Packet Analysis

Sukriti Gupta(2016CS50084) Chinmay Rai(2016CS50615)

1 FTP Codes and Response Codes

We observed the following FTP response and command codes:

1.1 Commands

- USER : Authentication username.
- PASS : Authentication password.
- OPTS : Select options for a feature
- SYST : Return system type.
- PWD : Print working directory. Returns the current directory of the host.
- NOOP : No operation (dummy packet; used mostly on keepalives).
- CWD : Change working directory.
- LIST : Returns information of a file or directory.
- PORT : Specifies an address and port to which the server should connect.
- QUIT : Disconnect.
- FEAT : Get the feature list implemented by the server.
- TYPE : Sets the transfer mode (ASCII/Binary).
- PASV : Enter passive mode.
- SIZE : Return the size of a file.
- RETR : Retrieve a copy of the file
- STAT : Returns information on the server status, including the status of the current connection
- REST : Restart transfer from the specified point.

- MDTM : Return the last-modified time of a specified file.
- ABOR : Abort an active file transfer.
- ALLO : Allocate sufficient disk space to receive a file.
- CDUP : Change to Parent Directory.
- EPRT : Specifies an extended address and port to which the server should connect.
- EPSV : Enter extended passive mode.
- HELP : Returns usage documentation on a command if specified, else a general help document is returned.
- SITE : Sends site specific commands to remote server
- MODE B,S : Sets the transfer mode (Stream, Block, or Compressed).
- NLST : Returns a list of file names in a specified directory.
- STOR : Accept the data and to store the data as a file at the server site
- STRU : Store file uniquely.

1.2 Responses

- 125 : Data connection already open; transfer starting.
- 150 : File status okay; about to open data connection
- 200 : The requested action has been successfully completed.
- 202 : Command not implemented, superfluous at this site.
- 211 : System status, or system help reply.
- 213 : File status.
- 214 : Help message.
- 215 : NAME system type. Where NAME is an official system name from the registry kept by IANA.
- 220 : Service ready for new user.
- 221 : Service closing control connection.
- 225 : Data connection open; no transfer in progress.
- 226 : Closing data connection. Requested file action successful

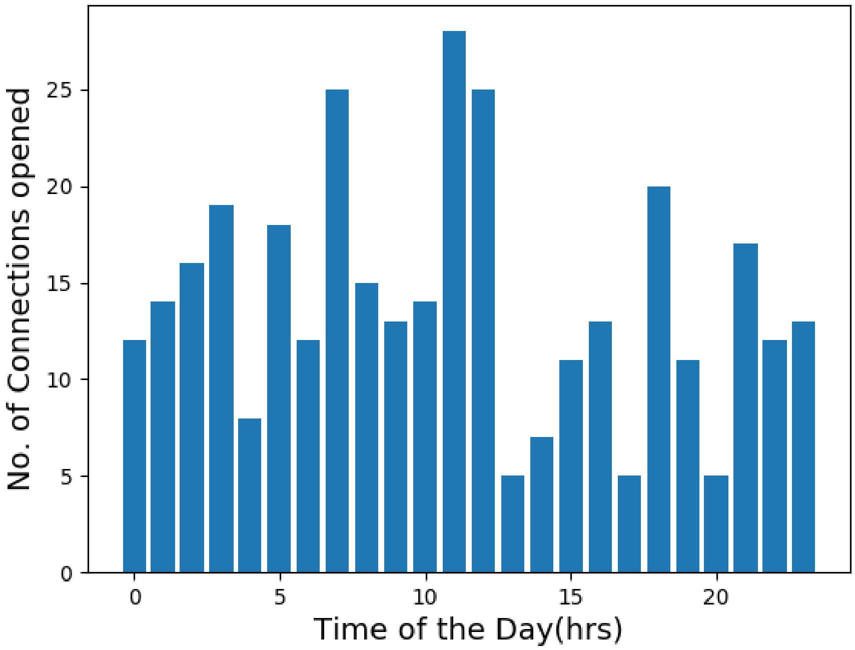
- 227 : Entering Passive Mode
- 230 : User logged in, proceed. Logged out if appropriate.
- 250 : Requested file action okay, completed.
- 257 : ”PATHNAME” created.
- 331 : User name okay, need password.
- 350 : Requested file action pending further information
- 400 : The command was not accepted and the requested action did not take place, but the error condition is temporary and the action may be requested again.
- 421 : Service not available, closing control connection.
- 425 : Can’t open data connection.
- 426 : Connection closed; transfer aborted.
- 451 : Requested action aborted. Local error in processing.
- 500 : Syntax error, command unrecognized and the requested action did not take place.
- 501 : Syntax error in parameters or arguments.
- 502 : Command not implemented.
- 504 : Command not implemented for that parameter.
- 530 : Not logged in.
- 550 : Requested action not taken. File unavailable
- 553 : Requested action not taken. File name not allowed.

2 Packet Trace Analysis

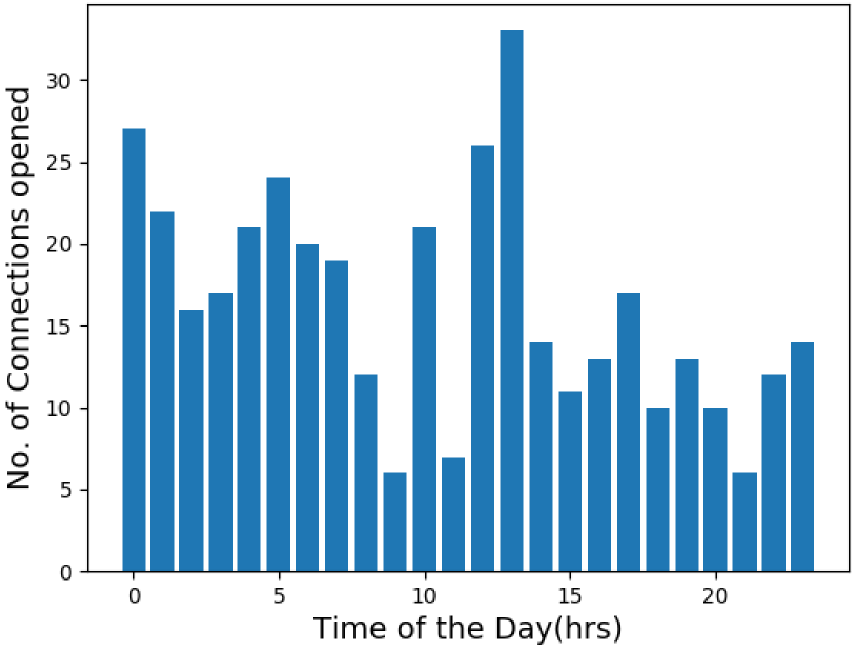
Interpretation: Counting has been done by SYN packets. May change slightly if [SYN,ACK] or other are used as incomplete connections exist. Uniqueness of TCP flow is defined as uniqueness of entire 4 tuple

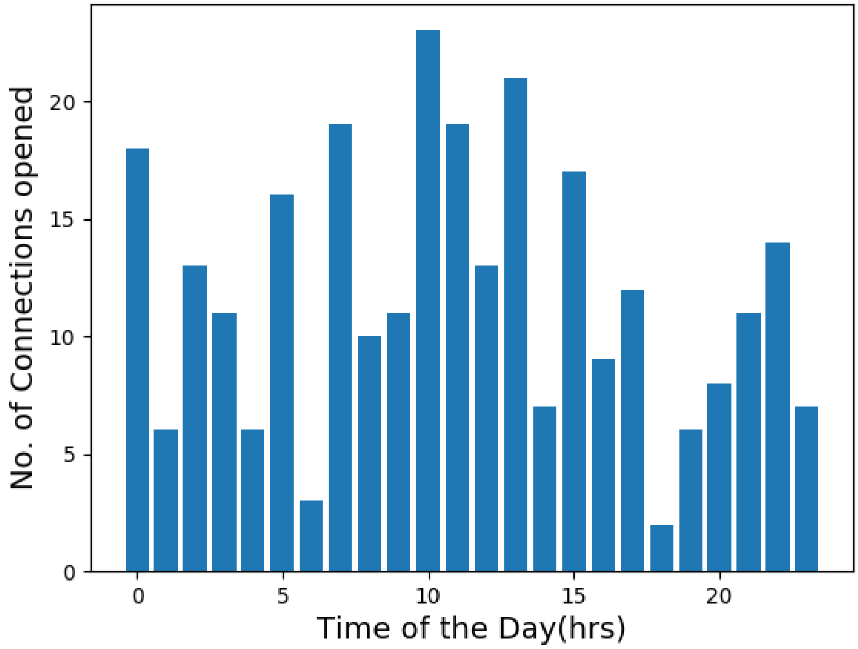
Question :	PART 1		PART 2
File1	#Servers: 45	#Clients: 526	#Flows: 3256
File2	#Servers: 50	#Clients: 945	#Flows: 5422
File3	#Servers: 89	#Clients: 519	#Flows: 3280

2.1 Daily Profile



Done for server:131.243.2.12

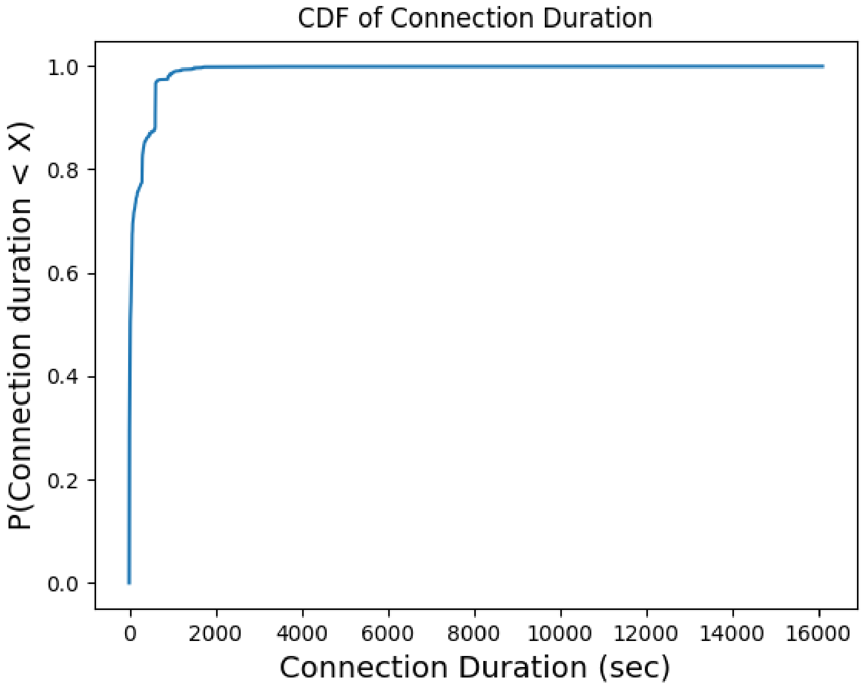




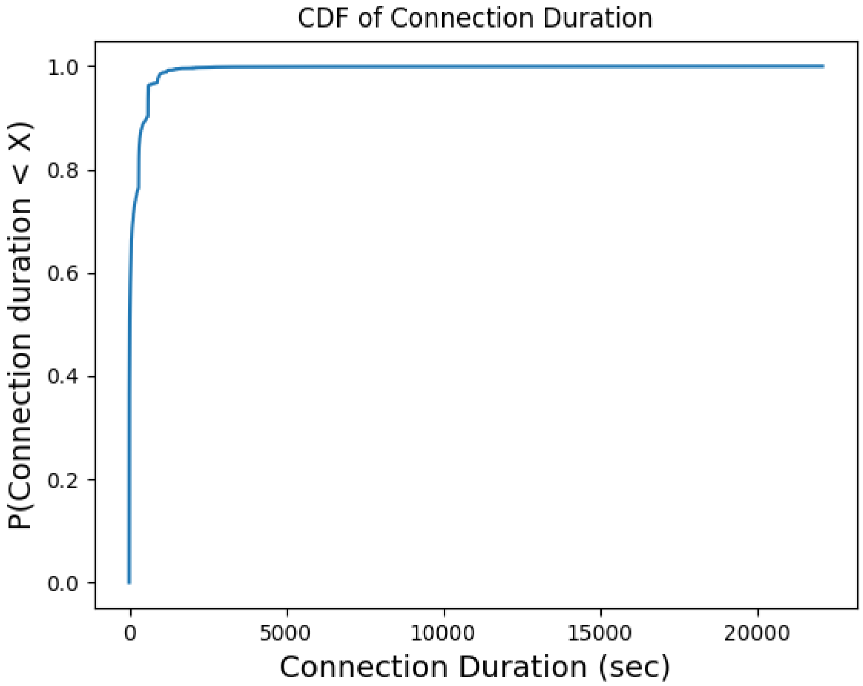
Anomally detection on this profile. DoS attack is seen as a very sharp peak in the Profile.

2.2 Connection Duration

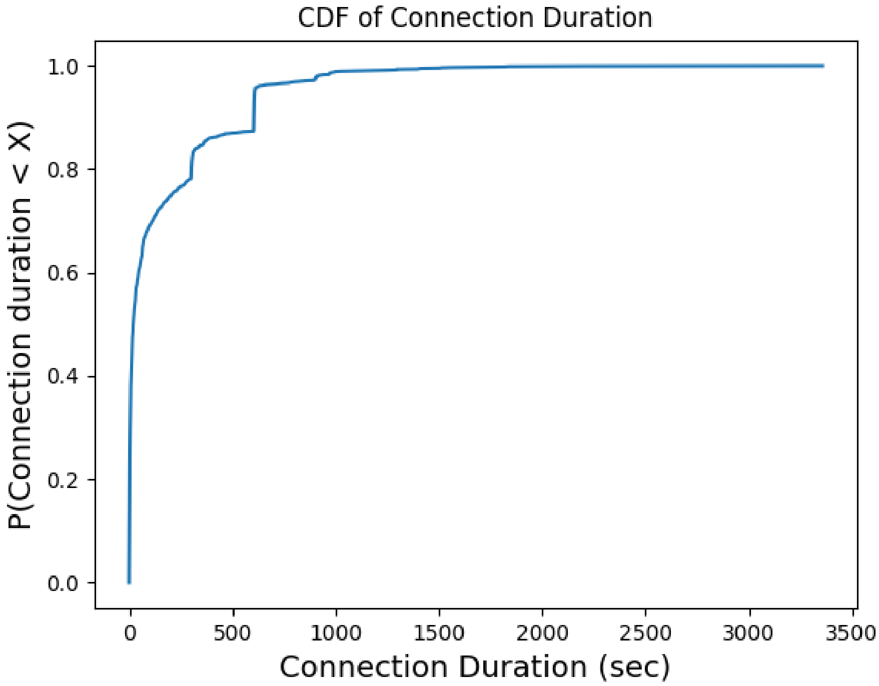
Mean = 160.02790 sec Median = 24.12750 sec



Mean = 166.68434 sec Median = 18.58785 sec



Mean = 153.98576 sec Median = 19.50087 sec



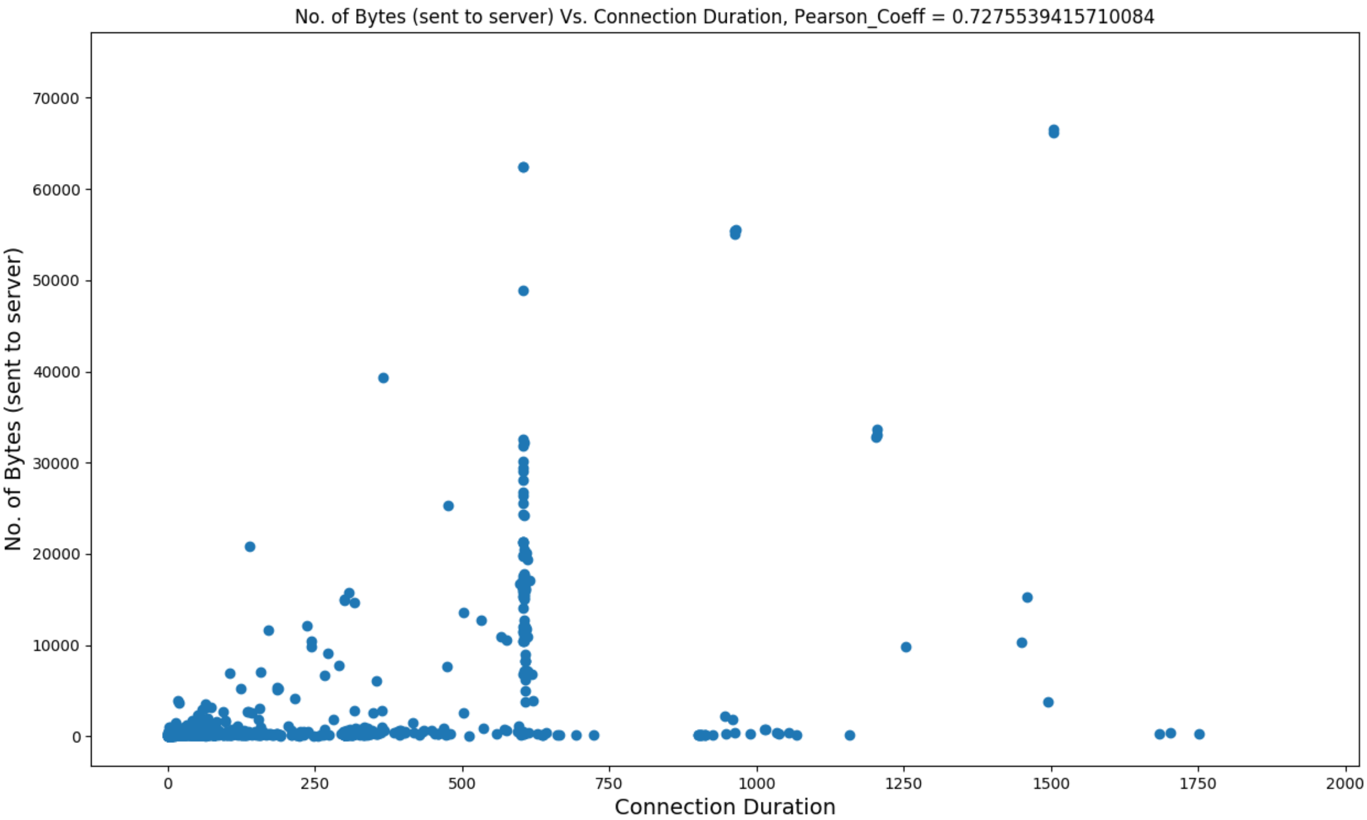
2.3 Data transfer Vs. Connection Duration

