Exam #3 for Computer Networks II (CIS 6930)

Spring 2003

>>> *SOLUTIONS* <<<

Welcome to Exam #3 in Computer Networks II (CIS 6930). You have 75 minutes. Read each problem carefully. There are seven required problems (each problem is worth 14 points - you get 2 points for correctly following instructions). You may have with you a calculator, pencils, eraser, blank paper, lucky rabbit's foot, and one 8.5 x 11 inch "formula sheet". On this formula sheet you may have anything you want (definitions, formulas, etc.) *handwritten by you*. You may use both sides. Computer generated text, photocopies, and scans are not allowed on this sheet. Please submit your formula sheet with your exam. Please start each numbered problem on a new sheet of paper and do not write on the back of the sheets (I do not care about saving paper!). Submit everything in problem order. No sharing of calculators. Good luck and be sure to show your work. A T-score table is included on the last page of the exam.

Problem #1

Answer the following short-answer questions about traffic characterization:

a) What is autocorrelation and what does it meausure?

`Autocorrelation measures the relationship between values in a time series serparated by a specified lag (number of values). Autocorrelation ranges from -1 for complete negative correlation to 0 for completely independent to 1 for complete positive correlation.`

b) What is the significance of autocorrelation in traffic to network performance?

`Autocorrelation affects queueing behavior. Typically an autocorrelated arrival process will result in much greater queue lengths (and thus greater amounts of packet overflows than expected at routers, etc.) than will an independent arrival process given the same moments (e.g., same mean rate of arrivals).`

c) What is the key contribution of the Jain "Packet Trains" paper?

The key contribution was to show that network packet traffic is not Poisson (i.e., interarrival times are not exponentially distributed). Instead, packets arrive in correlated bursts than Jain characterized as "packet trains".

d) What is the key contribution of the Leland et al. Bellcore paper ("On the Self-Similar Nature of Ethernet Traffic")?

`The key contribution was to show that network packet traffic was self-similar (long range dependent, fractal) and to show that such traffic does not "smooth out" when aggregated.`

Problem #2

Write a C function to return a random variable from an empirical distribution based on the following measurements.

```
3, 3, 4, 2, 3, 4, 4, 4, 1, 2, 3, 2

int func(void)
{
  double z;

  z = (double) rand() / RAND_MAX;

  if (z <= 1.0 / 12.0) return(1);
  if (z <= 4.0 / 12.0) return(2);
  if (z <= 8.0 / 12.0) return(3);
  else return(4);
}
```

Problem #3

Answer the following short questions about simulation and random number generation:

a) What is a simulation model?

A software program that models the imporant aspects of a system under study. Or, "the dicipline of designing a model of an actual or theoretical physical system, executing the model on a computer, and analyzing the execution output.: (Fishwick, 1989).

b) What are the three phases of a simulation study?

Design, Execution, and Output Analysis. Each phases is interelated with every other phase.

c) List five key components of a simulation. Describe each component in one or two sentences.

Components are (pick five):

- System state - variables to describe the system at a certain time
- Simulation clock - Veriable giving current simulation time
- Event list - List containing the next time when each type of event will occur
- Event routine - Subprogram that updates system state on each event occurrence
- Statistical counters - Variables used for storing statistical information about system performance
- Library routines - random variate generators
- Report generator - outputs a report
- Initialization - routine to initialize all variables
- Main program - invokes timing routing and report generation

d) List the four desired properties of a random number generator

The four desired properties are:

- Numbers produced must be uniformly distributed and exhibit no correlation
- Generator must be fast and not require much storage
- Random number stream must be reproducible
- Must be able toproduce separate streams of random numbers

Problem #4

Attachment #1 is a CSIM program. Give the execution output.

