CZ3006/CE3005: Netcentric/Computer Networks

Tang Zhengtian

Student Name : \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

TS4

Group : \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

21/10/2020

Date : \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# LAB 3: ANALZING NETWORK DATA LOG

You will be provided with the data file, in .csv format, in the working directory. Write the program to extract the following informations.

# EXERCISE 3A: TOP TALKERS AND LISTENERS

One of the most commonly used function in analyzing data log is finding out the IP address of the hosts that send out large amount of packet and hosts that receive large number of packets, usually know as TOP TALKERS and LISTENERS. Based on the IP address we can obtained the organization who owns the IP address.

List the TOP 5 TALKERS

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | IP address | # of packets | Organisation |
| 1 | 103.26.47.233 | 9646 | Malaysian Research & Education Network |
| 2 | 13.107.4.50 | 4950 | Microsoft Corporation |
| 3 | 155.69.160.78 | 4563 | Nanyang Technological University |
| 4 | 130.14.250.7 | 3914 | National Library of Medicine |
| 5 | 173.194.22.215 | 2896 | Google LLC |

TOP 5 LISTENERS

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | IP address | # of packets | Organisation |
| 1 | 103.22.221.73 | 9646 | National Information Society Agency, KR |
| 2 | 137.132.228.33 | 7835 | NUS information Technology |
| 3 | 137.132.228.29 | 5964 | NUS information Technology |
| 4 | 137.132.228.42 | 4987 | NUS information Technology |
| 5 | 103.37.198.100 | 3915 | A-STAR |

# EXERCISE 3B: TRANSPORT PROTOCOL

Using the IP protocol type attribute, determine the percentage of TCP and UDP protocol

|  |  |  |  |
| --- | --- | --- | --- |
|  | Header value | Transport layer protocol | # of packets |
| 1 | 6 | TCP | 155799 (76.83%) |
| 2 | 17 | UDP | 45377 (22.38%) |
| 3 | 47 | GRE | 891 (0.44%) |
| 4 | 50 | ESP | 643 (0.32%) |
| 5 | 1 | ICMP | 62 |
| 6 | 58 | IPv6-ICMP | 7 |
| 7 | 41 | IPv6 | 3 |
| 8 | 0 | HOPOPT | 3 |
| 9 | 2 | IGMP | 2 |
| 10 | 89 | OSPF | 1 |

# EXERCISE 3C: APPLICATIONS PROTOCOL

Using the Destination IP port number determine the most frequently used application protocol.

Page 3-1

CE3005: Computer Networks Laboratory 3

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | Destination IP port number | # of packets | Service |
| 1 | 443 | 42975 | HTTPS |
| 2 | 80 | 11960 | HTTP |
| 3 | 56800 | 3918 | Dynamic port/Ephemeral port |
| 4 | 15000 | 2697 | Registered ports |
| 5 | 44678 | 1158 | Registered ports |

# EXERCISE 3D: TRAFFIC INTENSITY

The traffic intensity is an important parameter that a network engineer needs to monitor closely to determine if there is congestion. You would use the IP packet size to calculate the estimated total traffic over the monitored period of 15 seconds. (Assume the sampling rate is 1 in 1000)

|  |  |
| --- | --- |
| Total Traffic( MB) | 189937.23MB |

# EXERCISE 3E: ADDITIONAL ANALYSIS (BONUS MARKS)

Please described additional analysis of the data and how it is useful. Please use a separate sheet to submit your new graphs and observations. Your report for this exercise is limited to 2 pages. The answer template and the two page additional analysis are to be submitted to your e-learning drive.

|  |
| --- |
| Examples   * Visulisation using scatter graph of port and IP address to determine if a specific node been port scanned by another node. * Which is the most popular node that provide service on port 80, port= 443 ?     You must analise and explain the graphs. Please do not be limited by the above examples. |

# EXERCISE 3F: SOFTWARE CODE

Please attach a softcopy of your code to the e-learning drive.

