JACET

J. ADV COMP ENG TECHNOL, 5(3) Summer 2019

Detecting Active Bot Networks Based on DNS Traffic Analysis

Zahra Nafarieh[1], Ebrahim mahdipur[2], Haj Hamid Haj Seyed Javadi[3]

1- Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

2- Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. (mahdipour@srbiau.ac.ir)

3- Department of Mathematics and Computer Science, Shahed University, Tehran, Iran.

Received (2018-08-04) Accepted (2019-03-07)

Abstract: *One of the serious threats to cyberspace is the Bot networks or Botnets. Bots are malicious software that acts as a network and allows hackers to remotely manage and control infected computer victims. Given the fact that DNS is one of the most common protocols in the network and is essential for the proper functioning of the network, it is very useful for monitoring, detecting and reducing the activity of the Botnets. DNS queries are sent in the early stages of the life cycle of each Botnet, so infected hosts are identified before any malicious activity is performed. Because the exchange of information in the network environment and the volume of information is very high, Storing and indexing this massive data requires a large database. By using the DNS traffic analysis, we try to identify the Botnets. We used the data generated from the network traffic and information of known Botnets with the Splunk platform to conduct data analysis to quickly identify attacks and predict potential dangers that could arise. The analysis results were used in tests conducted on real network environments to determine the types of attacks. Visual IP mapping was then used to determine actions that could be taken. The proposed method is capable of recognizing known and unknown Bots.*

Keywords: *Bot Networks, DNS Traffic Analysis, Fast Flux, Intrusion Detection, Network Security, Security Threats.*

How to cite this article:

Zahra Nafarieh, Ebrahim mahdipur, Haj Hamid Haj Seyed Javadi. Detecting Active Bot Networks Based on DNS Traffic Analysis. J. ADV COMP ENG TECHNOL, 5(3) Summer : 129-142

1. INTRODUCTION

One of the increasing threats on the Internet and computer networks that violate the principle of availability is the Botnet. Botnet is a software program that manipulates computers for malicious purposes, known as Bots. Bots are small scripts built to carry out specific automated tasks [1]. These Bots are controlled by one or a small collaborative group of attackers known as "botmaster" [2].

Botnets are considered a launching pad for a number of several illegal activities such as distributed denial of service (DDoS), click fraud [3], phishing, identity theft [4], spamming [5] and malware distribution [6]. Until now, there exists no permanent solution for the detection or mitigation of Botnets threats because their techniques and methods keep changing over time [7]. The Botnet detection process stands as an ongoing challenge for researchers and organizations [8].

Domain Name System (DNS) represents one of the core components of the Internet that facilitates decoupling of service names and hosting IP addresses[9]. However, in addition

to the pivotal role in functioning of the Internet, DNS is often abused by cyber criminals[9].Most of the malware such as trojan and other remote access tool often uses DNS especially Dynamic DNS (DDNS) to locate command and control server. The primary convenience of dynamic DNS is that, the user can change the domain to point to a new IP address at any time. There are plenty of dynamic DNS providers such as NO-IP and DynDNS, and most of them are free[10].

In order to access the command and control server, Bots ask the DNS to locate the command and control server which is typically hosted by the DDNS provider. therefore, detection of DNS Botnet traffic is possible by monitoring DNS and detecting DNS traffic abnormalities.

for example, Distributed denial of service (DDOS) attacks typically uses a hierarchical control network formed Botnet attacks and attacks [11]. A DDoS attack is designed to interrupt or shut down a network, service, or website [12] so that the victims were infected with the virus unknowingly and become part of the puppet. DDoS usually attack by a huge amount of packets. In recent years, the concept of big ideas and technical data develop rapidly and become popular. In 2015, Big Data has become part of many industries [13].

The purpose of Splunk engine collection is to deal with big data generated by these machine language, and made graphical reports with the data. We can observe the characteristics of DDoS attack with data visualization and use Splunk platform to do data analysis to determine the attacks quickly and predict possible hidden crisis [14].

In this research, Botnet discovery is performed on the basis of two criteria:

1. Signature Discovery: Signature-based detection techniques can be used to detect known Botnets; therefore, this solution does not work for unknown and new Botnets.

2. Abnormal behavior-based: Abnormal behavioral detection techniques attempt to detect network-based anomalies based on network traffic abnormalities such as high network latency, high traffic volume, traffic on unusual ports and abnormal behavior of the system, which can represent malicious Botnets in the network.

As shown in the flow of Fig. 1, and doing an analysis from the graphical reports generated by the network traffic, the attack can be detected. In the experiment, the network-traffic flow of the organization is transferred to Splunk platform to make it graphically and analyzed. Splunk analysis platform will start from the flow of traffic, and detect whether there are abnormal flow in and determine attacks.

In this paper, we propose a novel system placed at the network egress point that aims to efficiently and effectively detect malware infections based on traffic analysis. we received information about known Botnets from reputable resources that provide services under the CTI(cyber threat inteligence) name. These data are generated by the intrusion detection sensor, so that network traffic is given to the intrusion detection sensor and various logs are generated and registered by it. Splunk analysis platform stores and categorizes all generated logs and by generating dashboards, It's possible to monitor the network security status.

The main contributions of this paper are as follows:

- This method present a novel system placed at the network edge using a combination of malicious DNS detection technology and intrusion detection technology to detect malware infections inside the network.
- In this approach not only does not reduce the volume of network traffic which needs to be recorded and analyzed, but also improve the sustainability of the system.
- This method defined the specifics, and by focusing on these features, we identified compromised clients that have been remotely controlled.
- In this article, we provide a method for analyzing DNS logs and effectively detect fast-fluxing domain names and domain-fluxing and DDOS attack.
- In this method, signature-based and behavior-based analysis is used to identify Botnets.
- Momentary, periodic, and online monitoring of network security status and dynamic analytics to find malicious activities on the network.

