DETECTING NETWORK ANOMALIES IN ISP NETWORK USING DNS AND NETFLOW

Andreas Tedja, Charles Lim, Heru Purnomo Ipung

Swiss German University, Indonesia
Swiss German University, Indonesia
Swiss German University, Indonesia

andreas.tedja03@gmail.com

Abstract: The Internet has become the biggest medium for people to communicate with other people all around the world. However, the Internet is also home to hackers with malicious purposes. This poses a problem for Internet Service Providers (ISP) and its user, since it is possible that their network is compromised and damages may be done. There are many types of malware that currently exist on the Internet. One of the growing type of malware is botnet. Botnet can infect a system and make it a zombie machine capable of doing distributed attacks under the command of the botmaster. In order to make detection of botnet more difficult, botmasters often deploy fast flux. Fast flux will shuffle IP address of the domain of the malicious server, making tracking and detection much more difficult. However, there are still numerous ways to detect fast flux, one of them is by analysing DNS data. Domain Name System (DNS) is a crucial part of the Internet. DNS works by translating IP address to its associated domain name. DNS are often being exploited by hackers to do its malicious activities. One of them is to deploy fast flux.Because the characteristics of fast flux is significantly different than normal Internet traffic characteristics, it is possible to detect fast flux from normal Internet traffic from its DNS information. However, while detecting fast flux services, one must be cautious since there are a few Internet services which have almost similar characteristics as fast flux service. This research manages to detect the existence of fast flux services in an ISP network. The result is that fast flux mostly still has the same characteristics as found on previous researches. However, current fast flux trend is to use cloud hosting services. The reason behind this is that cloud hosting services tend to have better performance than typical zombie machine. Aside from this, it seems like there has been no specific measures taken by the hosting service to prevent this, making cloud hosting service the perfect medum for hosting botnet and fast flux services.

Keywords: fast flux, DNS, botnet

1. Introduction

The internet has now become the biggest medium for communication, entertainment, and information. Anything that people want to look for will most likely be on the internet. At the time of the creation, the main focus of the internet is to share information (Chellapa and Pavlou, 2002). They want to ensure that the internet can connect systems throughout the world. Security of the internet was not the main concern at the time. As the usage of the internet grows, threats to internet users start to appear. One of the earliest internet threat is the rabbit virus. Rabbit virus is a type of virus which will multiply itself in a system until the system’s performance deteriorates and eventually crashes (Snyder, 2010).

Internet security threat can be directed to any party that is using the internet, be it a person, a company, a government organization, etc. The software that is causing internet security threat is called malware, which stands for for malicious software. There are many types of malware which serves different purposes. One of the type of malware is botnet.

Botnet is a network of infected systems, also known as bots or zombies, which are remotely controlled by the owner using command and control (C&C) software. The botmaster can use the botnet for malicious activities, such as distributed denial of service (DDoS) attacks, sending spam, steal data, and take control over the zombie device. Botnet is considered to be one of the biggest threat to Internet security. Botnet's C&C servers may use domain names to lead victims into its servers which host the malicious codes and phishing sites that will infect the victim and possibly steal sensitive information. Spamhaus, an organization that tracks spam and cyber threats, released a Botnet threat report for 2017. Spamhaus's Botnet Controller List (BCL) saw more than (40%) increase of listing of IP addresses for botnet servers in 2017, and more than (90%) since 2014 (Spamhaus Malware Labs, 2018).

The spread of the malicious domains can be done through opening spam emails or visiting suspicious websites from search engines. This method of Botnet spreading can easily be prevented by informing users of the danger of opening malicious sites, however not all internet users have been educated regarding this problem, therefore the problem remains existing to this day. Some preventive actions have been done in order to decrease the spread, one of the most notable method is by providing blacklisting services. Blacklist is contructed of manual reports by users, honeypots, etc. (Ma, J. et al., 2009). One example of blacklisting service is from www.dnsbl.info. However, not all malicious sites can be blacklisted due to outdated detection method, the sites never being detected, or the sites using stealth methods to hide its malicious characteristics.