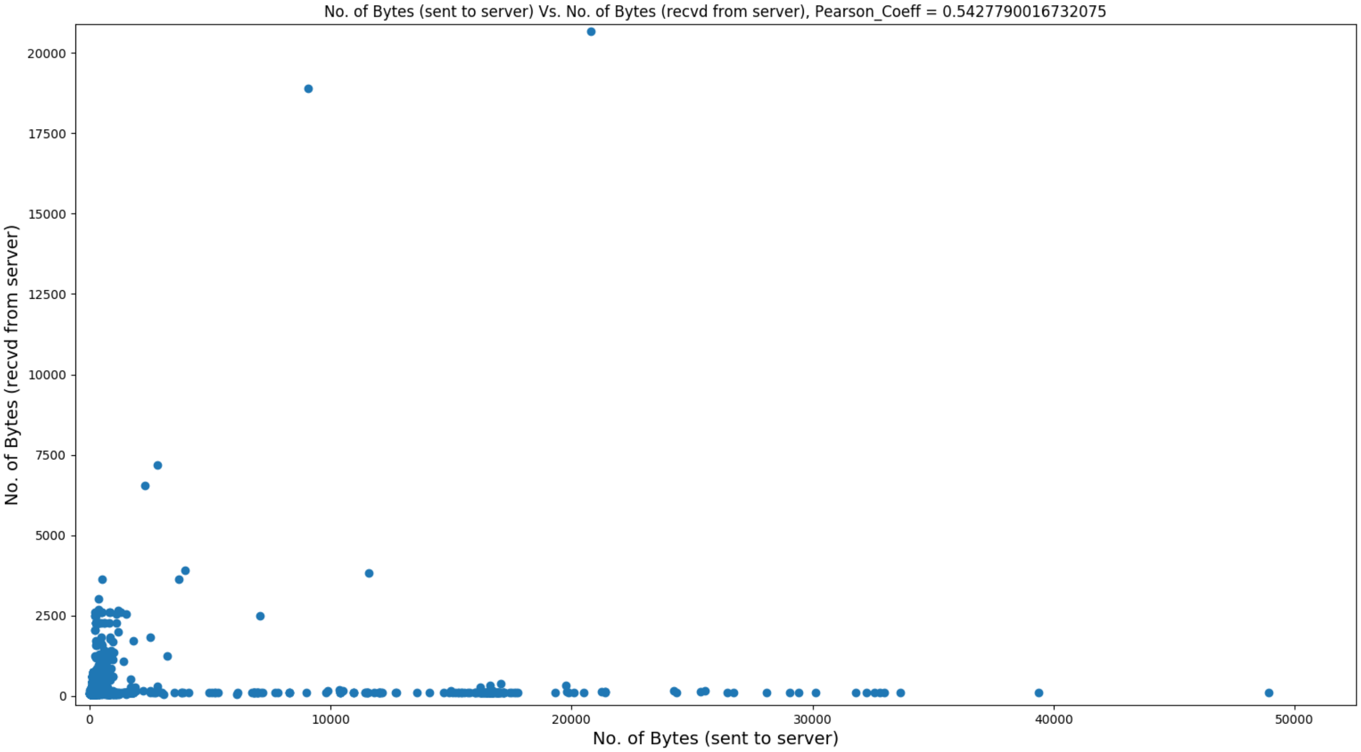
We have tried to correlate the :

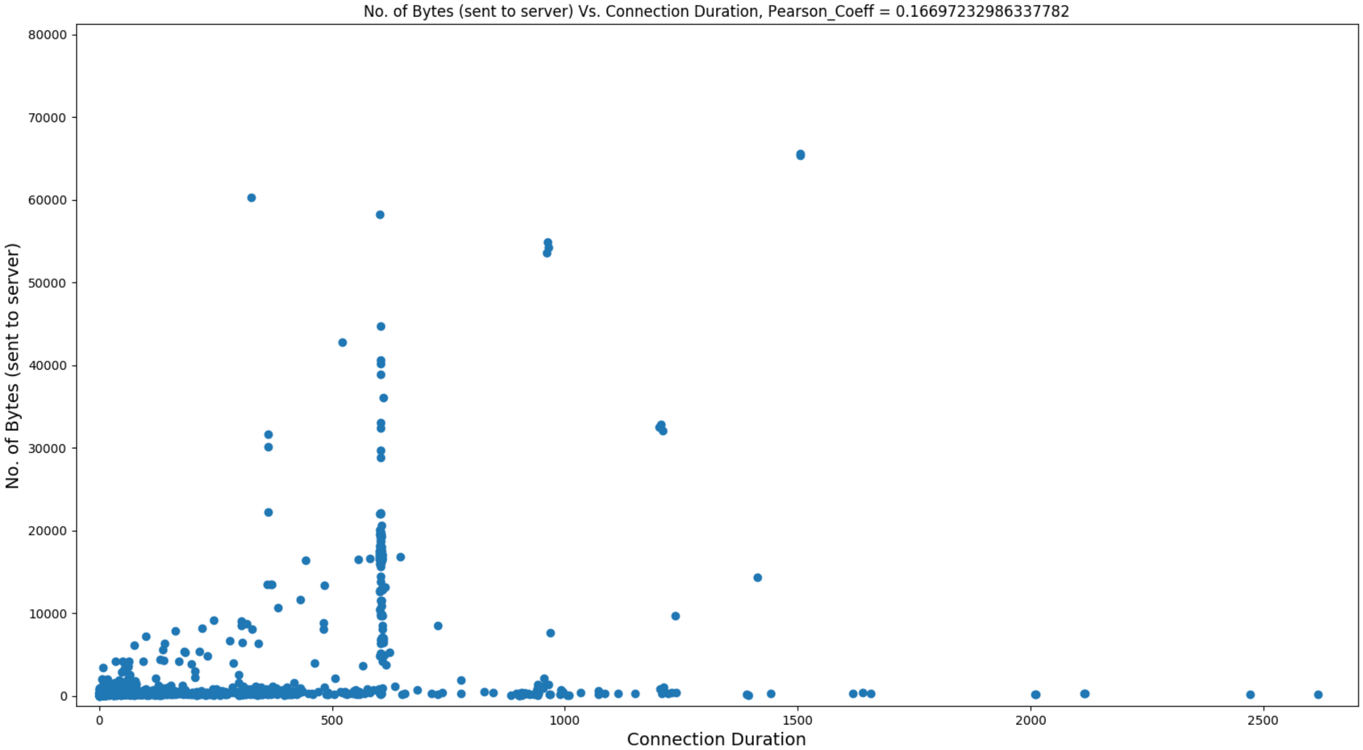
- No. of Bytes(Sent to server) Vs. Connection Duration

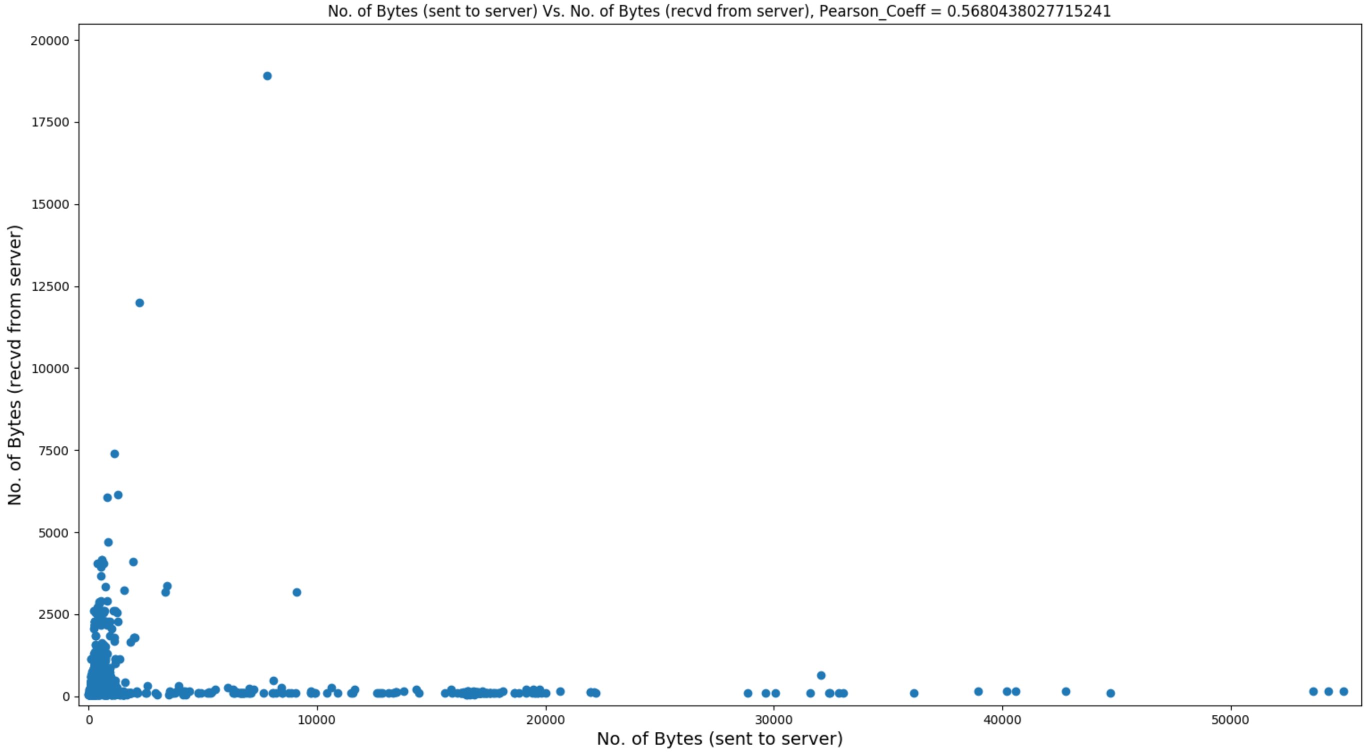
- No. of Bytes(Sent to server) Vs. No. of Bytes(Received from server)

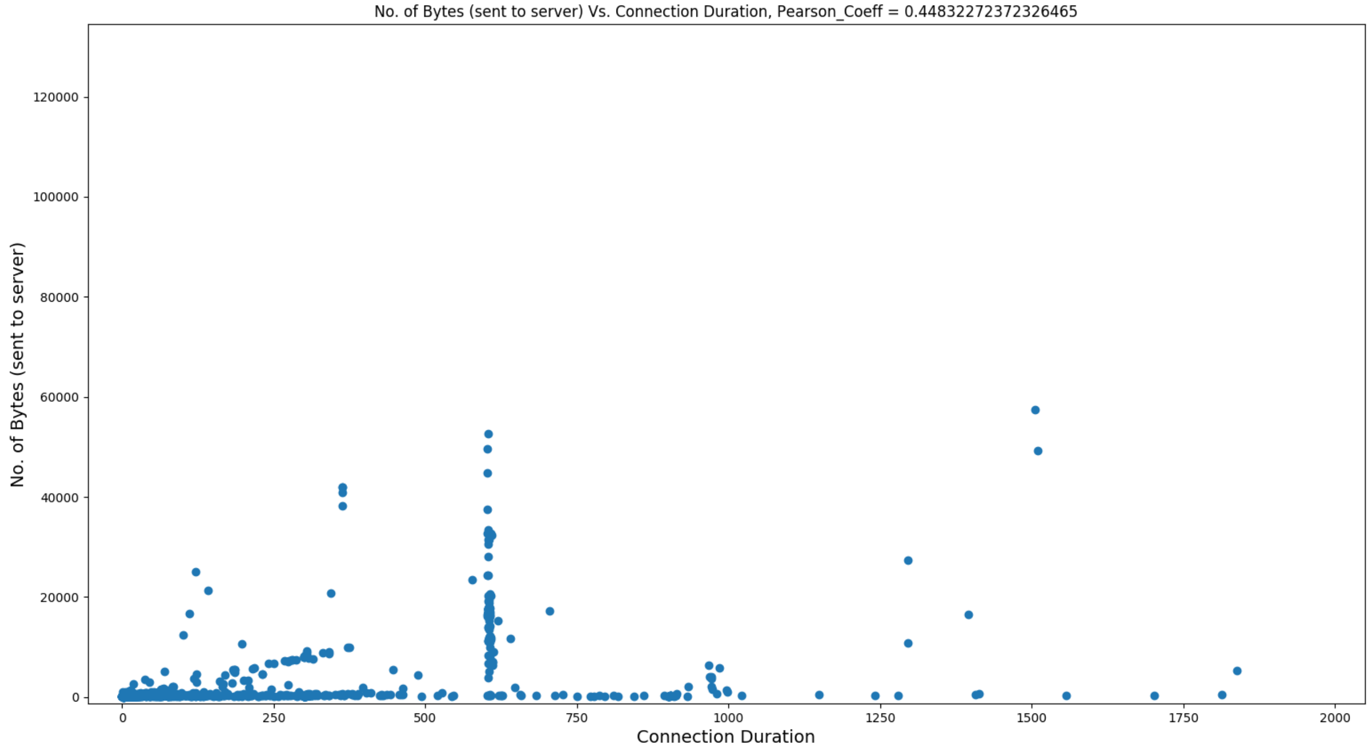
for the 3 days of data that we were provided. There does not seem to be too much of a correlation, though it is not too weak either. The results(in order of date on file) are:

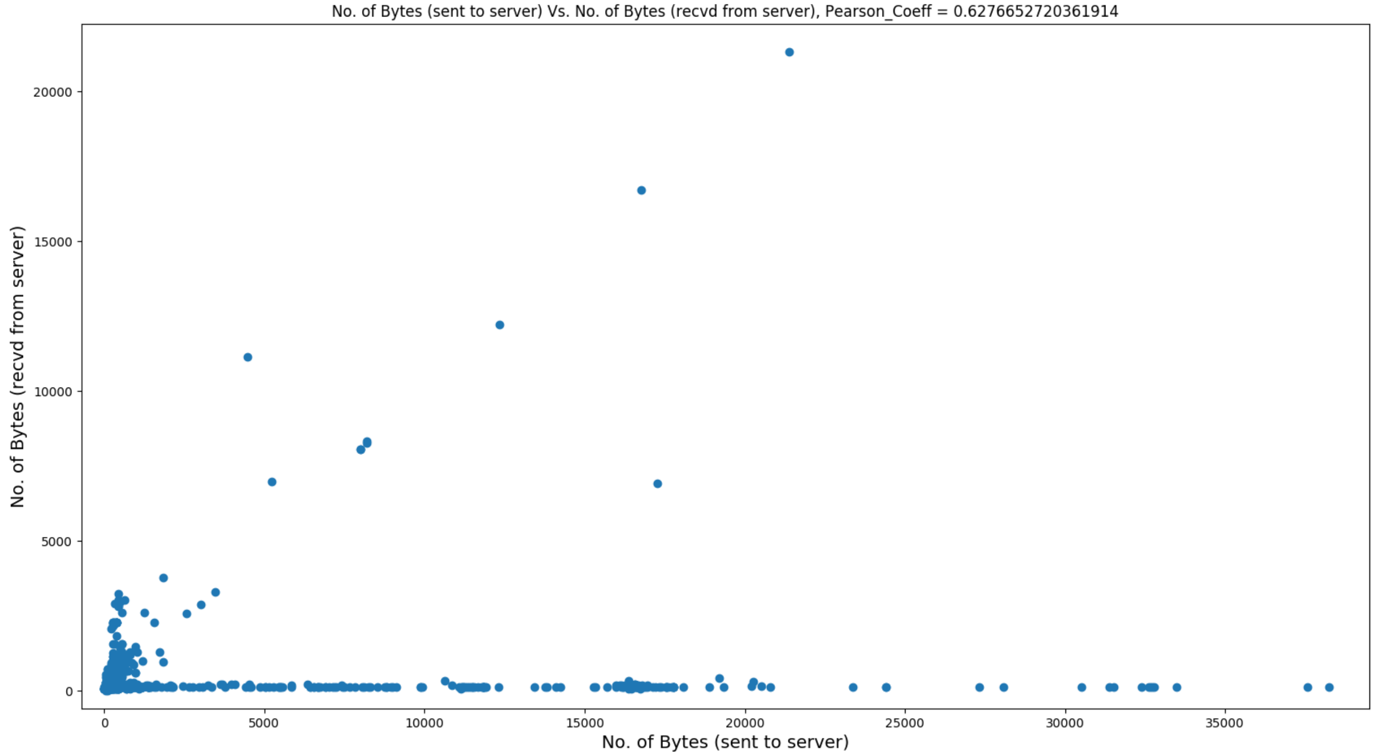

No. of Bytes (sent to server) Vs. Connection Duration, Pearson_Coeff = 0.7275539415710084
No. of Bytes (sent to server)
70000
60000
50000
40000
30000
20000
10000
0
0
250
500
750
1000
1250
1500
1750
2000
Connection Duration


