

*(Figure 0) Peer to peer botnet layout – pandasecurity*

Botnet Detection and Perimeter Defence

Niall A - 40417661

Coursework Submission for CSN08703

March 2019

# Section 1 - Research

## Botnets

It is known within the cyber security industry that in recent years, botnets have emerged as one, if not the, major threat to information security within businesses. This is primarily due to the fact that botnets, and subsequently the coders behind them, have been evolving in both size, sophistication, and knowledge of evasive techniques. *(Aussems, Noë, and Rivera, 2014, pp. 1-8)*

The purpose of botnets can be seen below:

*“Special Trojan viruses to breach the security of several users’ computers, take control of each computer and organise all of the infected machines into a network of ‘bots’ that the criminal can remotely manage.” (Kaspersky, 2019)*

A botnet works by having multiple bots (zombies) listening to a C&C channel, and upon receiving an instruction, will carry out a specific task. These tasks can vary from positive impacts to negative impacts on end uses and businesses. (The Honeynet Project, 2008)

The strengths of botnets consist of, but are not limited to, their wide range of capabilities; they can be used for phishing via spam email, as well as theft of end user confidential information, enabled by keystroke logging, as well as DDoS capabilities, provided that the botnet is large scale. In addition to this strength, botnets can adapt; one stage in the bot lifecycle (can be referred to as *egg downloading (Vuong and Alam, 2011, p.56)*)allows the bot to download new executables, for example a payload to deactivate the antivirus capabilities of a system. Thus, making botnets an extremely robust and sophisticated malware.

The weaknesses of a botnet consist of, but again are not limited to, how well it is implemented, and the topology used to create connections to the bots. The implementation, while in some cases being successful, can be a weakness; IDS’ are only as good as the botnets are bad, meaning that if a botnet is put together without planning or carefulness, it can become easy to spot and block, allowing it to be taken down. In addition to this, the topology used while setting up and, in turn, running the botnet can also be a weakness; more centralised structures, having a central C&C device connected to the array of bots, creates a single point of failure for the botnet – if the C&C device goes down, or is detected, it becomes easy to disrupt, and as a result destroy the botnet. *(Aussems, Noë, and Rivera, 2014, pp. 1-8)*

Botnets as a whole could become more sophisticated as time goes on; the introduction of more IoT devices offers a wider range of potential bot devices, as more and more brands are becoming commonplace, and it can be assumed that security is not as big of an issue as other features, such as usability and consumer cost in order to gain more market share; more security focus during development is likely to increase the cost, and subsequently drives down sales as the market for IoT is flexible. Due to this, it could be suggested that botnets can become more sophisticated to evade IDS’s through hiding activity via IoT protocols and listening to multiple C&C channels, such as RSS feeds, IRC Servers and HTTP pages whilst incorporating a distributed topology to avoid being destroyed from a single point of failure.

## IDS – Intrusion Detection Systems

An intrusion can be described as a potential possibility of a deliberate unauthorised attempt made in order to access/manipulate information, as well as render a user system unusable or unreliable. Subsequently, an IDS (intrusion detection system) can be described as a piece of hardware / software in place to detect abnormal activity on a device / network (HIDS / NIDS) which is created by an attempted intrusion. The purpose of an IDS is to point out when an intrusion occurs, and can categorise into 6 different types of intrusion:

|  |  |
| --- | --- |
| Attempted break ins | Denial of service |
| Leakage | Malicious use |
| Masquerade attacks | Penetration of the security control system |

An advantage of IDS systems is that through anomaly-based detection, the IDS can detect unseen attacks as its aim is to detect unusual (compared to normal) activity. This has a disadvantage however in that it will subsequently result in a higher volume of false positives as not all traffic will be malicious but it can still be classed as unusual. However, by using both signature detection and anomaly-based detection, the IDS can become very powerful. *(Chawla, Lee, Fallon and Jacob, 2018)*

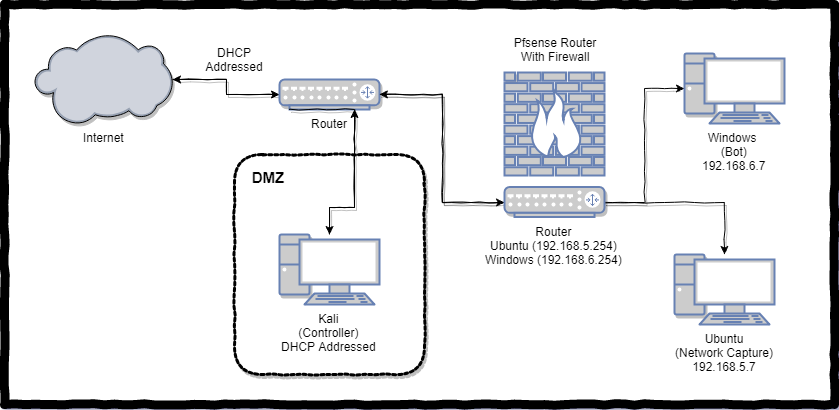
# Section 2 – Botnet Analysis

### Dynamic Analysis

Dynamic analysis, via packet sniffing, enables the botnets behaviour to be captured and interpreted, and, as is shown below, enables the viewing of the plaintext commands sent over the network to and from the bot, followed by the responses of the controller. Tools such as Wireshark can be used to capture live information flowing over the network and allow the user to follow TCP streams created. In order to dynamically analyse the botnet, a network layout and topology is required; each machine in use for this task has been configured with an IP address (as per the specification), a default gateway, and a DNS server which it can access. All of the default logins for VMs have been changed to secure the isolated network of the botnet. It should be noted that some of the testing is not taken on the vSoc environment provided, but from a virtual machine network configured using Oracle VM VirtuaBox.