Page 3-2

**3E. Additional Analysis:**

Top communication pairs using IP addresses:

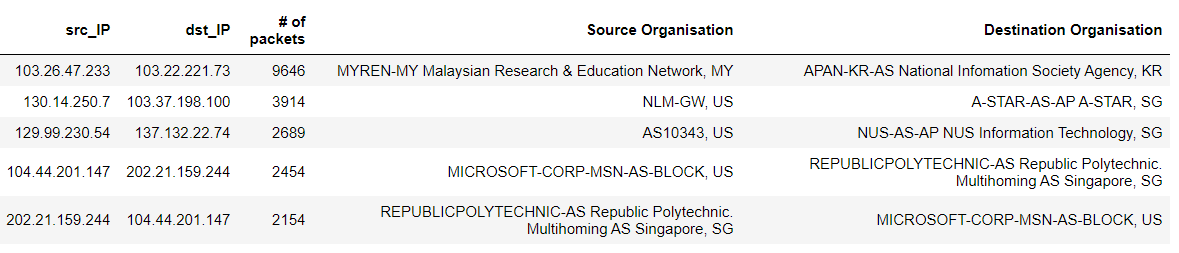


Fig.1 Top 5 communication pairs

From Fig.1 we can see the top 5 pairs of communication between organizations based on their IP addresses. This can tell us which organization is sending most number of packets to which other organization on the Network exchange point router. This can help us to identify organizations that are sending numerous number of packets of data and also possibly identify flooding of data packets sent by malicious organizations.

Top communication pair using IP addresses based on total packets size sent:

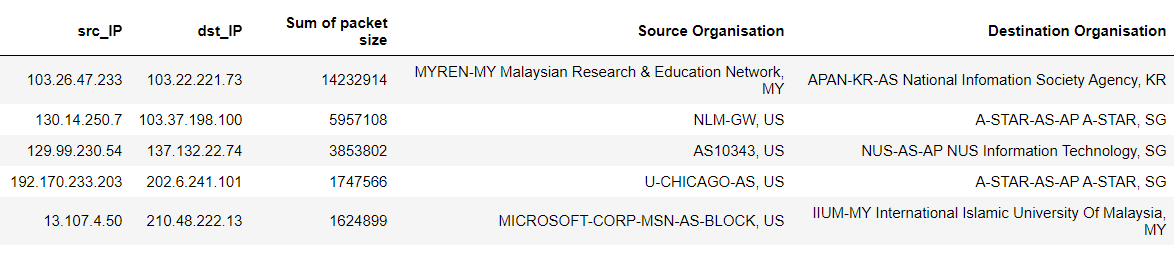


Fig.2 Top 5 communication pairs based on sum of packet size

From Fig.2 we can observe the total amount of data exchanged between organization. As the compared to just simply number of packets sent in Fig.1, Fig 2 can let us know which pair of organization is sending huge amount of data with each other. This can help to identify unusual load on the network.

Since both the top 5 communication pairs in terms of number of packets sent and total packet size sent are from legitimate organizations such as schools or research organization. We can assume that these communications are sent legitimately where organizations exchange research information to each other in huge amounts.

The top place in both Fig.1 and Fig.2 shows large and numerous amount of packets sent between Malaysian Research & Education Network with National Information Society Agency possibly due to their involvement in Asia@connect project implemented by TEIN\*CC to bridge digital divide across Asia-Pacific

Source: <https://apan.net/meetings/apan43/pres/2/20170215-Asi@Connect%20Keynote.pdf>

MAC address Analysis:

From the dataset itself there are a total of 96 unique MAC address nodes with 225 connections between them.

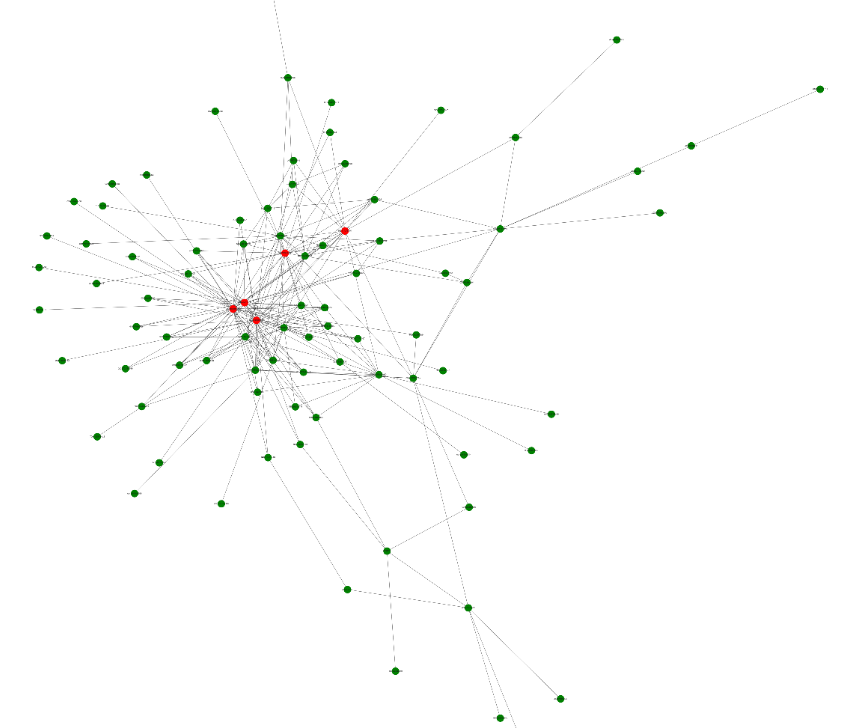
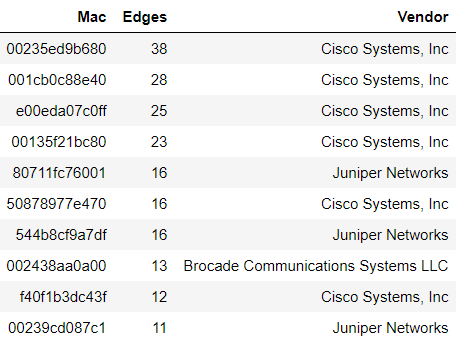


