Scientific and Practical Cyber Security Journal (SPCSJ) 4(2): 28-41 ISSN 2587-4667 Scientific Cyber Security Association (SCSA)

IMPROVED MIRAI BOT SCANNER SUMMATION ALGORITHM

Faisal A. Garba Department of Computer Science EducationSa'adatu Rimi College of Education Kano, Nigeria

ABSTRACT: Mirai is the most dangerous Distributed Denial of Service (DDoS)-capable IoT malware to date that is in the wild and yet very simple in nature. Mirai attack an array of Internet of Things (IoT) and embedded devices that ranges from Digital Video Recorders (DVRs), Internet Protocol (IP) cameras, routers and printers recruiting them to form a botnet. The biggest DDoS attack in history was executed using Mirai botnet. A recent study proposed the Mirai Bot Scanner Summation Prototype that analyzes the network traffic generated from Mirai bot host discovery. The Mirai Bot Scanner Summation Algorithm however, cannot recognize if a network traffic is truly Mirai bot host discovery traffic or not. Given any network traffic, the Mirai Bot Scanner Summation Prototype will proceed to summate and output number of bots, retransmission packets, number of packets and number of potential victim IoT devices using only the source Internet Protocol (IP) address and destination IP address of a packet without identifying if it is truly a Mirai bot host discovery packet or not. This paper present an Improved Mirai Bot Scanner Summation Algorithm that looks at the packet to determine whether it is a truly packet generated due to Mirai bot host discovery by looking at the TCP flag of the packet and the port number of the packet. To perform a host discovery Mirai bot sends out SYN packet over TELNET port 23 or 2323 to a randomly generated non-governmental IP addresses to establish a TCP 3-way handshake with a potentially vulnerable IoT device. The Improved Mirai Bot Scanner Summation Algorithm uses this condition to determine whether a packet is a Mirai bot host discovery packet or not. The Mirai Bot Scanner Summation Algorithm and the Improved Mirai Bot Scanner Summation Algorithm are evaluated using IoT Network Intrusion Dataset. The evaluation results have shown that the Improved Mirai Bot Scanner Summation Algorithm provides more accurate results than the Mirai Bot Scanner Summation Algorithm.

KEYWORDS: *Mirai, Internet of Things, botnet, Denial of Service Attack, cyber attack.*

INTRODUCTION

Internet of Things (IoTs) devices are a key targets for cyber attacks as a result of their fast growing number in smart cities, smart homes, smart hospitals etc and the quantity and sensitivity of the data they collected (Frank, 2019). Two issues highlighted by the IoT botnet are the reality that: a large number of IoT devices are easily reached over the Internet and most of the times security is a later addition to the design of most of the deployed IoT devices, that is if it has been given any consideration at all (Angrishi, 2017). One of the most predominant Distributed Denial of Service (DDoS) - capable IoT malware of the past few years is the Mirai malware which was discovered in the year 2016 and has changed the global view of IoT security since then (De-Donno et al., 2018). Mirai attack an array of IoT and embedded

devices that ranges from Digital Video Recorders (DVRs), Internet Protocol (IP) cameras, routers and printers (Antonakakis et al., 2017). The Mirai bot scanner generates a random non-governmental IP address and also creates a network socket and performs a TCP handshake (Frank, 2019). The biggest DDoS attack in history was executed using Mirai (York, 2016). This was achieved as a result of constructing a large Agent-Handler botnet that comprises of mini IoT devices taken over via a simple dictionary attack. Mirai malware is capable of carrying out a diverse number of attacks based on variety of protocols such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP) and Hypertext Transfer Protocol (HTTP) and can exploit devices that are based on different architectures. Mirai is the most dangerous DDoS-capable IoT malware to date that is in the wild and yet very simple in nature (De-Donno et al., 2018).

REVIEW OF RELATED WORKS

Kumar and Lim (2020) developed a network-based algorithm which can be used to detect IoT bots infected by Mirai and other malware in large-scale networks. They developed an algorithm that targets bots scanning the network for vulnerable devices to detect the bots before they launch an attack. They analyzed Mirai signatures to identify it's presence in an IoT device. They lay their claim that uninfected IoT devices are not expected to open TELNET connections to any device. Kumar and Lim (2020) work is aimed at detecting Mirai Botnet attack in progress.

Frank (2019) developed Mirai Scanner Summation Prototype. With the use of Python scripts, the Mirai Bot Scanner Summation Prototype searches through the Bot scanner dataset to sum up bots, potential new bot victims and network packet types including TCP SYN and retransmission packets and save result in a database. The Mirai Scanner Summation Prototype however does not look at the packets to ensure that they are really Mirai bot scanner packets. The Mirai Scanner Summation Prototype only looks at the source and destination IP of the packets and whether they are unique or non-unique packets before proceeding to calculate the number of bots and potential bot victims from the packets. Therefore given

any network traffic that are not Mirai bot scanner packets, the Mirai Bot Scanner Summation Prototype will still proceed to calculate the number of bots and potential bots victims out of the packets.

