Computer Communications and Networks (COMN) 2017/18, Semester 1

Assignment Part 1 Results Sheet

Forename and Surname:	Michael Michaelides
Matriculation Number:	s1447836

Question 1 – Number of retransmissions and throughput with different retransmission timeout values with stop-and-wait protocol. For each value of retransmission timeout, run the experiments for **5 times** and write down **average number of retransmissions** and **average throughput**.

Retransmission timeout (ms)	Average number of re-transmissions	Average throughput (Kilobytes per second)
5	2146.8	46.280554
10	1090.4	44.6270544
15	752.4	40.2782776
20	217.2	40.9302088
25	195.4	38.8791726
30	201.6	37.2055356
40	205.0	34.3086338
50	193.4	31.5868064
75	206.4	27.1261054
100	198.6	22.6624404

Question 2 – Discuss the impact of retransmission timeout value on number of retransmissions and throughput. Indicate the optimal timeout value from communication efficiency viewpoint (i.e., the timeout that minimizes the number of retransmissions and keeps the throughput as high as possible).

If the timeout is too low, i.e. lower than the RTT which is 20ms in this case, then a lot of premature retransmissions occur as we can see above. The sender starts the timer when the datagram is sent, but the reply needs at least RTT ms to arrive. Hence, if the timeout is less than that, then the sender will inevitably re-send datagrams whose ACKs are on the way back. This causes multiple unnecessary retransmissions and congestion on the network. However, the throughput is high because in the case of a lost datagram, the client doesn't wait too much.
On the other hand, if the timeout is too long - a lot longer than the RTT time - then in the case of a lost datagram the sender will wait too long for the timeout to occur before resending, which causes decrease of throughput. However, the retransmission rate is low because the sender avoids premature retransmissions and hence doesn't send unnecessary datagrams.

Hence, as retransmission timeout increases, throughput decreases but number of retransmissions decrease too and as retransmission timeout decreases, number of retransmissions and throughput increases (they're negatively correlated).
Thus, the ideal timeout value should be longer than RTT to avoid premature retransmissions but not too long to avoid waiting unnecessarily. Specifically, it is calculated as below:

Timeout = EstimatedRTT + 4 * VarianceRTT
where:
EstimatedRTT denotes an averaged measure of the RTTs accounting both for past and current RTTs. Specifically it is equal to: **0.875 * EstimatedRTT** (history of previous RTTs) **+ 0.125 * SampleRTT** (current sampled RTT of the ACK just received)
VarianceRTT denotes the spread around EstimatedRTT, i.e. the variance around the mean of the values. It is equal to **0.75 * VarianceRTT + 0.25 * |SampleRTT - EstimatedRTT|.**

In this specific case we explicitly configured EstimatedRTT to be 20ms and there's not much variation due to the sender and receiver being on the same machine. Hence, I will set the safety margin to 5ms and set timeout = 25 ms as the optimal timeout, because it reaches near-optimal number of retransmissions and the throughput is not much lower than the highest. Note that, the ideal timeout is constantly changing and we need to update after each successful transmission to reflect reality using the above formula.