The rest of the paper is organized as follows: Section 2 presents existing work on identifying malicious DNS traffic. Section 3 introduces a novel method for identifying Botnet activity

and potentially compromised clients within large-scale networks by monitoring DNS traffic. Section 4 illustrates an example of the DNS log reviewed in this paper. Section 5 examining two samples of Logs which is called Cyber Threat Intelligence (CTI). Section 6 presents the results of the proposed detection method. Finally, Sect 7 concludes the paper.

2. RELATED WORKS

Several methods have been proposed for detecting Botnets, Some of them are referred to below:

Choi et al. [15] presented a method for identifying Botnets based on anomalies. By controlling group activities in DNS traffic, this method detects Botnets at different stages of their life cycle. Identifying botnets should begin at the stage of their life cycle infection, or at least before the implementation of the command. This method detects the botnet at different stages of their life cycle by controlling group activities in DNS traffic.

Gu and colleagues [16] identify botnets by using cluster-based techniques during an attack phase.

In this way, similar traffic and similar malicious traffic are initially clustered in order to identify the host with a similar pattern to the malware activity and similar communication pattern. The major weakness of this approach is offline working that is considered a major drawback in the Botnet detection systems.

Yadav et al. [17] developed a recognition algorithm in order to detect new generation Botnets which uses characters in a DNS query.

In this way, high accuracy and low false positive rate depends on a large number of DNS queries, and there is a high rate of false positive due to the lack of attention to the history of the activities of the suspected group of hosts in the network. In this way, only known Botnets are detected, which is a major weakness in botnet detection methods.

Huang [18] has proposed a method for detecting Botnets in a local network. Failures are collected for each of the infected and non-infected hosts in TCP and UDP streams. Eventually, a number of feature vectors are extracted and classified as inputs to the algorithm, and a model is generated that is used to identify infected viruses. The main drawback of this method is that it relies on the failure of each host's traffic and it does not pay attention to the history of group activities and also has false negative.

fic. Kopis [19] proposed a detection method that analyzed DNS traffic for identifying malware domain names in DNS traffic. By analyzing domain names in DNS traffic, this method identifies malicious domain names that belong to top-level domain (TLD) such as *.ca. The main point of this approach is the fact that it does not cover Fast-flux and it works on the domain names with a particular TLD, it's difficult to identify the domain name of the TLD and according to the authors, it is difficult to identify DGA domain names.

Pleiades [20] is a technology which identifies domain names that encountered with NXDOMAIN(domain does not exist) by analyzing the queries from DNS traffic. In a huge amount of failures, since most of the domain names generated do not belong to registered domains, resulting in NXDOMAIN responses. The main weakness of the method is that it only considers the Domain-Flux and does not consider Fast-Flux.

Perdisci et al. [21] developed a large-scale system that analyzes DNS traffic and it placed at the edge of the network to detect Botnet infections. This approach identifies the clusters of malicious domain names using a supervised Machine Learning Algorithm (MLA). The primary drawback of this method is that it is more prone to generating false positives and is not able to identify Domain-flux.

Bilge et al. [22], in order to detect domain names associated with various types of malicious activities, proposed a large-scale system which analyzes DNS traffic. The detection system classifies domain names as malicious using supervised Machine Learning Algorithm (MLA) according to a collection of statistical features. This method is capable of detecting Domain-Flux contamination by analyzing DNS traffic and accepts passive DNS analysis techniques to identify malicious domain activity.

No previous work has tried to identify malicious domain names, IP-flux, Domain-flux and malware infections activities, online and

using Big Data. We place the system on the edge of the network and analyze network traffic to detect infected machines inside the network.

3. PROPOSED METHOD

Attackers can remotely control infected devices and steal sensitive information. DNS is popular for malware prevention because of its locating command and control server. To detect Bot networks, it can perform identification tasks at 2 points: 1) analysis within computers ¬ and 2) network traffic analysis. In this research, we propose a new system that is located at the edge of the network, which is the point of the network output that effectively detects malicious behavior by analyzing the network traffic. In this system, behavioral analysis and signature-based analysis are used to identify the Bot networks.

The experiment was conducted for two months in an extensive academic network. Using our approach, we did not only attempt to reduce traffic volume for analysis but also investigated all the features in the big data volume. One of the goals of this research was a behavioral diagnosis that relied on the DNS to identify the attacks. The proposed system is located at the edge of the network in order to record and monitor the network traffic. In the first stage, the network traffic was collected. Suricata, an Intrusion Detection Sensor, was used to generate logs for the network traffic. The Suricata sensor received the network traffic and then generated the DNS, HTTP, TLS, and NetFlow logs. In this study, the DNS traffic comprised only a little percentage of the total network traffic. In addition, we received information about known Botnets from external authoritative sources that provided services called Cyber Threat Intelligence (CTI). It is worthy to note that the PCAP (packet capture) traffic related to the infected traffic was integrated into the Splunk platform, and we saw the pattern of the attack types. Finally, the results of the network traffic and known malware information integrated into the Splunk platform were displayed in the form of dashboards.