Meidan et al., (2018) proposed N-Balot, A Network-Based Detection of IoT Botnet Attacks using Deep Autoencoders, proposed and empirically evaluated a novel network-based anomaly detection method which extracts behavior snapshots of the network and uses deep autoencoders to detect anomalous network traffic emanating from compromised IoT devices. The researchers claimed their work effectively detects attacks as they are being initiated from a compromised IoT device that are part of a botnet.

IMPROVED MIRAI BOT SCANNER SUMMATION ALGORITHM:

The Summation Algorithm used in the Mirai Bot Scanner Summation Prototype proposed by Frank (2019) cannot differentiate if the network traffic in the form of pcap files passed on to it is truly Mirai Bot Scanner network traffic. Looking at the Mirai Bot Summation Algorithm in Figure 1, the algorithm expects to have network traffic generated due to Mirai bot scanning activity. As pointed out by Frank (2019), the Mirai Scanner network traffic dataset (Internet Addresses Census dataset, IMPACT ID: USC-LANDER/Mirai-Bscanning-20160601/rev5870, 2016), consists of only SYN packets sent by a Mirai Bot over TELNET port 23 or 2323 to initiate a connection with an IoT device. The Summation Algorithm only looks out for the source IP address and the destination IP address of a packet (any packet) and if the packet is unique, the algorithm concludes that it is a Mirai Bot and if the packet is not unique, then the summation algorithms concludes the destination IP represents a non-vulnerable IoT device (Frank, 2019).

```
//Initialization
Total_Bots = 0, Total_Potential_New_Bot_Victims = 0, Total_SYN = 0
Total_Retransmission = 0, Total_Packets = 0, Starting_Time = 0, Ending_Time = 0
Packet_date = 0, L = [], S= [], SUBNETS = []
Starting_Time = now
Packet_date = date_from_filename(PCAP)
// Read the network packets of the PCAP file
Insert into list L the source and destination IP of each network packet
// Go thru each element of L
For i in L
    // Summate total packets
    Total_Packets = Total_Packets + 1

        // Determine subnet of destination IP
        Add Subnet(L[i].destination_IP) to SUBNETS
        // Unique source IP represents a Bot
        If the count(L[i].source_IP in L) == 1
            Total_Bots = Total_Bots + 1
        // Unique SYN packet
        If the count (L[i] in L) == 1
            Insert L[i].destination_IP into S
            Total_SYN = Total_SYN + 1
        // Retransmission packet
        If the count (L[i]) > 1
            Total_Retransmission = Total_Retransmission + 1

// Go thru each destination IP in S
For j in S
    // a unique destination IP in S represent a potential New Bot Victim
    If the count(L[j] in S) == 1
        Total_Potential_New_Bot_Vicitms = Total_Potential_New_Bot_Vicitms + 1

Ending_Time = now

//Insert summation results into the database
Insert Total_Bots, Total_Potential_New_Bot_Victims, SUBNETS
    Total_SYN, Total_Retransmission, Total_Packets,
    Starting_Time, Ending_Time, Packet_date
Into
Persistent Storage
```

Figure 1: Mirai Bot Summation Algorithm

There is a need to improve the Summation Algorithm to check a packet to determine whether it is actually a SYN packet and sent over port 23 or 2323 before assuming it is a SYN packet sent out by a Mirai Bot.

To validate this fact we ran the Mirai Bot Scanner Summation Prototype over mirai-ackflooding-n(1~4)-dec.pcap files from the IoT Network Intrusion Dataset made available by Kang et al., (2019). This file contains Mirai Bot ACK flood packets and benign packets (Kang et al., 2019). A description of the dataset is provided in Table 1 and the contents of the IoT Network Intrusion Dataset is shown in Figure 2.

Table 1: Description of the IoT Network Intrusion Dataset

Packet File Name	Description

benign-dec.pcap	Benign-only traffic
mitm-arpspoofing-n(1~6)-dec.pcap	Traffic containing benign and MITM(arp spoofing)
dos-synflooding-n(1~6)-dec.pcap	Traffic containing benign and DoS(SYN flooding) attack
scan-hostport-n(1~6)-dec.pcap	Traffic containing benign and Scan(host & port scan) attack
scan-portos-n(1~6)-dec.pcap	Traffic containing benign and Scan(port & os scan) attack
mirai-udpflooding-n(1~4)-dec.pcap	Traffic containing benign and UDP flooding of zombie pc compromised by mirai malware
mirai-ackflooding-n(1~4)-dec.pcap	Traffic containing benign and ACK flooding attack of zombie pc compromised by mirai malware
mirai-httpflooding-n(1~4)-dec.pcap	Traffic containing benign and HTTP Flooding attack of zombie pc compromised by mirai malware
mirai-hostbruteforce-n(1~5)-dec.pcap	Traffic containing benign and initial phase of Mirai malware including host discovery and Telnet brute-force attack

.