No. of Bytes (sent to server) Vs. No. of Bytes (recvd from server), Pearson_Coeff = 0.5427790016732075
No. of Bytes (recvd from server)
20000
17500
15000
12500
10000
7500
5000
2500
0
0
10000
20000
30000
40000
50000
No. of Bytes (sent to server)

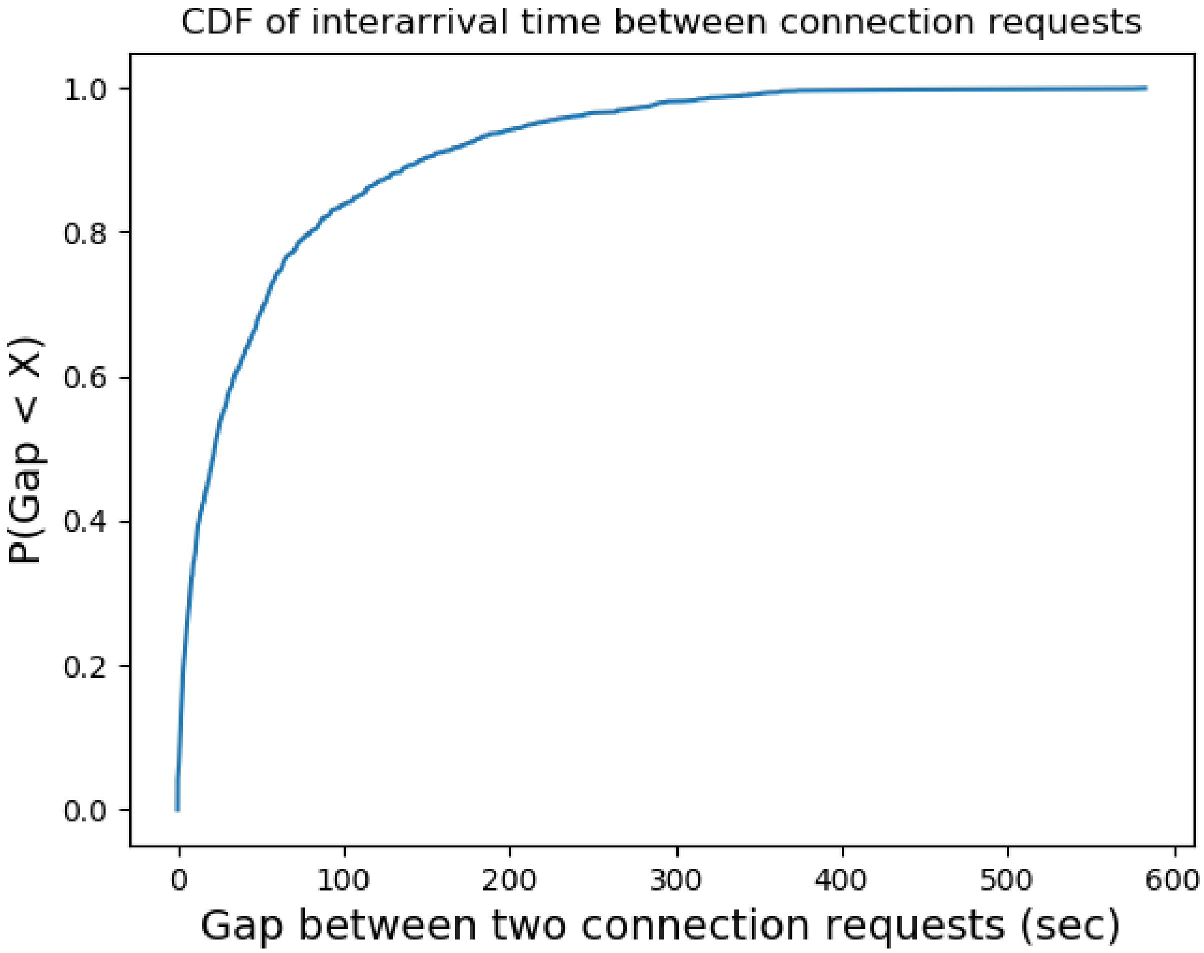

No. of Bytes (sent to server) Vs. Connection Duration, Pearson_Coeff = 0.16697232986337782


No. of Bytes (sent to server) Vs. No. of Bytes (recvd from server), Pearson_Coeff = 0.5680438027715241

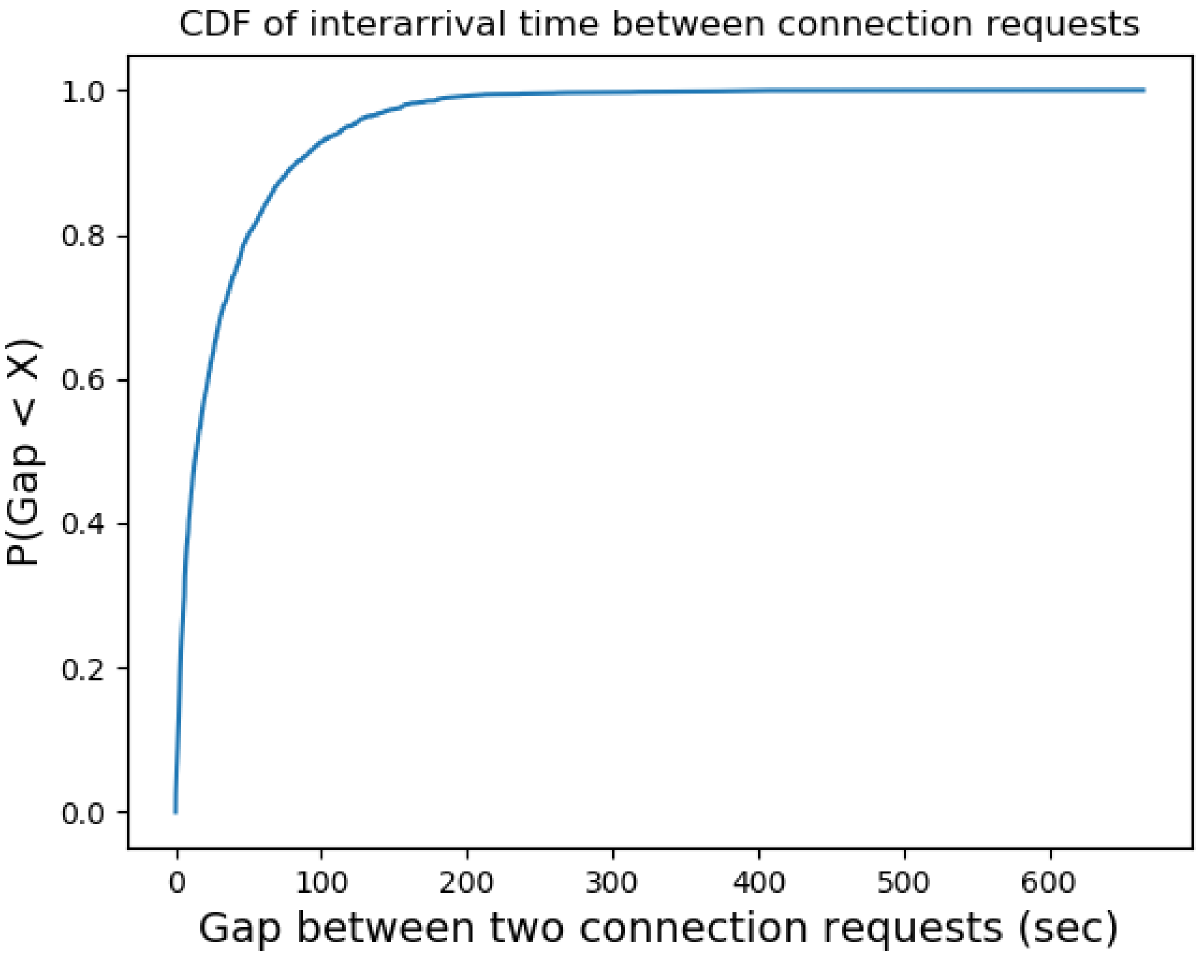

No. of Bytes (sent to server) Vs. Connection Duration, Pearson_Coeff = 0.44832272372326465
No. of Bytes (sent to server)
120000
100000
80000
60000
40000
20000
0
0
250
500
750
1000
1250
1500
1750
2000
Connection Duration


No. of Bytes (sent to server) Vs. No. of Bytes (recvd from server), Pearson_Coeff = 0.6276652720361914
No. of Bytes (recvd from server)
20000
15000
10000
5000
0
0
5000
10000
15000
20000
25000
30000
35000
No. of Bytes (sent to server)

3 Question 6

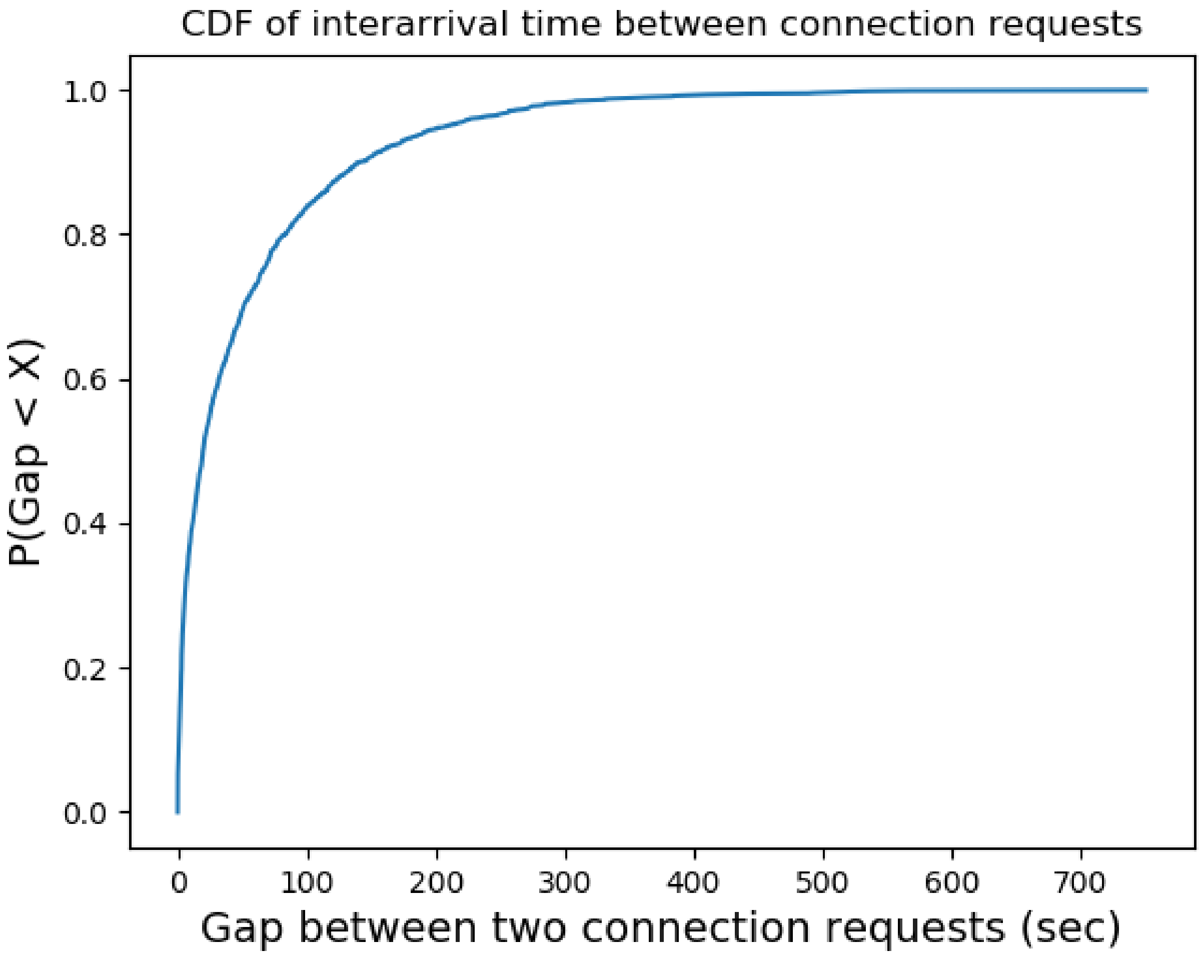


Day 1
Mean of Interarrival Time: 51.6563137835 Median of Interarrival Time: 22.2736705



