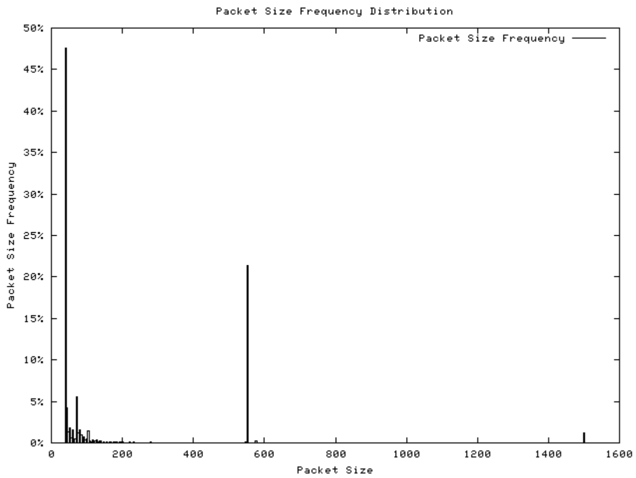
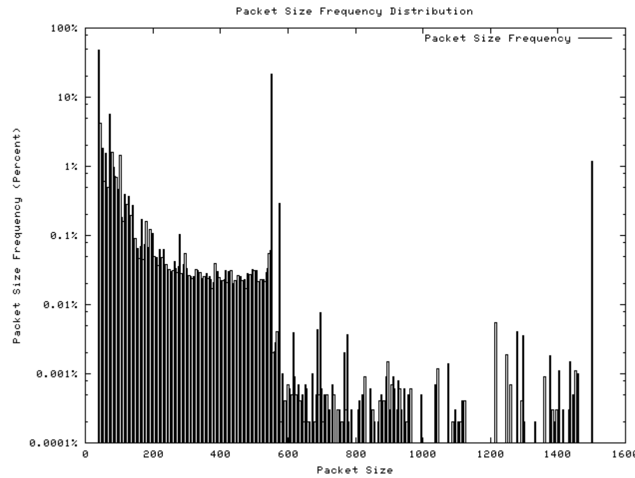
Wayne(Wenguang) Wang wew036@cs.usask.ca

CMPT 855 Assignment 1: TCP/IP Traffic Analysis

1. (3 marks) Plot a graph (histogram) of the TCP/IP packet size frequency distribution observed in this trace. Annotate your graph with a few short comments about the TCP/IP packet size distribution that you observe.



(a) Linear scale



(b) Logarithmic scale

Figure 1 Packet Size Frequency Distribution

Figure 1 shows that the size of the packets ranges from 40 bytes to 1500 bytes. Among all the packets, about 48% are 40 bytes, and about 21% are 552 bytes. The 40-byte IP packet consists of a 20-byte IP datagram header and a 20-byte TCP ACK packet (which is actually a TCP segment header). All the other packets (52% of all packet) besides the 40-byte ACK packets are data packets that carry useful data. Nearly half of them (21% to 52%) are 552 bytes, which carry about 512 bytes data per packet.

Nearly half of the packets (48%) are ACK, which means that for nearly every data packet, there is an ACK packet to acknowledge it. Therefore, although TCP connections are full-duplex, the traffic is asymmetric, i.e., data are transmitted mainly in one direction.

2. (3 marks) Plot a graph illustrating the distribution of TCP/IP packet traffic amongst the source hosts identified in the trace. Annotate your graph with some comments and observations about the distribution of TCP/IP packet traffic amongst hosts on this network. Suggest a hypothesis that might help explain the observation(s) that you made.