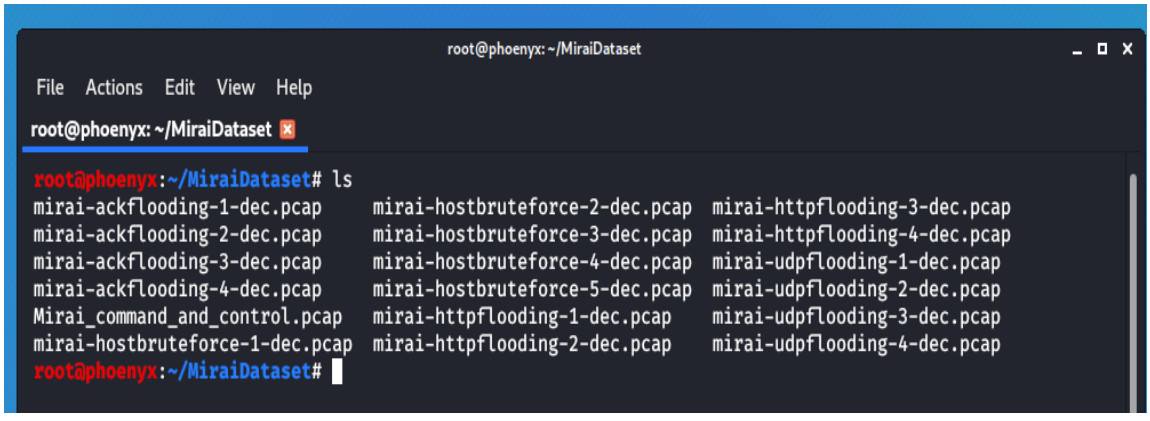

Figure 2: Contents of the IoT Network Intrusion Dataset

In Figure 3, we opened the mirai-ackflooding-1-dec.pcap using Wireshark. Next we apply the filter "tcp.flags.syn==1 and tcp.flags.ack==0 and (tcp.port==23 or tcp.port==2323)" to see if there are Mirai Bot Scanning packets in Figure 3.

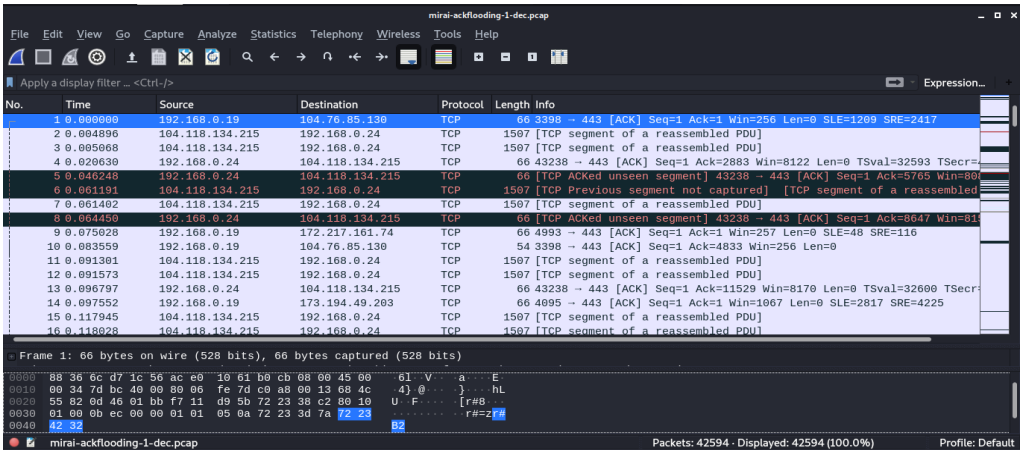

Figure 3: : Viewing mirai-ackflooding-1-dec.pcap File using Wireshark

As we can see in Figure 4, the filter returns no results. This shows that the mirai-ackflooding-1-dec.pcap contains no SYN packets sent over Telnet port 23 or 2323. However, running the Mirai Bot Scanner Summation Prototype over the mirai-ackflooding-1-dec.pcap, it couldn't identify that it is not a Mirai Bot Scanner packets, it went on straight ahead to summate the packets as if they were packets generated due to Mirai Bot scanning activity as seen in Figure 5.