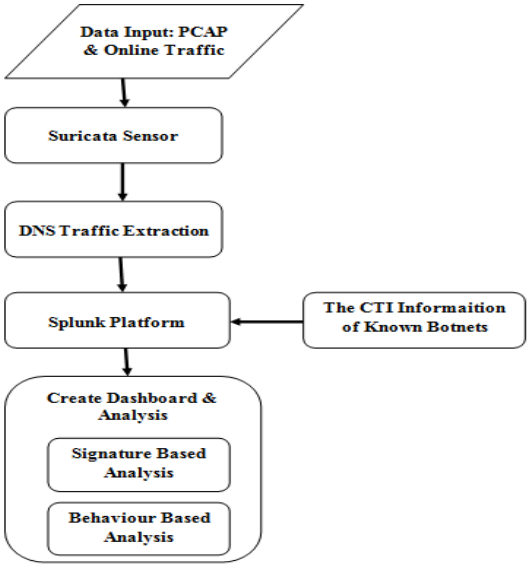

Fig.1. Flowchart of proposed methodology

The Splunk engine deals with the bulky data created by the network traffic and generated graphical reports of data. For example, we could see the features of the DDOS attack with the visual data and used the Splunk platform to analyze the data in order to determine the rapid attacks and predict the hidden crisis. If we can obtain useful information through data analysis, we can then prevent internal or external attacks resulting from deliberate or inadvertent manipulations.

In this research, we used some indicators to identify the various features of malware-related DNS, as well as features that could identify vulnerable clients:

1. Fast-flux or IP-flux: In the fast-fluxing technique, a Bot establishes communication with its command and control server, by using a single domain name that resolves to vary the IP addresses over a short time. This frequent change of IP addresses circumvents the traditional IP-based traffic filtering because it takes some time to detect a new IP as malicious. Attackers abuse this ability to change IP address information related to a host name by linking multiple IP addresses with it and rapidly changing the linked addresses over time. In traditional security systems, blocking domain names is generally more complicated compared to blocking IP addresses. Fast-fluxing can be understood as a behavior of the domain name system.

2. Domain-fluxing (domain generation algorithms): In this technique, instead of a fixed embedded domain name, the malware is equipped with an algorithm which is known as DGA (domain generation algorithm) and it generates domain names at runtime. The use of DGA instead of a list of hard-coded domains significantly strengthens the capability to evade detection and/or deletion. Domain names have two or more segments separated by a dot (.). Bot managers can, however, exploit the creation of sub-domains or second level domains (SLDs). A Bot can buy a domain name like malicious. com from a domain name registrar, and then build the subdomains "dga1.malicious.com" and "dga2.malicious.com" to save on its costs. Even if the traffic of a sub-domain is blocked, the traffic to other sub-domains within the same SLD will not be blocked. Continuous change of domain is a method by which a large number of domain names are mapped to the same address by DNS queries.

3. Suspicious group activity: One of the inherent features of the Botnets is the presence of command and control channels which make infected hosts with the same malicious code function as a coordinated group activity. Sending multiple DNS queries is called a group activity. In this research, the detection of online Botnets activity by using suspicious group activities in DNS traffic was investigated.

4. Suspicious failure: In a host that has been infected with malicious code, a predetermined algorithm in the Bot code generates a list of domain names of command and control servers dynamically. Algorithmic domain names will fail if they do not map to the address of a command and control server. A DNS query will fail if the domain name of the query is not mapped to the address of any server. Failures in a host DNS query are termed suspicious if their number exceeds a certain threshold. By examining the normal network traffic and the Botnet infected network traffic, it can be observed that the number of NXDomain responses in infected hosts is much higher than that of normal hosts. As a result, the number of NXDomain responses is calculated over a period of time, and those that are more than a threshold are considered suspicious.

5. Longevity of domain names: The attacker changes the address of the command and control servers dynamically by generating a large number of domain names and choosing a short lifespan for them. Therefore, , Botnet recognition systems

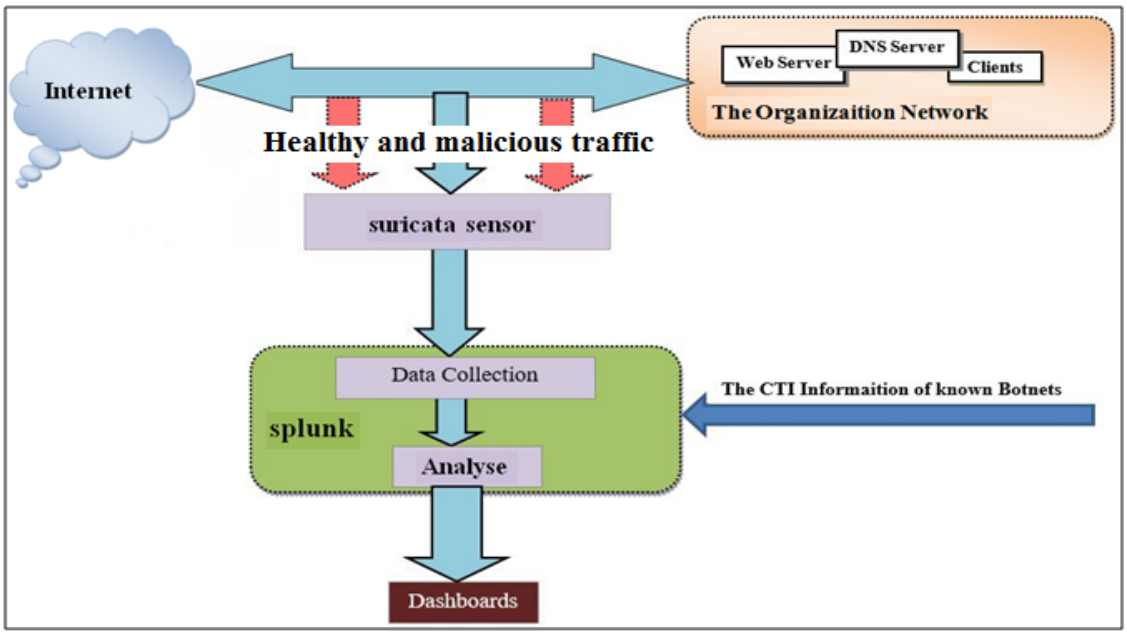

Fig.2. The architecture of the Botnet detection system is based on traffic analysis

should identify the domain names generated by the algorithm that has a short TTL (Time-To-Live).

