Indonesian Journal of Electrical Engineering and Computer Science
Vol. 18, No. 3, June 2020, pp. 1584~1595
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v18.i3.pp1584-1595

Characterising and detection of botnet in P2P network for UDP protocol

Noor Zuraidin Mohd Safar[1], Noryusliza Abdullah[2], Hazalila Kamaludin[3], Suhaimi Abd Ishak[4], Mohd Rizal Mohd Isa[5]
[1,2,3,4]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Malaysia
[5]Faculty of Defence Science and Technology, Universiti Pertahanan Nasional Malaysia, Malaysia

Article Info

Article history:

Received Sep 15, 2019
Revised Dec 16, 2019
Accepted Dec 30, 2019

Keywords:

Bot
Botmaster
Botnet
P2P
Peer to peer network
UDP
UDP botnet
UDP protocol

ABSTRACT

Developments in computer networking have raised concerns of the associated Botnets threat to the Internet security. Botnet is an inter-connected computers or nodes that infected with malicious software and being controlled as a group without any permission of the computer's owner. This paper explores how network traffic characterising can be used for identification of botnet at local networks. To analyse the characteristic, behaviour or pattern of the botnet in the network traffic, a proper network analysing tools is needed. Several network analysis tools available today are used for the analysis process of the network traffic. In the analysis phase, the botnet detection strategy based on the signature and DNS anomaly approach are selected to identify the behaviour and the characteristic of the botnet. In anomaly approach most of the behavioural and characteristic identification of the botnet is done by comparing between the normal and anomalous traffic. The main focus of the network analysis is studied on UDP protocol network traffic. Based on the analysis of the network traffic, the following anomalies are identified, anomalous DNS packet request, the NetBIOS attack, anomalous DNS MX query, DNS amplification attack and UDP flood attack. This study, identify significant Botnet characteristic in local network traffic for UDP network as additional approach for Botnet detection mechanism.

Corresponding Author:

Noor Zuraidin Mohd Safar,
Faculty of Computer Science and Information Technology,
Universiti Tun Hussein Onn Malaysia,
86400 Parit Raja, Batu Pahat, Johor, Malaysia.
Email: zuraidin@uthm.edu.my

1. INTRODUCTION

Computer networks are critical to modern society. An extensive range of business, infrastructure, and human needs, such as communications, utilities, banks, and leisure services presently delivered by systems that depend on secure and efficient operation of networks. As networks increase in size and complexity, a thorough understanding of their behaviour is crucial to protect them from security threats. One of the utmost substantial intimidations to the network today is the threat of Botnets.

Botnet is a group of inter-connected compromised hosts (also referred to as bots) that being control by an assailant known as the Botmaster [1]. The Botmaster capable to instruct commands to the bots to execute malicious activities. This actions including recruiting new bots, initiating distributed denial of service (DDoS) attack against some hosts, gaining profound information from the bot device, spam emails distribution and other threats [2]. Therefore, botnets have developed as a colossal risk to the web networks. The significant distinction between botnets and other security threats is that a Botmaster communicates frequently with the bots by centralized or decentralized network communications.

These bots perform any type of destruction on receiving the commands from the Botmaster. Another threat of botnet is when they are used as distributed supercomputers by attackers wishing to crack cryptographic keys [3]. It is likely more than six million computers worldwide are connected to a botnet being compromised. The proprietors of infected computers do not know that their computers had been attacked by botnet [4]. Most of the study related to the botnet threats are the preventive actions from the attack of the botnet, identifying the botnet characteristic, detection and the mitigation of the botnet [5]. Based on the current severity of the botnet, it is crucial to develop an agile botnet detector in combating the botnet attack. However, before any tools or mechanism can be developed, the characteristic of the botnet need to be studied thoroughly [6].

This study, will focus on the overall background how the botnet works, analysing the network traffic that suspiciously consist of botnet, study the botnet behaviour by comparing the network traffic analysis before and after the botnet is deployed. The network traffic analysis will determine the variant attack of the peer to peer botnet in User Datagram Protocol (UDP) traffic. The three-way handshakes network communication of Transmition Control Protocol (TCP) is discarded in the characterization process. Random Forest classifier is used to classify network traffic of UDP, TCP and Domain Name System (DNS) for botnet detection, the study capture the heuristics of botnet network activity [7]. Most of the characteristic and classification were used complex mathematical and algorithm, however in this study the analysis of the network traffic will be used to study the botnet behaviour, topologies, lifecycle events and action or the pattern of the botnet. The result of this study could lead to the development of the botnet detector and a future study of the botnet detection and mitigation process.This paper is structured as follows: section 2 describes the literature and previous study, section 3 describes the methodology, section 4 presents the results and discussions and section 5 is the conclusions.

2. LITERATURE AND PREVIOUS STUDY

Botnets are a relatively young malware category, and yet their evolution and usage has propagated at an unprecedented rate. They are a dynamic, constantly evolving threat [8]. They have continuously evolved since their establishment in Internet Relay Chat (IRC) maintenance aids into the current Mirai Botnet [9]. The main component of the botnet is depicted in Figure 1, the four main components are Botmaster, Infected Host or Bot (become zombie), Command and Control Channel (Server) and the Attack Victim. Botnet initializes the first attack through exploiting vulnerabilities in users' computers. Then the users' computer download the malicious binary file and locally executed. This program connects to the command and control server (C&C) and warned its host, known as Botmaster that the computer has become a bot.