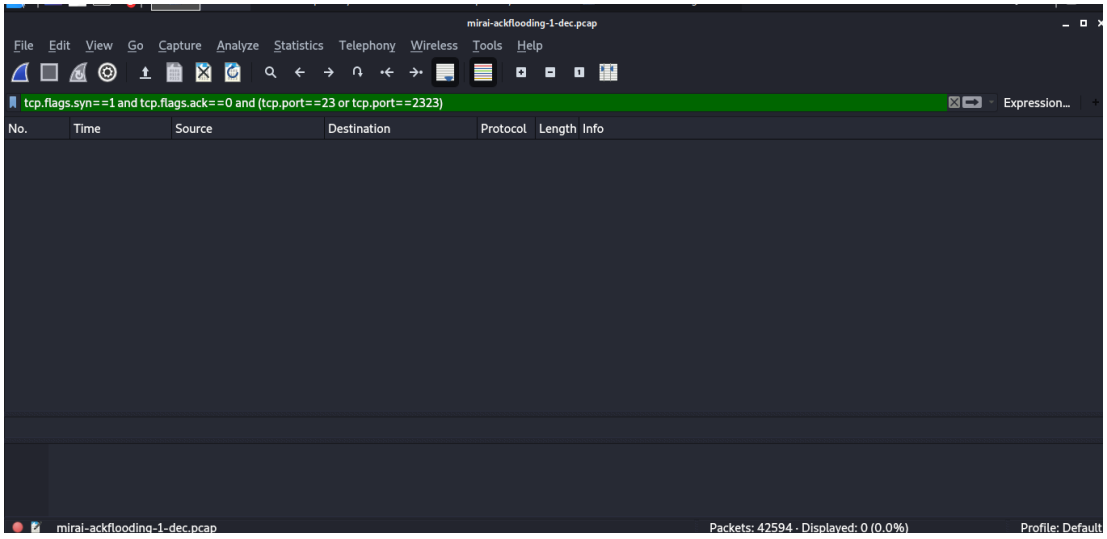

Figure 4: Applying filter to mirai-ackflooding-1-dec.pcap in Wireshark

Figure 6 presents some of the assessment results from the summation of Figure 5. The assessment results (although incorrect) is for each of the packets summated in the Mirai AckFlood folder which contains packet files generated as a result of Mirai Ack Flood DoS attack and benign packets as described in the dataset provided by Kang et al., (2019).