Day 2

Mean of Interarrival Time: 31.3369680945 Median of Interarrival Time: 14.351235

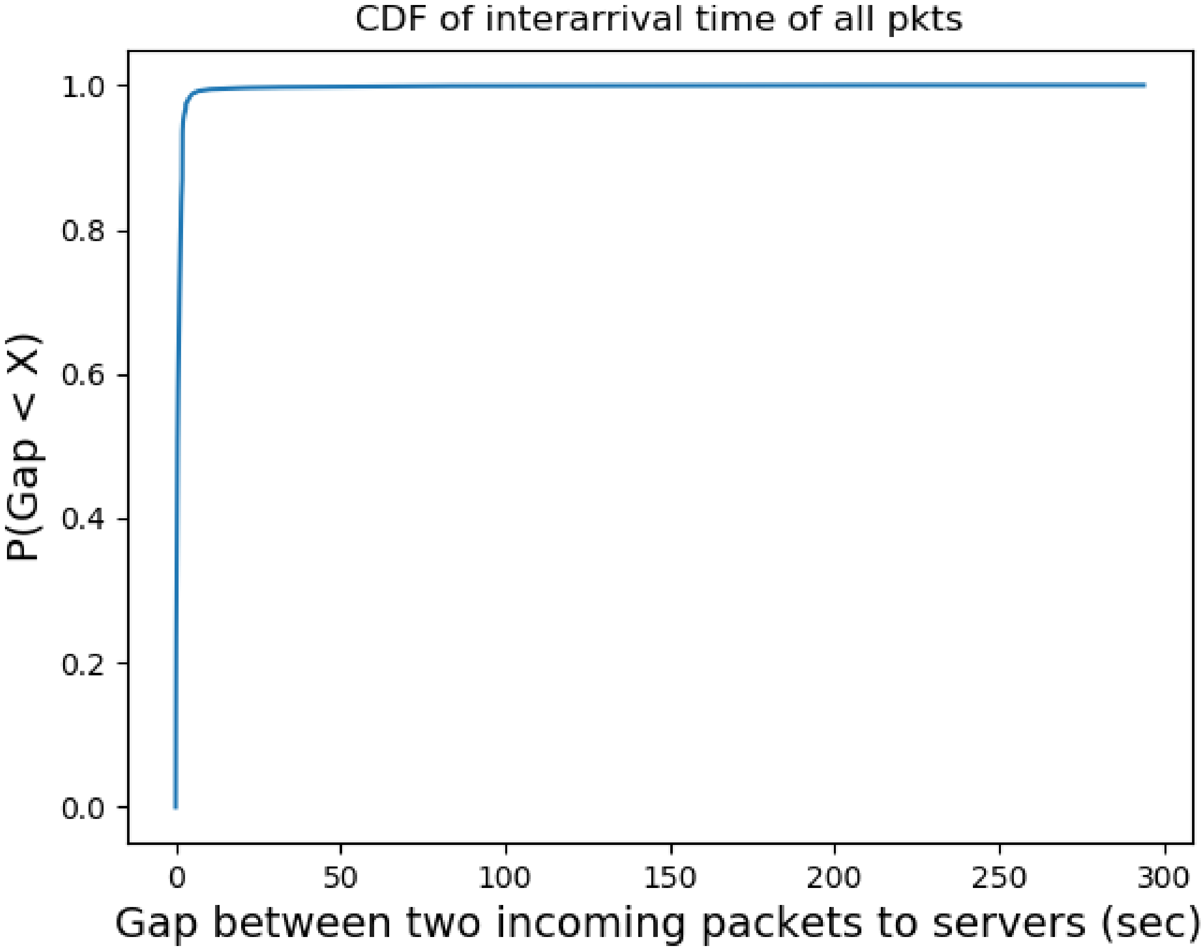


Day 3

Mean of Interarrival Time: 50.9182253282 Median of Interarrival Time: 19.582489

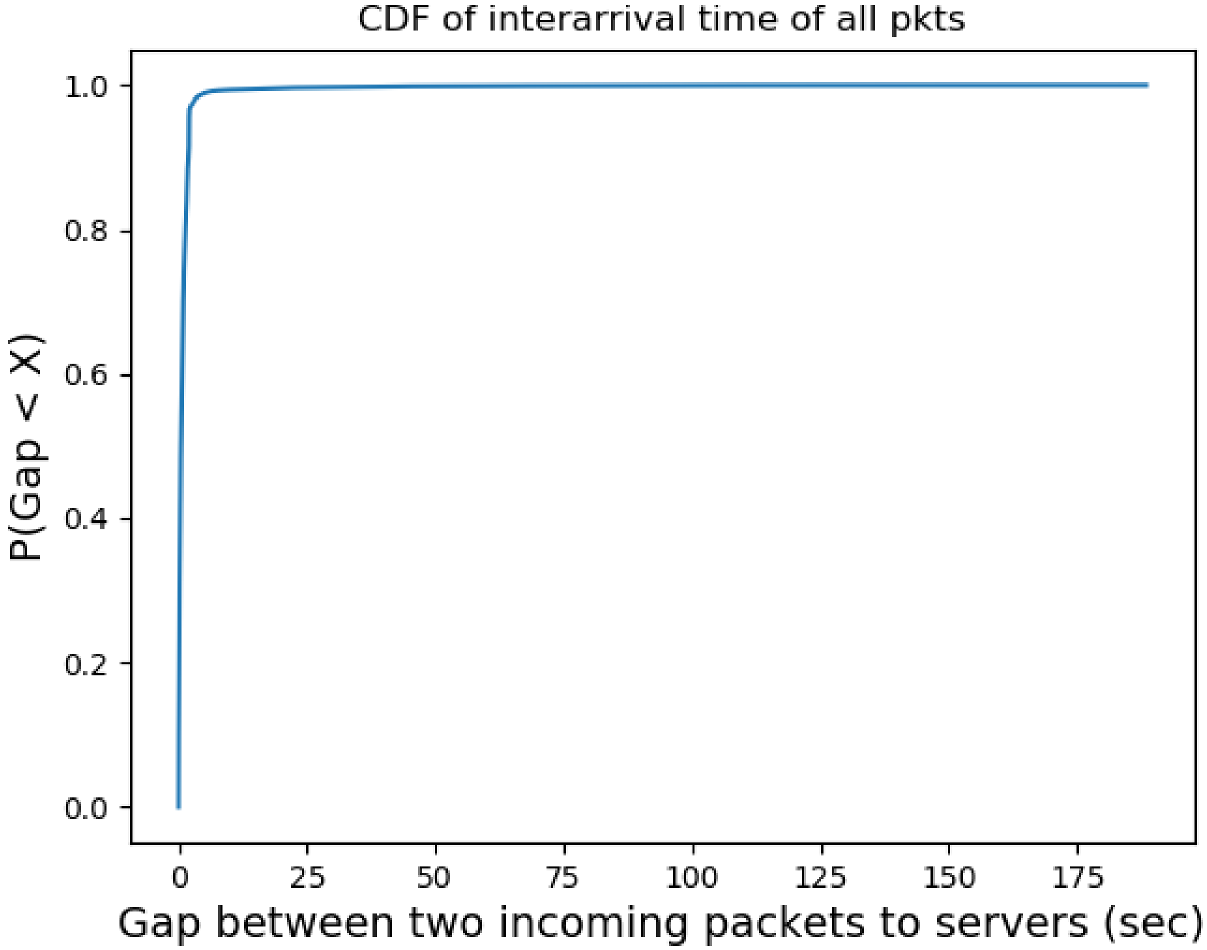
4 Question 7

Because this is a memoriless system. Arrival of one connection request is independent of the other and most likely follows exponential distribution. In general, servers would be receiving requests without too much gap but some times there would be large gaps also. This would not happen a lot though. Such a distribution leads to comparatively larger means but very small medians.



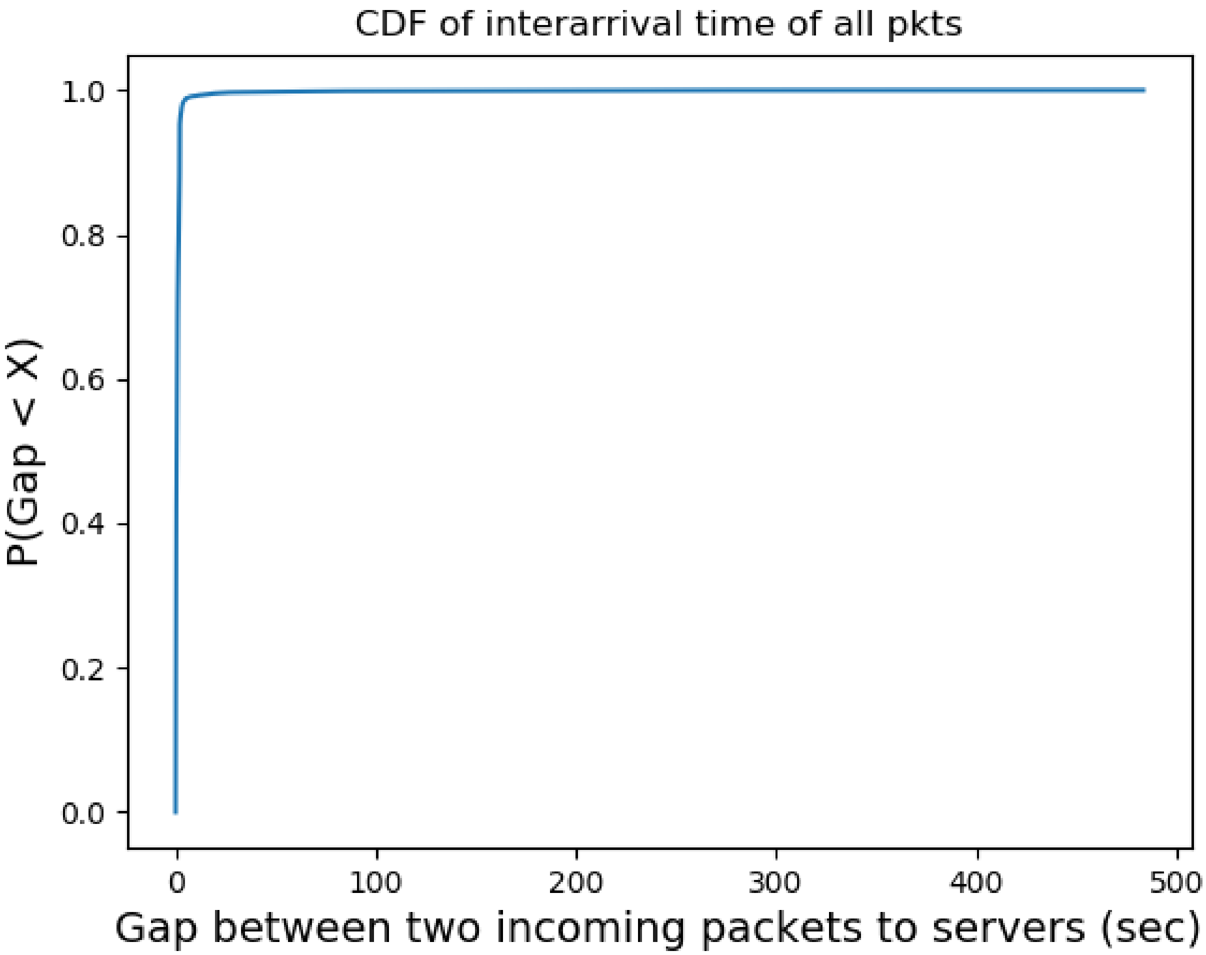
Day 1

Mean of Interarrival Time incoming packets to server: 1.1328663752065466 Median of Interarrival Time incoming packets to server: 0.5502000000033149



Day2

Mean of Interarrival Time incoming packets to server: 0.9471976691783004 Median of Interarrival Time incoming packets to server: 0.45764300000155345



Day 3

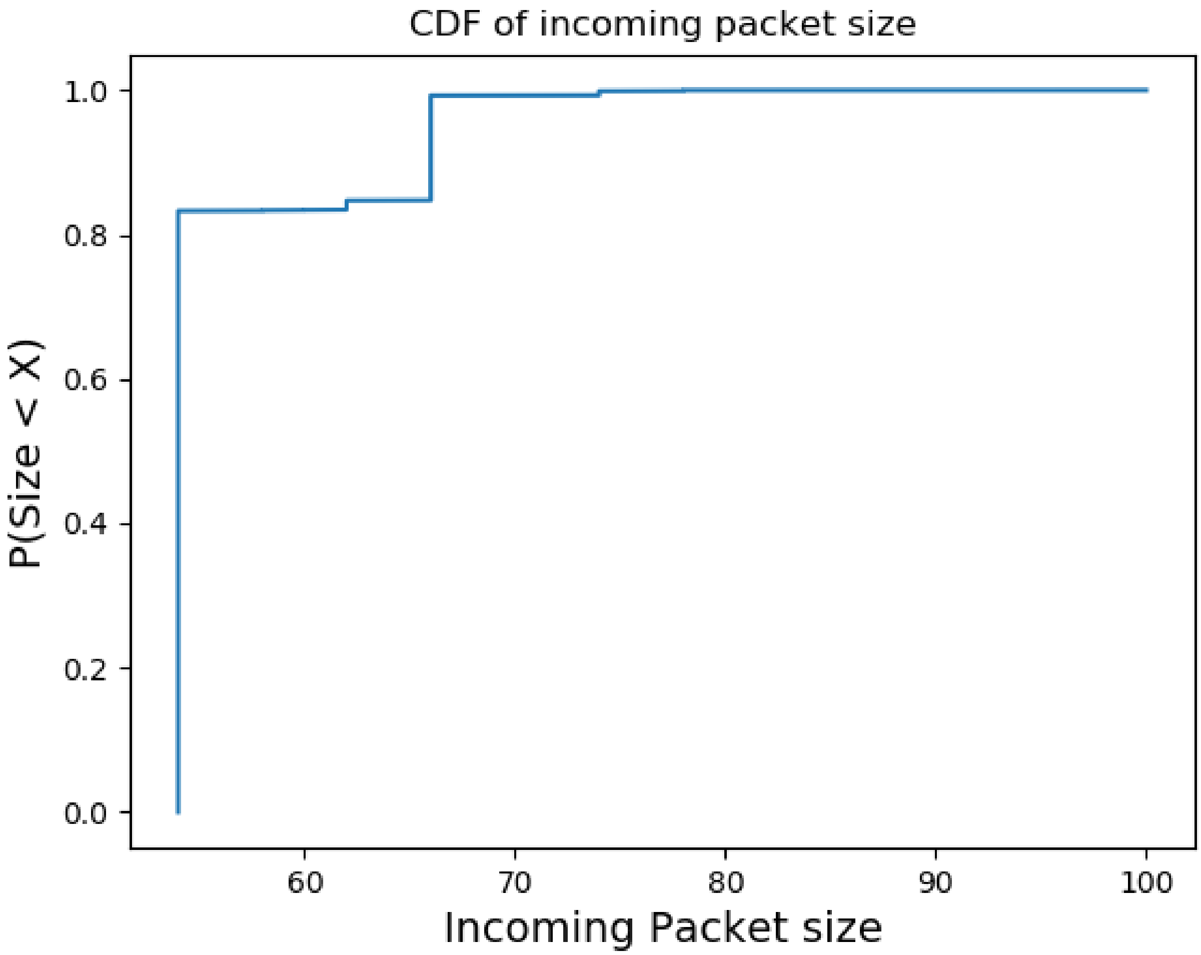
Mean of Interarrival Time incoming packets to server: 1.17464847025686 Median of Interarrival Time incoming packets to server: 0.553568000004816

5 Question 8

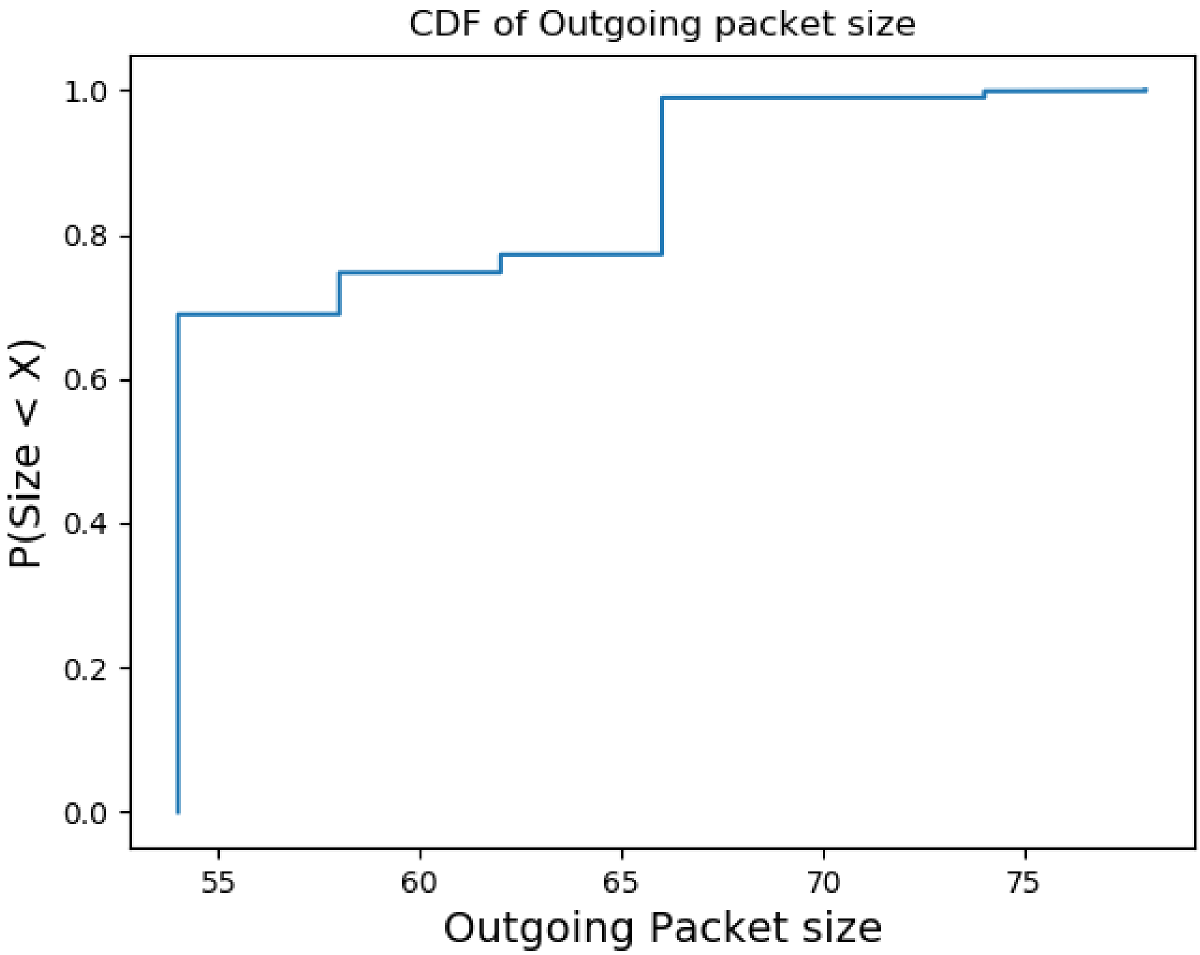
The CDF is probably clustered because the packets sizes are probably fixed to be some discrete values and not the entire range of continuous integers.