All VMs have connectivity with all of the other VMs used, tested with sending ICMP packets from each machine to every other machine. The firewall, upon later testing, will be configured in accordance to the basic rule-set provided again by the specification. As a result, not all traffic will be allowed throughout the network. This is later evidenced by screenshots of the pfsense router and its rules.

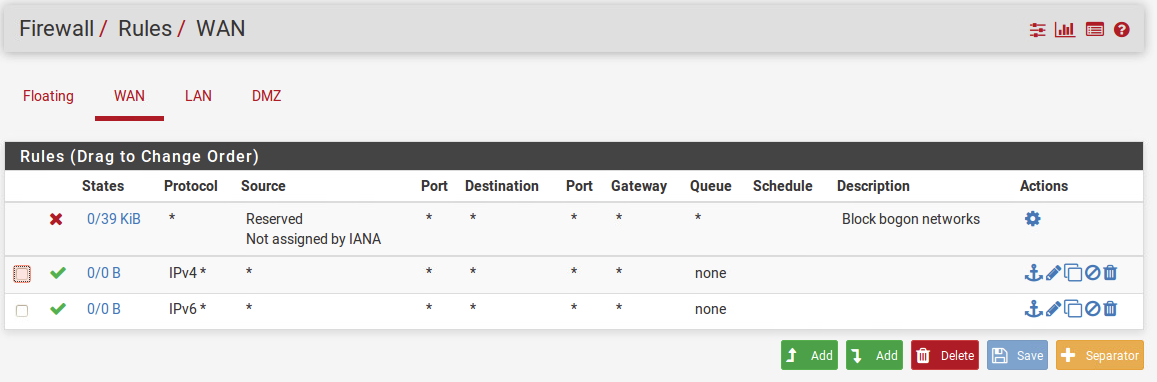
### IP Addressing

*(figure 1) This basic diagram outlines the addressing scheme*

As can be seen, the Windows and Ubuntu machines (representing the Public and Private areas of the network) are behind the pfsense router and firewall, and have IP addresses conforming to the specification. The DMZ area and machine will have a DHCP address assigned to it from the vsoc external router, which also provides internet access to the rest of the network (unless otherwise prevented).

### Basic Network Configuration and Connection Testing

Unlike with the basic firewall rules, this isolated testing network is used to find botnet traffic, and is not configured to prevent certain traffic from flowing through the network. With this basic setup, all network connections throughout the network topology are allowed. As a result, the botnet actions and behavior can be seen below, the supporting figures show the firewall rules used in pfsense to enable this.



### 

*(Figure 2, 3, 4) Pfsense firewall rules enabling network traffic through all interfaces*