The execution output is:

```
*** BEGIN ***
(1) at   1.00
(2) at   1.25
(2) at   1.50
(2) at   1.75
(1) at   2.00
(2) at   2.25
(2) at   2.50
(2) at   2.75
*** END (at time = 3.000000) ***
```

Problem #5

Attachment #2 is the generator process for the load balancing program we discussed in class. Add the necessary code to model displatch to shortest queue load balancing. In the case of a tie (i.e., both system1 and system2 have the same number of customers, you should randomly choose the destination queue).

The necessary code is bolded in the below…

```
//==============================================================================
//=  Process for a traffic source and load balancing                           =
//==============================================================================
void generate(double lambda, double mu1, double mu2)
{
  int   index;        // RR poller index
  int   len1, len2;   // Queue lengths

  create("generate");

  // Do forever, system() queues a customer (with random load balancing)
  index = 0;
  while(1)
  {
    hold(expntl(1.0 / lambda));
    if (BALANCE == RANDOM)              // Random load balancing
    {
      if (prob() < 0.50)
        system1(mu1);
      else
        system2(mu2);
    }
    if (BALANCE == RR)                  // Round robin load balancing
    {
      if (index == 0)
        system1(mu1);
      else
        system2(mu2);
      index = (index + 1) % 2;
    }
    if (BALANCE == SHORT)               // Shortest queue balancing
    {
      len1 = qlength(Server1) + num_busy(Server1);
      len2 = qlength(Server2) + num_busy(Server2);
      if (len1 < len2)
        system1(mu1);
      else if (len2 < len1)
        system2(mu2);
      else
      {
        if (prob() < 0.50)
          system1(mu1);
        else
          system2(mu2);
      }
    }
  }
}
```

Problem #6

You have simulated two systems for five replications each. The sample means for the response times (in millisec) from system #1 are 110, 105, 110, 108, and 102 and from system #2 are 110, 90, 100, 100, and 101. Can you with 95% confidence state that either system #1 or system #2 is better (better is lower mean response time)? Can you make such a statement with 90% confidence? Be sure to show your work.

We first find D to be 0, 15, 10, 8, and 1. The mean of D is 6.8 and the sample standard deviation of D is 6.301. From the T-score table for 4 degrees of freedom we find the $t_{\alpha/2}$ for 95% CI is 2.78 and for 90% CI it is 2.13. Thus, H95 (half-width for 95% CI) is 2.78 * (6.301 / sqrt(5)) = 7.833 and H90 = 2.13 * (6.301 / sqrt(5)) = 6.002. Thus, for 95% CI we have the true mean in a range of (6.8 - 7.833) to (6.8 + 7.833) which crosses zero, so we cannot say with 95% confidence that one system is better than another. For 90% CI we have the true mean in a range of (6.8 - 6.002) to (6.8 + 7.833) which is entirely above zero, so we can say with 90% confidence that system #2 is better (lower mean delay).

Problem #7

This semester you learned a toolkit essential for research in computer networks. List the key things that you have learned. Describe how they can be used in research - and also in industry - in the area of computer networks. Give the names of a few key people in the area of performance evalution of computer networks.

The key topics learned this semester were measurement and modeling in the scope of performance evaluation. In the area of modeling you learned both analytical (using operational laws and Markov modeling) and simulation methods. Queueing theory is at the heart of most of the tools. These tools are key to research and are used to evalute new protocols and architectures. Research can also be in the area of the toolkit itself, that is in finding way to create better methods of performance evaluation. In industry, performance evaluation is used in capacity planning to optimize the use of scare resources (not too much resource and wasted money, not too little resources and lost customers). Some key people include Kleinrock, Leland, and Jain. There are many others.

T-score table

$N-1$	t for $\alpha/2 = 0.050$	t for $\alpha/2 = 0.025$
4	2.13	2.78
5	2.02	2.57
6	1.94	2.45
7	1.90	2.37
8	1.86	2.31
9	1.83	2.26
10	1.81	2.23
11	1.80	2.20
12	1.78	2.18
13	1.77	2.16