast fluxing, the mapping between multiple IP addresses and one single domain name is frequently changing. This technique makes it more complicated to block or take down the C&C Server. Networks which apply fast fluxing techniques are called fast fluxing (FF) Network. By single flux, a domain may be resolved to different IPs in different time ranges. For example, a user accesses the same domain twice in a short while. For the first time, the user sends a DNS query to the DNS server, which resolves that the corresponding IP address is IP1 "255.155.162.110". With IP1, the user accesses the fluxing agent FA1, which redirects the request to the real server "mothership". This "mothership" then processes the request and responds to FA1. Finally, FA1 passes on the response back to the user. After a while, the user accesses the same domain again.

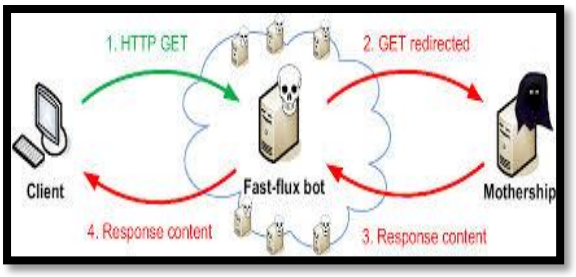


Fig2. Fast Flux

However, due to a short TTL of IP addresses,IP1 has expired; therefore, the user needs query the updated IP address again of that domain. Now, the DNS server responses to the user with a different IP address - IP2 "255.155..122.22". Then the user uses this new address IP2 to connect to another flux agent FA2 which redirects the user to the "mothership". Double-flux is a more sophisticated way of counter detection. It involves the repeated changing of both the flux agents and the registration in DNS servers. That is to say, in addition to fluxing their own agents, the authoritative domain name server is a part of fluxing as well. This provides an additional layer of redundancy within malware networks.

4. DOMAIN FLUX

Fast flux has a single point of failure due to the single domain flux, once fluxing is identified. A more survivable mechanism is domain flux (DF). It was found that some recent botnet programs such as Torpig resolved this problem by using DF. They discussed domain-generation techniques, and also provided a report on how they took over an advanced DF based botnet - Torpig .It explained the process of DF of taking over Torpig. DF is based on the idea of generating domains through a domain generation algorithm (DGA). Both C&C server and its bots follow the same algorithm fixed by the same value to obtain consistent dynamic domain names. Bots try to contact the C&C server and other servers controlled by the botnet master according to a domain list until one DNS query succeeds. If the current domain has been blocked or suspended by authorities, bots will try to calculate other domains by the DGA . The key idea is that the algorithm must make sure that all bots can generate domains by the same seed. Stone-gross et al. revealed that Torpig calculates sub-domains by using the current week and year first but independent of the current day, and then appends the Top Level Domain (TLD). The domains generated might be "xyz.com" or "xyz.biz", etc. Next, the bots will use these auto-generated domain names to

contact the C&C server. If failed, bots will use the day information to calculate the "daily domain", such as "day.com" or "day.net", etc. If all these domains cannot be resolved, bots will try to use the hard-coded domain names in a configuration file at the end.

5. BOTNET DETECTION TECHNIQUES

5.1. Honeypots and Honeynet

A honeypot is an "environment where vulnerabilities have been purposely introduced to moniter the attacks and intrusions"(Pouget & Dacier, 2004). It is a computer system that is used to draw the attention to attack this computer system. All Honeypots have their own concepts. They are computer systems that don't have any production value. All Honeypots have a unique concept. They have a strong ability to detect security threats, to collect malware signatures and to understand the motivation and technique behind the threat used by performer. For example, you may launch a honeypot web server in the DMZ in your network. The limitations are exposed to the attackers so that they take advantage of these limitations and get traced. So we can conclude that while using Honeynet for Botnet detection, we have to wait until one bot in the network infect our system then we can track or analyze the machine.

5.2. Signature Based Botnet Detection

Rule based intrusion detection systems like Snort are running by using known malware signatures.The signs of intrusions and other threats are monitered throughout the network. It is obvious that consignment information of network traffic is transformed and embedded into the signature or rule. The Intrusion Detection system functions to moniter the malicious activities running on the network. Gu et al. (2007) proposed a framework, "BotHunter", to correlate IDS based detection alerts.