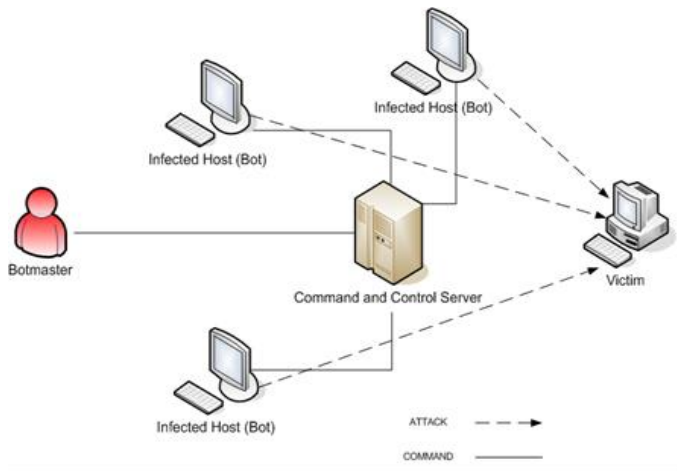

Figure 1. The main component of the botnet

The term of bots is commonly referred to a software application running automated tasks over the Internet. The running application normally consist the malicious software inside the affected computers. These computers are compromised and are being used without owner's knowledge. The compromised machines are called drones or zombies [10]. Now it can be used its influence to spread to additional

computers by echoing the identical process. The big dissimilarity between botnets and other security threats is a Botmaster communicates frequently with bots, either through the centralized or decentralized communication network. These bots will perform any type of harm after receiving instructions from the Botmaster. The Botmasters send the command, control all bots, then there were both unite to perform the attack to the victim [11]. Botnets are developing rapidly in the fast pace which makes it hard to identify and recover from their side effect.

Characterizing botnet by adopting anomaly based detection technique emphases on identifying traffic anomalies that are related to botnet processes [12]. The anomalies are varies from simply noticeable variations in network traffic level and latency to complex abnormalities in traffic flow patterns. One of the most conspicuous anomaly based approaches detect anomalies in packet payloads, DNS traffic and botnet group behaviour [13, 14]. Those characteristics of the network behaviour traffic in Peer to Peer (P2P) network are essential in botnet detection. Node based sampling for the treatment of packet capture mechanism based on P2P botnet process was tested for detection efficiency on the time window of feature extracting and sampling time interval was studied in [15] .DNS based detection methods are constructed on particular DNS information generated by a Botnet. Bots generally initiate association with C&C server to get commands. In order to access the C&C server bots carry out DNS queries to locate the particular C&C server that is typically hosted by a dynamic DNS (DDNS) provider. Thus, it is possible to detect botnet DNS traffic by DNS monitoring and detect DNS traffic anomalies [16]. Botnet C&C traffic is difficult to detect. In fact, since botnets utilize normal protocols for C&C communications, the traffic is similar to normal traffic and does not cause network latency [12]. Data mining base detection technique including machine learning, classification, and clustering can be used efficiently to detect botnet C&C traffic [17]. Botminer used data mining methods for botnet detection in C&C traffic. Botminer clustered related traffic with the similar malicious traffic. After clustering process ended, cross cluster correlation is implemented to Identify hosts who share similar patterns of communication and malicious activity. Botminer is an advanced technological botnet monitoring tool which entirely independent of the protocol and structure of the botnet. Moreover it has a capability to detect botnets including IRC based, HTTP based, and P2P botnets with a very low false positive rate [18]. In signature based botnet detection technique, botnet signatures and behaviours of existing botnets were used in identification and detection method [19]. However, signature based detection techniques only capable to detect of existing known botnets.

In this study, the main focus of the detection strategies is based on the signature, DNS and anomaly approaches. The core protocol in the network is relying on the UDP and Transmission Control Protocol (TCP), both protocols is in the Internet Protocol Suite. As mention earlier, this study will identify the botnet characteristic based on the UDP protocol only in the P2P network environment.

3. METHODOLOGY

3.1. Data Preparation

Real network traffic is used in this research. The real network traffic is captured in several days. The process of capturing is divided into two sessions. First, the normal network traffic is captured in five days and after deploying botnet to the network the network traffic capturing process run again for several days. For anomalous network traffic it took seven days to finally end the capturing process. Two group of dataset is selected in this study, the normal and anomalous network traffic. A server with Linux operating system is used to capture the network traffic. Tcpdump and cronjob run on server computer is responsible to do the capturing process for both type of data. Seven computers were used in this test bed setup. One of the computers is a Sun Blade server computer running Linux operating system and the remaining of the computers is the desktop computer with Windows operating system. All of the nodes are interconnected and accessible to the internet. The capture of network traffic is divided into two separated sessions, starting with normal traffic then followed by the network traffic with a various botnet was executed where each of them established connections on several protocols and performed different actions. Normal network traffic can be defined as a functional network behavior free from anomalies characteristic, meanwhile anomalous network traffic is a deviation in the normal behavior of a network in the presence of any intruder in a network. These anomalous events disrupt the normal functionality of network services [20].

3.2. Normal Network Traffic

In normal traffic, data is captured in a good health of the network environment. During this period the Kaspersky antivirus software is installed as a protection to the any malicious activity in the computers and the network. Torrent application is run to enable the P2P activities throughout the capturing process. Figure 2 shows the network design for capturing the network traffic. The static IP address is set on each of the host. A host use to capture the traffic data is a Sun Blade Server with Linux base operating system.

Tcpdump is used to capture the traffic data. The tcpdump is managed by the crontab services to run the cron job. The crontab process is written in the script name dumpit. After a few days of capturing the network traffic, the data is collected and store in the form of compress data with the following format: <filename>.tcpdump.gz.