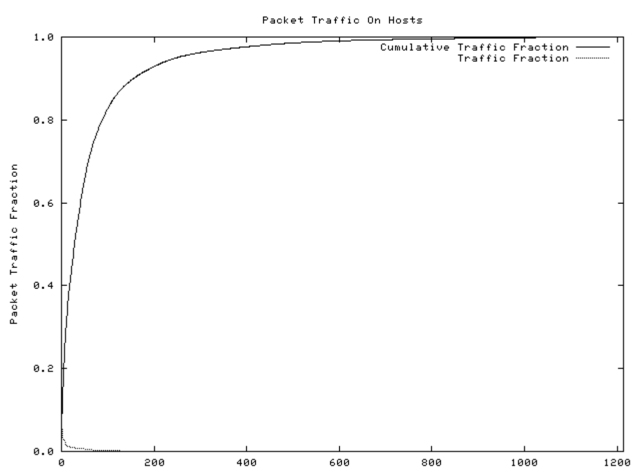


Figure 2 Packet Traffic On Hosts

In figure 2, there are 1214 different source hosts. Among them, about 2.4% send out 50% packets, and about 12.9% send out 90% packets. Therefore, the packet traffic is non-uniformly distributed amongst the hosts on the network. Some dominant hosts, which are rather popular sites, generate a majority of the packets on the network.

3. (3 marks) Repeat the analysis of (2) above, but on a byte count basis, rather than on a packet count basis. Annotate your graph with any additional comments and observations about the distribution of TCP/IP byte traffic amongst hosts on this network. Suggest a hypothesis that might help explain the observation(s) that you made. State whether your observations here are consistent with those from part (2).

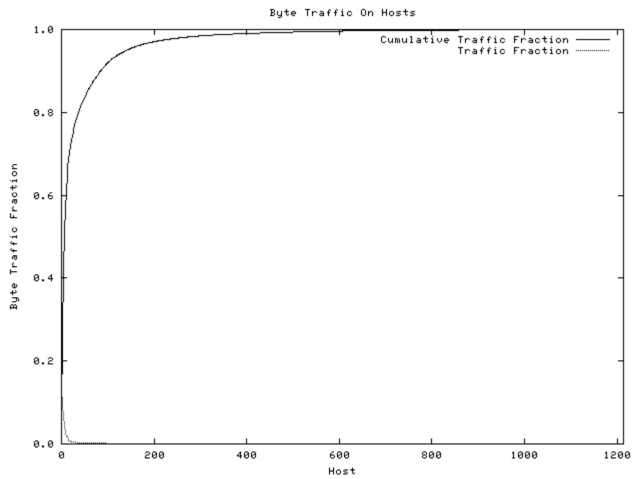
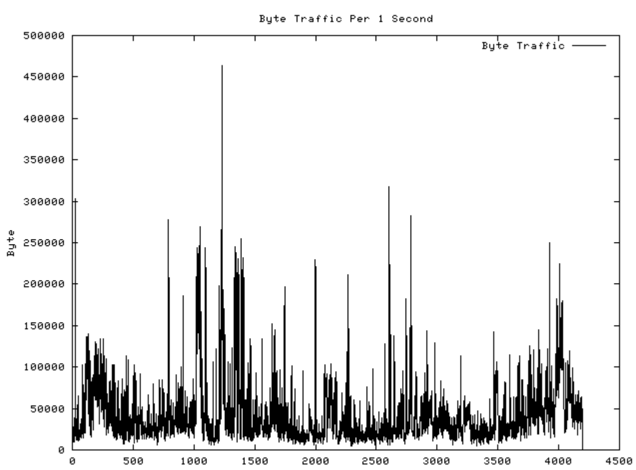