5.1 Day 1

Incoming Packet Size CDF

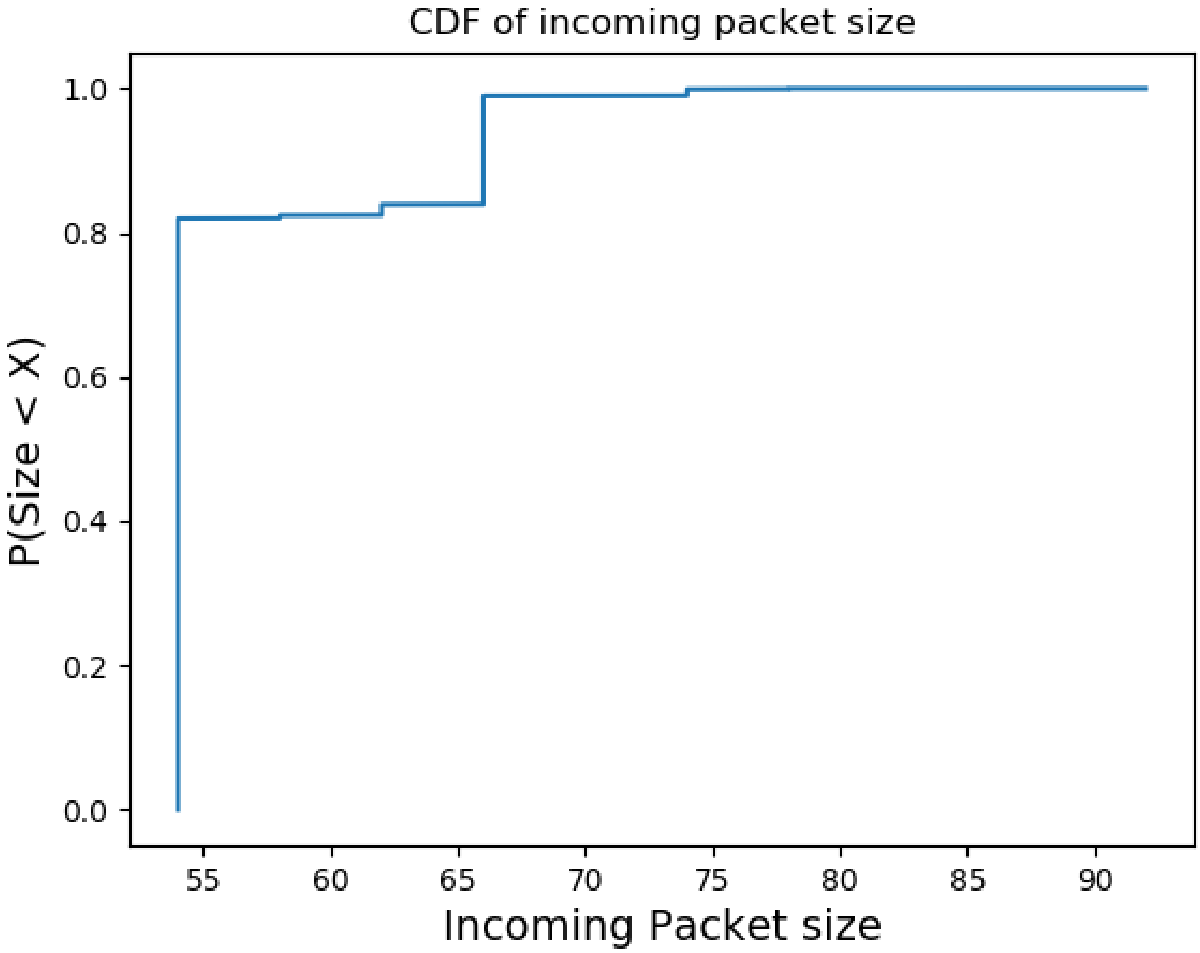


Outgoing Packet Size CDF

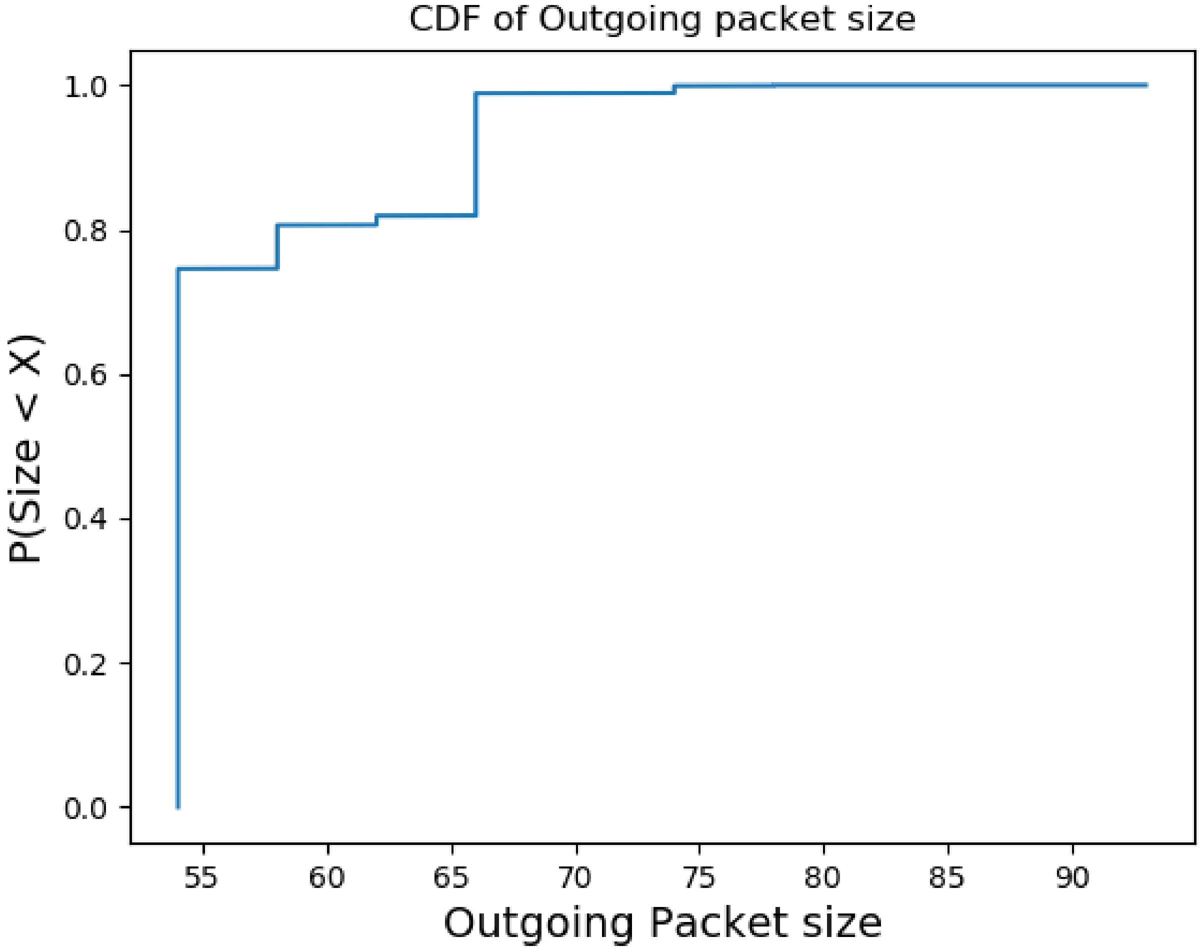


5.2 Day 2

Incoming Packet Size CDF

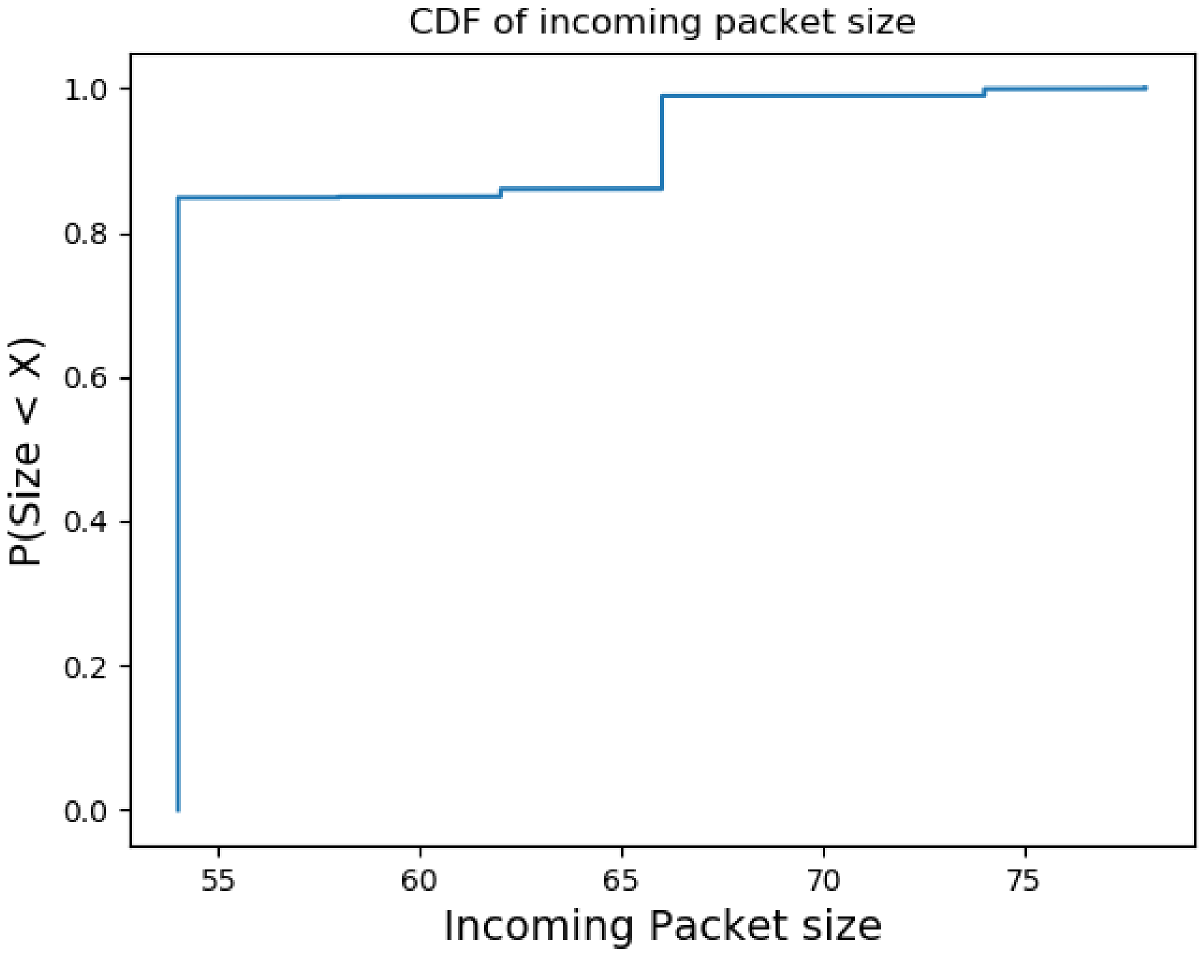


Outgoing Packet Size CDF

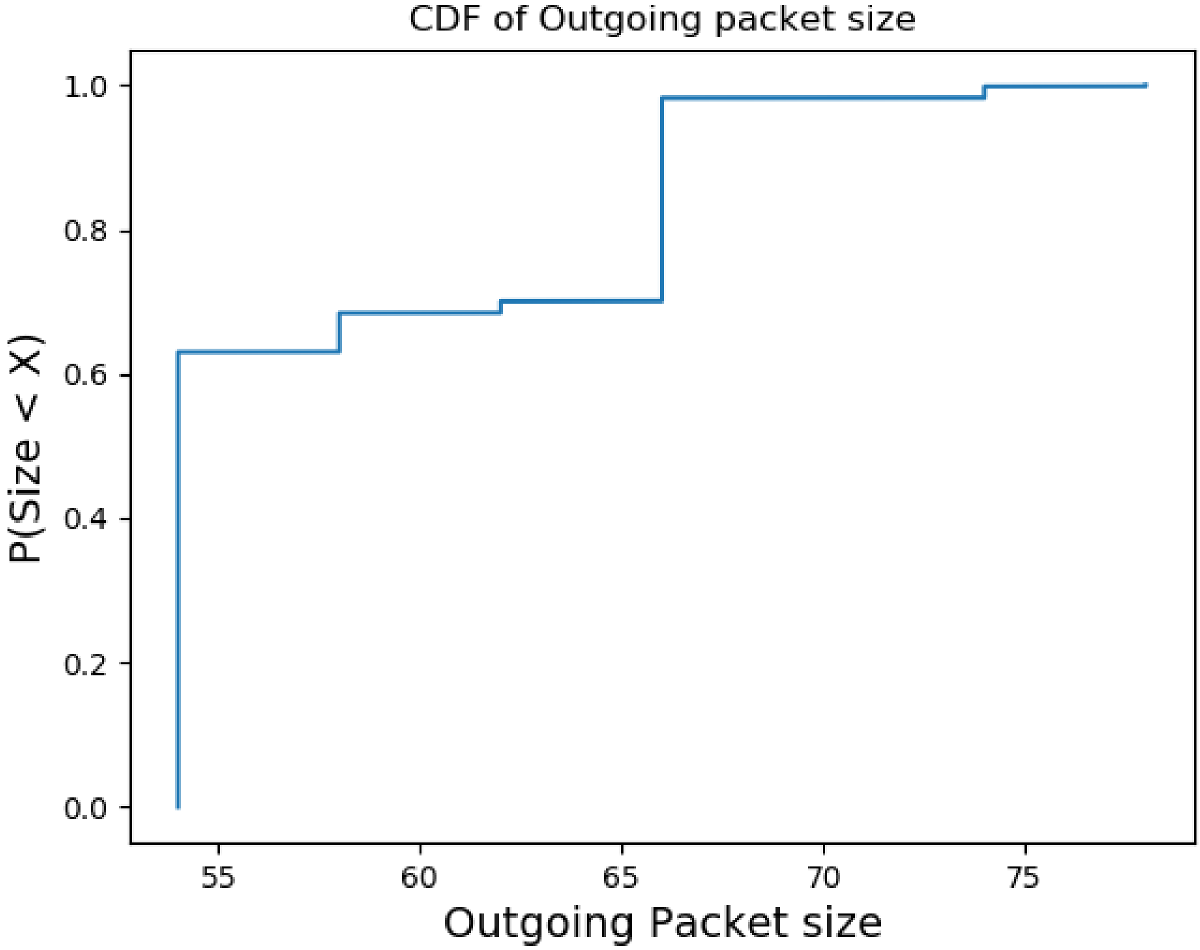


5.3 Day 3

Incoming Packet Size CDF

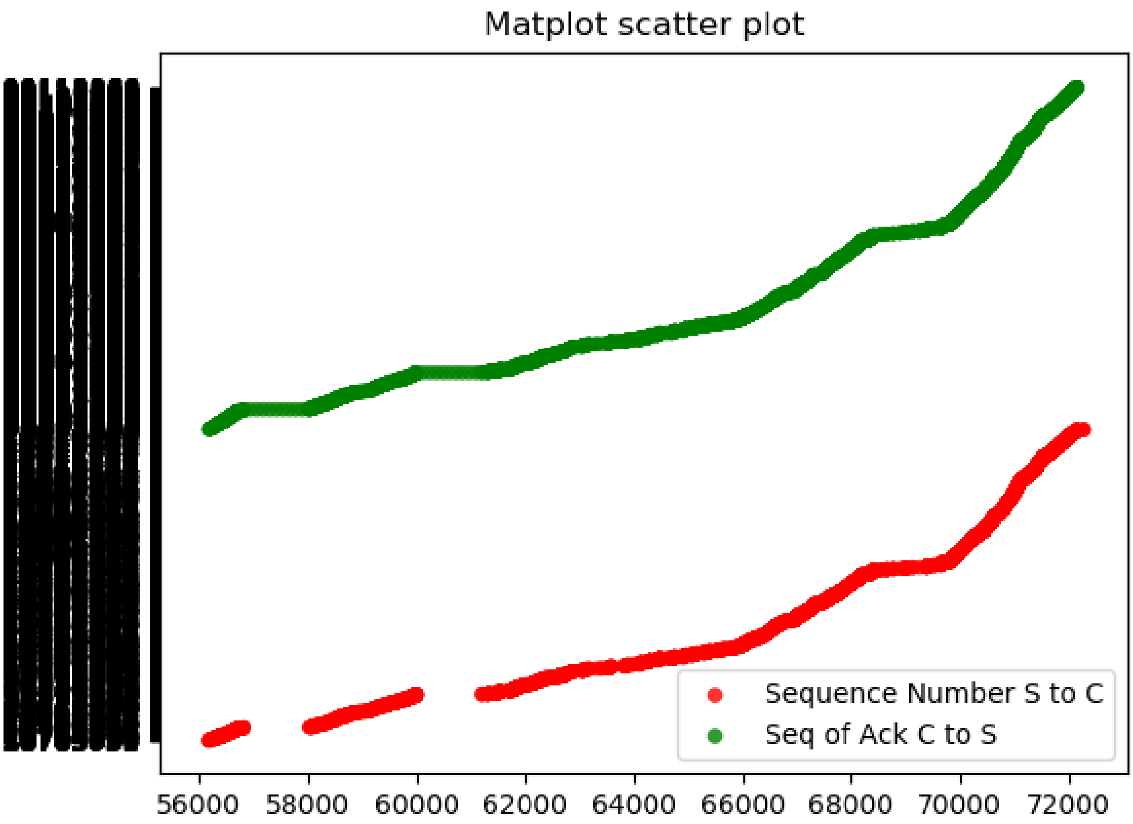


Outgoing Packet Size CDF

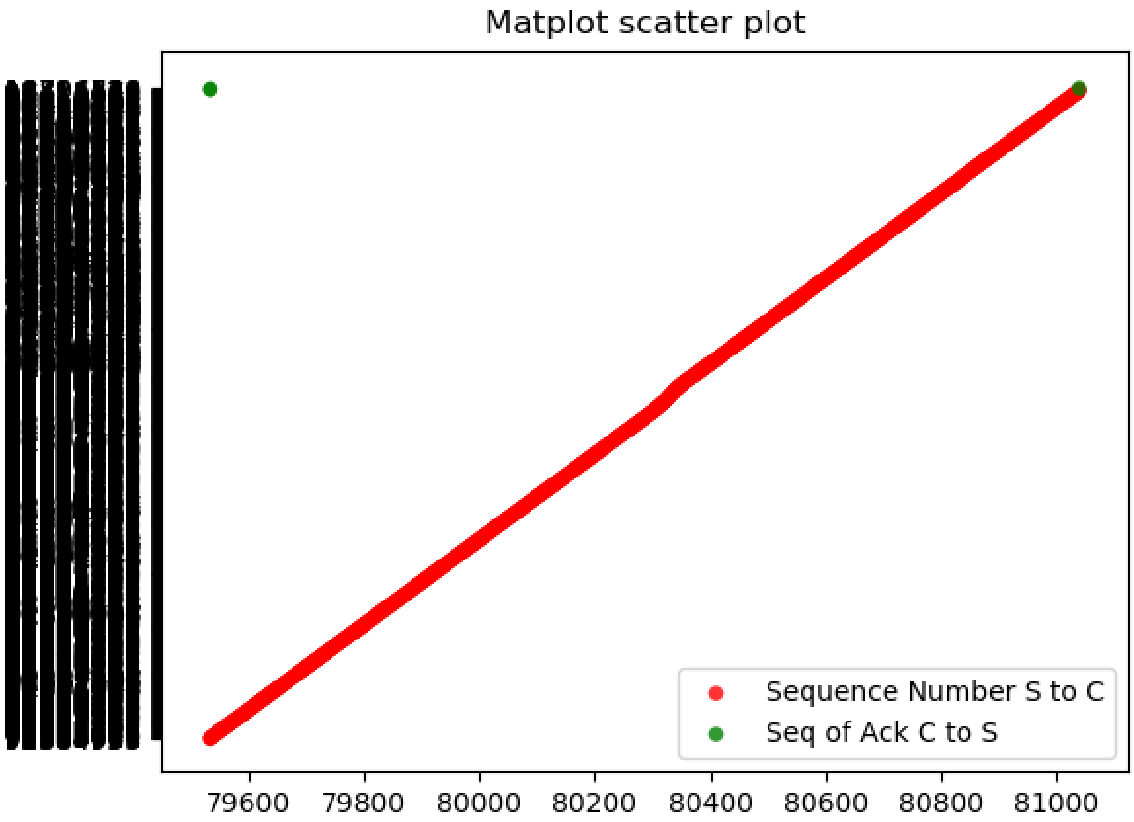

CDF of Outgoing packet size
P(Size < X)
0.0
0.2
0.4
0.6
0.8
1.0
55
60
65
70
75
Outgoing Packet size

6 Question 9

6.1 Data Intensive Connections

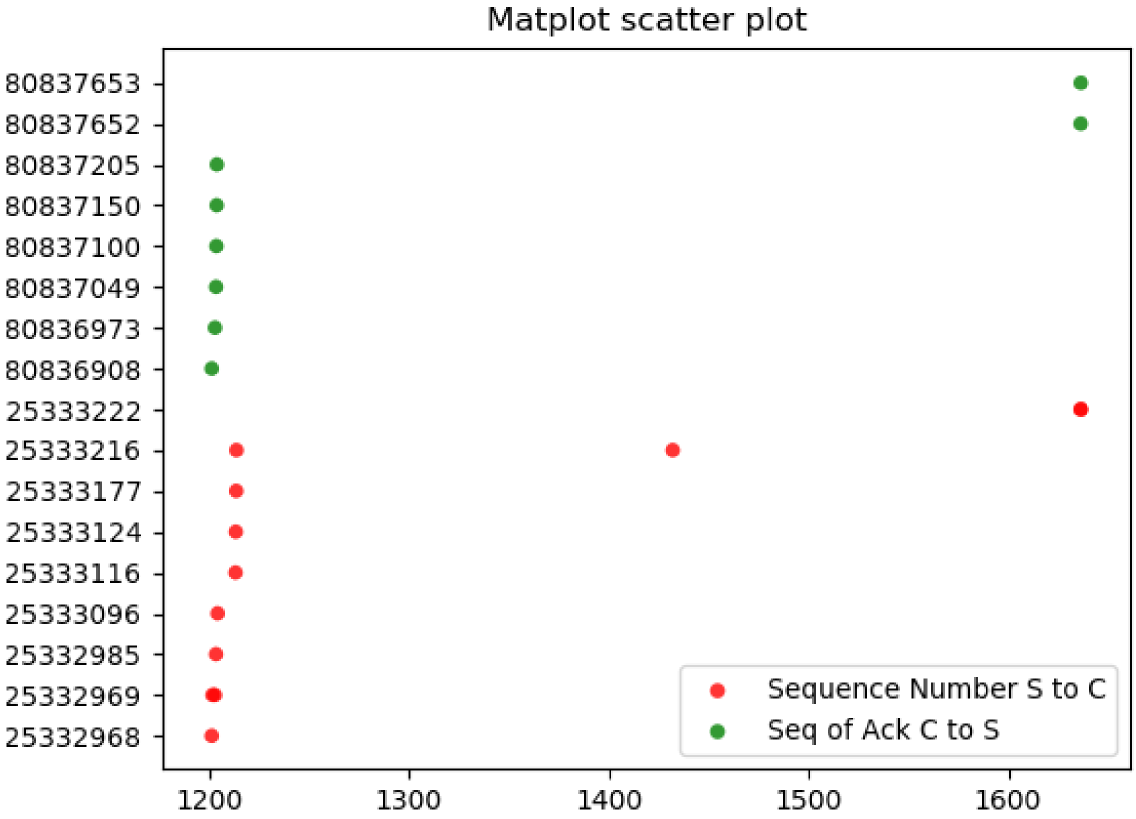


