

Student Name : Hendy

Group : A33

Date : 21/03/2023

LAB 4: ANALZING NETWORK DATA LOG

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

EXERCISE 4A: 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	193.62.192.8	3041	European Bioniformatics Institute
2	155.69.160.32	2975	Nanyang Technological University
3	130.14.250.11	2604	National Library of Medicine
4	14.139.196.58	2452	Indian Institute of Technology
5	140.112.8.139	2056	National Taiwan University

TOP 5 LISTENERS

Rank	IP address	# of packets	Organisation
1	103.37.198.100	3841	A*STAR
2	137.132.228.15	3715	National University of Singapore
3	202.21.159.244	2446	Rpnet
4	192.101.107.153	2368	Pacific Northwest National Library
5	103.21.126.2	2056	Indian Institute of Technology Bombay

EXERCISE 4B: TRANSPORT PROTOCOL

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

	Header value	Transport layer protocol	# of packets (% of traffic)
1	6	TCP	56063 (80.819%)

2	17	UDP	9462 (13.640%)
---	----	-----	----------------

EXERCISE 4C: APPLICATIONS PROTOCOL

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

(For finding the service given the port number <https://www.adminsub.net/tcp-udp-port-finder/>)

Rank	Destination IP port number	# of packets	Service
1	443	13423	HTTPS/SSL
2	80	2647	HTTP
3	52866	2068	Dynamic Port
4	45512	1356	Unassigned
5	56152	1341	Dynamic Port

EXERCISE 4D: TRAFFIC

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)	7.722
--------------------	-------

EXERCISE 4E: ADDITIONAL ANALYSIS

Top 5 Communication Pairs (Sender->Receiver):

#	Source IP	Source Org	Destination IP	Destination Org	Num of Packets
1	193.62.192.8	European Bioinformatics Institute	137.132.228.15	National University of Singapore	3041
2	130.14.250.11	National Library of Medicine	103.37.198.100	A*STAR	2599
3	14.139.196.58	Indian Institute of Technology	192.101.107.153	Pacific Northwest National Library	2368
4	140.112.8.139	National Taiwan Univeristy	103.21.126.2	Indian Institute of Technology Bombay	2056
5	137.132.228.15	National University of Singapore	193.62.192.8	European Bioinformatics Institute	1910

Visualisation of Captured Network using igragh ([web hosted image for better visual clarity](#)):
Nodes are the hosts and edges are the connection, shows top 100 communication pairs

EXERCISE 4F: SOFTWARE CODE

Please also submit your code to the NTULearn lab site.