Botmasters will try to keep their botnet network running for as long as possible while maintaining low detection rate to avoid being detected. According to Mahmoud and Matrawy (2015), one of the ways botmasters do this is by implementing Fast Flux Service Network on their botnet network. Domains that implements fast flux have multiple IP addresses associated to it and have the IP addresses constantly change in high frequency. This allows the botmaster to do attacks while hiding their server.

1.1. Research Problem

Internet Service Providers (ISPs) will send and receive packets of data as part of its function. However, not every packet that goes through the ISP is legitimate. One malicious traffic hidden between millions of legitimate packets will be hard to detect. The effect of the malicious traffic is yet to be known, but it is certain that there will be damage done to either or both the ISP and the customer. Botnets are also often using fast-flux domains to mask its malicious websites and the C&C servers from being detected, making it harder for detection system to track the malicious networks. Therefore, a countermeasure is needed to handle the threats.

1.2. Related Works

Previous researches on fast flux detection has proposed a number of methods that uses different sets of features in its detection system. The first research on fast flux detection Nazario and Holz, (2008) proposed a detection system using spam trap analysis which collects URLs from spam emails to then be analyzed. They used ATLAS system from Arbor Networks to identify and track new fast flux networks. The features used in this research are A record, NS record, and SOA records from DNS. They also applied blacklist and whitelist to filter out benign domains.

In Mahjoub (2013), the author introduced fast flux botnet monitoring using recursive and passive DNS information. The research focuses on studying the Kelihos botnet. The method starts by taking a set of domains confirmed to be fast flux and hosted by Kelihos botnet. From the malicious domains, the author build a training set and attribute features to each individual hostnames based on passive DNS, WHOIS¡ and geolocation information. Then, a detection model can be built and used in live networks.

The authors in Bilge, L. et al. (2011) introduced EXPOSURE, a malicious domains detection system which uses passive DNS analysis. They used 15 features in total for their detection system, some of which have been used in previous researches. EXPOSURE managed to characterize different properties of DNS names and the ways the properties are queried. EXPOSURE managed to automatically identify unknown malicious domains which are used in malicious activities, such as for botnet C&C server.

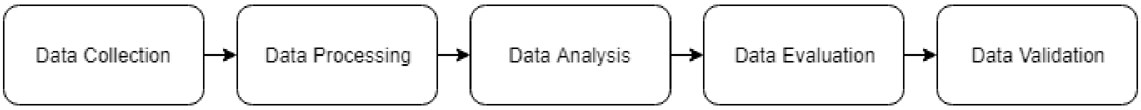
In Holz, T. et al. (2008), they used flux scoring based on the extracted DNS features, such as IP address diversity, A record count, NS record count, etc. To validate the fast flux service networks (FFSN), they used empirical measurements on 128 verified FF domain and 5,803 benign domains as the input, then performed two consecutive DNS lookups of all domains. On doing the DNS lookup, they waited TTL +1 second between the two lookups as to not get cached response on the second lookup. A

10 fold cross validation is then done on the set of labeled fast flux domains and benign domains. Using this methoed, they managed to achieve a 99.98% accuracy.

Another type of detection system is proposed in Chen, Huang, and Ou (2014), in which the authors used only the A record and NS record of the passive DNS information to determine if the FQDN (Fully Qualified Domain Name) is legitimate or malicious. For the detection system, they used dissimilar ASNs, the reverse lookup of the DNS, and the time of the domain registration as the detection features. The A and NS records can be collected by using dig command in Ubuntu. They used Step function on the detection system in order to classify the domains as legitimate or malicious. A step function threshold was determined from experiments they have done previously.

2. Methodology

2.1. Research Framework



The research framework shown in Figure 1 is based on Marchal, S. et al. (2012). The first step of this research is to collect the data required for the research. This is done by deploying a sensor in the ISP network in order to capture the network traffic. After having the required data of network traffic, the data is then stored in a server for further processing. In the processing step, the data will be filtered to reduce the amount of data that needs to be analyzed. Then, manual labeling will be done by human operator to classify a domain to suspected fast flux domain or benign. Suspected fast flux domains will then be validated to decide if it is truly malicious or not.