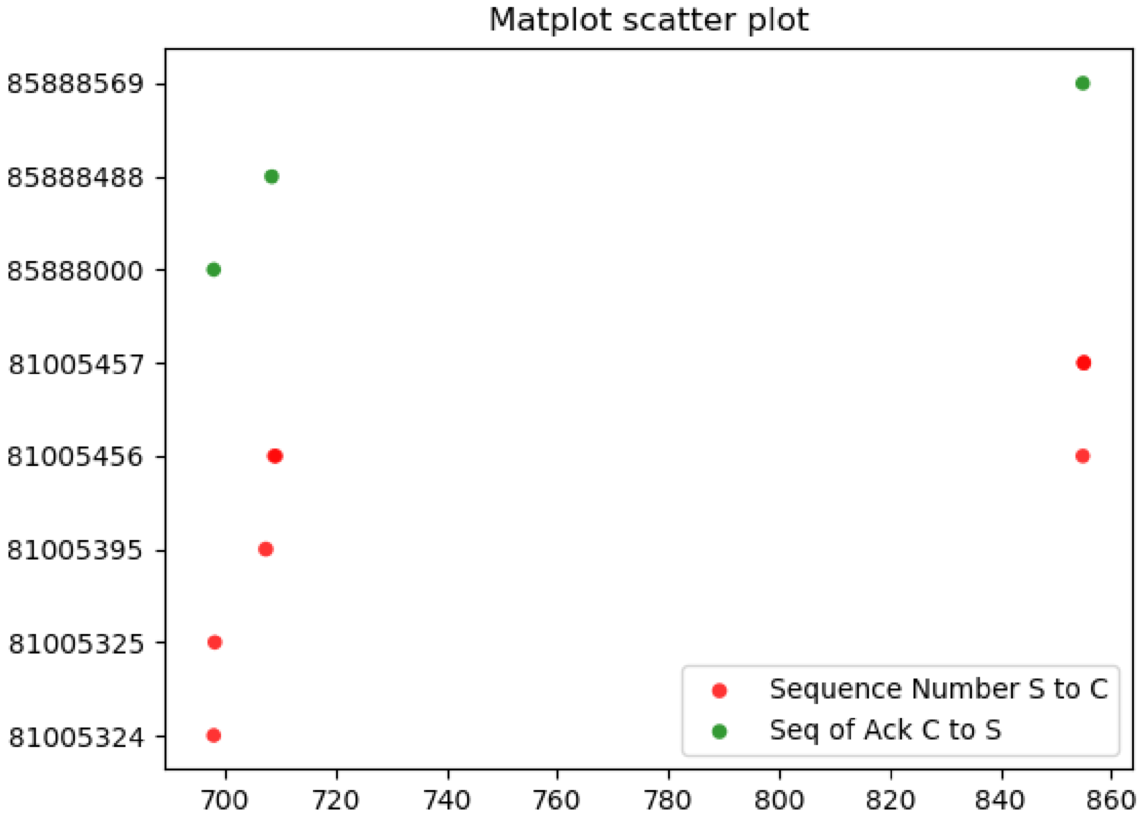
Probably a large file was requested.



6.2 Retransmission

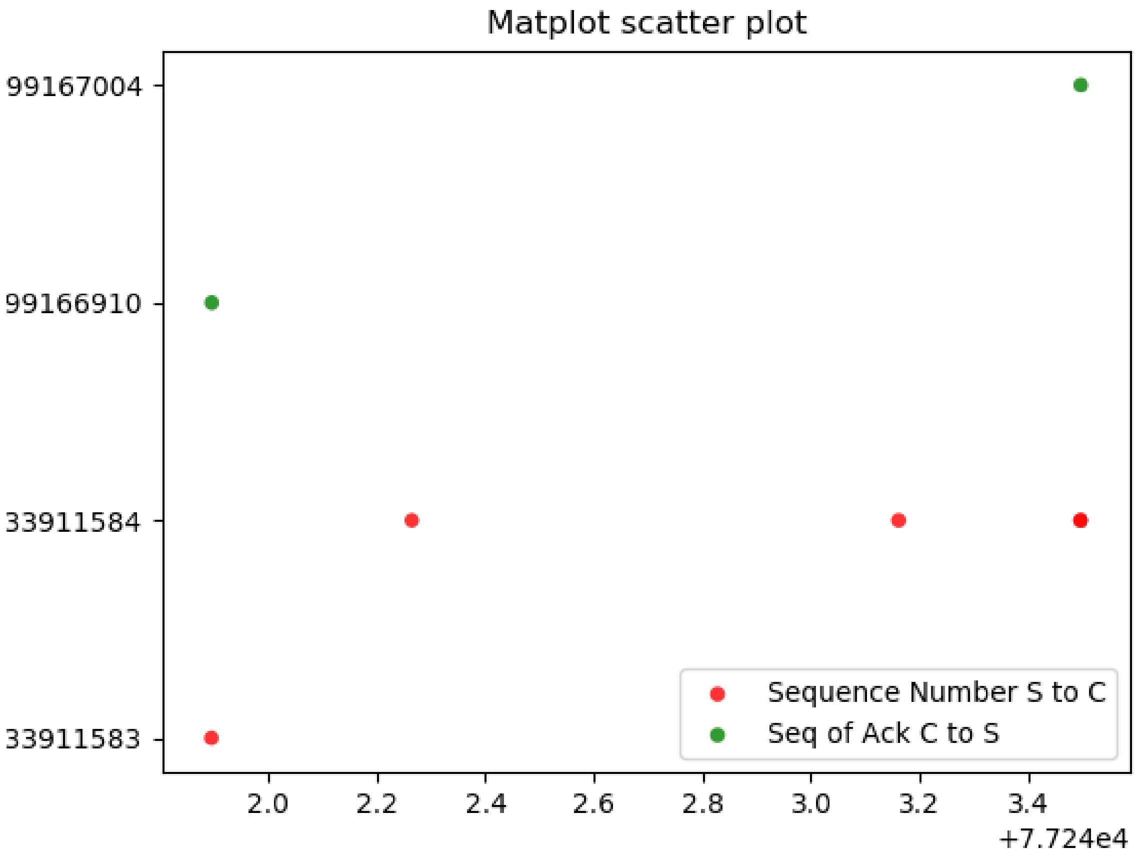
Same sequence number is seen multiple times: two data points at the same y


Matplot scatter plot
80837653
80837652
80837205
80837150
80837100
80837049
80836973
80836908
25333222
25333216
25333177
25333124
25333116
25333096
25332985
25332969
25332968
1200
1300
1400
1500
1600
● Sequence Number S to C
● Seq of Ack C to S

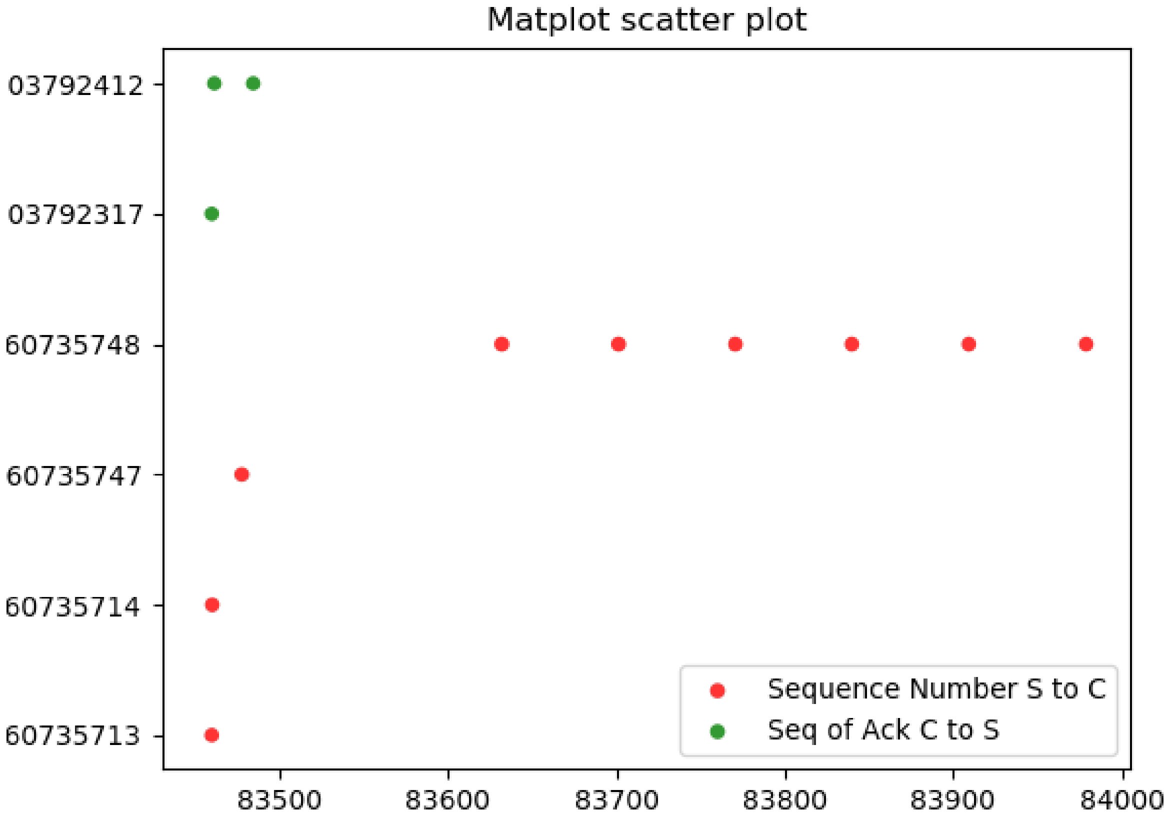


6.3 Spurious Retransmission

Probably the Ack was received almost at the same time as the timeout.

