Packet Data Analysis: Idle Laptop

[0.5 pts] Group members: Leonel Covarrubias

[0.5 pts] Goal of the project:

The Goal of the project was to analyze the incoming and outgoing data from my laptop while it is idle; as well as how it differs from different times of day. The project is broken up into 3 Sections of analysis. The first Section finds which IP addresses are interacting with my laptop the most. Section 2 then gets these addresses and finds the IP Owners, which will help me find potential applications on my laptop that are communicating to the network in the background. The third section will be for analyzing the percentage of TCP and UDP packets in the data captures, and the percentage of successful acknowledgments; as well as how it compares to a collection of packets that were obtained while using my laptop normally (idle vs non-idle).

[4 pts] Details of Code: Provide details of how to run your code and how to reproduce your work, and any other comments that you have about your code.

- **How to run:** In order to run the code, type in 'python3 .\packet_percentage_generator.py' and '.\packet_analysis_generator.py' into the terminal.

- **Environment:** The only library needed is the 'scapy' library. Visual Studio Code was the environment this project was conducted in.

- **Input/Output:** For running the Python script, the capture files will be used to generate statistics. Below is an example output for analyzing a Wireshark capture packet.

  Expected Output for 'python3 .\packet_percentage_generator.py':

  Night 1:

  Top 5 Source IPs when 192.168.1.223 is the destination:
  Source IP Address: 192.168.1.1, Occurrences: 632822
  Source IP Address: 34.104.35.123, Occurrences: 5125
  Source IP Address: 13.107.42.12, Occurrences: 4597
  Source IP Address: 72.21.81.240, Occurrences: 1017
  Source IP Address: 192.168.1.189, Occurrences: 644

  Top 5 Destination IPs when 192.168.1.223 is the source:
  Destination IP Address: 13.107.42.12, Occurrences: 12146
  Destination IP Address: 34.104.35.123, Occurrences: 705
  Destination IP Address: 72.21.81.240, Occurrences: 390
  Destination IP Address: 20.190.151.133, Occurrences: 253
  Destination IP Address: 151.139.71.24, Occurrences: 217

Expected Output for '.\packet_analysis_generator.py':

Night 1:

Percentage of TCP packets: 4.824%
Percentage of UDP packets: 95.060%
Percentage of successful Acknowledgments: 99.301%

- **Extra comments:** At the bottom of the python files, the actual function calls to run the analysis on my data are commented out, as running the entirety of the script at once takes time. In replication, ensure the python files have access to the data.

[3 pts] Details of data:

- **Pointer to data:** The data is available for download via the Dropbox link:

  https://www.dropbox.com/t/01eo4uAlXYUnV1Jr

- **Cleaning and processing:**

  For capturing the data, I placed the filter 'ip.addr == 192.168.1.223', on Wireshark, which only captures the packets associated with my laptops IP address. Once these packets were captured, they were processed further in the Python code that was provided. Please refer to the Methodology and Evaluation sections to see exact processing done on the packet data.

- **Any extra comments:**

  Please email me at lcovarrubias@nevada.unr.edu if you have any problems retrieving the data from the Dropbox link. Dropbox was used as GitHub would not allow me to upload larger collections of data.

[2 pts] Methodology:

Firstly, was the collection of the data packets. These packets were manually collected via Wireshark. 10 captures were made at 3 different parts of the day. 3 captures were made in the morning, 3 were made in the afternoon (day), and 3 were made in the night. These 9 captures ran with no windows open on my computer. This way we can capture data when the laptop is idle. One capture was taken under normal usage when I was doing homework and was used to compare during Section 3's evaluation. To capture data that was only associated with my personal laptop, the filter, 'ip.addr == 192.168.1.223', was used to capture the data. Each capture was run for about 30 minutes.

For the first section of analysis, I wrote a python script using the 'scapy' library. The purpose of this first script was to display the most common IP addresses interacting with my computer. The function in the, 'packet_analysis_generator.py' file takes in a pcapng file and an IP address. It then displays the most common destination IP addresses, when the specified IP address is the source, and vice versa. This script ran on all 9 idle pcapng captures.

Using those outputs created, I combined the outputs based on the time of day and placed these into Excel. Once having the combined outputs in Excel, I combined the IP addresses using longest prefix match, using a 24/ subnet mask, as well as combining the number of occurrences. This was done to associate similar IP addresses to the same owner. I then removed IP addresses having less than 100 occurrences as IP addresses having more than 100 occurrences were considered significant. These were entered into the tables in Section 1 of my evaluation. To find the owners of the IP addresses, the website, "www.whois.com", was used. This site allows you to enter an IP address and return the information of the owner of the IP address. The site does not return information pertaining to private IP addresses. The results were then entered on the table.