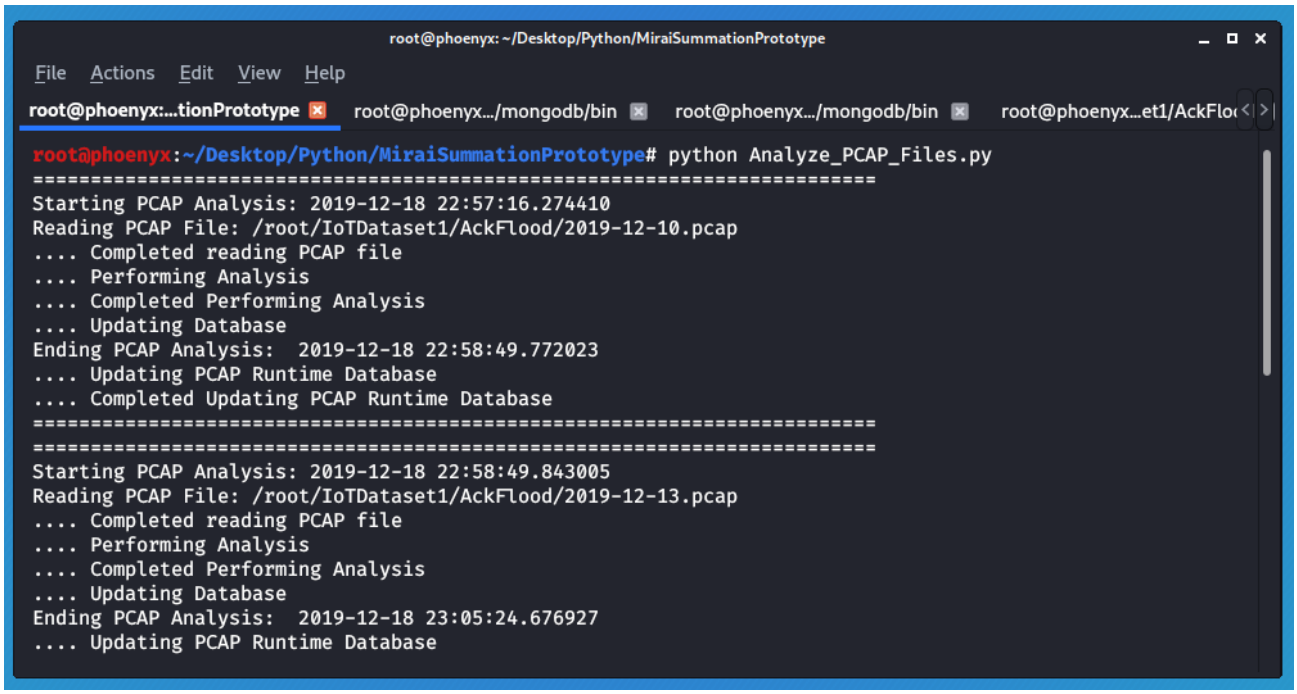

Figure 5: Summating AckFlood Files with Mirai Bot Scanner Summation Prototype

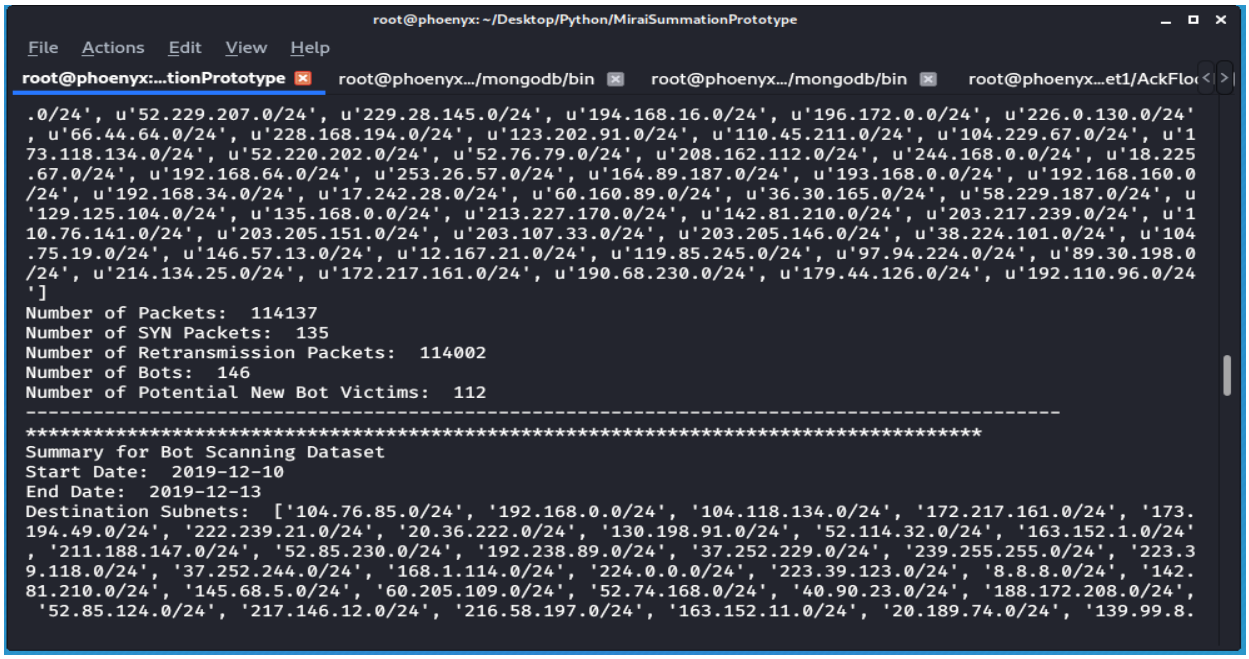

Figure 6: Assessment Results

Figure 7 presents the Improved Mirai Bot Summation Algorithm. Line 13 adds the check for the packet to ensure it is a SYN Telnet packet. Next we run the Improved Mirai Bot Summation Algorithm which is implemented in the BotScanner.py component of the Mirai Bot Scanner Summation Prototype developed by Frank (2019). The Mirai Bot Scanner Summation Prototype is available for download from its Github repository[1]

Figure 8 is a code snippet that implements the line 13 of the Improved Mirai Bot Summation Algorithm that checks for a packet being a SYN Telnet packet. The code snippet is implemented in the function analyze_pcap_file(). Variable SYN has been initialized outside the function with the value 0x02, which is the hexadecimal representation of the TCP SYN flag (Chandel, 2018).

[1] https://github.com/infosecchazzy/Mirai_Bot_Scanner_Summation_Prototype

```
//Initialization
Total_Bots = 0, Total_Potential_New_Bot_Victims = 0, Total_SYN = 0
Total_Retransmission = 0, Total_Packets = 0, Starting_Time = 0, Ending_Time = 0
Packet_date = 0, L = [], S= [], SUBNETS = []
Starting_Time = now
Packet_date = date_from_filename(PCAP)
// Read the network packets of the PCAP file
Insert into list L the source and destination IP of each network packet
// Go thru each element of L
For i in L
    // Summate total packets
    Total_Packets = Total_Packets + 1
    If  i == Telnet SYN packet
        // Determine subnet of destination IP
        Add Subnet(L[i].destination_IP) to SUBNETS
        // Unique source IP represents a Bot
        If the count(L[i].source_IP in L) == 1
            Total_Bots = Total_Bots + 1
        // Unique SYN packet
        If the count (L[i] in L) == 1
            Insert L[i].destination_IP into S
            Total_SYN = Total_SYN + 1
        // Retransmission packet
        If the count (L[i]) > 1
            Total_Retransmission = Total_Retransmission + 1

// Go thru each destination IP in S
For j in S
    // a unique destination IP in S represent a potential New Bot Victim
    If the count(L[j] in S) == 1
        Total_Potential_New_Bot_Vicitms = Total_Potential_New_Bot_Vicitms + 1

Ending_Time = now

//Insert summation results into the database
Insert Total_Bots, Total_Potential_New_Bot_Victims, SUBNETS
    Total_SYN, Total_Retransmission, Total_Packets,
    Starting_Time, Ending_Time, Packet_date
Into
Persistent Storage
```

Figure 7: Improved Mirai Bot Summation Algorithm

```
119
120         try:
121             # check if it is a bot SYN packet
122             flag = each_packet.getlayer(TCP).flags
123             dest_port = each_packet[TCP].dport
124
125             if (flag == SYN) and (dest_port == 23 or dest_port == 2323):
126
```

Figure 8: Code Snippet Containing the Check for the SYN Telnet Packet

In both implementations of the Mirai Summation Algorithm and that of the Improved Mirai Summation Algorithm we had to rename the packets in the mirai-ackflooding-n(1~4)-dec.pcap folder with dates from 10-12-2019 to 13-12-2019 as seen in Figure 9. This is in order to comply with the function pcap_file_date() in BotScanner.py component of the Mirai Bot Scanner Summation Prototype which requires the name of the packet files to be in date format as seen in Figure 10.