Fig.3 Top MAC address and its edges and their Vendor Fig.4 Graphical representation of MAC connection cluster

From Fig.3 we can see the MAC addresses/interfaces with descending number of connections based on Edges. The interfaces with the most connections are probably most important and should be kept available. The graph in Fig.4 also shows that majority of the interfaces(nodes) are connection closely to each other in a cluster with the top 5 interfaces being colored in red which represents about 30% of all connections.

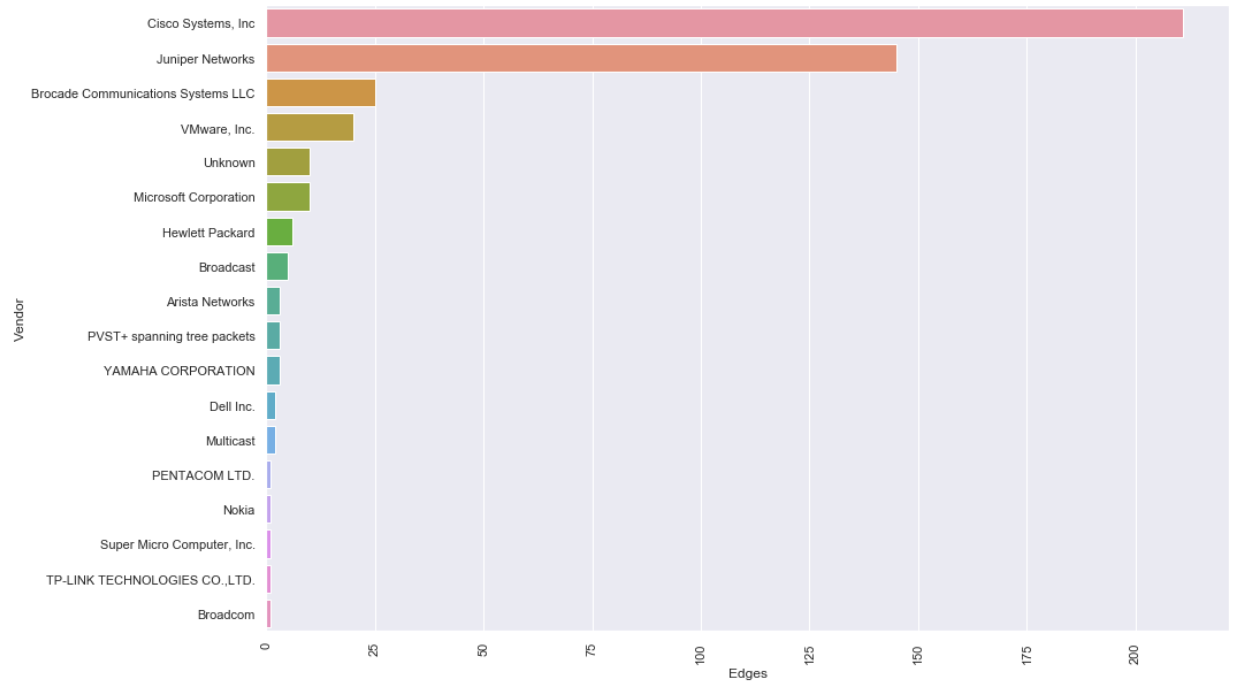


Fig.5 MAC vendor and their respective connections

From table in Fig.5 we can observe that most (50%) of connections on the network between interfaces involves router supplied by Cisco Systems Inc and 32% of network interfaces are from Juniper Network. This is in line with them being the top players in network technologies.