For the Section 3 analysis, the Python script 'packet_percentage_generator.py' was created. This file contains a function that takes in a pcapng file and finds the percentage of TCP and UDP files; as well as the percentage of successful ACKs. This was run on all 10 Wireshark captures. This data was then entered into Excel, and the averages were taken by the time of day. These were then averaged out together, to be able to compare the percentage between idle and normal usage. These findings were then entered into the table in Section 3.

SECTION 1: Most Common IP Addresses

Goals of Analysis:

This section will be analyzing the Top 5 IP addresses when My IP address is the Source IP, and the top 5 IP addresses when my IP address is the destination IP addresses.

Observations:

- Most common IP address is shared across all days and times.
- Disproportionate number of occurrences in the first place Source IP addresses
- Disproportionate number of Source occurrences from the Source IP addresses compared to the destination IP addresses.
- Lots of IP addresses that are shared with the same \24 subnet mask as my IP address.
- Not too diverse
- Different times of day do not mean more diverse results.
- Day 3, Night 3, and Night 1 had destination occurrences that are significantly more than any other destination IP address.
- The Night IP addresses have More occurrences than other times of days.

```
Morning 1:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 308104
Source IP Address: 204.79.197.200, Occurrences: 1430
Source IP Address: 192.168.1.189, Occurrences: 616
Source IP Address: 52.113.194.132, Occurrences: 436
Source IP Address: 162.159.134.234, Occurrences: 102

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 204.79.197.200, Occurrences: 249
Destination IP Address: 192.168.1.189, Occurrences: 166
Destination IP Address: 162.159.134.234, Occurrences: 101
Destination IP Address: 40.83.247.108, Occurrences: 92
Destination IP Address: 52.113.194.132, Occurrences: 92

Morning 2:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 296780
Source IP Address: 69.197.43.26, Occurrences: 429
Source IP Address: 20.106.86.13, Occurrences: 169
Source IP Address: 192.168.1.189, Occurrences: 157
Source IP Address: 20.125.63.4, Occurrences: 82

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 20.106.86.13, Occurrences: 163
Destination IP Address: 192.168.1.1, Occurrences: 129
Destination IP Address: 20.125.63.4, Occurrences: 74
Morning 3:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 155022
Source IP Address: 192.168.1.189, Occurrences: 634
Source IP Address: 205.196.6.214, Occurrences: 126
Source IP Address: 142.250.68.67, Occurrences: 99
Source IP Address: 34.198.235.52, Occurrences: 73

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 192.168.1.189, Occurrences: 195
Destination IP Address: 205.196.6.214, Occurrences: 130
Destination IP Address: 40.83.247.108, Occurrences: 99
Destination IP Address: 52.239.235.196, Occurrences: 99
Destination IP Address: 142.250.68.67, Occurrences: 85
```

Figure 1Most Common IPs in the Morning

Analysis:

Considering that the most common IP address is shared by all days and times, I think that this is my local Gateway router that it is communicating with. Given that there are significantly more occurrences in the Source IPs, it tells me that my laptop is receiving much more information than it is sending. Having multiple IP addresses in my IP's subnet shows that my laptop is communicating locally to a nearby device. There were also a lot of similarities in the IP Addresses across the days and time, meaning that the same background processes are running when idle. However, I am not sure which applications caused substantially more occurrences in Day 3, Night 3, and Night. A possible explanation for more occurrences in destination IP addresses during the night is potentially leaving background processes running after using my laptop all day.

```
Day 1:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 144075
Source IP Address: 192.168.1.189, Occurrences: 320
Source IP Address: 20.60.194.195, Occurrences: 118
Source IP Address: 52.137.108.250, Occurrences: 72
Source IP Address: 20.190.190.195, Occurrences: 54

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 20.60.194.195, Occurrences: 377
Destination IP Address: 192.168.1.189, Occurrences: 90
Destination IP Address: 205.196.6.215, Occurrences: 75
Destination IP Address: 40.83.240.146, Occurrences: 74
Destination IP Address: 20.190.190.195, Occurrences: 59

Day 2:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 427332
Source IP Address: 192.168.1.189, Occurrences: 662
Source IP Address: 13.71.55.58, Occurrences: 104
Source IP Address: 20.190.190.194, Occurrences: 82
Source IP Address: 20.190.151.9, Occurrences: 75

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 192.168.1.189, Occurrences: 183
Destination IP Address: 40.83.240.146, Occurrences: 87
Destination IP Address: 13.89.178.26, Occurrences: 82
Destination IP Address: 239.255.255.250, Occurrences: 76
Destination IP Address: 205.196.6.214, Occurrences: 69

Day 3:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 284532
Source IP Address: 72.21.81.240, Occurrences: 120055
Source IP Address: 151.139.71.3, Occurrences: 5138
Source IP Address: 151.139.51.179, Occurrences: 1871
Source IP Address: 192.168.1.189, Occurrences: 633

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 72.21.81.240, Occurrences: 19989
Destination IP Address: 151.139.51.179, Occurrences: 1184
Destination IP Address: 151.139.71.3, Occurrences: 948
Destination IP Address: 192.168.1.189, Occurrences: 165
Destination IP Address: 52.165.164.15, Occurrences: 145
```