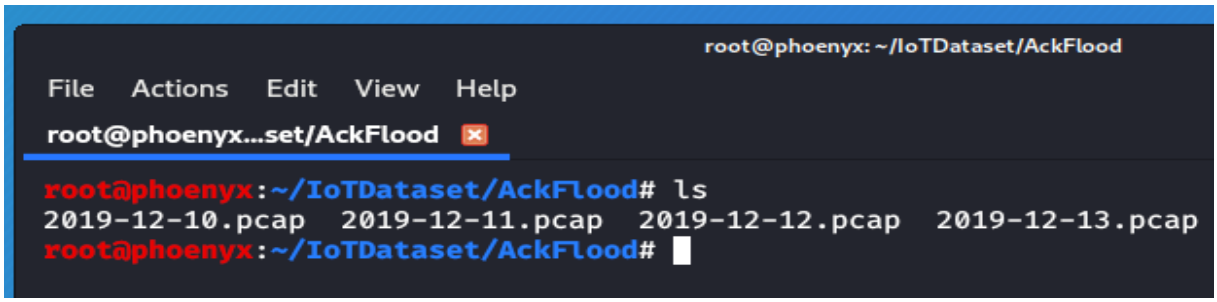

Figure 9: Renaming the Files in mirai-ackflooding-n(1~4)-dec.pcap

```
##
# This function returns the PCAP file date
# Parameter:
# pcap_file - pcap file
##
def pcap_file_date(pcap_file):
    # Initialize
    packet_date = ""

    # split the file name that includes the complete path
    packet_date = pcap_file.split("/")

    # Split to get the file name without the extension, which is the date
    packet_date = packet_date[(len(packet_date) - 1)].split(".")

    # return the date
    return packet_date[0]
```

Figure 10: Packet Files Required to be in Date Format

Next we proceed to run the Improved Mirai Bot Scanner Summation Prototype over the AckFlood packets by invoking the Analyze_PCAP_Files python script as seen in Figure 11.

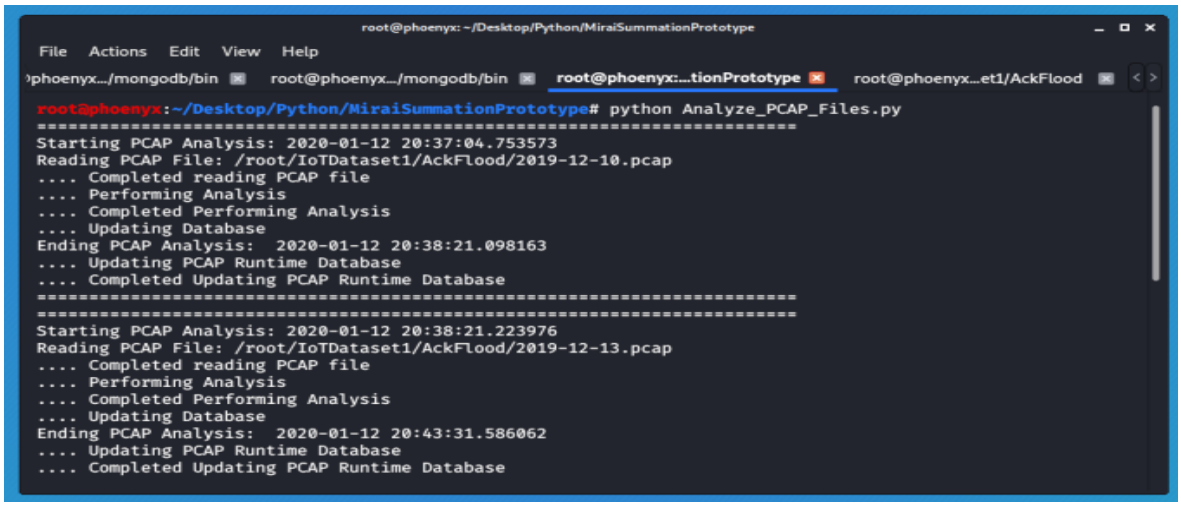

Figure 11: Running the Analyze PCAP Files Python Script

When we invoke the Answer_Research_Questions.py python scripts, a component of the Mirai Bot Scanner Summation Prototype that parses through the MongoDB to tabulate the number of bots, potential bot victims and number of packets, it didn't shows that there are no bots and potential bots victims which is right, since the AckFlood dataset contains AckFlood packets and not Mirai bot scanner packets. This is seen in Figure 12.