Time-To-Live (TTL) is set by an authoritative name server for a DNS record. TTL means how long a resolver may cache the response result for a domain.

6. Alternate periodic activities: Hosts infected with Bots periodically and almost at all time windows and alternately are connected to the command and control server and receive Bot commands.

7. Similar suspicious activities at different time windows: Due to the fact that infected hosts execute the same malicious code, it is expected that at different time windows, the participating hosts will have a small variation in suspicious activities. In other words, if there are different hosts in suspicious activity at different time windows, there is a little possibility that these hosts are infected with Bots.

4. A SAMPLE OF THE DNS LOG

This log is a .txt format and is related to the organization's network traffic of which Part is visible in Figure 3. The first part displays the date and time. The second part specifies whether a client has been sent a request or the client has received a response. The third part displays the result of this request such as NXDOMAIN or URL address. The fourth part displays the destination and source address.

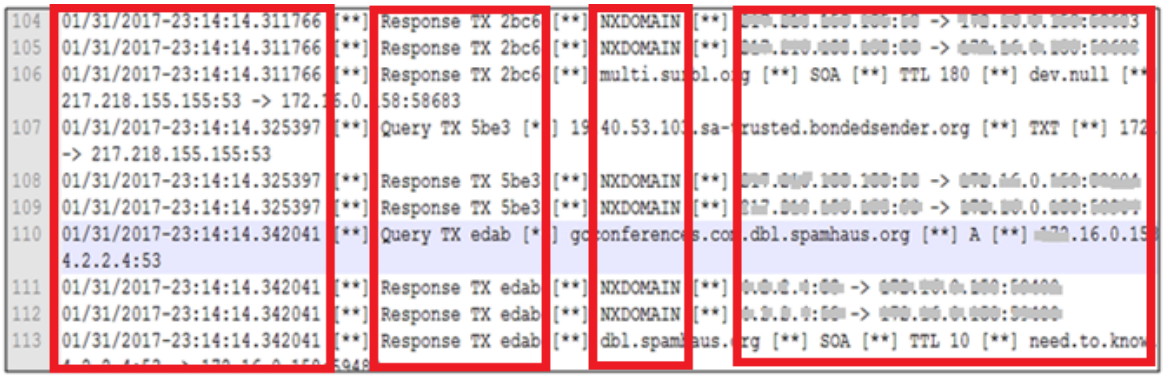

Fig.3. An example of a DNS Log

5. EXAMINING TWO SAMPLES OF CTI LOG

In order to detect Bots, information about known Botnets was collected from trusted sources that provide services under the CTI name, for example from malicious logs (which contained infected URLs), as well as the Controller log containing the hash of the infected malware, IP, port, domain address, and other malware information. This information was extracted as fields and by comparing them with the DNS logs that were available from the network traffic, detection was done.

5.1. A Sample of the Malicious Logs

This log is a .txt format and contains information about the known malware, for example, infected URLs. This information was extracted from the malicious log as fields and the diagnosis was performed by comparing these fields with the DNS logs that were available from the network traffic.

```
26  http://blog-rye.blogspot.com/www.facebook.com/plugins/likebox.php?href=https://www.faceb
    &amp;show_faces=true&amp;colorscheme=light&amp;stream=false&amp;show_border=true&amp;hea
27  http://apkhotgames.blogspot.com.br/2013/10/dead-effect-autohealing-v102-apk.html
28  http://noticias-do-moment.blogspot.com/2013/07
29  http://notivideox.blogspot.mx/2011/07/el-video-y-fotos-de-juanita-viale-y.html
30  http://googletrendsonly.blogspot.com.es/2012/11/imogen-thomas.html
31  http://netflixmagazine.blogspot.com.es/2008_07_01_archive.html
32  http://duakitabi.blogspot.com.es/2011/07/ya-fettah.html
33  http://malaysianinfonews.blogspot.ru/2014/07/kejam-gaza-di-robek-israel-lagi.html
34  http://freegayclip.blogspot.fr/2012/01/andy-roddick-david-archuleta-and-joe.html
35  http://downloadvectordep.blogspot.com/2010_08_01_archive.html
36  http://estelaraziel.blogspot.in/2012/05/anushka-boobs-exposed_7.html
37  http://apivones.blogspot.com/2011_09_01_archive.html
38  http://gamezaya.blogspot.com.es/2014/03
39  http://bbvaball.blogspot.co.uk/2011/10/fernando-torres-el-nino-is-back-pictures.html
40  http://hot-fuck-guys.blogspot.com/2012/07/today-jack-off_18.html
41  http://listfamousactresses.blogspot.mx/2011/10/latest-pics-of-deanna-russo-photos.html
```

Fig.4. An example of a Malicious Log


```
"malware": [
    {
        "sha1": "8fae27750683e23b1c38645a2bb4e9573861f4fc",
        "md5": "38d6e6d0830906d60165b4f11fc75e02"
    },
    {
        "sha1": "d060d7ac04e4dc14d4d6a39d5b1dfe409914dadb",
        "md5": "40a6c73cbb5b614f9a9d2d445ab683d5"
    }
],
"protocol": "6",
"mask": "78.58.230.174",
"port": "1604",
"controller": [
    {
        "last_verified": "2016-02-26 17:16:55",
        "first_active": "2014-09-11 04:16:48",
        "went_offline": "2016-02-26 17:17:25",
        "came_online": "2016-02-26 17:13:38",
        "active": "0",
        "ip": "78.58.230.174",
        "confidence": "100"
    }
],
"threat": "darkcomet",
"resolves": "0",
"active": "0",
"first_seen": "2014-09-11 04:16:48",
"type": "tcp",
```


Fig.5. An example of a Controller log

5.2. A Sample of the Controller Log

This log has a JSON format and contained information about the infected malware. The log comprised of information such as the malware name, malware hashes which are mostly presented in two sha1 and md5 formats, malware activity; either active or not active, confidence, and other malware information. Figure 5 shows part of this log.