Enormous size of data is captured for several days. Most of the data size is larger than ten gigabyte. The length of the captured traffic is about sixty minutes to a hundred eighty minutes of network traffic. Three dataset were selected in this study for analysis which are namely as Normal_1, Normal_2 and Normal_3. Table 1 is the summarization of the selected captured data.

Table 1. Summarization of the Selected Captured Data for Normal Network Traffic

Data Set	Length in second	Tcpdump size in kilobyte
Normal_1	5050	976564
Normal_2	4689	976563
Normal_3	4731	976562

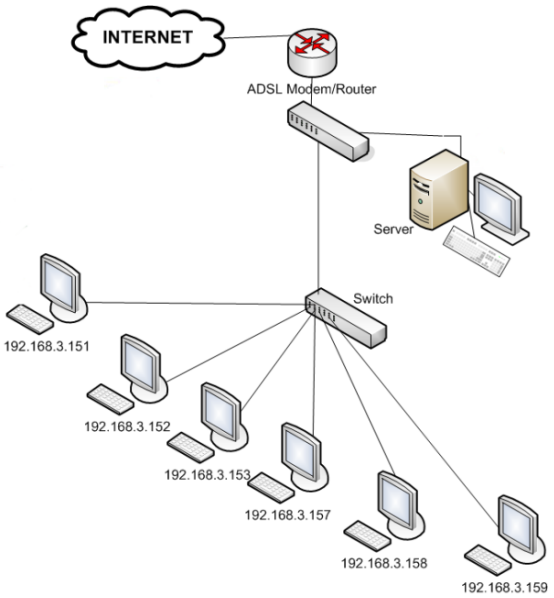

Figure 2. Network Design for the normal traffic

3.3. Anomalous Network Traffic

In anomalous data traffic, data is captured in unhealthy network environment. During this period the Kaspersky antivirus software is uninstalled. Torrent applications remain active to enable P2P activities throughout the capturing process.

The static IP address is set on each of the host and remains unchanged from the previous setup. The configurations of capturing data in normal traffic is remain and were used again for the anomalous traffic, except the Kaspersky antivirus is removed. The data capturing process also remain the same. In anomalous traffic data, two phases of capturing data were involved; the first phase only BitTorrent torrent application run on the network and the second phase, few more torrent applications were installed. Three dataset were selected in this study for analysis which are namely as Anomalous_1, Anomalous_2 and Anomalous_3. Table 2 is the summarization of the selected captured data for anomalous traffic.

Table 2. Summarization of the Selected Captured Data for Anomalous Network Traffic

Data Set	Length in second	Tcpdump size in kilobyte
Anomalous_1	5700	976562
Anomalous_2	4980	781251
Anomalous_3	4800	781250

3.4. Hardware and Software Configuration

There are seven host connected with the two switches. One of the switches is connected to the ADSL modem/router to establish the internet connectivity. Figure 2 show the same network design and connected nodes. The same computers and peripherals setup is use for both data collection. Server with Linux base operating system is use to capture and store the data. Tcpdump and cronjob are responsible to manage the entire task of capturing, storing and extracting of the data. P2P torrent application and Kaspersky antivirus software were installed in each of the desktop computer. Most of the torrent application is in the fair

size and it is freely available online. Kaspersky antivirus somehow need to be uninstalled prior to the anomalous traffic capturing process begin. The following P2P torrent applications are installed; BitTorrent, uTorrent, BitComet, Deludge and Lime Wire. During the process of capturing network traffic data for the anomalous traffic environment several botnet variants were deployed in the host PC. The variants are listed in Table 3. Wireshark and Cascade Pilot network analyzing tools were used for data analysis.

Table 3. List of the Botnet Variant and Malware Name

List of the Botnet Variant (MD5)	Malware name and alias
4a270b9e3b708a55639a531de71c7af4	Net-Worm.Win32.Kido.ih
	Win32/ConfickW32/Conficker.C.worm
	Win32:Rootkit-gen
	Trojan.Win32.Generic.521F1AAC
3509cafa6887aa1cde4c880cb7c595ea	Net-Worm.Win32.Kido.ih
	Win32/Conficker.X
	W32/Conficker.C.worm
	Win32:Rootkit-gen
	Hack.Exploit.Win32.MS08-067.fd
7021c2547f82120e069db074552042b7	Net-Worm.Win32.Kido.ih
	avariantofWin32/Conficker.X
	W32/Conficker.C.worm
	Win32:Rootkit-gen
	Trojan.Win32.Generic.5208D020
865915650a85e7c27cdd11850a13f86e	Win32/IRCBot.worm.variant
	Worm/Rbot.246784.1
	Backdoor/Win32.Rbot.gen
	Win32:Rbot-GKN
	Win32:Rbot-GKN
	IRC/BackDoor.SdBot2.HHB
	Worm.Generic.278963
	Backdoor.Rbot.bqj
	Trojan.Mybot-5073
	W32/Sdbot.OTR
	Backdoor.Win32.Rbot
	BackDoor.IRC.Sdbot.3172
	Backdoor.Rbot!IK
	Win32.Rbot.gen
	Win32/Rbot.GQP
kido	Worm/Kido.BO
	W32/Conficker.worm.gen.a

4. RESULTS AND DISCUSSIONS

The network behaviour analysis tools are used to analyze the network traffic. In this study, the well-known open source traffic analyzer is recommended such as Wireshark or Ethereal. The visualization analyzer tools like Cascade Pilot is considerable to achieve the objective of the study. The analysis of the botnet characteristic is relying on the collected data as mention in previous section.

Figure 3 shows the network analysis process. The anomaly signature is defined in the early step followed by the traffic characteristic then the comparison between normal and anomalous traffic data will produce the result of the botnet characteristic. As mentioned previously in literature, there are several techniques and method to identify the network. In this study, passive analysis is used to identify the botnet behaviour will be perform into two approaches Signature base detection and DNS anomaly base detection. Passive analysis mean the analysis of the network traffic is non-real time and the analysis take part after the data is collected.