**

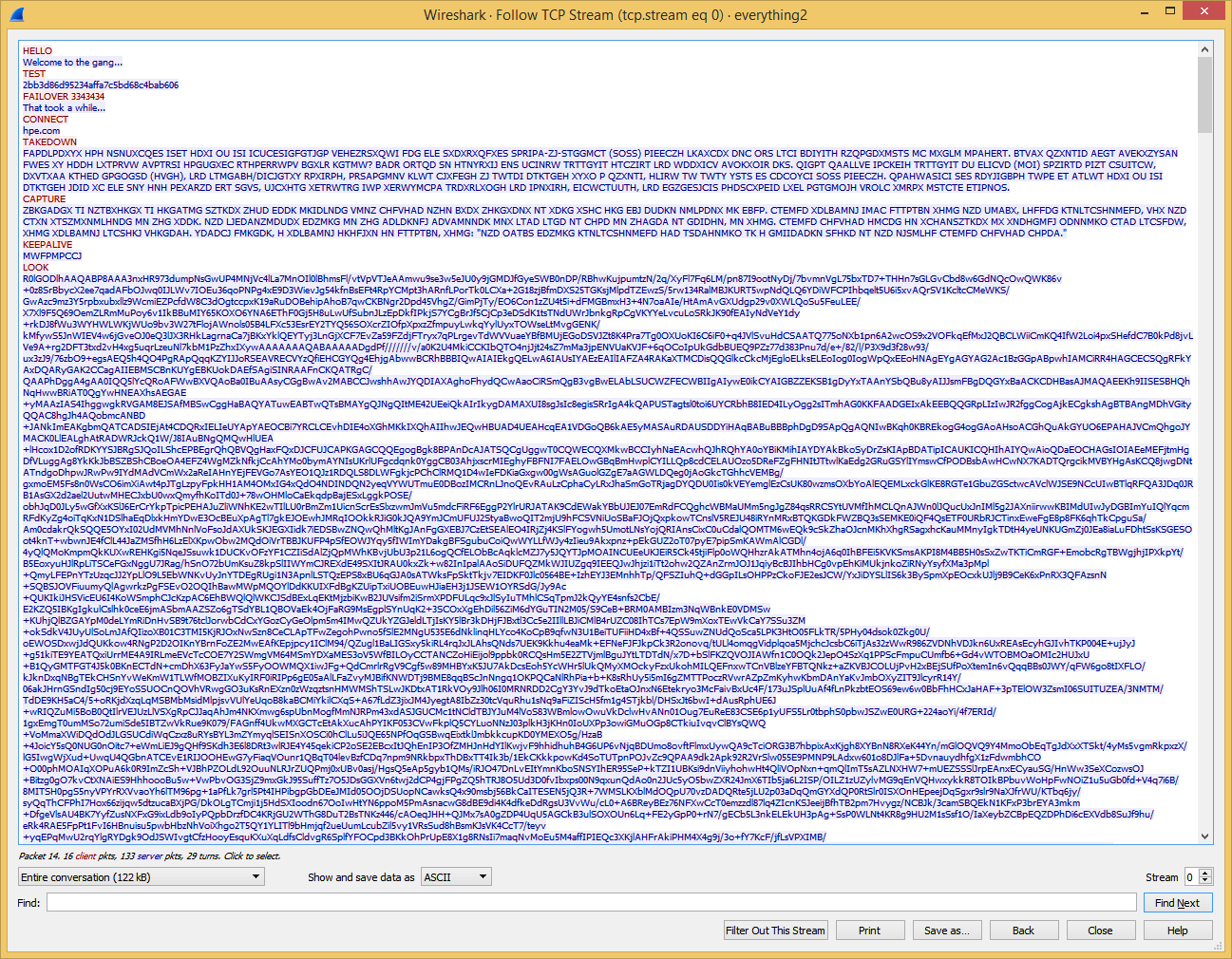
*(figure 5,6) Windows connectivity testing to Ubuntu, Kali, Internet and router interfaces)*

*See Appendix X for other connectivity checking*

### Traffic Analysis

One of the initial tasks when analyzing a botnet and its associated traffic, is identifying the botnet in questions network connections and, by proxy, traffic generated. This can be done through eavesdropping / monitoring network activity, and filtering out unrelated traffic using an appropriate tool, such as Wireshark. The analysed network traffic and an interpretation of what it does can be found below.

In order to observe the botnet operations, the bot will be run on the Windows VM, found on the private network, connecting to the controller found on the Kali VM, found on the DMZ (public) network. Below will highlight some of the botnet traffic found in Wireshark, followed by sending custom commands through the use of telnet via the Windows machine.

**

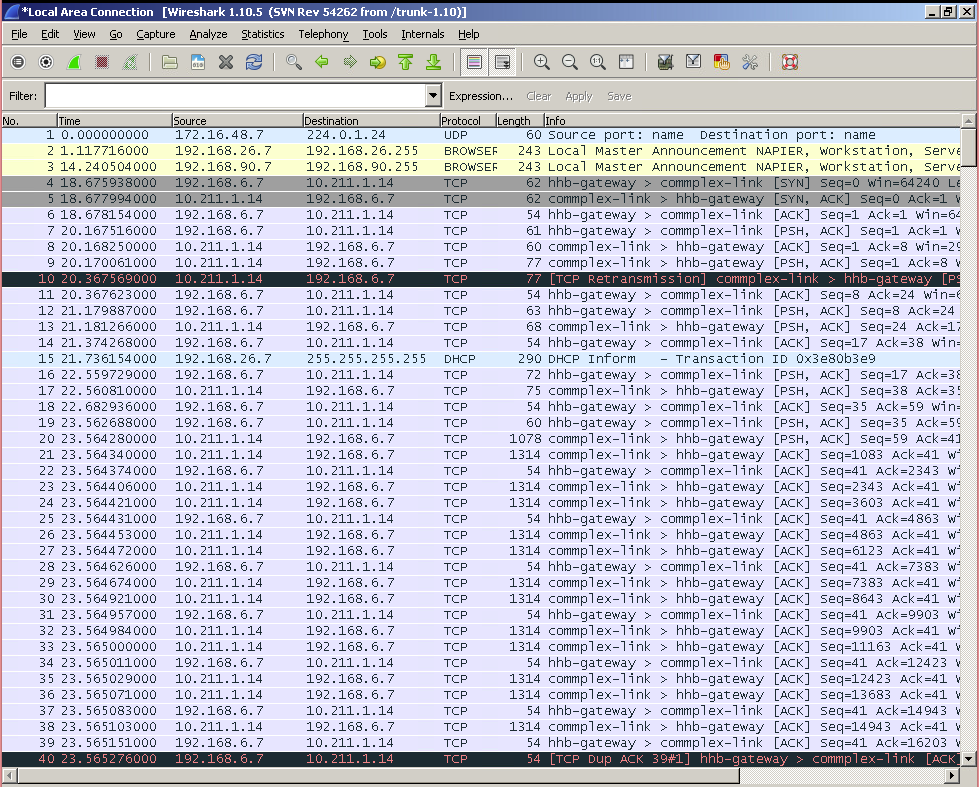
As can be seen, the Wireshark capture contains the bot’s communication to the controller, followed by the controller’s response. The bot will evidently run random commands, presumably from a list, and will always stop after the ‘Goodbye’ command receives a reply of ‘sleep...’