6. MONITORING AND RESULTS

This proposed system uses signature-based detection and anomaly-based detection together, to provide the maximum defense for the monitoring network. The signature-based technique detects only known bots through signature matching using the IDS (intrusion detection system). The anomaly-based or behavior-based method, on the other hand, tries to detect bot networks through analyzing group activity of bots, network traffic, traffic to unusual ports, sudden amount of network traffic, anomalous behavior and high network latency that may show the existence of bots in the network. These approaches have the ability to identify new bots [8].

The anomaly-based method is categorized into host-based and network-based [8]. DNS-based diagnostic methods do not require any basic knowledge around protocols of bot networks, communications, or botnet signatures because most of the bot networks use the domain name to locate their command and control servers (C & C) and to collect information which is stolen from infected hosts [23]. Thereupon, the DNS traffic of a Botnet can be identified and monitored by detecting anomalies in the DNS traffic [6]. The proposed method identifies the infected host by monitoring the DNS traffic.

6.1. View the Source and Destination IPs on the Google Map

By specifying the source and destination IPs on a Google map, we could get a good understanding of the network traffic according to the location of the source and destination IPs. Once the number of source IPs referred to a destination address, for example, a DNS server will rise to a certain degree, and depending on the circumstances, it could be suspected of a DDOS attack and on the Google map, we could see the volume of requests from the source IP to the destination IPs. For example, in figure 6, an alert that indicated a redCount of 2,545 times was raised referring to a suspected DDOS attack on a destination address located in San Francisco.

6.2. The location of the Clients with NXDOMAIN Response on a Google Map

Using the Google Map in the Splunk platform allowed us to track the Source and Destination IPs. That is, in this map, we could determine the sources from where most requests were sent and also the targeted destinations. Figure 7. illustrates an example of tracking the source IP to the destination, and looking at the figure, we could easily identify the sources from where information was sent and their targeted destinations. The geolocation of the clients whose requests were encountered with the NXDOMAIN(domain does not exist) response is displayed on the Google map.

6.3. Checking the Traffic of Queries from the DNS Server

Most of the Bots used DNS in their communication and the DNS traffic had unique features defined as a group activity. Therefore, by using the group activity property of Botnet DNS traffic, the Botnet was detected. Due to the fact that the Bots were grouped and almost simultaneously send their requests to the DNS for access to the command and control server, we saw a group activity and traffic requests to the DNS server. In Figure 8, the traffic which was sent to the DNS server is displayed with a time chart and the changes can be seen online in this chart. This abnormal increase of the traffic can be suspected of Bot activity. According to figure 8, at a particular time, the DNS traffic went up above 1200, indicating a suspicion of Bots activity on the network.

6.4. Checking the Number of Referrals to a Specific Domain in DNS Traffic

Figure 9 actually represents a DDOS attack. As we can see from the figure, a referenced traffic count of 4,326 times within a specified time period was made to a domain address. Since it came from different clients, it was suspected of