The data will be collected using passive DNS replication method rather than using active probing approach. Weimer (2004) describes Passive DNS Replication as a technique to reconstruct partial view of the data available in global DNS to central database where it is indexed and queried. The sensors will be deployed in the ISP node in order to be able to capture both customer and internal network. Perdisci, R. et al. (2009) explained that by using active probing, it will be easily detectable by the botmaster. If the botmaster detects any attempt to track the malicious flux network, it may stop responding to the queries from the probing system. This is unlikely to happen using passive DNS replication because there is no direct communication with the flux network, therefore it is stealthier. When connected to the internet, the ISP will have network for its internal use and network for the customer. Both network will be mirrored to a storage that stores the DNS traffic and Netflow traffic. The information from both networks will then be analyzed.

Data captured from the ISP is then stored in a server. In the server, the data will be processed before a human operator does manual labeling. The processing stage begins with feature selection. Then, whitelist from Majestic Million will be applied to the data to eliminate traffic from known legitimate domains. This process is done using a Python script. After whitelist is applied, further filtering is done according to the fast flux features.

Analysis stage will start after all the processing is done. To start the analysis stage, the data needs to be indexed in a search engine and then to a visualization tool to further ease the process of analysis. The analysis is done by a human operator since there is no machine learning applied in this experiment. The result of this stage is a list of domains that are suspected to be part of Fast Flux Service Network.

Data Evaluation stage will further evaluate suspected malicious domains from the analysis stage to confirm the maliciousness of the domain. A function will be applied to the DNS information to help deciding the maliciousness of the domain:

$$w1nA+w2nASN=w3eD+w4\Delta RTT=\beta$$

and a linear function is then used to determine the maliciousness of the domain:

fx-wTx- β

Further validation is done to the suspected domains. The validation stage is done to add additional information of the suspected malicious domains. The features used in validation stage are features of fast flux domains which are not necessarily available in the passive DNS record. The features used in validation stage includes domain age, registration country, IP change of rate, and the result of a reverse dns lookup. This stage is based on Perdisci et al. (2009). This is to ensure that no legitimate domain is misclassified as malicious.

3. Experiment Result

3.1. Analysis

During the four weeks of data captured in the ISP of Alam Sutera, the detection system in this research managed to detect two fast fux domains. The domains detected by the system are: diamongsl.info and spendentaly.info.

Spendentaly.info has TTL value of 179, 4 unique A records, 4 unique NS records, and 1 ASN. After being evaluated with the evaluation function, the domain turned out as malicious. Upon further validation, the domain turned out to be created on 13-03-2018. This means that at the time of this research, the domain is less than one year old. The domain is also registered in Ukraine and has 1 to 5 days IP change rate, which is much faster than normal domains.

Diamongsl.info was detected on June 7 2018. It has TTL value of 300, 5 unique A records, 4 unique NS records, and 1 ASN. After being evaluated with the evaluation function, the domain also turned out as malicious. The validation process specified that the domain is created on 02-08-2018, which at the time of this research, is less than 1 year old. it is also registered in Ukraine and has 1 to 5 days IP change rate.

Both domains are then checked on virustotal.com. It turned out that both diamongsl.info and spendentaly.info shared IP addresses on multiple occassions. This means that it is possible that both domains are part of the same fast flux network. The IP addresses are not only associated to these two domains, but also to several other domains which are also suspected to be part of the same fast flux network. Some of the domains are: entionale.info, lekhands.info, and gleaminist.info. The domain entionale.info. Further analysis were done to these domains. Domain entionale.info turned out to be communication with an executable file called KLIULIANGBAO%5B1%5D.EXE. This file is flagged on 54 detection engines as malicious because the file contains Virut, which is known to be a botnet malware. The executable file is also communicating with the domain liuliangbao.cn, which is flagged as a known infection source in virustotal.com.

Upon further checking of the executable file, it is revealed that it also communicates with another domain, which is casinosmart.info. This domain only has one associated IP address. However, when cheking the IP address, the IP address turns out to resolve to multiple domains. The fact that many of these domains seem to be created randomly, and that the IP resolves to multiple domains per day, as many as 6 domains per day, shows that this is a possible Domain Generation Algorithm (DGA).