4.1. Anomalous DNS Packet

The initial phase of analysis, the comparison between normal and anomalous traffic is based on the total number of packet carry out by the DNS protocol for all types of DNS Record. After filtering the DNS request for both normal and anomalous network traffic data, the DNS record type A and MX are existed. One of the aggregate features from the analysis of the DNS in network traffic data is its frequency, suspicious behaviour, such as DNS cache poisoning attack and DNS amplification attack, frequently generates many DNS queries in a short time. Subsequently, when the DNS queries are higher the total number of packet carry out by the DNS in anomalous traffic is also higher [21].

Table 4 and 5 show the traffic flow for both normal and anomalous data traffic. The total number of packets for DNS is shows a huge different amount between normal and anomalous traffic. The average

number of the total packets for DNS request in normal traffic data is 105.67. In contrast, the average number of the total packets for DNS request in anomalous traffic data is 2121. In anomalous DNS request the transmission rate of the packet is 0.68 packets per second, which is higher compared to the average of 0.05 packets per second in normal traffic data. It is clearly indicated that many DNS queries is generated in anomalous traffic data within a short time. Thus the irregularity of the frequency creates suspicious behaviour. The initial observation glued the comparison between the normal and anomalous traffic present the differences, within the approximately same amount of time and size of the pcap file between both data there is some anomalous behaviour in the traffic.

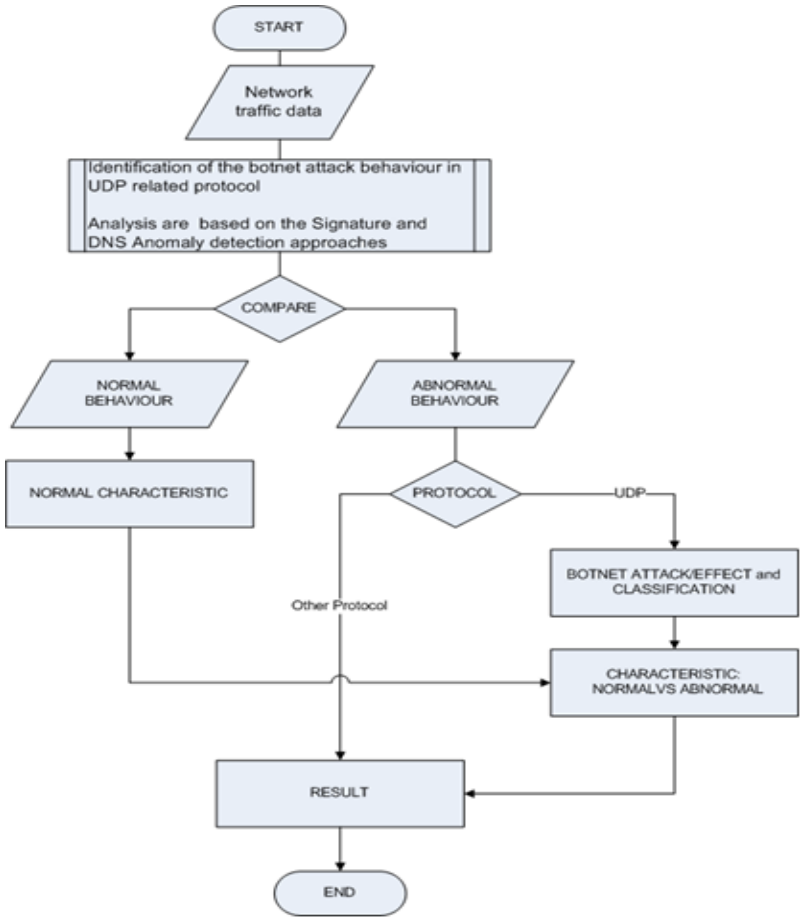

Figure 3. Network Traffic Analysis Process

Table 4. Normal UDP by Port 53 and DNS Protocol

Data Set	Protocol	Port	Packets/s	Total Packets
Normal_1	DNS	53	0.05	126
Normal_2	DNS	53	0.04	85
Normal_3	DNS	53	0.05	106
Average	DNS	53	0.05	105.67

Table 5. Anomalous UDP by Port 53 and DNS Protocol

Data Set	Protocol	Port	Packets/s	Total Packets
Anomalous_1	DNS	53	1.28	3559
Anomalous_2	DNS	53	0.4	1612
Anomalous_3	DNS	53	0.35	1192
Average	DNS	53	0.68	2121

4.2. NetBIOS Attack

Conficker is one of the botnet that already harm the internet user in past few years. The variant still exist and the algorithm of the botnet will be used by the future attacker to create similar attack. Conficker characteristic can be defined by several signatures one of them is through the NetBIOS attack [22]. Conficker works by exploiting the NetBIOS vulnerability in Windows operating system and uses many new techniques, including a domain generation algorithm, self - defense mechanisms, web and P2P updates and efficient local propagation and brute force attack of NetBIOS admin shares [23]. All Conficker versions used pseudo

random domain generation with known set of Top Level Domains (TLD) to obtain updates. This mechanism is called domain flux and the URL which is contacted is called the HTTP rendezvous point [23] and the domain name is generating by Conficker through its pseudo random domain generation. Table 6 shows the suspicious domain generated by the Conficker after filtering the NBNS protocol in Wireshark.