Fig.6. View the number of referrals to Destination IPs on the Google map

Fig.7. Location of the clients with NXDOMAIN requests

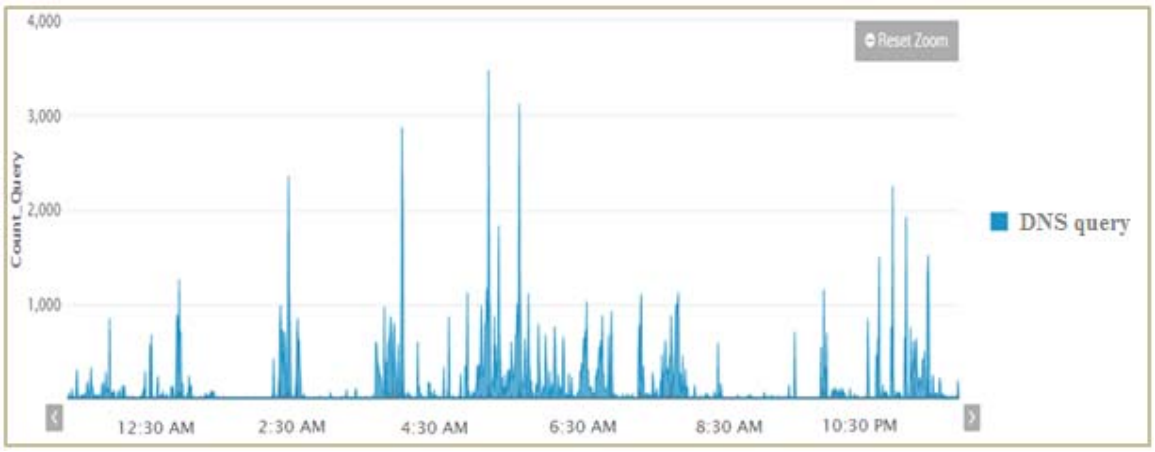

Fig.8. Shows the traffic of queries from the DNS Server

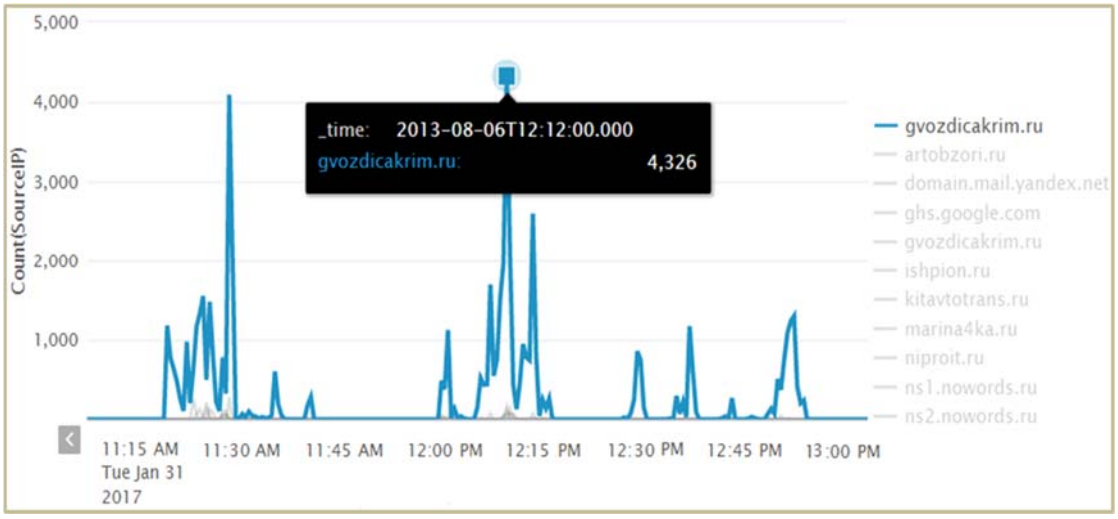

Fig.9. Number of referrals to a specific domain in DNS traffic

a DDOS attack. An increase in traffic associated with this type of request within a specified time period can be suspicious of a Bot activity in the organization's network.

6.5. Detecting Domain-Fluxing in DNS Traffic

In the Domain Flux technique, several domain names were assigned to an IP address which was constantly changing. One of the latest and most known changes to the Domain-Flux was based on Domain-Generating Algorithms (DGA). A DGA is an algorithm that can generate a large set of domain names and create a large number of NXDOMAINs, because most of the DGA generated domain names do not belong to registered domains. [9]. This method generates a dynamic list of several Domain names, some of which are already registered,and the Bots then communicate with the broker via command and control. Generated domains have a short life span of one day and expire quickly. As a result, finding and tracking such domains is difficult and, in some cases, impossible. Figure 10 is an example of a Domain Flux attack on a DNS. As can be seen in this anomaly, for a specified IP

address, "74.125.232.195", there are different domain addresses and their TTL values are also low, indicating a short life span.

6.6. Detecting fast-fluxing domain names in DNS Traffic

In the Fast-flux or IP-flux technique, the IP addresses are frequently changed. Attackers, by linking multiple possible IP addresses with a single domain name, exploit this ability to change an IP address related to a hostname, thereby rapidly changing the linked addresses over time. The fast-flux technique uses a combination of round-robin allocations and very short time-to-live (TTL) values, by registering multiple IP addresses to evade detection and blocking [9]. In the IP Flux technique, mapping of the command and control server at short intervals execute changes to the IP address in order to prevent its detection by a Botnet. By investigating this type of Botnet there is an indication that a new set of IP addresses is assigned to the domain names using IP Flux every 3 to 10 minutes. Therefore, this technique prevents a large number of Bots from connecting to a unique IP address and reveals their group behavior. Figure 11 is a sample of a

DNS_RequestType	DNS_Question_DomainName	DNS_TTL	Dns_Response
Response	zfxqvbd.akazxy.net	TTL 27	74.125.232.195
Response	zfxaqwd.akazxy.net	TTL 19	74.125.232.195
Response	zfxwert.akazxy.net	TTL 11	74.125.232.195
Response	zfxwer.akazxy.net	TTL 32	74.125.232.195
Response	zfxoio.akazxy.net	TTL 25	74.125.232.195
Response	zfxrt.akazxy.net	TTL 24	74.125.232.195
Response	zfxhjlk.akazxy.net	TTL 22	74.125.232.195
Response	zfxzdfh.akazxy.net	TTL 20	74.125.232.195
Response	zfxzcd.akazxy.net	TTL 30	74.125.232.195
Response	zfxdf.akazxy.net	TTL 28	74.125.232.195

Fig.10. Checking Domain Flux misuse in DNS Traffic

DNS_RequestType	DNS_Question_DomainName	DNS_TTL	Dns_Response	DNS_Question_Type
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.36	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.35	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.34	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.41	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.40	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.39	A
Response	any-world.ngd.ysm.yahoodns.net	TTL 11	217.163.21.38	A

Fig.11. Checking Fast Flux misuse in DNS Traffic

Fast Flux attack on a DNS traffic, having different IP addresses assigned to a given domain address in this anomaly and their TTL values are also low, indicating they have a short life span.

6.7. The Periodic Activity of Bot's Communications

In a vulnerable network where one of the hosts has been infected with malicious code, it is likely that the Bot will be released to other hosts in that network. Generally, Bots need to communicate periodically with their command and control server. This periodic behavior in infected logs was discovered according to the output charts.

Figure 12 represents the periodic behavior in DNS traffic. Three different IP addresses related to three different clients infected with the Bot code which periodically sent its requests to the DNS server at specific intervals is displayed.

As shown in the figure13, the intended Bot related to the client with IP '192.168.89.2' which is infected with the Bot code sent a request to DNS at intervals of 15 minutes.