After inspection of multiple instances of the bot running, it can be concluded that some of the commands will yield multiple responses. This needs to be further analysed to discover what is happening.

Additionally, it appears that the bot will send ping requests to every a.a.a.(x) address in range 0-255 for x, and follow the SNOOP

*(figure 7) pcap file created when network traffic was being monitored in Wireshark.*

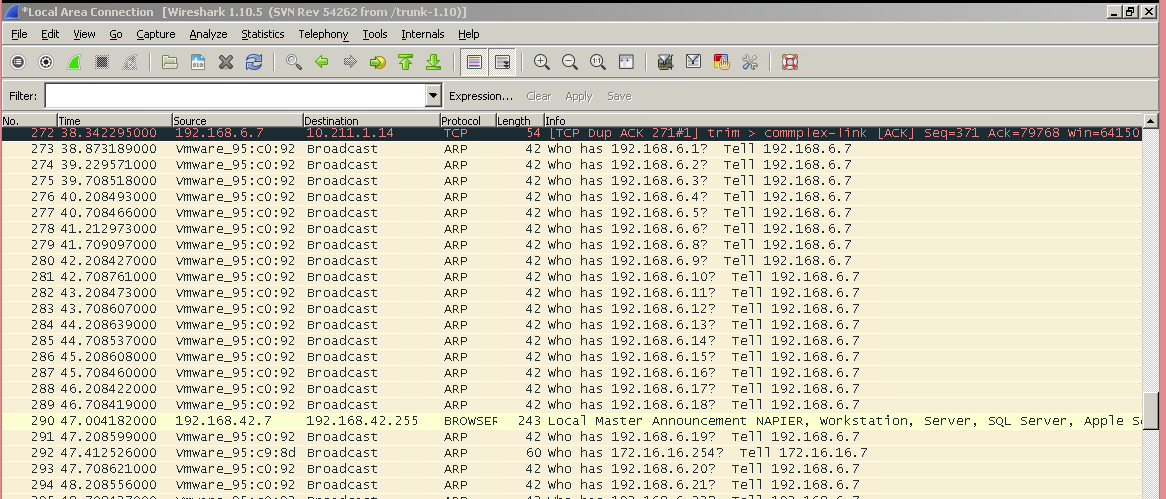
command with a list of IP addresses found on the network. (figure 9)

*(figure 8) .pcap showing src and dest addresses*

This capture (figure 8) shows the connectivity and activity between the two devices situated at opposite ends of the network. As the specification states, the windows machine (bot) is at 192.168.6.7, and the Kali machine (controller) is in the DMZ (10.211.1.14).

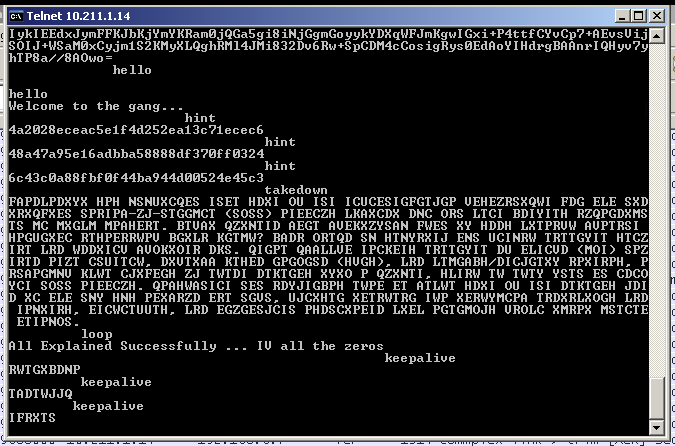
It is found that content is sent through [PSH, ACK] packets, and the introduction is like any handshake ([SYN], [SYN, ACK]), and that the connection ends based on ‘sleep...’ being received, followed by a [FIN, ACK] packet.

Some commands will trigger the bot to perform network activities, such as pinging every address in the network, with the range 0-255 for the last octet of the address. This is picked up as many ARP requests on some machines, and many ICMP requests on others. It can be believed that the SNOOP command will cause this.