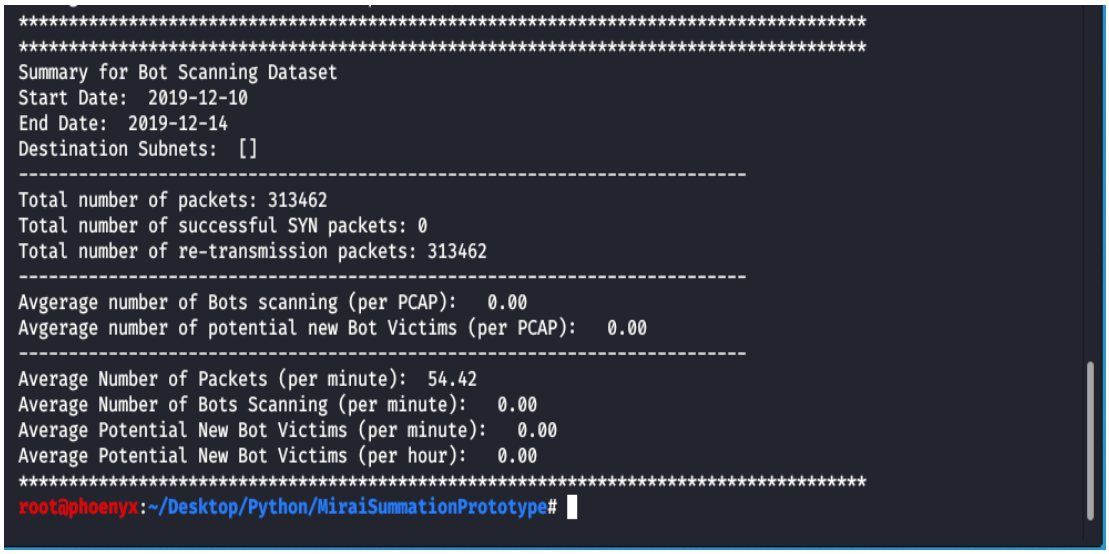

Figure 12: Assessment Results for AckFlood Packets

The Mirai HostBruteforce pcap file contains benign and initial phase of Mirai malware including host discovery and Telnet brute-force attack according to the dataset description provided by Kang, et al., (2019). So we expect to see Mirai bot scanner SYN packets sent over Telnet port 23 or 2323. We also validate that using Wireshark and the TCP filter "tcp.flags.syn==1 and tcp.flags.ack==0 and (tcp.port==23 or tcp.port==2323)". From Figure 13 we can see that mirai-hostbruteforce-1-dec.pcap contains Mirai bot scanner packets.

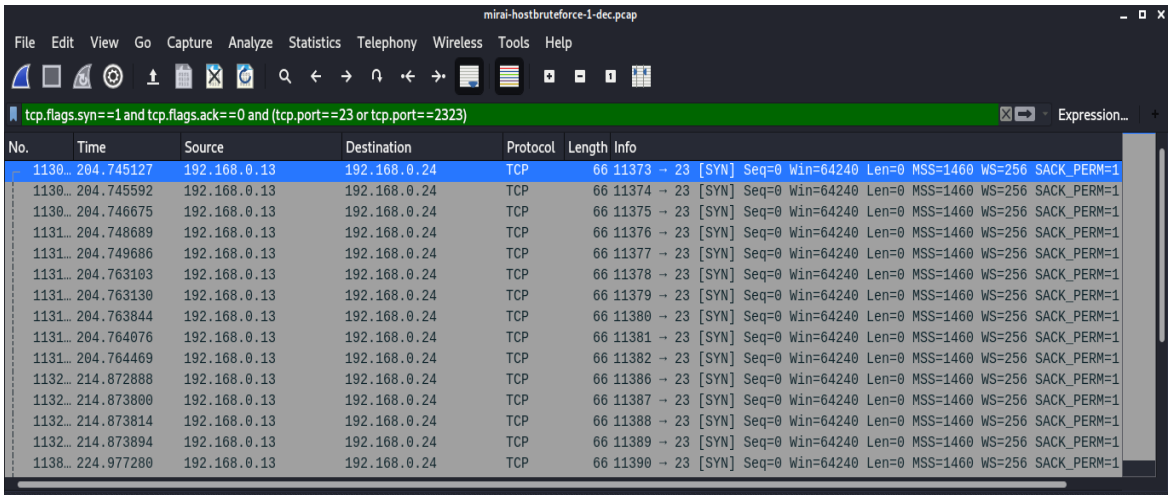

Figure 13: Viewing mirai-hostbruteforce-1-dec.pcap with Wireshark

Next we proceed to run the Mirai Bot Scanner Summation Prototype against the mirai-hostbruteforce-1-dec.pcap by invoking the Analyze_PCAP_Files.py after which we invoke the Answer_Research_Questions.py python script to carry out assessment of the summated data in the MongoDB database. The assessment result is presented in Figure 14. From Figure 14 we can see that there is a single Mirai bot per pcap file (for the 4 pcap files passed to the program), the destination subnet is 192.168.0.0/24 and there are 453,355 total packets.).