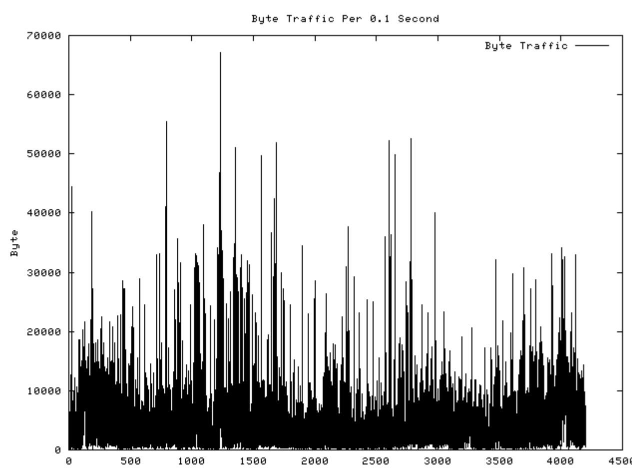
Figure 3 Byte Traffic Per Host

In figure 3, about 0.6% hosts send 50% data volume (in bytes) on the network, and about 7% hosts send 90% data volume (in bytes) on the network. Therefore, the byte traffic is non-uniformly distributed amongst the hosts. Some dominant hosts, which are rather hot sites, generate most of the data traffic. This result is consistent with that of the question 2.

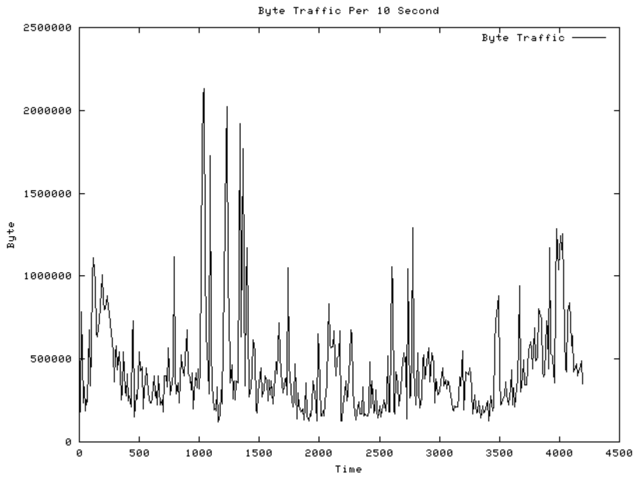
4. (3 marks) Define a time interval I = 1.0, 0.1, and 10 seconds, and process the trace to count the number of TCP/IP bytes transmitted on the network in each such non-overlapping consecutive time interval in the trace. Annotate one of your graphs with your comments and observations about the TCP/IP traffic profile.



(a) Time unit = 1 second



(b) Time unit = 0.1 second



(c) Time unit = 10 seconds

Figure 4 Byte traffic on different time scales

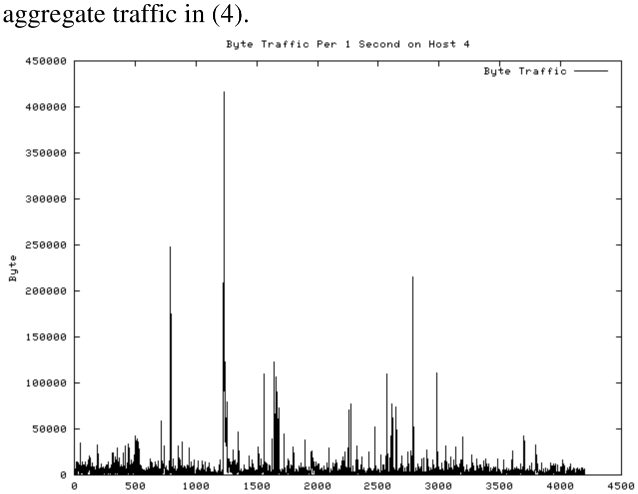
Figure 4 (a), (b), and (c) present the network byte traffic on time intervals 1.0, 0.1, and 10 seconds respectively. The average traffic, peak traffic and the peak to average ratio are listed in Table 1. Table 1 shows that the traffic is bursty across different time scales (0.1 – 10 seconds).