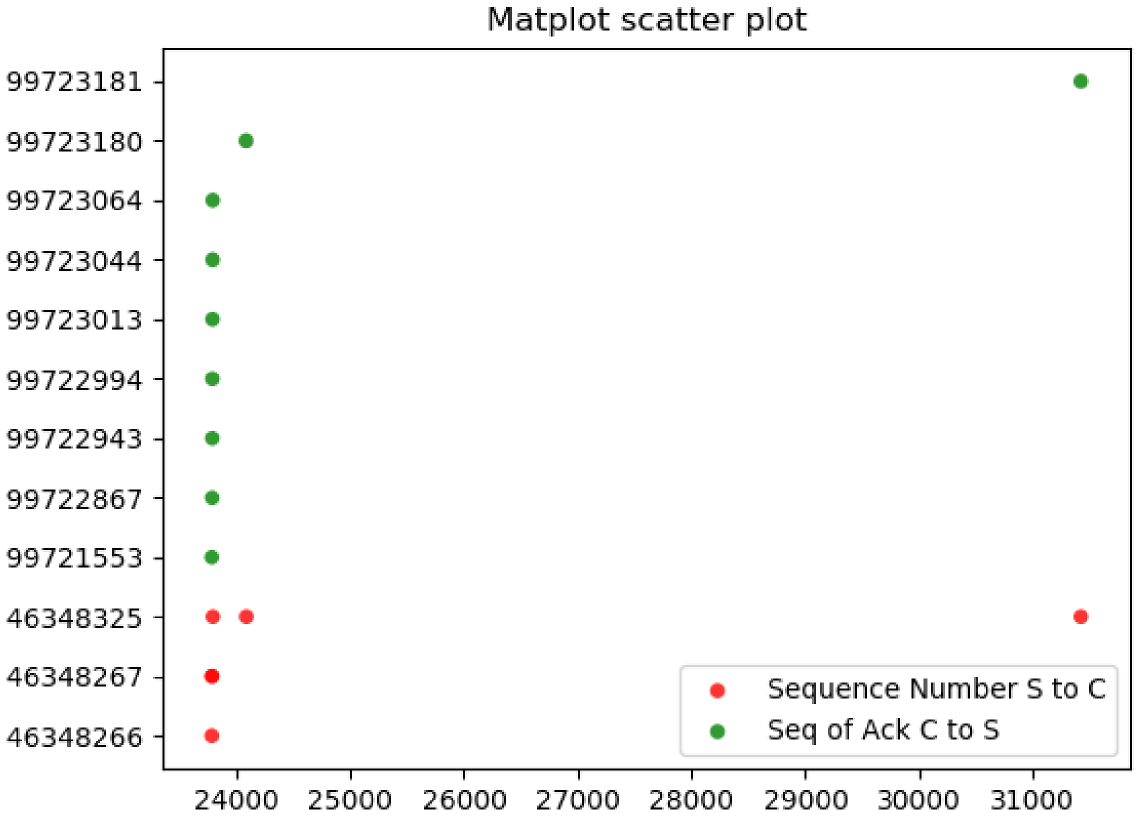


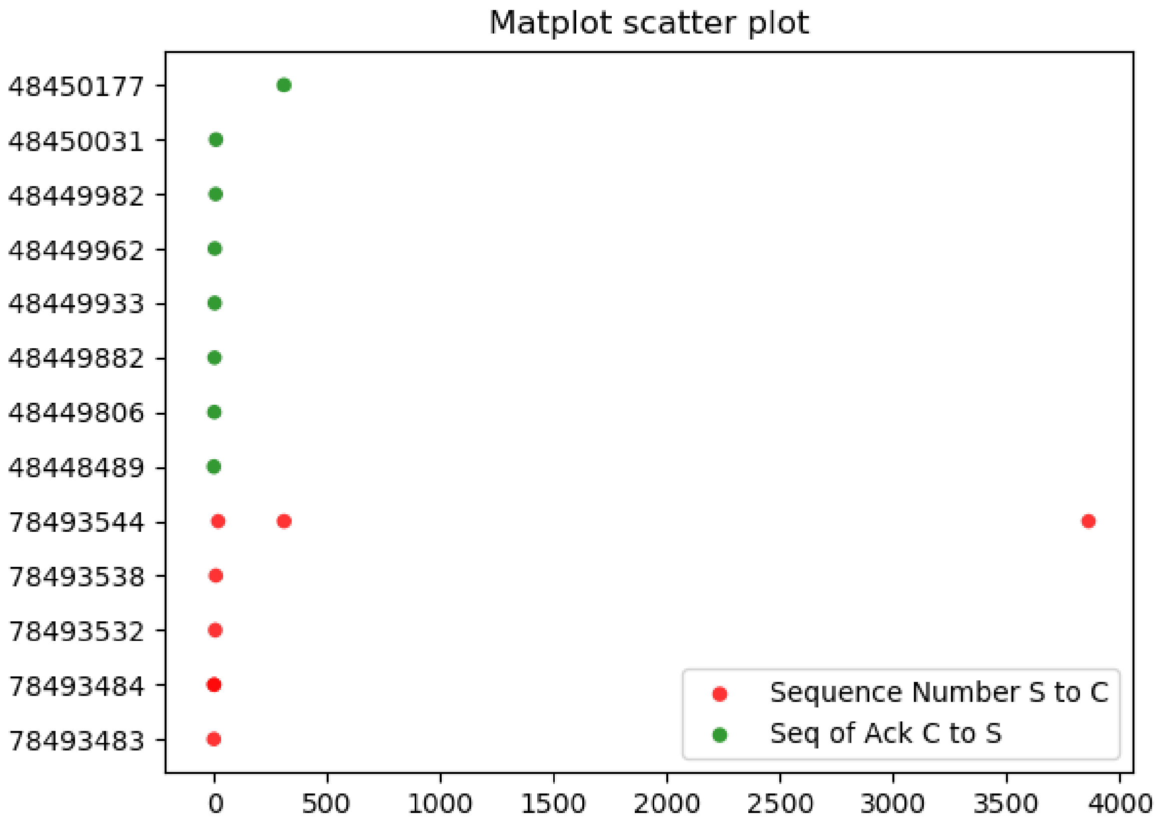
Probably the Ack was corrupted and a data point is missing.



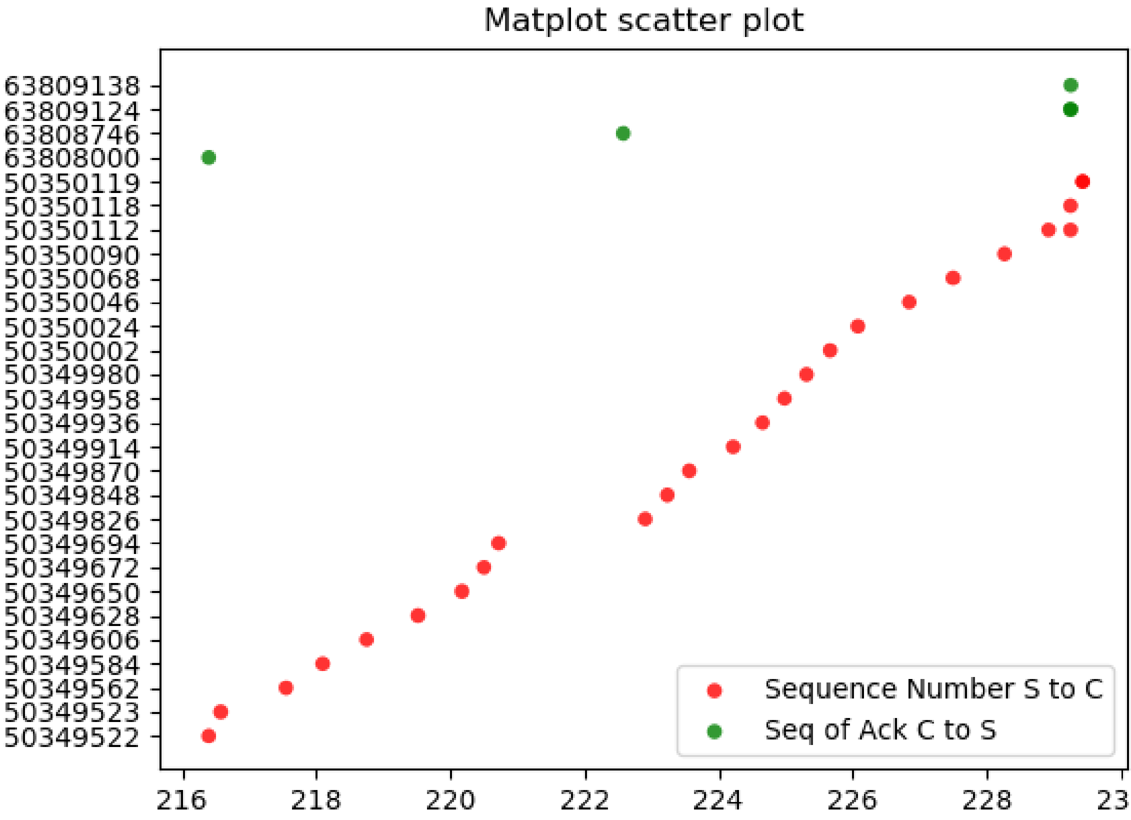
6.4 Duplicate Acks

Probably client wants to send more data packets to the server


Matplot scatter plot
99723181
99723180
99723064
99723044
99723013
99722994
99722943
99722867
99721553
46348325
46348267
46348266
24000
25000
26000
27000
28000
29000
30000
31000
● Sequence Number S to C
● Seq of Ack C to S


Matplot scatter plot
48450177
48450031
48449982
48449962
48449933
48449882
48449806
48448489
78493544
78493538
78493532
78493484
78493483
0
500
1000
1500
2000
2500
3000
3500
4000
● Sequence Number S to C
● Seq of Ack C to S

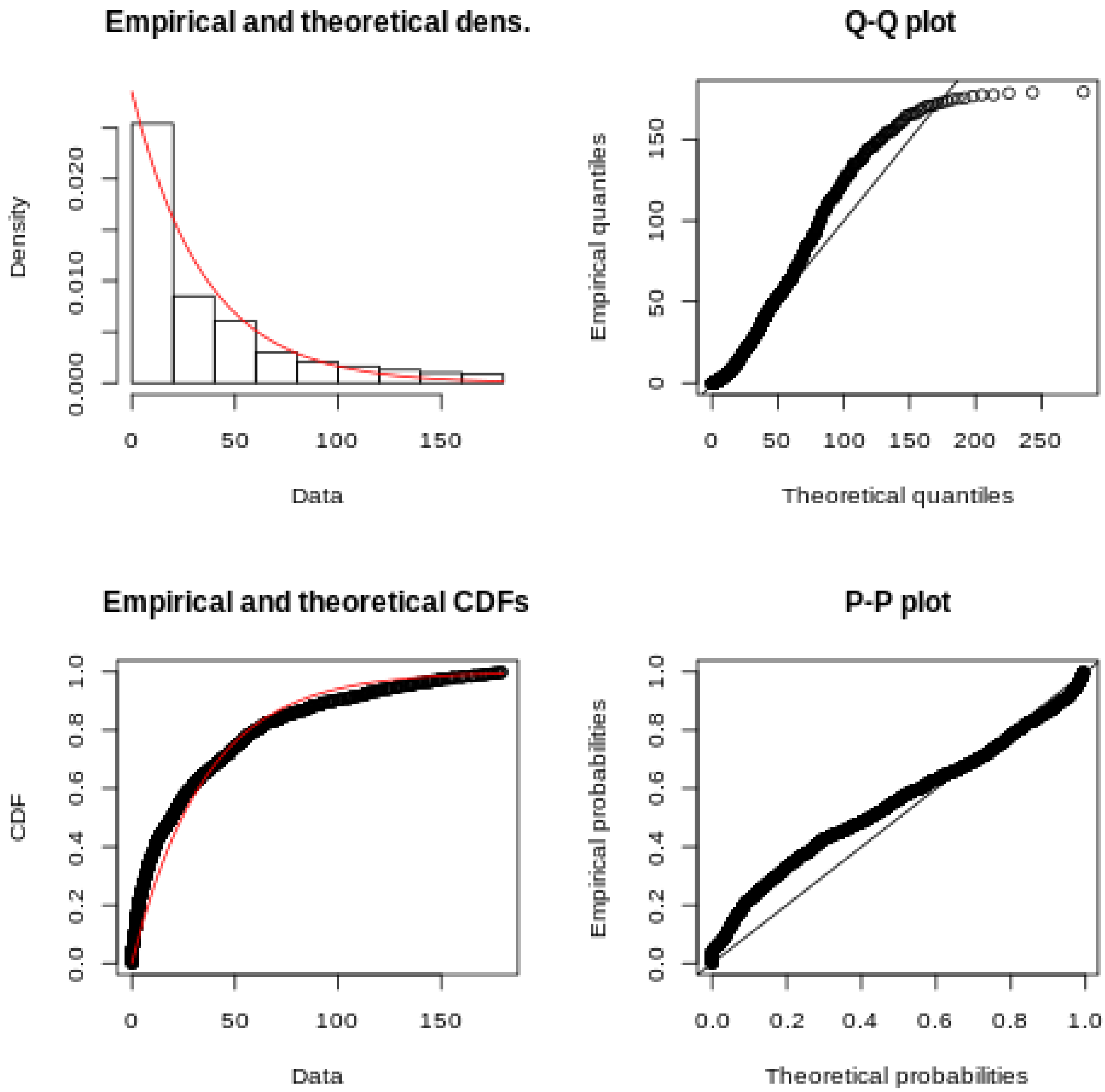
6.5 Out of Order Delivery

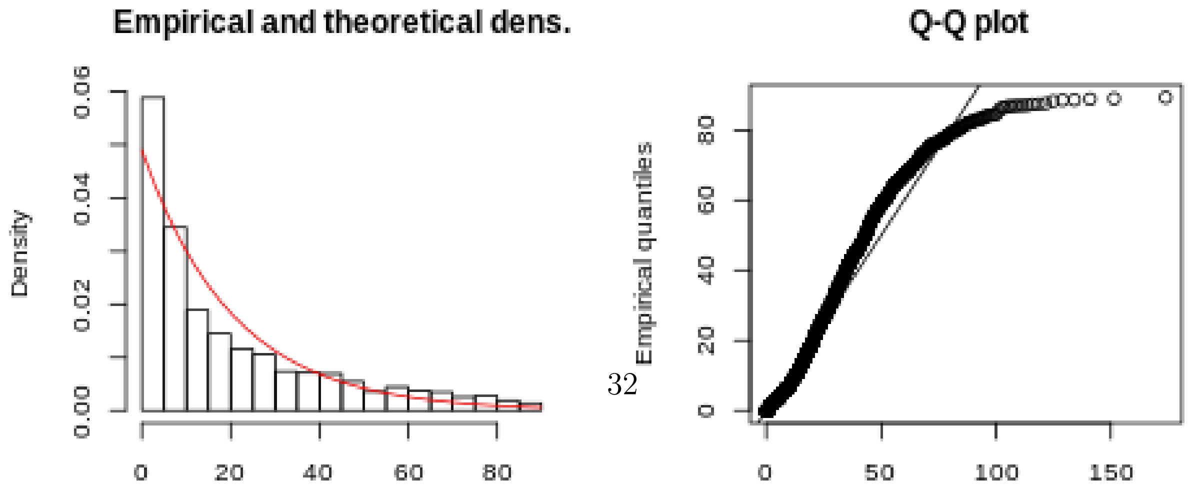


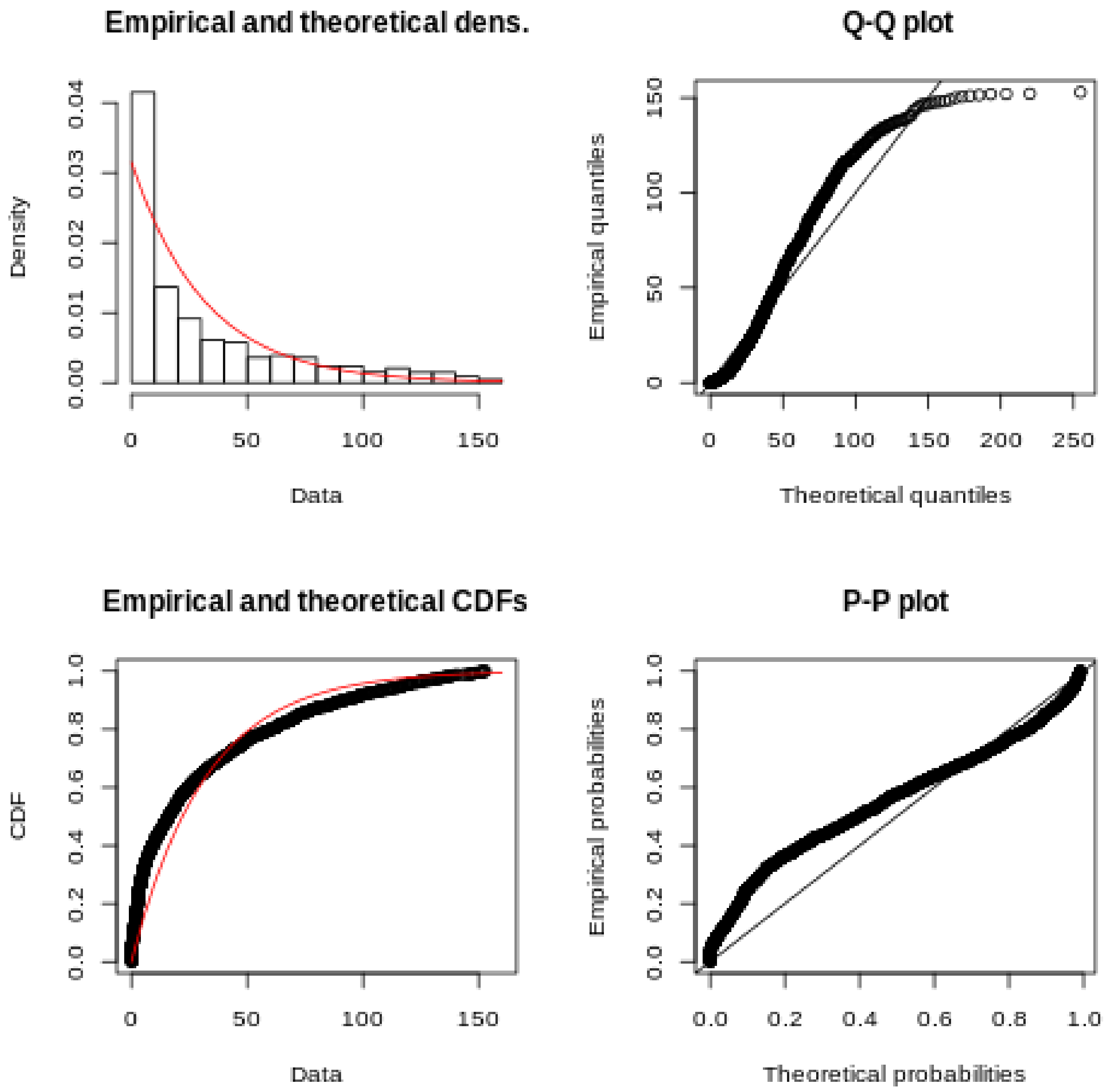
We could only find one such flow (on day 2). The out of order delivery seems to be caused due to retransmission of data from the server to the client