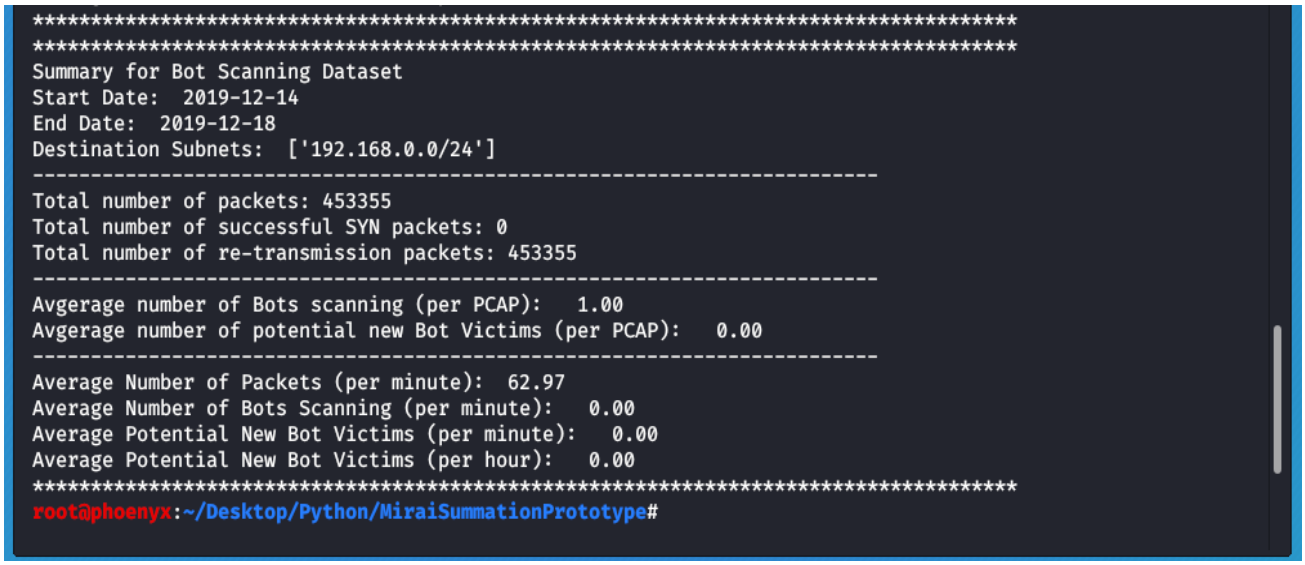

Figure 14: Assessment Results for Mirai HostBruteforce Packets

CONCLUSION AND FUTURE WORK

This paper have presented Improved Mirai Bot Scanner Summation Algorithm which is an improvement upon the Mirai Bot Scanner Summation Algorithm proposed by Frank (2019). The paper has evaluated both algorithms with IoT Network Intrusion Dataset provided by Kang, et al., (2019). Evaluation results have shown that the Improved Mirai Bot Scanner Summation Algorithm analyzes Mirai bot scanner pcap files more accurately. Future work will extend the Mirai Bot Scanner Summation Prototype developed by Frank (2019) to address those component of Mirai Malware operations not handled by the Mirai Bot Scanner Summation Prototype which include Mirai command and control, Mirai bruteforce login and Mirai Denial of Service (DoS) attacks.

REFERENCES

1. A. Kumar and T. J. Lim, "Early Detection of Mirai-Like IoT Bots In Large-Scale Networks Through Sub-Sampled Packet Traffic Analysis," in *Advances in Information and Communication. FICC 2019. Lecture Notes in Networks and Systems*, San Francisco, CA, USA , 2020.
2. C. Frank, "Mirai Bot Scanner Summation Prototype," Masters Theses & Doctoral Dissertations, 2019.
3. H. Kang, D. H. Ahn, G. M. Lee, J. D. Yoo, K. H. Park and H. K. Kim, "IoT network intrusion dataset," 09 September 2019. [Online]. Available: https://ieee-dataport.org/open-access/iot-network-intrusion-dataset.
4. K. Angrishi, "Turning Internet of Things(IoT) into Internet of Vulnerabilities (IoV) : IoT Botnets," *arXiv*, pp. 1-16, 2017.
5. K. York, "Read Dyn’s Statement on the 10/21/2016 DNS DDoS Attack," 2 October 2016. [Online]. Available: https://dyn.com/blog/dyn-statement-on-10212016-ddos-attack/.
6. M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, A. J. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C.

Seaman, N. Sullivan, K. Thomas and Y. Zhou, "Understanding the Mirai Botnet," in *26th USENIX Security Symposium*, Vancouver, BC, Canada, 2017.

7. M. De-Donno, N. Dragoni, A. Giaretta and A. Spognardi, "DDoS-Capable IoT Malwares: Comparative Analysis and Mirai Investigation," *Hindawi Security and Communication Networks,* pp. 1-30, 2018.
8. R. Chandel, "Nmap Scans using Hex Value of Flags," 31 January 2018. [Online]. Available: https://www.hackingarticles.in/nmap-scans-using-hex-value-flags/.
9. Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai and Y. Elovici, "N-BaIoT: Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders," *IEEE PERVASIVE COMPUTING*, vol. 13, no. 9, pp. 1 - 6, 2018.