*(figure 9) pcap showing ARP requests to find active hosts*

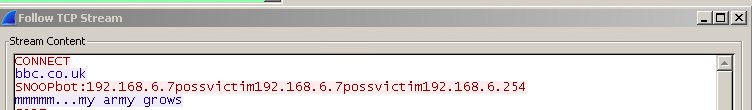
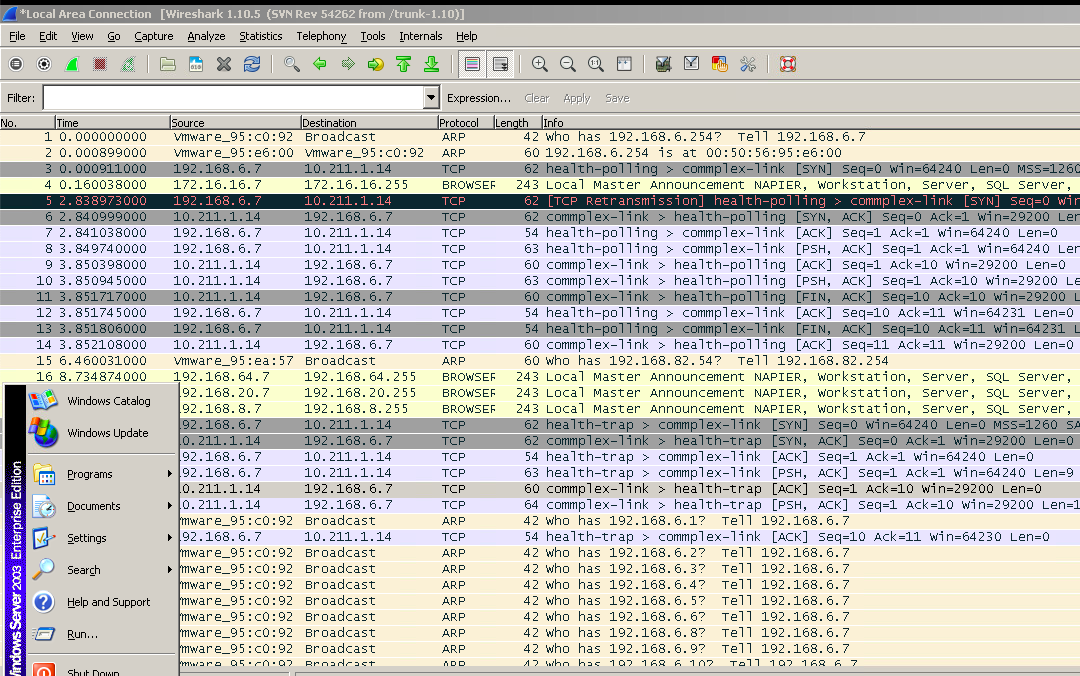
The controller is also able to respond to custom sent commands; messages not from the bot as its source. Through the command ‘telnet 10.211.1.14 5001’ on the Windows machine, a telnet prompt opens and it is possible to send over commands and view the results in the command line window, as is shown below.



*(figure 10) telnet prompt showing sent messages from the host machine, and a response from Kali (controller)*

Building on the first task, there is the requirement to identify types of traffic, reconnaissance / command and control traffic, which is shown below and has supporting evidence in the appropriate appendices.

As previously stated, there are multiple forms of traffic that are generated when the bot is running and connects to the controller. This is shown below (figure 11), and will be interpreted and analysed accordingly.

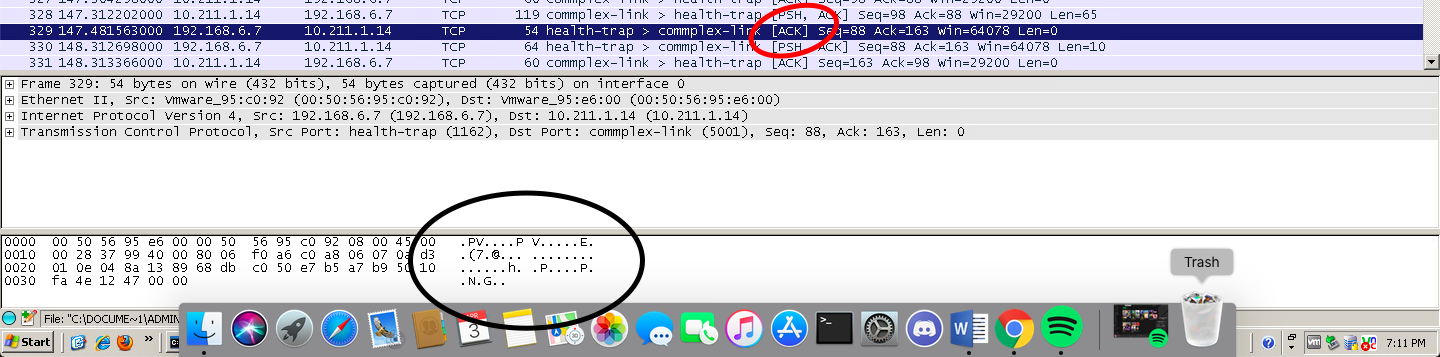
It can be assumed that the traffic involving ARP (other versions of this have picked this up as ICMP traffic) packets are in relation to a sniffing (reconnaissance) function that the bot triggers, possibly in order to tell the controller what machines are on the network and could be targeted. This claim is backed up upon closer inspection of the [PSH, ACK] packets following the ARP sniffing.