Conficker use NBNS (NetBIOS Name Service or netbios-ns) protocol to propagate itself into network [24]. Filtering NBNS protocol in Wireshark will show the information carry out by the NBNS protocol. NBNS will read the hostname which is tried to attach by Conficker botnet. The hostname will indicate which hostname or computer attach by Conficker. The source IP is the infected computer and attempt to distribute and update itself. Figure 4 shows the Conficker attack by filtering the NBNS in Wireshark, meanwhile in Figure 5 only a few NBNS name query and none of them queries the suspicious domain name.

Table 6. Suspicious Domain

Protocol	Port	Suspicious Domain
netbios-ns	137	ANTVJRFYTIC.CC
netbios-ns	137	DAJSRMDWHV.COM
netbios-ns	137	ELBMESUSAJ.NET
netbios-ns	137	HFGSGRPFQIID.CC
netbios-ns	137	MQGAWG.YI.ORG
netbios-ns	137	MZMJWPY.YI.ORG
netbios-ns	137	TBMRPQ.YI.ORG
netbios-ns	137	TMLHKYZHLZEC.CC
netbios-ns	137	UEAQSAPMFBFF.CC
netbios-ns	137	VMITRPYHOZI.COM
netbios-ns	137	YFZIVIUX.COM
netbios-ns	137	ZVQGMWIAV.COM

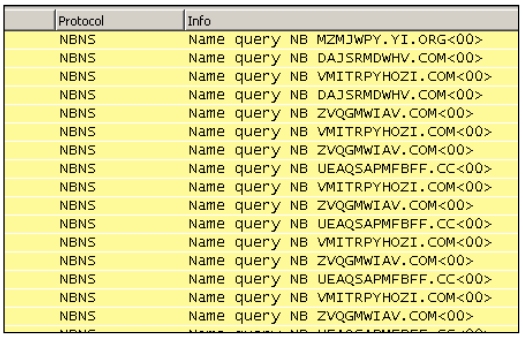

Figure 4. NBNS filtering in anomalous behaviour

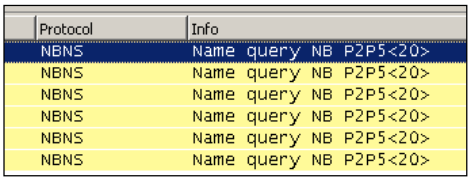

Figure 5. NBNS filtering in normal behaviour

The following result from the filtering NetBIOS show the differences of the NBNS activities between normal and anomalous data. Table 7 and Table 8 indicate the total packets for normal data is less than anomalous data. The analysis of the normal traffic data after filtering the NetBIOS protocol produce the average of the total packets is 19.67. However, in anomalous traffic data the average is 6209.67, this is the query create by the host that was infected with the Conficker botnet. The infected hosts run the Conficker pseudo random domain generation, to create the suspicious domain name.

Table 7. UDP Traffic by Port 137 for Normal Traffic

Data Set	Protocol	Port	Total Packets
Normal_1	netbios-ns	137	31
Normal_2	netbios-ns	137	17
Normal_3	netbios-ns	137	11
Average	netbios-ns	137	19.67

Table 8. UDP Traffic by Port 137 for Anomalous Traffic

Data Set	Protocol	Port	Total Packets
Anomalous_1	netbios-ns	137	6178
Anomalous_2	netbios-ns	137	6728
Anomalous_3	netbios-ns	137	5723
Average	netbios-ns	137	6209.67

The analysis of the network traffic by comparing the total packets between normal and anomalous traffic determine there is one characteristic of Conficker botnet exist in the network. The characteristic find in the network analysis is the NetBIOS attack. From the signature based detection, Network administrator or user can use this characteristic to determine the existence of the Conficker botnet in the network.

4.3. Anomalous DNS MX Query

In legitimate emails delivery and based on Simple Mail Transfer Protocol (SMTP), uninterrupted delivery is completed by the Mail Transfer Agent (MTA), not by the end users or their Mail User Agent (MUA). A client or MUA usually connects to a server or MTA for sending an email. MTA organizes delivery by transferring it to another MTA. In many cases the second MTA is actually the Mail Exchange (MX) server of the recipient. This practice can be illustrated in Figure 6.

Spam generated by Botnet or known as spambot does not attempt to travel in this legitimate way; undisclosed methods are used to anonymize the spam source. MTAs will query the DNS MX domain registry of the recipient and then send the message directly to the server. Spambot must obtain these servers ' IP addresses for each email address in the list. It can query the MX DNS of all the domains of the recipients, but this slows the spamming process. A more advanced way is to download MX records from the botnet servers or peers along with email lists [23]. This will need an extra efforts for maintaining and updating these lists by botnet controllers. Simple solution suggested by [5] is to reconfigure default mail server port to another port number. Botnet generated spam can be delivered directly to the Mail MX server of the recipient's domain as depicted in Figure 7.