Time interval (second)	Average traffic (bytes)	Peak traffic (bytes)	Peak to Average ratio
0.1	4479	67098	14.98
1	44783	363789	8.12
10	447730	2129840	4.76

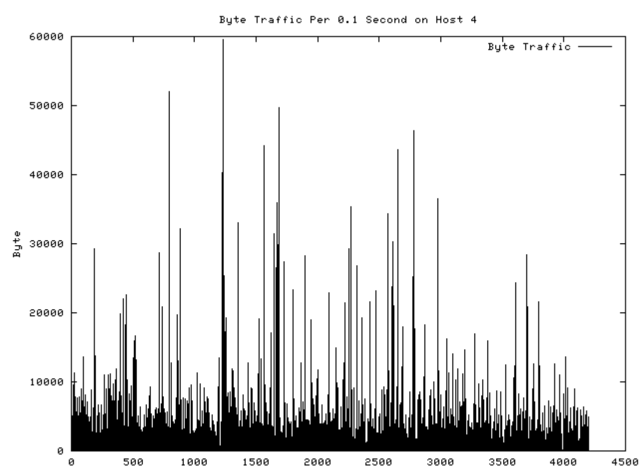
Table 1 Byte traffic on different time scales

Figure 4 (a), (b), and (c) are very similar regardless of the time scales (from 0.1 second to 10 second). This indicates that the traffic is self-similar. The buffer management of the self-similar traffic is quite difficult because no matter how big the buffer is, the buffer overflow would happen under the traffic burst on a longer time scale.

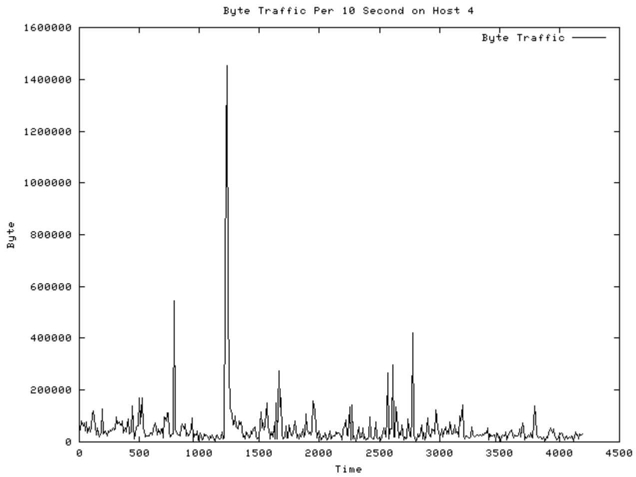
5. (3 marks) Repeat the analysis in (4) but using only the TCP/IP packet traffic generated by the most active host. Plot the resulting three traffic profile plots, for I = 10.0 seconds, I = 1.0 seconds, and I = 0.1 seconds. Annotate one of these three graphs with your comments and observations about the TCP/IP traffic profile for this host, and its relationship (if any) to the aggregate traffic in (4).



(a) Time unit = 1 second



(b) Time unit = 0.1 second



(c) Time unit = 10 seconds

Figure 5 Byte traffic on different time scales for host 4

Figure 5 (a), (b), and (c) show the network byte traffic for host 4 on time intervals 1.0, 0.1, and 10 seconds respectively. The average traffic, peak traffic and the peak to average ratio are listed in Table 2. Table 2 shows that the traffic of host 4 is bursty across different time scales (0.1 – 10 seconds).

Time interval (seconds)	Average traffic (bytes)	Peak traffic (bytes)	Peak to Average ratio
0.1	518	59616	115.09
1	5186	416779	80.37
10	51847	1451368	27.99

Table 2 Average traffic and peak traffic for host 4

Figure 5 (a), (b), and (c) are very similar regardless of the time scales (from 0.1 second to 10 seconds). This indicates that the traffic of host 4 is self-similar. Comparing Figure 4 with Figure 5, the burst time of host 4 is consistent with that of the whole network traffic. Therefore, the active hosts in a network can affect the traffic characteristics of the whole network.