5.3. Anomaly Based Botnet Detection

This technique tries to detect Botnet depending on the network traffic parameters such as network latency, high volumes of traffic, unusual system behavior, and traffic on unusual ports that could show existence of bots in the network. Unknown botnets can be detected through this technique. It can be classified into Host based and Network Based Detection. In host based detection technique, a detection strategy which monitors and analyzes the internals of a computer system instead of network traffics on its external interfaces. Limitation with this system is high false positive. A network-based technique is a detection strategy which tries to detect Botnets by monitoring network traffics. We can classify Network-based techniques into two categories: Active monitoring and Passive monitoring. In Active monitoring, it injects test packets in network to measure the reaction of network such that gaining extra traffic on network. The paper proposed Botsniffer that uses network-based anomaly detection to identify Botnet C&C server in a local area network.

5.4. DNS Based Detection Technique

The bots use DNS queries in order to locate the C&C server hosted by the Dynamic DNS provider. Monitoring the traffic and the DNS makes it pretty easy to detect the botnet and DNS traffic irregularity. This is most famous and easy technique to botnet detection but it will be tough to detect recent advanced botnet through this technique.

5.5. Data Mining Based Detection Technique

Data mining puts its emphasis on recognizing useful patterns to get the regularities and irregularities in large data sets. Packet flow provides full information of flow data but in large file type. Anomaly based techniques are mostly based on network behavior anomalies such as high network latency, activities on unused ports. Data mining technique can be applied for optimization purpose. It enables to extract sufficient data for analysis from network log file. Most useful data mining techniques includes correlation, classification, clustering, statistical analysis, and aggregation for efficiently knowledge discovery about network flows. Flow connection algorithms are useful to compare flow objects based on some characteristic other than packet content. When the content of the packet is not encrypted,this technique may not prove that efficient. These kinds of algorithms utilize the characteristic values as inputs into one or more functions to create a metric used to decide if the flows are correlated.

Classification algorithms assume that incoming packet will match one of the previous patterns. Therefore, it is not an appropriate approach to detect new attacks. Clustering is a well-known data mining technique where data points are

clustered together based on their feature values and a similarity metric. Clustering differs from classification, in that there is no target variable for clustering. Clustering algorithms divide the entire data set into subgroups or clusters containing relatively identical features. Thus, clustering provides some significant advantages over the classification techniques, since it does not require a labeled data set for training. To find particular pattern from large dataset is known as aggregation method, collecting and analyzing several types of records from different channels simultaneously. Association rule is to find the correlation of different items appeared in the same event. Association rule mining is to derive the implication relationships between data items under the conditions of a set of given project types and a number of records and through analyzing the records, the commonly used algorithm is Apriori algorithm.

6. TECHNIQUES TO DETECT/TRACE MORE SOPHISTICATED BOTNET

6.1. Fast Flux Botnet Detection Techniques

6.1.1. Flux score

The Fast Flux Service Network is a network of computer connected to the internet with frequently changing ip address and the Fast Flux Attack Network is a network of bots that take the advantage of the frequency of change in ip address and perform malicious activities. Flux score is a metrix for detecting a Fast Flux Service Network(FFSN) based on the Parameters: the number of IP-domain mappings in all DNS lookups, the number of name server records in one single domain lookup, the number of autonomous system in all IP-domain pairs. All these parameters of the network are monitored todetect whether the network is being used for malicious activities.

6.1.2. Recursive DNS

The Recursive DNS technique monitors DNS traffic and stores information from FF domains into a centralized data collector. Network is considered malicious Fast Flux network based on four characteristics- short TTL, the change rate of the set of resolved IPs returned from each query, a large number of resolved IPs, and resolved IPs scattered across many different networks.All these participate in detecting a Fast Flux Attack Network.

6.2. Domain Flux Botnet Detection Techniques

6.2.1. Batch Method

One method for detecting the Domain Flux Botnet is to keep track of the various domain name properties to detect any unauthorized domain that is generated using the Domain Flux algorithms.This method keep track of the parameters like WHOIS (registrars) properties, domain name properties (TTL, etc.), and geographic properties are used to detect the malicious websites.This information prevents to detect websites to be launched from unauthorized host.