6.8. Checking the A Records in DNS Traffic

In the Bot networks, the network members try to communicate with the command and control server to receive attack commands. For this reason, they start sending requests of A record to the DNS system. The DNS system returns in response to these requests the corresponding A record and if the domain name does not exist, they receive NXDomain. Therefore, in the network traffic infected with Bots, there would be an increase in the number of requests of A record

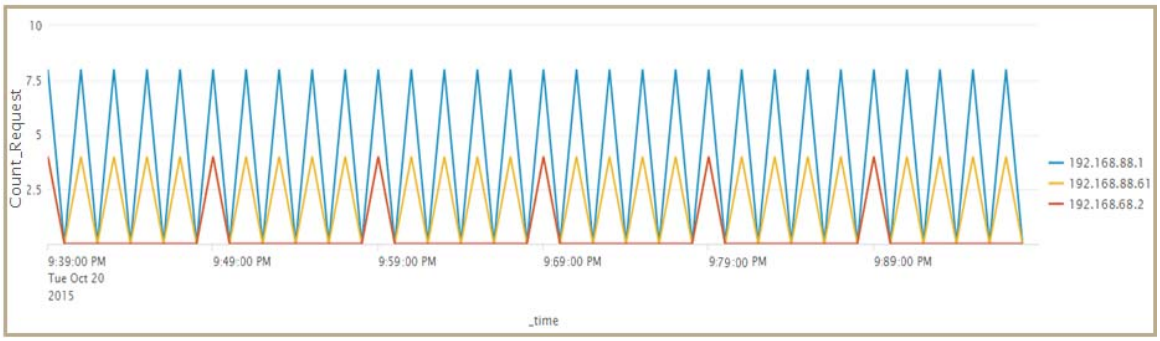

Fig.12. Periodic activity in the Bots communication

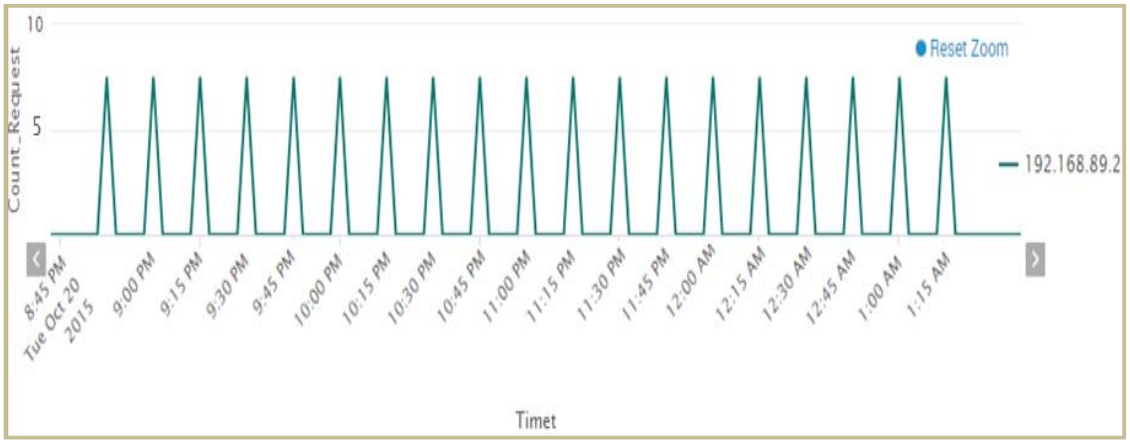

Fig.13. The Bot activity

and subsequently, an increase in unsuccessful responses or NXDomain, and eventually, an increase in DNS traffic volume over the entire network traffic. In Figure 14, the requests of A record to the DNS system is displayed from 10:15 PM to 11:17 PM. Increasing traffic associated with this type of request within a specified time period can be a suspicion of Bot activity in the organization's network.

DNS traffic analysis alone can not produce high-accuracy results. However, it can be used as a supplement to integrate its results with other traffic and traffic analysis methods to enhance the accuracy of Botnets identification. Due to the difference between the networks and the use of different Bots in the experiment, a complete comparison between the methods could not be made. In Table 1, the proposed method was compared with five different features with a number of available methods for detecting new generation Botnets. In fact, in this proposed method, more queries increased the accuracy of the response in the output.

7. CONCLUSION

Due to the massive amount of information from network traffic and the amount of malware that is added every day, Using a system that uses one of the big data tools to analyze security logs is important. In this paper, with behavioral and signature-based analysis, we proposed a method for detecting online Botnet activity by monitoring DNS traffic. This novel system placed at the network egress points to detect malware infections inside the network by using Splunk big data technology and DNS traffic analysis. The results of momentary, periodic, and online monitoring of network security status and dynamic analytics helped to identify security problems before an

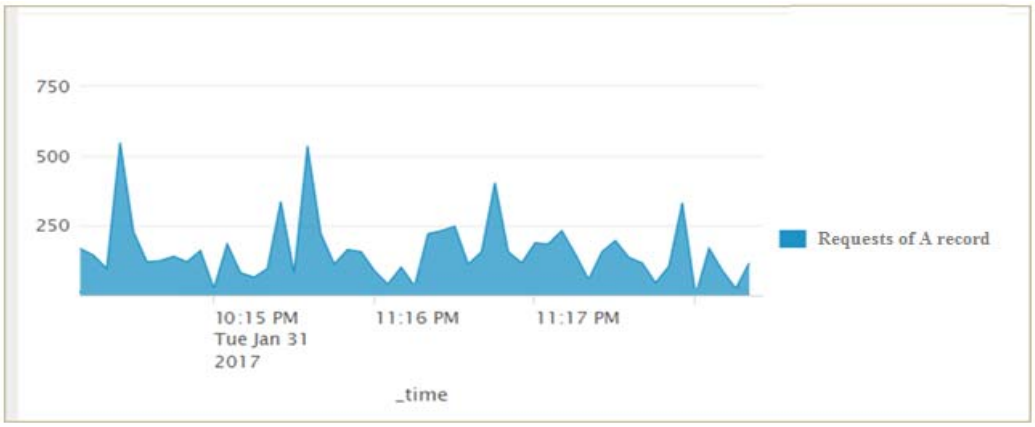

Fig.14. Checking the A records in DNS traffic

Table.1. Comparison of methods for detecting new generation Botnets.