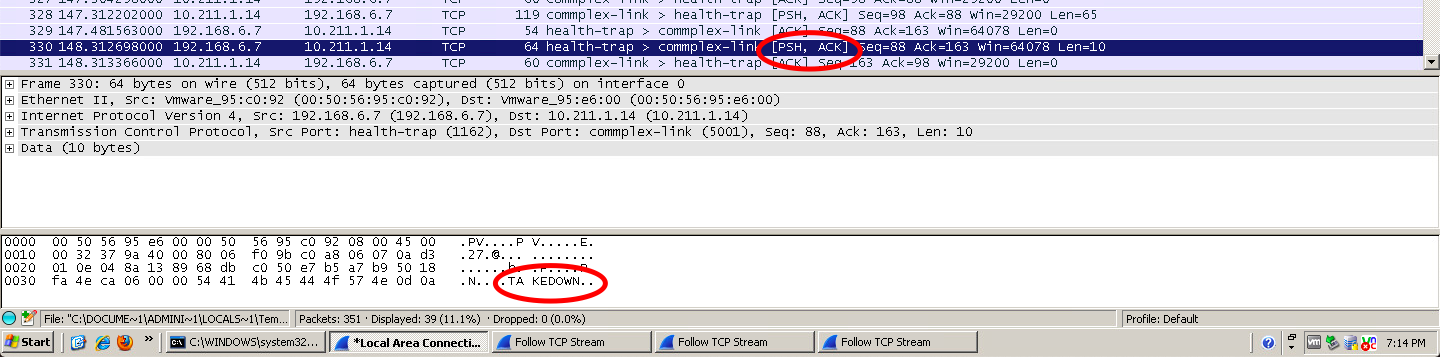
*(figure 11) Sample TCP, ARP, [ACK], [PSH, ACK] packets sent to and from the bot*

When an IP tested returns the ping, it is later pinged by the bot, presumably to ensure the ARP sniffing was accurate.

*(figure 12) TCP stream showing after results of the ARP sniffing*

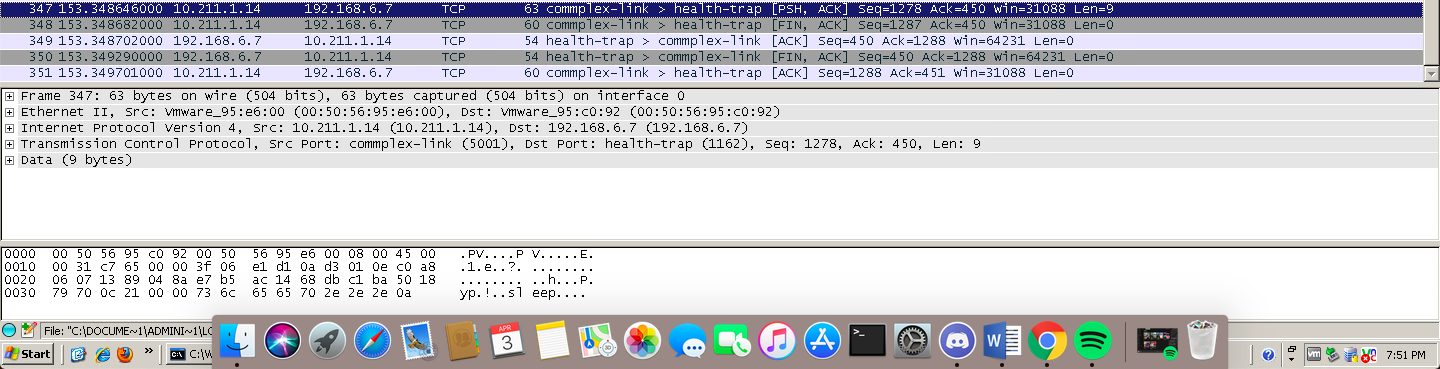
In addition to this, it becomes clear that the command and control traffic is sent in [PSH, ACK] packets, and following the TCP stream, as shown above, facilitates viewing the plaintext commands. Below shows a TCP stream, including content of an [ACK] packet compared to a [PSH, ACK] packet.

*(figure 13) [ACK] packet and contents*



*(figure 14) [PSH] packet and content*

As can be seen, the [ACK] packet content has no clear content (for human reading), whereas a [PSH, ACK] packet will hold readable content, thus making it command and control traffic. This traffic will originate from the bot host (Windows in this case), and will take the form of an [ACK] packet, followed by a [PSH, ACK] packet containing a command. From here, the controller (Kali in this instance) will reply with its own [ACK] packet, followed by its response inside a [PSH, ACK] packet.

It can be observed that the handshake originates from the bot ([SYN] packet from bot, [SYN, ACK] from the controller) and the connection is closed with both the bot and controller sending [FIN] packets followed by [FIN, ACK] packets, as seen below.

*(figure 15) [FIN] and [FIN, ACK] packets sent from bot and controller*

The task to identify and analyse specific botnet commands and response behaviour is further showcased below, again with supporting evidence in the appendices.

