COS 116
The Computational Universe
Laboratory 9: Virus and Worm Propagation in Networks

You learned in lecture about computer viruses and worms. In this lab you will study virus propagation at the quantitative level. We can also study, to a certain extent, the spread of human disease since it has similar properties as computer viruses. You will first use a simple simulator application to gather data and then use Microsoft Excel to visualize it.

The application can model a network of interconnected computers similar to the Internet but with fewer (up to 100,000) computers. When you set all relevant parameters and click "Run Simulation," it simulates the spread of a virus/worm in the network and outputs some numbers for you to analyze.

Hand in your lab report at the beginning of lecture on Tuesday, April 22.

Turn in only the Excel charts that are followed by a star (*). Do not turn in the charts without a star; they are just for you to visualize the simulation. **Make sure each chart is labeled with the Experiment number and Question number, as well as any relevant additional information (e.g. particular options). You will not receive credit for charts that are unlabeled. Also, turn in answers to all questions posed in the body of the lab and in the "Additional Questions" section.**

Introduction: Using the Simulator

1. **Download this file to your Desktop:**
http://www.cs.princeton.edu/courses/archive/spring08/cos116/lab9_files/vsim.exe

2. **Double-click the file to run the simulator. Ignore any security warnings.** The simulator has a variety of options to adjust. These options are explained below.

Note that the simulator uses randomnes, especially when using the social network settings, so your results may not look exactly like your neighbor's.

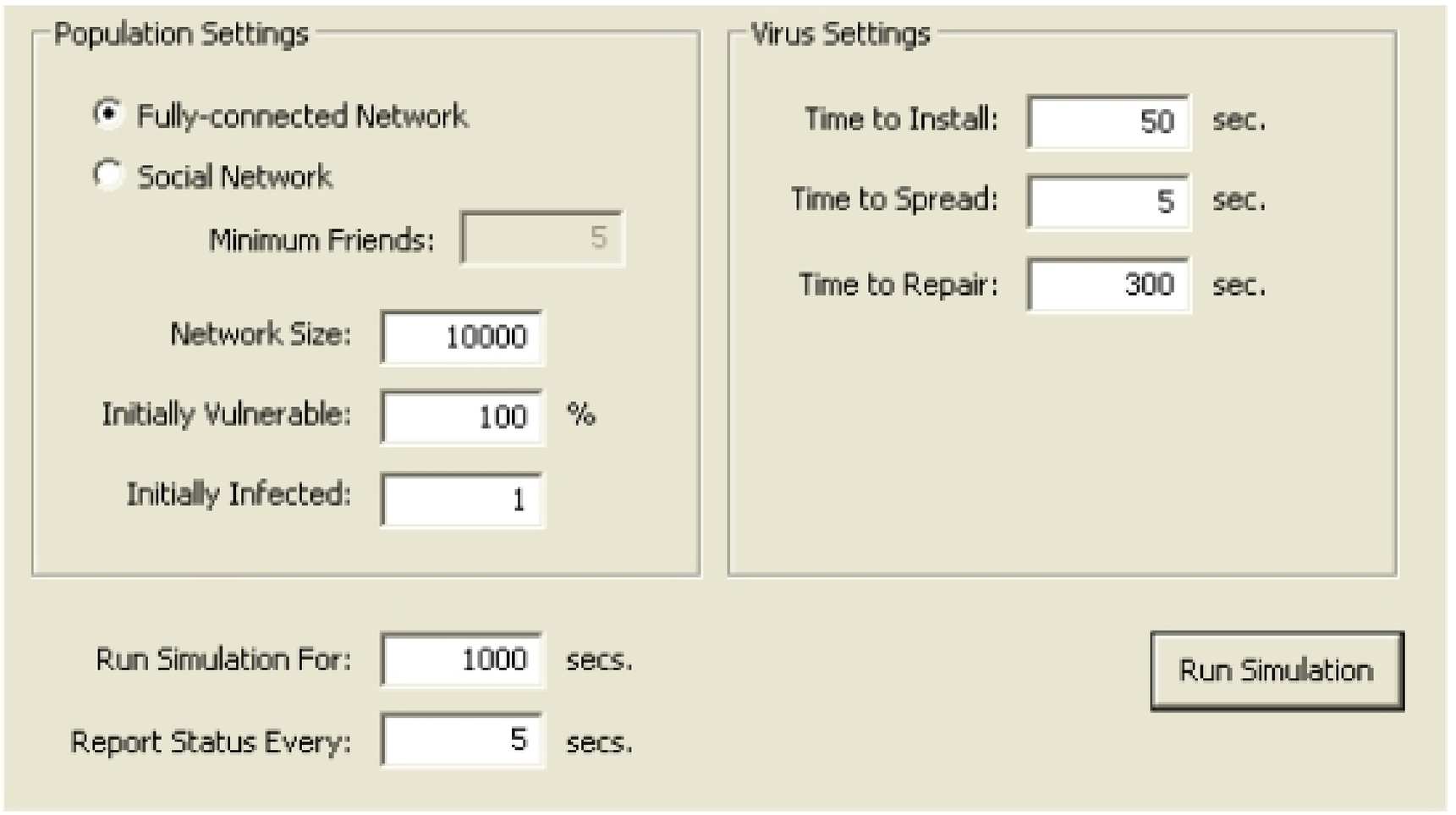

Figure 1: Simulator interface

Here is a brief explanation of all adjustable parameters.

