CS 536 (Park) **Assignment X** **Due: Nov. 11 & 14 (Tue. & Fri.), 1997**

PROBLEM 1 (Due 11/11.)

Redo the midterm exam (the questions are posted on the course homepage), giving a comprehensive/complete treatment of the problems as mentioned in class. Do *not* solve the bonus question, Problem 4 (b).

A score of 80 pts or higher is required for the scores to count. A score *lower than* 80 will count as 0. This is a hard threshold. The redone midterm exam must be submitted by office closing time on Nov. 11 (Tuesday). All other problems are to be submitted on Nov. 14, the regular time.

PROBLEM 2

(a) From first principles, using the steps followed in class for the M/M/1 (i.e., M/M/1/∞) queueing system, derive the queue length distribution $X_i(t)$ for the M/M/1/n queue, i.e., the Markovian system with finite queue length or buffer size n. Your derivation must be complete stating all the assumptions that you make. Draw the state transition diagram of your system.

Is the assumption $\lambda < \mu$ needed as in M/M/1? If not, what happens if $\lambda \geq \mu$? Explain your reasoning.

(b) How does π_i behave as a function of ρ and i? Give your interpretation of the queue length distribution based on plotting π_i as a function of $0 \leq i \leq n$ for different values of ρ, say $\rho = 0.8, 0.9, 1.0, 1.1., 1.2$. Pick $n = 30$.

Derive the mean queue length L of M/M/1/n and plot L as a function of ρ for $n = 10, 20, 50, 100$. Give an interpretation of your results. Have things fundamentally changed for the finite buffer system? Explain.

PROBLEM 3

(a) As an extension of Problem 4 (b), Assignment IX, rerun the experiments for $1/\lambda_Y = 30$ ms and $1/\lambda_S = 20, 25, 30, 35, 40$ ms to get the queue length traces for each (λ_S, λ_Y) combination, however, now transmitting 1000 packets instead of 500. For each queue length trace, compute the empirical queue length distribution $\hat{X_0}(t_{1000})$, $\hat{X_1}(t_{1000})$, ..., $\hat{X_{30}}(t_{1000})$ where t_{1000} is the ending time of the experiment of transmission 1000 packets. That is, $\hat{X_i}(t_{1000})$ is the *observed* fraction of the time that the queue length equals i during an experiment.

Compare that against the analytically derived formula for the corresponding M/M/1/30 queueing systems in equilibrium with the specified λ_S, λ_Y values—i.e., $\pi_0, \pi_1, \ldots, \pi_{30}$—by plotting them, for each (λ_S, λ_Y) combination, on the same graph. How accurate are the predicted values (in the long run) vis-à-vis the measured ones? What about the observed and predicted mean queue lengths? Give an interpretation of the results.

(b) What happens if you use the predicted values for $\hat{X_i}(t_{1000})$, $i = 0, 1, \ldots, 30$, and the mean queue length for the M/M/1/∞ queueing system? How accurate is the latter compared to that obtained in part (a)? Plot the corresponding graphs and give an interpretation of the results.

As you vary which system parameter will the M/M/1 system give a more accurate picture of the M/M/1/n system? How does ρ come into play here? Which result that we derived can you use to partially substantiate your claims?

PROBLEM 4

Assume you are the CEO (and sole employer/employee) of a network service provider company called *Network-QoS-'R-Us*. Assume you live in a dream world where your network can be accurately modeled as a M/M/1 queue with rates λ and μ. Suppose that as a service provider, the variable or resource under your control is μ, the service rate or link bandwidth of your system. Assuming the customers (with aggregate arrival rate λ) to your system are network audio applications that require a delay less than 200 ms (assume packets are fixed size and rates are denoted in units of "z packets per second"), for a given λ, what is the service rate μ needed to satisfy your customers' requests *on average*? Show your calculation?

Why might your customers not be satisfied with your service delivered above thus moving to another service provider? How would you approach to solve this problem? Can you compute a μ value such that you can proclaim services such as "no more than α % of your packets will experience delays longer than 200 ms"? If so, how to compute it?

(*Hint: Think of worst-case guarantees vs. average-case guarantees; for delay guarantees, try to compute the area of the tail portion of the queue length distribution and use that for your calculations.*)