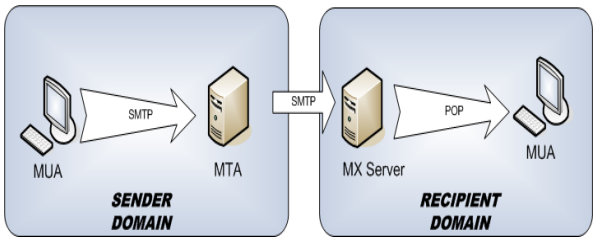

Figure 6. Legitimate email delivery

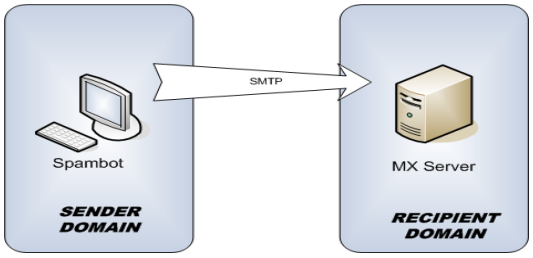

Figure 7. Spambot can be delivered directly to the recipient's domain MX server

From the analysis of the network traffic, the anomalous behaviour of DNS requesting the MX standard query shows irregularity. There was a lot of DNS MX record type from one of the node in the network. The data can be obtained using Cascade Pilot by drill down the DNS record type. From the filtering of the DNS query, only A (address record) and MX (mail exchange) record type is present, other DNS record like AAAA (IPv6 address record), PTR (pointer), CNAME (Canonical name) or others are not detected. The analysis also shows the anomalous DNS A record type. Table 9 and Table 10 shows the anomaly between normal and anomalous traffic referring to the amount of request for DNS A and DNS MX record type. There is no DNS MX record type for normal traffic.

Table 9. Number of DNS Queries in Normal Traffic

	Number of DNS Queries	
DNS Record Type	A	MX
Normal_1	74	0
Normal_2	48	0
Normal_3	57	0
Average	59.67	0

Table 10. Number of DNS Queries in Anomalous Traffic

	Number of DNS Queries	
DNS Record Type	A	MX
Anomalous_1	927	463
Anomalous_2	4170	296
Anomalous_3	1040	44
Average	2045.67	267.67

The number of request for DNS query relatively smaller compared to the anomalous traffic. In normal traffic data the total number of queries is 59.67 in average for DNS A record type. In fact there is no DNS MX record type in normal traffic. In comparison, for the anomalous network traffic the total of the DNS queries is 2313.34 in average. The DNS MX record type is 267.67 queries. This kind of behaviour is the process of the spambot try to query the DNS for MX record of all the recipients' domains.

The behaviour of the anomalous DNS MX query above is the process of spambot to obtain the IP address for the MX servers for each email address on it list. The botnet generated spam can query the DNS for MX record of all the recipients' domains. There are a lot of MX queries but it only contributes small percentage of the total traffic but the number of MX queries is higher than regular basis. By observing this behaviour, these queries are suspicious.

4.4. UDP Flood and DNS Amplification Attack

Botnets are regularly used to mount DDoS attacks in the Internet. Systematic review on aspects of DDoS detection is conducted and it was found that each technique used for the detection of attacks exploits certain characteristics of the network traffic, user requests and explicit tools [24, 25] proposed a technique for the forensics of UDP flood attack that capable to identify the source of UDP flooding bot attack. The UDP flood attack sends an enormous number of UDP packets to any randomly selected ports of a server to block other authentic traffic to the network port. The random ports could be echoed, daytime, or Character Generator Protocol (CHARGEN). The earliest utilisation of a botnet is to launch a DDoS attack that can causes a loss of service to users. A DDoS attack is an attack on a computer system or network that causes users to lose service, usually the loss of network connectivity and services by using the victim network's bandwidth or overloading the victim system's computer resources.

Most implementations of the DDoS attacks are categorised as TCP SYN and UDP flood attacks. In TCP, the most common technique is called TCP SYN flooding and involves creating a large number of half-open TCP connections on the target machine to exhaust kernel data structures and prevent the machine from accepting new connections. UDP's flooding technique is to overrun the target machine with a large number of UDP packets. Then the bandwidth of the network gradually exhausts and ends with the termination of service.

Botnets use several thousand machines repeatedly and allow an attacker to seriously damage the computer network. Botnets are commonly used for DDoS attacks because their bandwidth combined exceeds the bandwidth available on most target servers or systems. In addition, several thousand compromised machines can generate so many packets that the target unable to fulfill that many requests. Since a botnet often contains thousands of bots, the Botmaster can give instructions to a particular server or system for all online bots flooding packets. [26]. These packets will consume the bandwidth of the victim's network, overload the victim's computer resources or even contain the general Internet traffic to cause massive public damage. Several bots capable to send a report back to their controller and tell the botmaster how many size they flooded the victim. From the scenario above, Botnets are a serious threat in DDoS attacks capacity. DDoS attacks try to crush the target by sending an overwhelming amount of traffic from host computers, usually botnet members, which are distributed across different locations, making it difficult or impossible to simply block traffic by source address. There are several types of traffic: TCP SYN, UDP and ICMP packets floods. Attacks on services on a higher network layer are also common, in particular DNS, HTTP and SMTP. Botnet DDoS attack a large number of hosts will try to push a DNS server of the Internet. The characteristic of an attacking Botnet DDoS attack is by increasing amount of requests. During a request the Botnet will be sending a massive amount of requests to the DNS server to overwhelm the available resources.

From the captured data there were some activities related to anomalous DNS request. The comparison between the normal and anomalous traffic base on its DNS request, Table 11 and Table 12 shows the different of the total packets. The analysis of the data is based on the selected data set as mention earlier in implementation chapter. The average length of the network traffic is about eighty minutes for the normal traffic and the average of the length for the anomalous traffic is approximately eighty six minutes. The total number of packet in DNS protocol for anomalous network traffic is 2121 and the normal network traffic is 105.67 in average. Thus, it is indicate the different amount on which the anomalous network traffic carry the large number of total packets then the normal network traffic.

