ECE 567 COMMUNICATION NETWORK ANALYSIS **FALL 2006**
PROBLEM SET 5 **Due Wednesday, November 1**

Topics: Equivalent bandwidth and multiplexing bursty traffic streams, reversibility and circuit switched network models, deterministic delay calculus

Reading in notes: Sections 2.5, 2.7, 4.1, 4.2 and Chapter 5
More material on reversibility can be found in Appendix A, Chapter 3 of Bertsekas and Gallager, F. Kelly, *Reversibility and Stochastic Networks*, Chapters 1 (especially 1.2,.16 and 1.7), 2, 3 (especially 3.1) and 4 (especially 4.1) and Kleinrock, *Volume II*, Sections 4.11-4.12 (product form applied to computer time-sharing), 5.6-5.9 and 6.1.
More material on the deterministic delay calculus can be found in Cheng-Shang Chang, *Performance Guarantees in Communication Networks*, Springer, 2000 (on reserve in Grainger Engineering Library) Sections 1.1-1.4, 1.6, 1.8, 2.1, 2.2.1, 2.3.1, 2.3.2, 2.3.5.

1. Effective bandwidth, bufferless link

Consider a bufferless link of fixed capacity $C = 200$ serving connections of 2 types. The data rate of each connection is random, but constant in time. The data rate required by a connection of type 1 is uniformly distributed over [0,2], the data rate required by a connection of type 2 is exponentially distributed with mean 1, and requirements for different connections are mutually independent. (a) Calculate the effective bandwidth functions $\alpha_1(s)$ and $\alpha_2(s)$. (b) Find the largest integer n so that the Chernoff inequality implies that the blocking probability is less than or equal to 0.001 for a nominal load of n connections of each type. Show your work for full credit! (Hint: Write a short computer program for this problem. For a given value of n the Chernoff bound is computed by a minimization over s. Then n can be adjusted by binary search – increasing n if the corresponding overflow probability is too small, and decreasing n if the corresponding overflow probability is too large.) (c) Compute the effective bandwidths $\alpha_1(s^*)$, $\alpha_2(s^*)$, and the effective capacity $C + \frac{\log(0.001}{s^*}$. Which of the two effective bandwidths is larger? Sketch the corresponding acceptance region $A(s^*)$. Here s^* denotes the optimizing value of s in the Chernoff bound. (Hint: The nominal load point (n,n) should be contained in the region $A(s^*)$. Which of the two effective bandwidths is larger?)

2. Effective bandwidth for a buffered link and long range dependent Gaussian traffic

Consider a link with buffer size $B = 100$ and fixed capacity $C = 200$, serving connections of 2 types. The amount of data offered by a connection of type i over an interval of length t is assumed to be Gaussian with mean t and variance $V_i(t) = t^{2H_i}$, where H_i is the Hurst parameter. Assume $H_1 = 0.5$ and $H_2 = 0.9$. (So the class 2 connections have long-range dependence.) (a) Compute the largest value n so that the Chernoff approximation for the overflow probability for n connections of each type (simultaneously) is less than or equal to $\exp(-\gamma) = 0.001$. (Hint: Write a short computer program for this problem. For a given value of C and t the approximation is computed by a minimization over s, which can be done analytically. The maximization over t can be done numerically, yielding the approximate overflow probability for the given value of C. Finally, C can be adjusted by binary search as in the previous problem. Also, compute the effective bandwidths $\alpha_1(s^*, t^*)$, $\alpha_2(s^*, t^*)$, the critical time t^*, and the effective capacity $C_{eff} = C + \frac{B}{t^*} - \frac{\gamma}{s^* t^*}$. Sketch the corresponding acceptance region $A(s^*, t^*)$. (Hint: The nominal load point (n, n) should be contained in the region $A(s^*, t^*)$. Which of the two effective bandwidths is larger?) (b) Redo problem (a) for overflow probability $\exp(-\gamma) = 0.00001$. and comment on the differences between the answers to parts (a) and (b).

3. Time reversal of a simple continuous time Markov process
Let $(X(t) : t \geq 0)$ be a stationary time-homogeneous, pure-jump Markov process with state space $\{1, 2, 3, 4\}$ and Q matrix

$$Q = \left(\begin{array}{cccc} -1 & 1 & 0 & 0 \\ 0 & -1-\alpha & 1 & \alpha \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right),$$

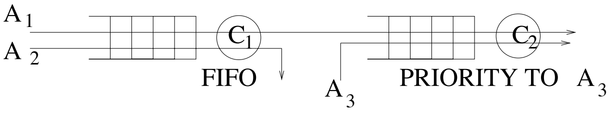
where $\alpha > 0$. Find the Q matrix of the time-reversed process. Sketch the rate transition diagram of both X and the time-reversed process.