Make note of which commands return an encrypted value

1. “Hello”

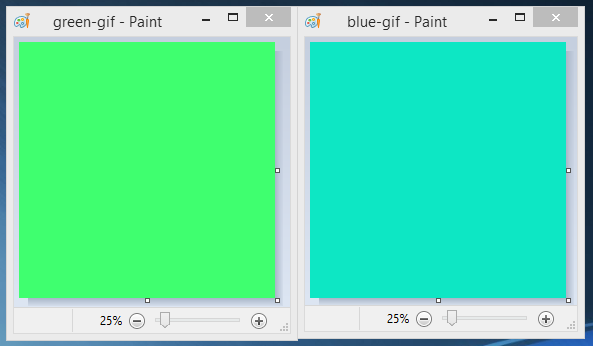
*  
(Figure 16) Get generated image*

* 1. Upon analysis, the hello command when sent from the bot, as well as via telnet, will return "Welcome to the gang…". This is shown in figure 10 (above).

1. “Get”
   1. Upon analyis of packets, the Get command when sent from the bot, as well as telnet, will return a base64 string that can be rendered into an image (figure 16). It can be suggested that the bot sends only one base64 string, however static analysis shows that two base64 strings are parsed together, using a space as a delimiter. This is shown in Appendix Botnet.6.
2. “Test”
   1. Upon analysis, the Test command when sent from the bot, as well as telnet, will return either an MD5, SHA1 or SHA256 encrypted string of a random word from the following “the, owls, are, not, what, they, seem”. In order to decrypt these, tools such as cyberchef (from GCHQ’s Github.io page) can be used. The process of decrypting / decoding these messages is shown below.
   2. This random is decided by the result of a modulus calculation; if divisible by three completely it will use an SHA1 hash, if MOD 1 is the result then SHA256 will be used, otherwise MD5 is used. The proof of this can be seen in figure X.
3. “Failover”
   1. Through dynamic analysis, it can be concluded that the Failover command generates a random integer and sent it along with the Failover keyword. The failover command also creates ‘out.dat’ on the host system, based on the controller response, containing a base64 string which can be rendered into an image (figure 17), but also has a hidden value stating *“create your own bot to access extra cmds ...run cmds many times”*.

*  
` ( (Figure 17) Failover generated image*

*  
(figure 18) Decyphered keepalive command responses*

*  
(figure 19) extracted Gif files from the look command output*

* 1. Static analysis proves this as the set of code run within shows a function that chooses a random integer between 5000 and 4000000, and for each element in a new array, with length of the random integer, assign said array element the value of its corresponding integer. The original number generated is then concatenated to the end of the Failover command string, and sent to the controller.

1. “Connect”
   1. Upon analysis, the Connect command when sent from the bot, as well as telnet, will return one of 6 sites (apple.com, microsoft.com, ibm.com, twitter.com, hpe.com, and bbc.co.uk). It is not clear if anything is done with these URL’s other than being sent from the controller.
2. “Takedown”
   1. Upon analysis, the Takedown command when sent from the bot, as well as telnet, will return one of three ciphers; mentioning Flashpoint in relation to the Mirai Malware, mentioning both types of malware (lockscreen and encryption), and mentioning Cryptolocker ransomware. The keys are Apple, Orange and Peach respectively.
3. “Capture”
   1. Upon analysis, the capture command will receive a string with an alphabet cipher to applied to it from the controller. There are 5 different messages, each of the items can be seen deciphered in Appendix A.
4. “Keepalive”
   1. keepalive, as dynamic analysis suggests, will return one of 68 fruits, each of which would have had a random Caesar Cipher applied. A collection of all fruits used by the botnet upon running the command repeatedly via telnet can be found, deciphered, in figure 18.
5. “Look”
   1. ‘Look’, will, upon analysis, return two separate strings concatenated into one, and upon further analysis is two gif images, seen in figure 19. There could be embedded files / information within each, as they contain headers for different file types after the GIF header. This command will store the controllers response in a file called ‘look.txt’, which can be used to render the two images.
6. “Code”
   1. This command will return “Calvin never forgets transmogrification animal ... may be key...” from the server, and could suggest information on decoding other messages.
7. “Generate”
   1. Upon analysis, the generate command will pull host information and send it to the controller. This could be beneficial in deciding which exploits to side-load onto the machine or to gather capabilities based on OS and versions. The controller will respond with “bgure pbzznaqf znl or ninvynoyr ... uvag uvag”, which can be deciphered to “other commands may be available ... hint hint”, with a Caesar cipher of 13. Based on static analysis, it becomes evident that the generate command will collect the machine name, current directory of the bot, OS version of the host (including service pack number), and username of the logged in user.
8. “Snoop”
   1. Upon analysis, the snoop command will cause the bot to ping all addresses on the network (from x.x.x.0 – x.x.x.255), if on an internal network, and send ip addresses of other internally connected machines, stating that they are potential victims. This, on some instances of Wireshark, is shown via a flood of ARP requests (others have purely ICMP traffic), followed by ICMP traffic for addresses found, followed by the snoop command with addresses of possible victims concateneted to it. The response from the controller says “mmmmm...my army grows”.If the bot machine was not on an internal network, the ping would be sent to 8.8.8.8 once. Static analysis has discovered that the command will test for presence on an internal network, and return the address if so, but will return “-” if not. The code then checks for if the returned statement is a “-” or an address. If a “-”, the bot will ping 8.8.8.8 and do nothing but send the ‘snoop’ command to the controller. If however the code returns an address that starts with a 10, 172 or 192, it will remove the element after the last ‘.’, and concatenate from 0 to 255 on the end of the string (in order) and ping the address. If the address is used, the ping should return true and the address used will be added to the text that is sent across the network.
9. “Loop”
   1. The loop command, as found by dynamic analysis, will send one of 2 AES encrypted strings. These strings can be decrypted, method can be found below, to say:
      1. Subjugating the enemys army without fighting is the true excellence --Sun Tzu...
      2. Its a mistake to think you can solve any major problems with just potatoes --Douglas Adams......
   2. This is done through a function producing a 0 or a 1, and the bot concatenating the element stored at the random number’s location within an array (size 2) holding the AES strings. Through static, another loop listener, which due to the layout of the code will never be run, that opens a new TCP session with 10.211.3.250 on port 53. This however, even if run forcefully by editing the source code, will return in an error “Connect failed”.
10. “Goodbye”
    1. Upon inspection, the goodbye message (sent from the bot) will have a return message (from the controller) of sleep, which writes “+]” to the command line, and initiates a [FIN, ACK],[ACK] sequence to close the connection. sets bool\_1 as true, bool\_0 as false, and closes the TCP connection that was established.
11. “Hint”
    1. This is not found in the dynamic analysis; the only way to send this command is through either utilising Telnet / Netcat to act as a bot client (tricks the controller into responding) and sending the command by typing. The returned result is an MD5 hash of a random element in an array containing the words “animal, Hobbes, is, KEY, what Calvin, pet”. This clearly hints to the keys used for the AES decryption, paired with the (A)ll (E)xplained (S)uccessfully response from the ‘loop’ command.