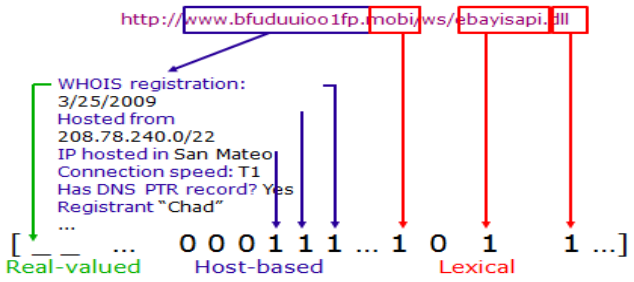


Fig3: Parameters of websites

6.2.2 DNS Failure Graph

DNS Failures method is the simplest and yet efficient method for detecting the attackers network.DNS failure are rare to occure in any network,but in attackers network the graph of DNS failure rises while generating new malicious websites. This become a way through which the attackers network can be traced. This method studies the DNS faliure graph to detect the attackers network.

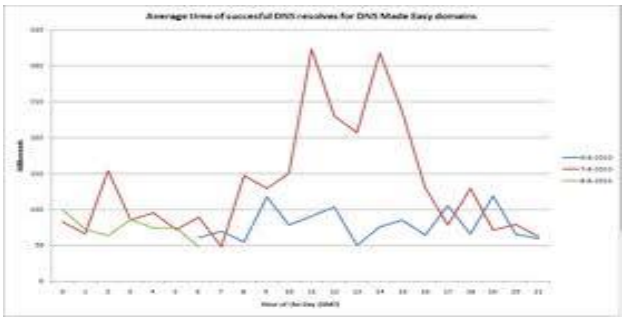


Fig4: DNS Failure Graph

The figure above shows the difference between the DNS failure graph of a normal network and a attackers network.

6.2.3 Phishnet

Phishnet is somewhat similar to traditional URL blacklisting where the phishing sites and the malicious URL were blacklisted and were monitered.This includes two components: URL prediction component.and an approximate URL matching component.

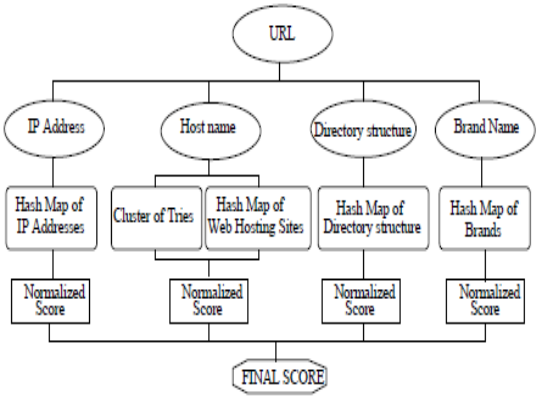


Fig5:URL Prediction Component

The URL forecast component generates the random URLs that are matched with the blacklisted URLs for approximation. If a predicted URL matches the blacklisted URL, then that particular URL is blocked to detect the attacker. A randomly generated URL is shown in following figure:

