CDA6530: Performance Models of Computers and Networks (Fall 2013)

Project 3: Worm Propagation Discrete-time Simulation
(Assigned Oct. 24th; due: Nov. 11th midnight via webCourse)

You are required to simulate simple worm propagation in a medium-scale network by using discrete-time simulation technique.

Assume that in an isolated network with Ω =100,000 IP address space (IP addresses can be treated with value from 1 to 100,000), there are N = 1000 vulnerable computers to a particular worm in this network. These vulnerable computers have specific IP addresses: 1, 2, 3,…, 10,   1001, 1002, …., 1010,   2001, 2002, …, 2010, ….. In other words, each cluster of 10 computers with the consecutive IPs are vulnerable to the worm, and in every 1000 IP addresses there will be one cluster of 10 vulnerable computers (so there will be 100 clusters overall).

Now the worm starts its infection within this network from 1 initially infected machine (randomly picked from those vulnerable computers). At each discrete time unit, a worm-infected computer can scan η randomly picked IP addresses within this network (the network has Ω IP addresses). If it finds a vulnerable computer, it infects the vulnerable computer and this newly infected computer can start infecting others from the next discrete time (no delay is considered, or we can say the delay is assumed to be one time unit).

For such a worm propagation, we have introduced that it can be modeled by:

$$dI(t)/dt = \frac{\eta}{\Omega} I(t)[N - I(t)]$$

Where I(t) is the number of infected computers at time t.

Your assignments are:

1). Simulate a random-scanning worm propagation.
Simulate a worm propagation with η =2.   You need to simulate the worm propagation for 100 runs in order to get the average values for I(t) for each discrete time t. Each of your simulation run should end when all vulnerable machines have been infected.

a). Draw a figure to compare the I(t) derived from the simulations (averaged value or called sample mean) and the above differential equation (i.e., the figure contains two I(t) curves). They should be matched with each other (with some statistical errors, and the average I(t) curve from simulation should be a little slower than the analytical curve). The numerical result of the differential equation above can be derived by Matlab Simulink.

b). Draw a figure shows the I(t) from the first 3 simulation runs. This figure will exhibit the statistical variance in worm propagation process (in each simulation run the worm's propagation dynamic is slightly different).

2). Simulate a local-preference scanning worm propagation.
The above random scanning has been used by most Internet worms, but some other worms have used local-preference scanning (such as Code Red II). You need to simulate such a worm: for an infected computer with IP value x, each time it picks an IP y trying to infect, the value y is computed as:
(1). With probability p = 0.4, it picks a value y such that y $\in$ [x-10, x+10] and y $\neq$ x.
(2). With the remaining probability 0.6, it picks a random value y between 1 to 100,000.

Again simulate the worm propagation with η =2 for 100 runs to obtain the average result I(t), then draw one figure showing the comparison between this I(t) with the I(t) of the random-scanning worm simulation from Question 1. Which worm spreads faster? Can you explain the reason for the speed difference?

Submission: Please submit your report document (a word file or PDF file), and your simulation codes. You can use Matlab or C or java to program your simulation code. Your report should explain how you design your simulation, what important variables you used in your code. Explain the meaning of each figure you draw in your report document. Also specify what each simulation code does.

In addition, your result figures should be clearly readable in black-white printout (as I explained in class), which will be counted in your grade!