Finally, there is/are encoded/encrypted botnet traffic/files. Below will show the process of cracking the messages and summarise the activity of the bot.

The commands which return an encoded / encrypted string are, through dynamic analysis, as follows:

## Static Analysis

Through static analysis, it is possible to gain an overview of all the bots' capabilities. In this case, using tools such as dnSpy, and de4dot allowed the source code of the bot and controller to be deobfuscated and reversed so the contents of the source code for each can be seen. An issue that occurred however was the fact that some string tokens were still obfuscated, however using de4dot’s command line argument “-–strtyp delegate” allowed further deobfuscation so that all commands and methods were made clear.

Bot Controller

The bot controller when statically analysed is easy to follow. At the start of its execution, it creates a new instance of “class1” uaing port 5001 to act as a listener for the bots that will try to connect to it, this is done with the “IPAddress.Any” argument meaning it will listen to any IP trying to connect. From this, Method0 is called, which is responsible for printing to output the connection made, accepting the client and string Method1. Method1 is responsible for creating the strings that will be asent over the network (responses to commands) by using streamWriters and Readers bound to the tcp client that was created on bot connection (ASCII encoded). Arrays created here are array, array2, array3 and array4. (Appendix botnet.1) It then goes to interpret any communications from the bots by using a variable called text, which is assigned the value of the streamReader of the tcp client. Each response specified by the controller will result in a certain response from the controller, using method5 to transmit the signal back to the bot. (Appendix botnet.2 – botnet.3). Below can be found the 15 commands for the bot, proving the dynamic analysis by showing the responses, and subsequently the functions behind them.

# Section 3 – Prototype Defences Implementation and Testing

* From your botnet analysis, create and test a basic prototype detection system for the Botnet agent and controller using an IDS sensor. Create IDS rules/signatures to detect the bot activity and not excessive many false positives. This section could show the Snort rules with descriptions of how they work, and screen shots of the testing/outputs and discussion on this.
* Create a closed perimeter firewall configuration to prevent/highlight future communications for this particular botnet, but allow certain valid traffic (specified in requirements spec’). Again show the configuration/rules and testing using screen shot snippets with short explanation, and any interesting discussion on the findings/outputs.

SNORT RULE CHECKING ALL TRAFFIC ON PORT 5001 → CHECK CONTENT > IF ON LIST THEN BLOCK

# Section 4 - References

# Bibliography

Aussems, E., Noë, B. & Rivera, N., 2014. *Botnets - A tenacious Web Technology,* Leiden: Leiden University.

Chawla, A., Lee, B., Fallon, S. & Jacob, P., 2018. *Host based Intrusion Detection System with,* Ireland: Athlone Institute of Technology.

Kaspersky, 2019. *What is Botnet? | Preventing Botnet Attacks | Kaspersky Lab UK.* [Online]   
Available at: https://www.kaspersky.co.uk/resource-center/threats/botnet-attacks

The Honeynet Project, 2008. *The Honeynet Project, 2008. Uses of Botnets. [Online].* [Online]   
Available at: Available at: https://www.honeynet.org/node/52  
[Accessed 14 March 2019].

Vuong , S. & Alam , M., 2011. Advanced Methods for Botnet Intrusion Detection Systems. *Intrusion Detection Systems,* pp.

### Appendix

Appendix A