Figure 2: Most Common IPs in the Day

```
Night 1:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 632822
Source IP Address: 34.104.35.123, Occurrences: 5125
Source IP Address: 13.107.42.12, Occurrences: 4597
Source IP Address: 72.21.81.240, Occurrences: 1017
Source IP Address: 192.168.1.189, Occurrences: 644

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 13.107.42.12, Occurrences: 12146
Destination IP Address: 34.104.35.123, Occurrences: 705
Destination IP Address: 72.21.81.240, Occurrences: 390
Destination IP Address: 20.190.151.133, Occurrences: 253
Destination IP Address: 151.139.71.24, Occurrences: 217

Night 2:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 377606
Source IP Address: 192.168.1.189, Occurrences: 725
Source IP Address: 52.113.194.132, Occurrences: 261
Source IP Address: 205.196.6.214, Occurrences: 136
Source IP Address: 3.93.253.89, Occurrences: 117

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 192.168.1.189, Occurrences: 163
Destination IP Address: 205.196.6.214, Occurrences: 136
Destination IP Address: 40.83.247.108, Occurrences: 121
Destination IP Address: 3.93.253.89, Occurrences: 115
Destination IP Address: 239.255.255.250, Occurrences: 93

Night 3:

Top 5 Source IPs when 192.168.1.223 is the destination:
Source IP Address: 192.168.1.1, Occurrences: 548820
Source IP Address: 151.139.51.187, Occurrences: 4978
Source IP Address: 192.168.1.189, Occurrences: 752
Source IP Address: 184.28.98.100, Occurrences: 185
Source IP Address: 72.21.81.200, Occurrences: 126

Top 5 Destination IPs when 192.168.1.223 is the source:
Destination IP Address: 151.139.51.187, Occurrences: 2186
Destination IP Address: 192.168.1.189, Occurrences: 185
Destination IP Address: 205.196.6.214, Occurrences: 113
Destination IP Address: 142.250.188.227, Occurrences: 104
Destination IP Address: 192.168.1.1, Occurrences: 94
```

Figure 3: Most Common IPs in the Night

SECTION 2: IP OWNER ANALYSIS

Goals of Analysis:

This section will be analyzing the owners of the IP addresses found in Section 1, and potentially find a background process associated with those IP addresses.

Results:

MORNING CAPTURES		
IP Addresses:	**Occurrences:**	**Potential Owners:**
192.168.1.0/24	760906	Local Gateway Router
204.79.197.0/24	1679	Microsoft Corporation (MSFT)
52.113.194.0/24	528	Microsoft Corporation (MSFT)
205.196.6.0/24	455	Valve Corp.
69.197.43.0/24	429	StackPath, LLC.
20.106.86.0/24	332	Microsoft Corporation (MSFT)
40.83.247.0/24	291	Microsoft Corporation (MSFT)
162.159.134.0/24	203	Cloudflare, Inc.
142.250.68.0/24	184	Google LLC
20.125.63.0/24	156	Microsoft Corporation (MSFT)

Table 1: Most popular IPs and their owners for Morning Captures

DAY CAPTURES		
IP Addresses:	**Occurrences:**	**Potential Owners:**
192.168.1.0/24	858270	Local Gateway Router
72.21.81.0/24	20137	Edgecast Inc.
151.139.71.0/24	6086	StackPath, LLC
20.60.194.0/24	495	Microsoft Corporation (MSFT)
20.190.190.0/24	190	Microsoft Corporation (MSFT)
40.83.240.0/24	161	Microsoft Corporation (MSFT)
52.165.164.0/24	145	Microsoft Corporation (MSFT)
205.196.6.0/24	144	Valve Corp.
13.71.55.0/24	104	Microsoft Corporation (MSFT)