6.6 Distribution of Inter-Arrival Times :

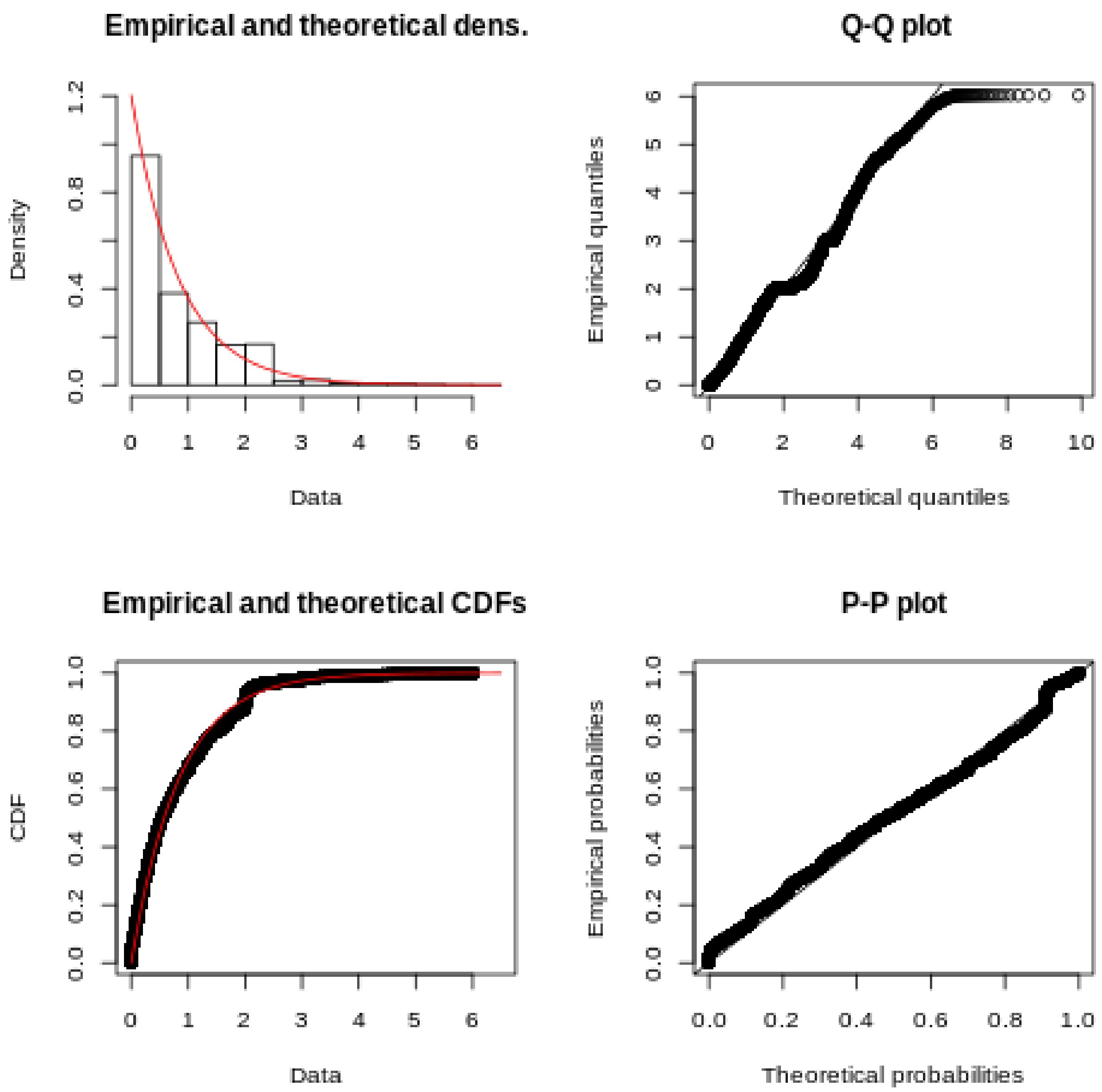
6.6.1 Two consecutive connections being opened

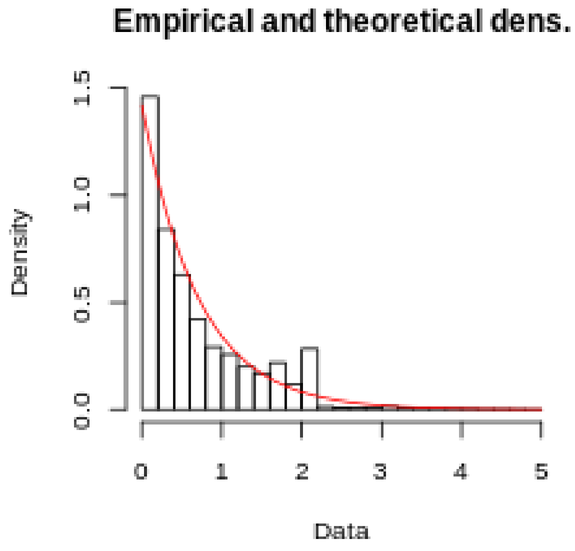


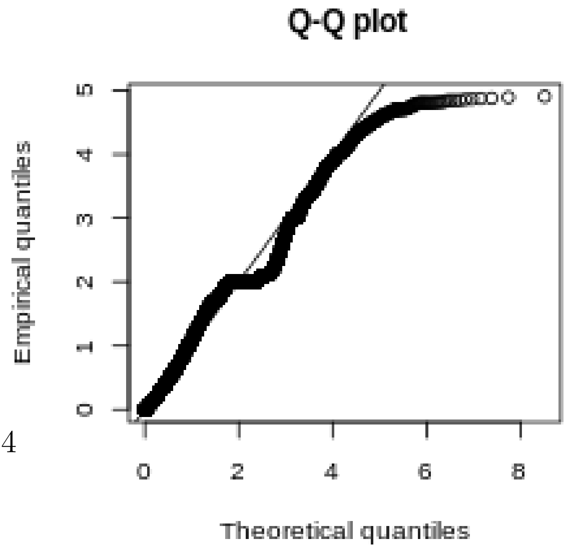


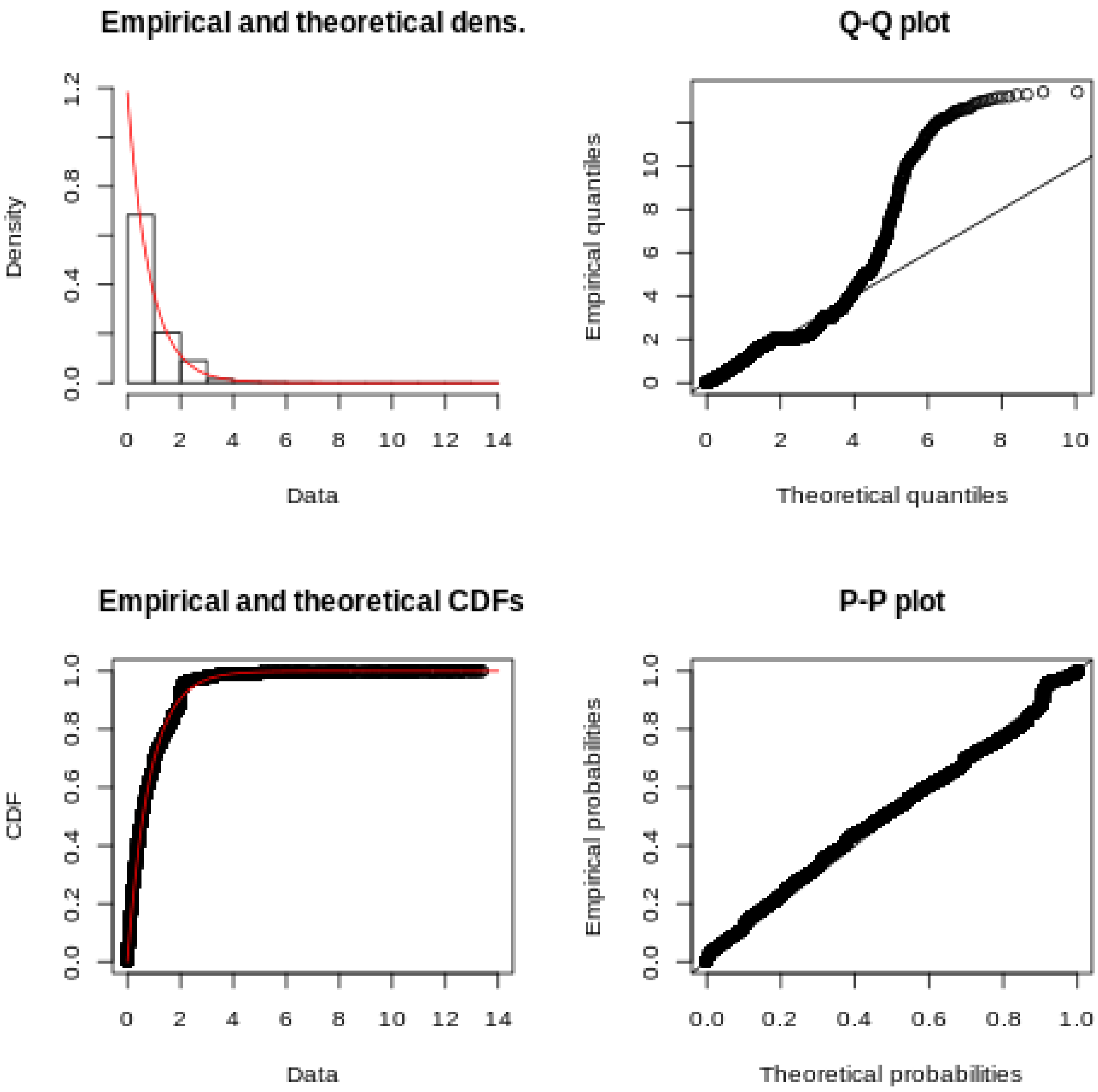

Empirical and theoretical dens.
Density
0.00
0.01
0.02
0.03
0.04
0
50
100
150
Data
Q-Q plot
Empirical quantiles
0
50
100
150
0
50
100
150
200
250
Theoretical quantiles
Empirical and theoretical CDFs
CDF
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
Data
P-P plot
Empirical probabilities
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Theoretical probabilities

6.6.2 Two consecutive incoming packets to the servers









Rate (Expected No. of packets per second) Parameter of the Exponential Function Fitting:

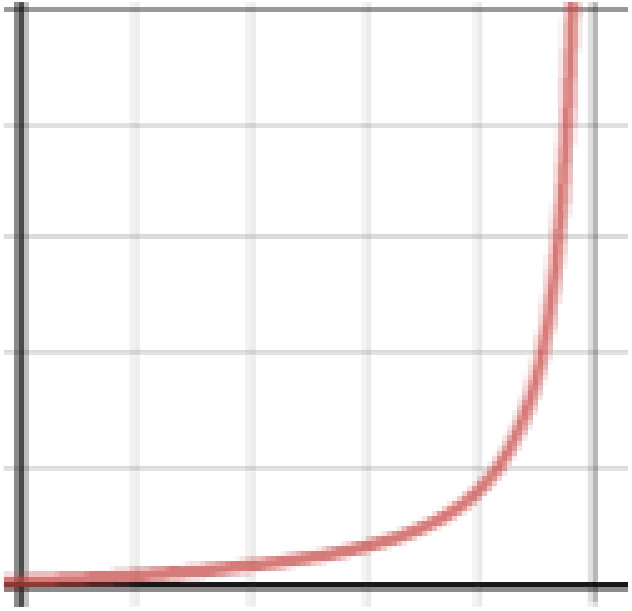
Rate :	PART 6	PART 7
File1	0.03404845	1.118508
File2	0.04576802	1.337741
File3	0.03390657	1.183119

7 Question 11

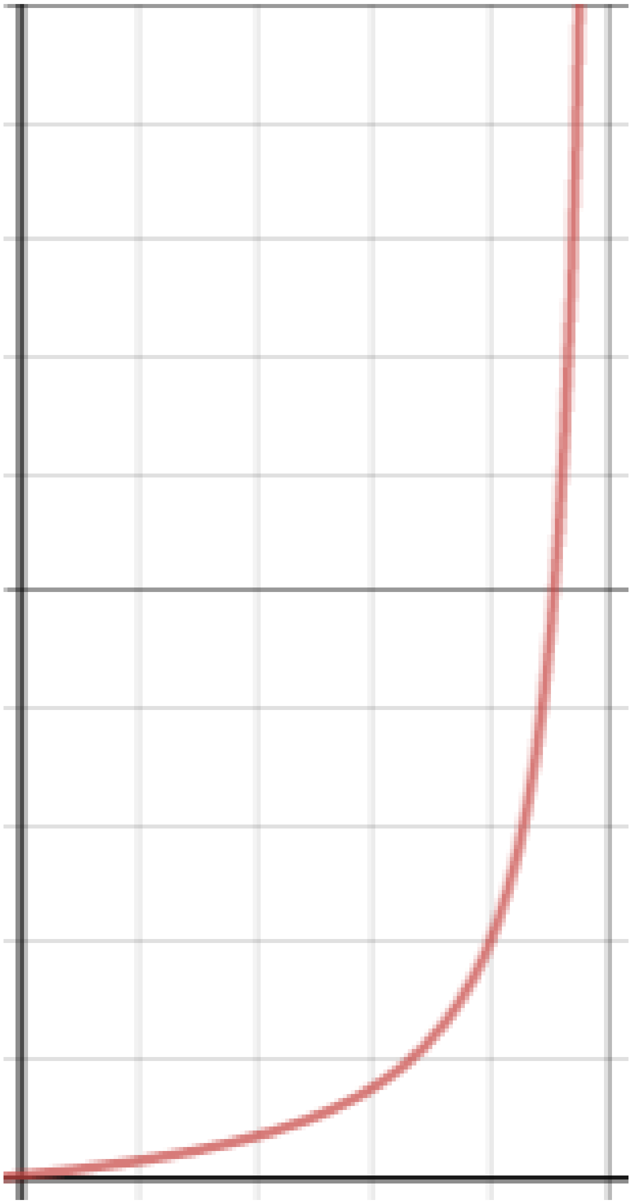
Interarrival Time between Packets

Rate :	PART 6	PART 7
File1	29.37	0.894
File2	21.85	0.747
File3	29.50	0.845

Mean packet size=56byte
Value of mu = 285 packets per second
Utilization factor rho on the 3 days:
F1=0.0039 F2=0.0047 F3=0.0041
Average Queue Size: F1=0.0039 F2=0.0047 F3=0.0041
Average Waiting Time: F1=13 microseconds F2=16 microseconds F3=14 microseconds
Waiting time vs lamba (for constant mu)



As can be seen the waiting time and the queue shoot up exponentially as lambda (arrival rate of packets) gets closer to mu



Average Queue Size vs Lambda