- **Network size**: The number of nodes in the network. This can be interpreted as the number of computers in a computer network or the number of people in a social network.
- **Network type**: In a fully-connected network, every node is directly connected to every other node. In a social network the connections are generated according to a randomized scheme so that the resultant network resembles actual social networks that have been studied. Though the details of the social network are interesting (feel free to ask your TA questions) they are not relevant for this lab. You just need to know that each node is directly connected to at least the number specified in "Minimum Friends" though some nodes may be connected to much more than this minimum. (In social networks, these are the popular people.)
- **Initially vulnerable**: The percentage of computers that is initially vulnerable to infection. This models the possibility that some computers may already be patched against a weakness, or that some people are immune to a disease.
- **Initially infected population**: The number of computers that are initially infected.
- **Time to install**: The amount of time before an infected computer begins to spread the virus to other computers.
- **Time to spread**: The amount of time between infection attempts by an infected computer that is spreading the virus.
- **Time to repair**: The amount of time from infection until the virus is removed and the computer is immunized.

How infection works:

An infected computer spreads the virus by randomly choosing a computer to which it is directly connected, every "Time to Spread" seconds. If that computer is vulnerable, then it will begin spreading the virus "Time to Install" seconds later. However, the good guys may rush software patches to remove the virus/worm. The "Time to Repair" is the time after which the computer is considered repaired -- no longer spreading the virus/worm, nor susceptible to infection again.

Interpreting the results using Excel:

The simulator outputs periodic counts of vulnerable, infected, and repaired computers. You can save this output to a file of type CSV ("Comma Separated Values"), which is a standard format for data files. To graph the data from a CSV file in Excel:

1. Double-click the CSV file to open it in Excel.
2. Use the mouse to select all the columns and rows that have data (or, to do this with the keyboard, press Ctrl + A).
3. From the Insert menu, click the Scatter chart type "with only markers."
4. Select File … Save As, and save the file as a *.xls file (not as a CSV file).

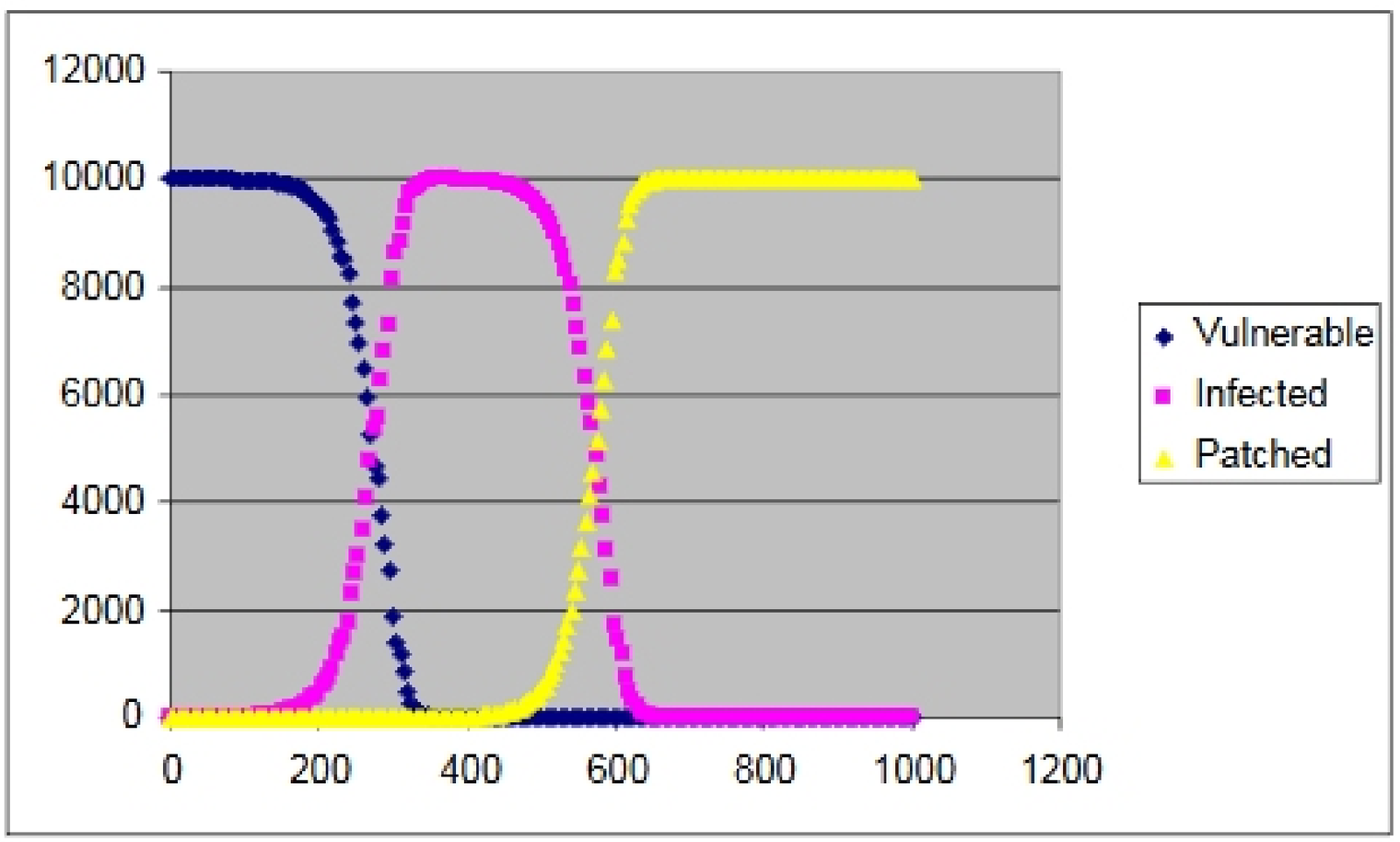

Figure 2: Example chart