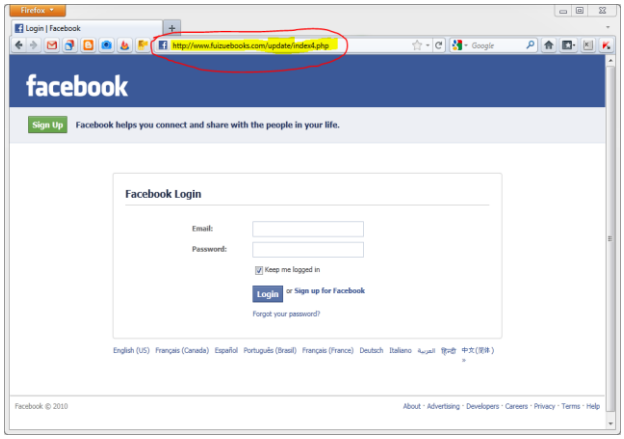


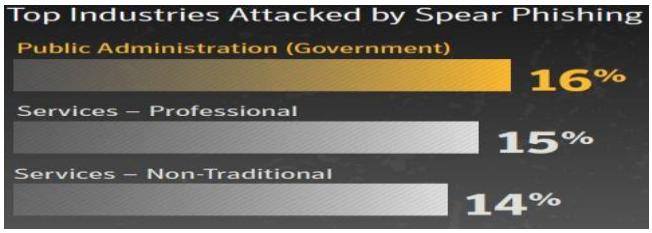
Fig:Phishing URL

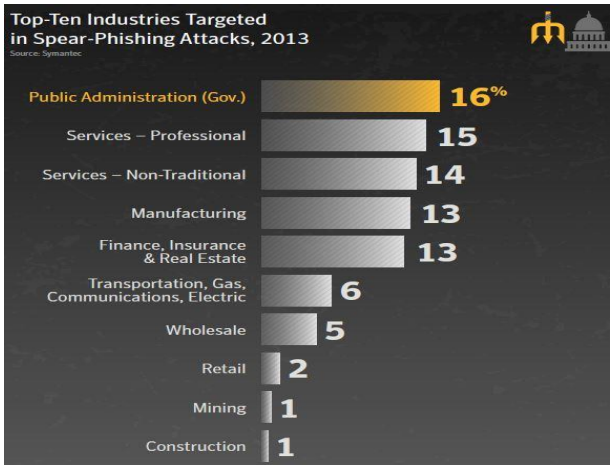
7. CONCLUSION

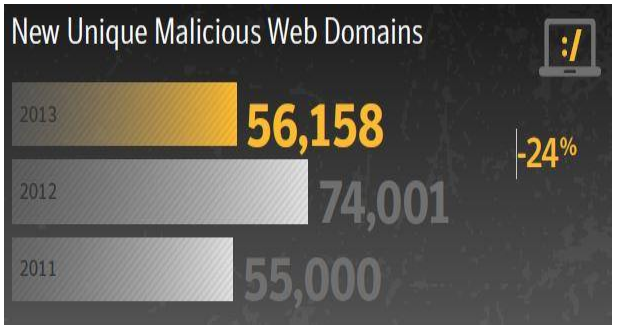
In this paper, readers can gain a deep understanding of the necessity for investigating fluxing features of botnet. This paper can help researchers to understand how botmasters can employ FF and DF features to evade detection. This paper includes the advantages and disadvantages, thus our multiple evaluation criteria provide a clear perspective of their features. The aim of this paper is also to raise more awareness from botnet researchers to develop more efficient models. Finally, the future work provides a brief view that what can be done in future to fight against botnets.

8. Some Recent Facts

The following statistics shows how the threats to internet world are increasing and the areas where the botmasters have their main emphasis:







9. REFRENCES

[1] H. Zeidanloo, M. Shooshtari, P. Amoli, M. Safari, and M. Zamani, "A taxonomy of botnet detection techniques," in Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on, vol. 2, july 2010, pp. 158 –162.

[2] Chung-Huang Yang , Kuang-Li Ting. *Fast Deployment of Botnet Detection with Traffic Monitoring*, Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pages 856-860, 2009

[3] Symantec, "Cutwails bounce-back; instant messages can lead to instant malware," http://www.messagelabs.com/mlireport, 2009.

[4] P. Bacher, T. Holz, M. Kotter, and G. Wicherski, "Know your enemy: Tracking botnets," http://www.honeynet.org/papers/bots, 2008.

[5] Y. Tang and S. Chen, "Defending against internet worms: a signaturebased approach," in INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, vol. 2, march 2005, pp. 1384 – 1394 vol. 2.

[6] C. Li, W. Jiang, and X. Zou, "Botnet: Survey and case study," in Innovative Computing, Information and Control (ICICIC), 2009 Fourth International Conference on, dec. 2009, pp. 1184–1187.

[7] M. Feily, A. Shahrestani, and S. Ramadass, "A survey of botnet and botnet detection," in Emerging Security Information, Systems and Technologies, 2009. SECURWARE '09. Third International Conference on, june 2009, pp. 268 –273.

[8] S. Yu, S. Zhou, and S. Wang, "Fast-flux attack network identification based on agent lifespan," in Wireless Communications, Networking and Information Security (WCNIS), 2010 IEEE International Conference on, june 2010, pp. 658 –662.