4. Time reversibility of an M/GI/1 processor sharing queue
A processor sharing server devotes a fraction $1/n$ of its effort to each customer in the queue whenever there are n customers in the queue. Thus, if the service requirement of a customer arriving at time s is x, then the customer departs at the first time s' such that $\int_s^{s'} 1/N(t)dt = x$, where $N(t)$ denotes the number of customers in the system at time t. It turns out that if the queue is stable then N is time-reversible, and its equilibrium distribution depends only on the load ρ. In particular, the departure process is Poisson. This result is verified in this problem for a certain class of distributions of the service requirement X of a single customer. (a) Is N Markovian? If not, explain what additional information would suffice to summarize the state of the system at a time t so that, given the state, the past of the system is independent of the future. (b) To arrive at a Markov model, restrict attention to the case that the service requirement X of a customer is the sum of k exponentially distributed random variables, with parameters μ_i for $1 \leq i \leq k$. Note that $E[X] = 1/\mu_1 + \cdots + 1/\mu_k$ and $\rho = \lambda E[X]$. Assume $\rho < 1$. Equivalently, X is the time it takes a pure birth Markov process to move from state 1 to state k+1 if the birth rate in state i is μ_i. Thus, each customer moves through k stages of service. Given there are n customers in the system, a customer in stage i of service completes that stage of service at rate μ_i/n. Let $N_i(t)$ denote the number of customers in the system that are in stage i of service at time t. Let $\Theta(t) = (N_1(t), \ldots, N_k(t))$. Then Θ is a Markov process. Describe its transition rate matrix Q.
(c) The process Θ is not reversible. Show that the equilrium distribution of Θ is given by $\pi(n_1, \ldots, n_k) = (1-\rho)\rho^n \binom{n}{n_1,\ldots,n_k} a_1^{n_1} \cdots a_k^{n_k}$, where $n = n_1 + \cdots + n_k$ and $a_i = (1/\mu_i)/(1/\mu_1 + \cdots + 1/\mu_k)$. At the same time, show that the time reverse of Θ has a structure similar to Θ, except that instead of going through the stages of service in the order $1, 2, \ldots, k$, the customers go through the stages of service in the order $k, k-1, \ldots, 1$, with the service requirement for stage i remaining exponentially distributed with paramenter μ_i. (It follows that then number of customers in the system in equilibrium has distribution $P[N(t) = n] = (1-\rho)\rho^n$.) (d) Explain why $N(t)$ is reversible for the class of service times used. (Note: If $\mu_i = k$ for $1 \leq i \leq k$, then as $k \to \infty$, the service time requirement approahes the determinitic service time $X \equiv 1$. By using a more general Markov process to model X, any service time distribution can be approximated, and still N is reversible.)

5. A sequence of symmetric star shaped loss networks
Consider a circuit switching system with $M \geq 2$ links, all sharing an endpoint at a central switching node. Suppose the set of routes used by calls is the set of all (unordered) pairs of distinct links. There are thus $\binom{M}{2}$ routes. Suppose that calls arrive for any given route according to a Poisson process of rate $\nu = 4/(M-1)$. Suppose that the call holding times are independent and exponentially distributed with mean one. Finally, suppose that each link has capacity $C = 5$, so that a call offered to a route is blocked and discarded if either of the two links in the route already has 5 calls in progress. (a) What Markov chain describes the system? Describe the state space and transition rates. Is the chain reversible? (b) What is the conservative upper bound on the blocking probability for a given link, and what is Whitt's lower bound on the call acceptance probability. Consider the cases $M = 2, 3, 4$ and $M \to \infty$. Is the bound exact for any of these values? (c) Derive the reduced load approximation for the blocking probability of a link. Making an independence assumption, derive an approximation for the probability that a call is accepted. Again, consider the cases $M = 2, 3, 4$ and $M \to \infty$. Comment on the accuracy of the approximation for different values of M.

6. Deterministic delay constraints for two servers in series

Consider the network shown. There are three arrival streams. Suppose that A_i is (σ_i, ρ_i)-upper constrained



for each i and that $\rho_1 + \rho_2 \le C_1$ and $\rho_1 + \rho_3 \le C_2$. The first server is a FIFO server, and the second server gives priority to stream 3, but is FIFO within each class. For all three parts below, express your answers in terms of the given parameters, and also give numerical answers for the case $(\sigma_i, \rho_i) = (4, 2)$ for $1 \le i \le 3$ and $C_1 = C_2 = 5$.

(a) Give the maximum delay d_1 for customers of the first queue. (Your answer should be finite even if $\rho_1 + \rho_2 = C_1$ due to FIFO service order). (b) Find a value of σ such that B_1 is (σ, ρ_1)-upper constrained. (Seek the smallest value you can.)

(c) Find an upper bound on the delay that stream 1 suffers in the second queue.

7. Calculation of an acceptance region based on the SCED algorithm

Let

$$\begin{array}{ll} g_1(t) = 20 + t & f_1(t) = 5(t-4)_+ \\ g_2(t) = 8 + 4t & f_2(t) = 6(t-20)_+ \end{array}$$

(a) Verify that

$$g_i^*(t) = \begin{cases} 0 & \text{if } t = 0 \\ g_i(t) & \text{if } t \ge 1 \end{cases}$$

(b) Sketch $g_i^* \star f_i$ for $i = 1, 2$ on two separate graphs. These functions are piecewise linear. Label the breakpoints (places the slopes change), the values of the functions $g_i^* \star f_i$ at the breakpoints, and the values of slope between breakpoints.

(c) Suppose that $n_1 + n_2$ streams share a constant bit rate link of capacity $C = 100$ such that for $i = 1, 2$, there are n_i streams that are g_i-upper constrained, and that each of these n_i streams require the link to appear as an f_i-server. Using Service Curve Earliest Deadline first (SCED), it is possible to accommodate the streams if (n_1, n_2) satisfies $\sum_i n_i g_i^* \star f_i \le Ct$ for all t. Find and sketch the region of admissible (n_1, n_2) pairs. (Hint: It is enough to check the inequality at the breakpoints and at $t \to \infty$. Why?)