Table 2: Most popular IPs and their owners for Day Captures

NIGHT CAPTURES		
IP Addresses:	**Occurrences:**	**Potential Owners:**
192.168.1.0/24	2914413	Local Gateway Router
13.107.42.0/24	17959	Microsoft Corporation (MSFT)
34.104.35.0/24	5835	Google LLC
72.21.81.0/24	3049	Edgecast Inc.
151.139.71.0/24	979	StackPath LLC.
205.196.6.0/24	272	Valve Corp.
52.113.194.0/24	261	Microsoft Corporation (MSFT)
20.190.151.0/24	253	Microsoft Corporation (MSFT)
40.83.247.0/24	242	Microsoft Corporation (MSFT)
184.28.98.0/24	185	Akamai Technologies, Inc.
3.93.253.0/24	117	Amazon Technologies Inc.
142.250.188.0/24	104	Google LLC

Table 3: Most popular IPs and their owners for Night Captures

Observations and Analysis:

This section was able to confirm my observation from Section 1 about the most common Ip belonging to my local gateway router. Microsoft is the most common IP owner; because my laptop is a windows laptop, I will attribute these to low level applications that help maintain the security, personal data, and other windows specific processes. After researching: Akamai Technologies, Edgecast, StackPath, seem to be common data delivery services that are used by many applications. This can be the same for Amazon and Google, however given that these are well known companies with services that I use from them, these deserve their own distinction. I believe these may be the result of Discord and Google Chrome running in the background. For Valve Corp, I am confident this is present due to Steam, the gaming platform. The final observation was that the Night captures have more than the other tables, potentially confirming my observation from Section 1: the laptop has more background processes due to me using it throughout the day.

SECTION 3: PACKET PERCENTAGE ANALYSIS:

Goals of Analysis:

This section will be analyzing the pcapng files and find the percentage of TCP and UDP packets in the entire capture. We will also be analyzing the successful ACK percentage of the data captures. These results will be compared to a packet that was captured under normal laptop-usage conditions.

Results:

Capture Statistics:

	TCP %	UDP %	Successful ACK %
Morning Captures	1.33%	97.19%	96.95%
Day Captures	2%	96.31%	97.43%
Night Captures	2.47%	97.41%	98.15%
Normal Usage Captures	11.29%	3.27%	98.20%
Idle Usage Captures	1.93%	96.91%	97.51%
(Normal – Idle)	+9.36%	-93.64%	+0.69%

Table 4: Percentage of UDP, TCP, and successful Acknowledgements

Observations and Analysis:

The most glaring statistic is the extremes in the UDP percentage between normal captures and idle captures. This shows that during normal conditions, low level background processes that use UDP are not being used as much. These processes may be taking advantage of the fact that the laptop is idle to execute their processes. Something interesting I saw is the trend of packet loss throughout the day. My roommates go to sleep early and wake up early, thus this may be attributed to the success of the packet given the network traffic; however, I was expecting more packet loss from the normal data. There is also a trend in TCP % throughout the day, adding further support for my reasoning of: the laptop has more background processes due to me using it throughout the day. As expected, the normal data has more TCP protocols, as this is used more for standard applications and such. My last observation was how much more UDP protocols were present compared to TCP during idle conditions.

[0.5 pts] Workload:

I was solely responsible for all work done on this project.

[0.5 pts] Tools that you used:

- Wireshark: Data Capture
- Python: writing scripts
- Scapy Python Library: Packet analysis
- Excel: Simple sorting and computations
- www.whois.com: Finding the owners of the IP addresses

[0.5 pts] Challenges:

One of the biggest challenges I faced was having to wait for the scripts to be completed. I am not sure if it was a performance issue on my computer, or the data size was truly too large, but checking my work was difficult as running the code took minutes. Also, when it came to the analysis of my data, I had to trust that the data was captured correctly, especially since I was a novice with Wireshark when the data was collected.

[0.5 pts] Future directions:

Some potential changes I would capture more data and ensure it was collected correctly. I would also capture more normal usage packets to get a fairer approximation. I would also flesh out the idle vs. non-idle analysis more. I also would change the intervals to 10 minutes, but much more captures, so my computer's performance isn't affected in analyzing the data.