Botnet detection method	The required number of DNS queries	Correlation of different logs and error notification	Diagnosis based on suspicious activity history	Detection based on failure in DNS traffic	Diagnosis based DNS group activity
Yadav [17]	High	×	×	✓	✓
Choi [15]	High	×	×	×	✓
Huang [18]	medium	×	×	✓	×
suggested method	High	✓	✓	✓	✓

attack happens and help managers to reduce the effects of intrusion into the organization network. The experimental results show that our security approach is good for detecting known and unknown malware infections and is feasible for improving the sustainability of the system. In the future works, we are going to combine the results of the DNS traffic analysis with other traffic and traffic analysis methods to increase the accuracy of the botnets identification.

Acknowledgment

The author would like to thanks the security team and their guideline, suggestions and technical support during the work.

REFERENCES

1. Alomari, E., Manickam, S., Gupta, B.B., Karuppayah, S. and Alfaris, R., 2012. Botnet-based distributed denial of service (DDoS) attacks on web servers: classification and art. arXiv preprint arXiv:1208.0403.

2. Lu, W., Rammidi, G. and Ghorbani, A.A., 2011. Clustering botnet communication traffic based on n-gram feature selection. Computer Communications, 34(3), pp.502-514.

3. Almomani, A., Gupta, B.B., Wan, T.C., Altaher, A. and Manickam, S., 2013. Phishing dynamic evolving neural fuzzy framework for online detection zero-day phishing email. arXiv preprint arXiv:1302.0629.

4. Al-Momani, A., Wan, T.C., Al-Saedi, K., Altaher, A., Ramadass, S., Manasrah, A., Melhim, L.B. and Anbar, M., 2011. An online model on evolving phishing e-mail detection and classification method. journal of applied science, 11(18), pp.3301-3307.

5. Alieyan, K., ALmomani, A., Manasrah, A. and Kadhum, M.M., 2017. A survey of botnet detection based on DNS. Neural Computing and Applications, 28(7), pp.1541-1558.

6. Zeidanloo, H.R., Shooshtari, M.J.Z., Amoli, P.V., Safari, M. and Zamani, M., 2010, July. A taxonomy of botnet detection techniques. In Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on (Vol. 2, pp. 158-162). IEEE.

7. Karim, A., Salleh, R.B., Shiraz, M., Shah, S.A.A., Awan, I. and Anuar, N.B., 2014. Botnet detection techniques: review, future trends, and issues. Journal of Zhejiang University SCIENCE C, 15(11), pp.943-983.

8. Alieyan, K., ALmomani, A., Manasrah, A. and Kadhum, M.M., 2017. A survey of botnet detection based on DNS. Neural Computing and Applications, 28(7), pp.1541-1558.

9. Stevanovic, M., Pedersen, J.M., D'Alconzo, A. and Ruehrup, S., 2017. A method for identifying compromised clients based on DNS traffic analysis. International Journal of Information Security, 16(2), pp.115-132.

10. Zhao, G., Xu, K., Xu, L. and Wu, B., 2015. Detecting APT malware infections based on malicious DNS and traffic analysis. IEEE Access, 3, pp.1132-1142.

11. Das, S., Mukhopadhyay, A. and Shukla, G.K., 2013, January. i-HOPE framework for predicting cyber breaches: a logit approach. In System Sciences (HICSS), 2013 46th Hawaii International Conference on (pp. 3008-3017). IEEE.

12. Bhandari, A., Sangal, A.L. and Kumar, K., 2016. Characterizing flash events and distributed denial-of-service attacks: an empirical investigation. Security and Communication Networks, 9(13), pp.2222-2239.

13. Woodie, A., 2015. Why Gartner dropped big data off the hype curve.

14. Marty, R., 2009. Applied security visualization (p. 552). Upper Saddle River: Addison-Wesley.

15. Choi, H. and Lee, H., 2012. Identifying botnets by capturing group activities in DNS traffic. Computer Networks, 56(1), pp.20-33.

16. Gu, G., Yegneswaran, V., Porras, P., Stoll, J. and Lee, W., 2009, December. Active botnet probing to identify obscure command and control channels. In Computer Security Applications Conference, 2009. ACSAC'09. Annual (pp. 241-253). IEEE.

17. Stone-Gross, B., Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., Kruegel, C. and Vigna, G., 2009, November. Your botnet is my botnet: analysis of a botnet takeover. In Proceedings of the 16th ACM conference on Computer and communications security (pp. 635-647). ACM.

18. Huang, C.Y., 2013. Effective bot host detection based on network failure models. Computer Networks, 57(2), pp.514-525.

19. Antonakakis, M., Perdisci, R., Lee, W., Vasiloglou, N. and Dagon, D., 2011, August. Detecting Malware Domains at the Upper DNS Hierarchy. In USENIX security symposium (Vol. 11, pp. 1-16).

20. Antonakakis, M., Perdisci, R., Nadji, Y., Vasiloglou, N., Abu-Nimeh, S., Lee, W. and Dagon, D., 2012, August. From Throw-Away Traffic to Bots: Detecting the Rise of DGA-Based Malware. In USENIX security symposium (Vol. 12).

21. Perdisci, R., Corona, I. and Giacinto, G., 2012. Early detection of malicious flux networks via large-scale passive DNS traffic analysis. IEEE Transactions on Dependable and Secure Computing, 9(5), pp.714-726.

22. Bilge, L., Sen, S., Balzarotti, D., Kirda, E. and Kruegel, C., 2014. Exposure: A passive dns analysis service to detect and report malicious domains. ACM Transactions on Information and System Security (TISSEC), 16(4), p.14.

23. Kang, B.B.H., 2011. DNS-based botnet detection. In Encyclopedia of Cryptography and Security (pp. 362-363). Springer, Boston, MA.