Based on the analysis, both domains project a characteristic unseen onp revious fast flux domains. While previous fast flux domains have multiple unique ASN due to the spread of the bot's geolocation, these domains have a single unique ASN. This may be caused by the botnet controller being hosted on Amazon Web Service, which practice has been rising since early 2017.

4. Conclusion

Botnet is a type of malware that is growing and increasing in size and damage. The existence of fast flux helps botnet maintain stealthiness and causes botnet to be able to do attacks without being detected. However, die to its nature, fast flux can still be detected and in turn, prevented.

DNS is one of the most important function of the Internet. Its purpose is to resolve domain name to its associated IP address. Because of its importance and design, hackers can exploit DNS to do malicious activities. One of the exploits of DNS is fast flux.

Fast flux can be detected from its features, which is significantly different from normal Internet traffic. However, there are also a few legitimate Internet services which have similar characteristics as fast flux, namely CDN and Round Robin DNS. When detecting fast flux, one needs to consider the existence of both as to not generate false positives.

This research also found a new trend of fast flux domains, which is to use cloud hosting service. This is believed to be done because the cloud hosting service provides significantly greater performance than normal zombie machines, which normally consists of home desktops. This is beneficial for botmasters because it provides better scalability. Aside from that, it seems that there is no particular steps done by the cloud hosting service to prevent its service from hosting malware, so for the time being, the bots hosted on cloud hosting services are safe.

References

Chellappa, R. K. and Pavlou, P. A. (2002) Perceived information security, financial liability and consumer trust in electronic commerce transactions, Logistics Information Management, 15(5/6), pp. 358368.
doi: 10.1108/09576050210447046.

Snyder, D. (2010) The very first viruses: Creeper, Wabbit and Brain. Available at: http://infocarnivore.com/the-very-first-viruses-creeperwabbit-and-brain/.

Spamhaus Malware Labs (2018) Spamhaus Botnet Threat Report 2017. Available at: https://www.spamhaus.org/news/article/772/.

Ma, J. et al. (2009) Beyond Blacklists : Learning to Detect Malicious Web Sites from Suspicious URLs, World Wide Web Internet And Web Information Systems, pp. 12451253. doi: 10.1145/1557019.1557153.

Mahmoud, M., Nir, M. and Matrawy, A. (2015) A Survey on botnet architectures, detection and defences, International Journal of Network Security, 17(3), pp. 272289.

Nazario, J. and Holz, T. (2008) As the net churns: Fast-flux botnet observations, 3rd International Conference on Malicious and Unwanted Software, MALWARE 2008, pp. 2431. doi: 10.1109/MALWARE.2008.4690854.

Mahjoub, D. (2013) Monitoring a fast flux botnet using recursive and passive DNS: A case study, eCrime Researchers Summit, eCrime. doi: 10.1109/eCRS.2013.6805783.

Bilge, L. et al. (2011) EXPOSURE : Finding Malicious Domains Using Passive DNS Analysis, Ndss, pp. 117. doi: 10.1145/2584679.

Holz, T. et al. (2008) Measuring and Detecting Fast-Flux Service Networks, Ndss, pp. 2431. doi: 10.1.1.140.188.

Chen, C.-M., Huang, M.-Z. and Ou, Y.-H. (2014) Detecting hybrid botnets with web command and control servers or fast flux domain, Journal of Information Hiding and Multimedia Signal Processing, 5(2), pp. 262273.

Marchal, S. et al. (2012) DNSSM: A Large Scale Passive DNS Security Monitoring Framework. doi: 10.1109/ICCCN.2006.286248.

Weimer, F. (2004) Florian Weimers Home Page. Available at: http://www.enyo.de/fw/.

Perdisci, R. et al. (2009) Detecting malicious flux service networks through passive analysis of recursive DNS traces, Proceedings - Annual Computer Security Applications Conference, ACSAC, pp. 311320. doi: 10.1109/ACSAC.2009.36.

Chen, C.-M., Huang, M.-Z. and Ou, Y.-H. (2014) Detecting hybrid botnets with web command and control servers or fast flux domain, Journal of Information Hiding and Multimedia Signal Processing, 5(2), pp. 262