Table 11. Total Packets for DNS Request in Normal Traffic

Data Set	Length in second	Tcpdump size in kilobyte	Protocol	Port	Total Packets
Normal_1	5050	976564	DNS	53	126
Normal_2	4689	976563	DNS	53	85
Normal_3	4731	976562	DNS	53	106
Average	4823.3	976563	DNS	53	105.67

Table 12. Total Packets for DNS Request in Anomalous Traffic

Data Set	Length in second	Tcpdump size in kilobyte	Protocol	Port	Total Packets
Anomalous_1	5700	976562	DNS	53	3559
Anomalous_2	4980	781251	DNS	53	1612
Anomalous_3	4800	781250	DNS	53	1192
Average	5160.00	846354.33	DNS	53	2121

The massive amount of packets is occurred in anomalous traffic (846354.33 kb). This is one of the indication of the DNS amplification attack. The attacker generates a large number of packets with flood attacks aimed at the target sources. Flood attacks for IP components including UDP SYN floods and ICMP echo packets are typical type of attack initiated by the Botmaster. The flood attack were used multiple distributed sources, such as botnets, a variant of DNS amplification is a favourite among the operators of commercial DDoS botnet operators. Both characteristic bases on the UDP flooding attack and DNS amplification attack can be determined as one of the possible communication between bormaster and bot to leverage the attack. Botmaster can launch the attack or recruiting the new candidate for the new bot by requesting the IP of the machine.

The DNS anomaly detection approaches produce the promising result in research finding. Anomalous DNS packet request, the NetBIOS attack, DNS MX query on Spambot, DNS amplification attack and UDP flood attack is identify in the analysis. Based on the result and analysis in this study, it will help network administrator or user alert with the behaviour and characteristic might be exist in their network. The characteristic of the botnet can be used for future element in identifying and combating the botnet.

5. CONCLUSION

Botnets are evolving rapidly; all elements of communication protocols initiation mechanisms, attacks on the rallying mechanism constantly evolve and difficult to detect. The current works in botnet detection strategies is discussed in the literature review which is to study the topology, classification, behaviour of the botnet and to identify the related works in characterizing and detection of botnet base on the passive method through analyzing traffic flow and anomaly signature. The objective of this study is to characterize the botnet in P2P environment based on the UDP protocol. Torrent application is run throughout the capturing process to create the live P2P network environment. UDP carry out most of the data transfer in network, it offers a marginal transport services, a nonguaranteed datagram delivery and provides applications direct access to the datagram service of the Internet Protocol (IP) layer. UDP is used for applications not requiring a TCP standard of service. Most of the botnet use TCP as a protocol to do the attack. However, UDP is still available for attackers to manipulate data carried by UDP to do the attack.

DNS protocol is selected in this study because it uses UDP protocol to carry out its messages. In this study, the detection of the botnet is centered on signature and DNS anomaly approaches. In signature approach, the studied shows one of the well-known botnet name Conficker is present in the network based on the NetBIOS attack. DNS anomaly approach is use to analyze the behaviour of the DNS. Based on the analysis in this study there were a lots of DNS abnormality exist in the anomalous network traffic. Base on the DNS anomaly strategy, the analysis of the network defines the characteristic of the botnet appeared in the network. The huge amount of DNS packet in anomalous network traffic can be used as an indicator for the existence of the botnet. The analysis of the DNS protocol by comparing the normal and anomalous traffic produce the identification of the DNS amplification attack, UDP flood attack and spambot activities in the network. The spambot characteristic is defined when there is an anomalous DNS MX query in the network. The analysis of the botnet from the captured data can be used for the future works to identify the existence of the botnet in particular network. Therefore, this study meets the objective to identify the botnet characteristic and behaviour from the analysis of the good and bad network traffic through network behaviour analysis tools and to analyze, examine and identify botnet characteristic in attacking victims.

ACKNOWLEDGEMENT

This work is sponsored by Universiti Tun Hussein Onn Malaysia under TIER1 vot: U897.

REFERENCES

[1] S. Chang, L. Zhang, Y. Guan, and T. E. Daniels, *"A Framework for P2P botnets,"* Proc. - 2009 WRI Int. Conf. Commun. Mob. Comput. C. 2009, vol. 3, pp. 594-599, 2009.

[2] M. Bailey, E. Cooke, F. Jahanian, Y. Xu, and M. Karir, *"Al survey of botnet technology and defenses,"* Proc. Cybersecurity Appl. Technol. Conf. Homel. Secur. CATCH 2009, pp. 299-304, 2009.

[3] C. Xiang, F. Binxing, Y. Lihua, L. Xiaoyi, and Z. Tianning, *"Andbot: towards advanced mobile botnets*," 4th USENIX Conf. Large-scale Exploit. emergent Threat., p. 11, 2011.
[4] A. A. Obeidat, "Analysis the P2P botnet detection methods," vol. 4, no. 3, pp. 1-11, 2016.
[5] A. V. Arzhakov and D. S. Silnov, "Analysis of Brute Force Attacks with Ylmf-pc Signature*," Int. J. Electr. Comput. Eng.*, vol. 6, no. 4, pp. 1681-1684, 2016.
[6] Y. Sharma, "Botnet Detection by Network Behavior Analysis," *Glob. Res. Dev. J. Eng.*, vol. 2, no. 11, pp. 34-40, 2017.
[7] M. Stevanovic and J. M. Pedersen, *"An analysis of network traffic classification for botnet detection,"* 2015 Int. Conf. Cyber Situational Awareness, Data Anal. Assess., pp. 1-8, 2015.
[8] M. N. Sakib and C. T. Huang, *"Using anomaly detection based techniques to detect HTTP-based botnet C&C traffic*," 2016 IEEE Int. Conf. Commun. ICC 2016, 2016.
[9] K. Alieyan, A. ALmomani, A. Manasrah, and M. M. Kadhum, "A survey of botnet detection based on DNS," *Neural Comput. Appl.*, vol. 28, no. 7, pp. 1541-1558, 2017.
[10] R. L. Joffe, "Method and system for detecting network compromise." Google Patents, 2016.
[11] G. Kirubavathi and R. Anitha, "Botnet detection via mining of traffic flow characteristics," *Comput. Electr. Eng.*, vol. 50, pp. 91-101, 2016.
[12] H. Lashkari, G. D. Gil, J. E. Keenan, K. F. Mbah, and A. A. Ghorbani, *"A Survey Leading to a New Evaluation Framework for Network-based Botnet Detection,"* Proc. 2017 7th Int. Conf. Commun. Netw. Secur. - ICCNS 2017, pp. 59-66, 2017.
[13] J. Wang and I. C. Paschalidis, "Botnet Detection Based on Anomaly and Community Detection," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 2, pp. 392-404, 2017.
[14] S. A. Raihana, F. A. Mohd, A. M. N. Zul, Z. M. Mohd, R. S. Siti, and Y. Robiah, "Revealing the Criterion on Botnet Detection Technique," *IJCSI Int. J. Comput. Sci. Issues*, vol. 10, no. 2, pp. 208-215, 2013.
[15] C. Yin, R. Sun, L. Yang, and D. Iko, "Node-based Sampling P2P Bot Detection," *TELKOMNIKA Indones. J. Electr. Eng.*, vol. 10, no. 5, pp. 1117-1122, 2013.
[16] G. Khehra, "*BotScoop : Scalable detection of DGA based botnets using DNS traffic*," 2018 9th Int. Conf. Comput. Commun. Netw. Technol., pp. 1-6, 2018.
[17] L. Mathur, M. Raheja, and P. Ahlawat, *"Botnet Detection via mining of network traffic flow,"* Procedia Comput. Sci., vol. 132, pp. 1668-1677, 2018.
[18] G. Gu, R. Perdisci, J. Zhang, and W. Lee, *"BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection,"* in 15th ACM Conference on Computer and Communications Security, 2008, pp. 139-154.
[19] W. Ye and K. Cho, "P2P and P2P botnet traffic classification in two stages," *Soft Comput.*, vol. 21, no. 5, pp. 1315-1326, 2017.
[20] M. Al-kasassbeh, G. Al-naymat, and E. Al-hawari, "Towards Generating Realistic SNMP-MIB Dataset for Network Anomaly Detection," no. December, 2016.
[21] Aizuddin, M. Atan, M. Norulazmi, M. M. Noor, S. Akimi, and Z. Abidin, *"DNS amplification attack detection and mitigation via sFlow with security-centric SDN,"* in Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication, 2017, p. 3.
[22] G. Gross, "Detecting and destroying botnets," Netw. Secur., vol. 2016, no. 3, pp. 7-10, 2016.
[23] R. Vinayakumar, K. P. Soman, P. Poornachandran, and S. Sachin Kumar, "Evaluating deep learning approaches to characterize and classify the DGAs at scale," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3, pp. 1265-1276, 2018.
[24] S. Bravo and D. Mauricio, "Systematic review of aspects of DDoS attacks detection," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 1, p. 155, 2019.
[25] A. Bijalwan, M. Wazid, E. S. Pilli, and R. C. Joshi, "Forensics of Random-UDP Flooding Attacks," *J. Networks*, vol. 10, no. 5, 2015.
[26] S. Behal and K. Kumar, "Characterization and comparison of DDoS attack tools and traffic generators - a review," *Int. J. Netw. Secur.*, vol. 19, no. 3, pp. 383-393, 2017.

BIOGRAPHIES OF AUTHORS

Noor Zuraidin Mohd Safar is a lecturer at Department of Information Security and Web Technology, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. He received BSc in Computer Science from University of Tulsa Oklahoma, MSc in Internetworking Technology from Universiti Teknikal Malaysia and Ph.D. from University of Portsmouth, United Kingdom. His research focuses on the area of machine learning, soft computing in environmental, meteorological data in tropics, computer network security and web technology.

Noryusliza Abdullah is a faculty member at Faculty of Computer Science and Information Technology at Universiti Tun Hussein Onn Malaysia. She received her B.Sc. (Computer Science) with Honours from Universiti Sains Malaysia in 2002 and M.Sc. (Information Technology) from University Teknologi Malaysia in 2006. She received the Ph.D degree in Information Technology from Universiti Tun Hussein Onn Malaysia in 2015. She had an industrial experience for 14 years as a System Analyst and Senior System Analyst. She is currently a Head of Department in Information Security and Web Technology. Her research interests include Decision Support System, Information System, Information Retrieval and Internet of Things.

Hazalila Kamaludin is a faculty member at Faculty of Computer Science and Information Technology at Universiti Tun Hussein Onn Malaysia. She received her B.Sc. (Computer) with Honours from Universiti Teknologi Malaysia in 2002 and M.Sc. (Information Technology) from University of Teknologi Mara in 2005. She received the Ph.D degree in Information Technology from Universiti Tun Hussein Onn Malaysia in 2018. Her research interests include RFID data management, data trustworthiness and Internet of Things.

Suhaimi Abd Ishak is a lecturer with Universiti Tun Hussein Onn, Malaysia. His research interests include energy-aware embedded systems, pervasive computing, and IoT energy harvesting. He received the Ph.D. degree in computer science and engineering from the University of New South Wales, Australia.

Mohd Rizal Mohd Isa received the BSc in Data Communication and Networking and MSc in Information Technology degrees respectively, from the Universiti Teknologi MARA (UiTM), Malaysia. In 2017, he received Ph.D degree at the University of Portsmouth, United Kingdom in Computer Engineering. He is now with Universiti Pertahanan Nasional Malaysia as a Director of Information Technology Centre, lecturer and researcher. His research interests include biometric systems, network security and information hiding.