Network Science (Spring 2017)

Homework Assignment #4

Due: 08:00 @ June 4 (Sunday), 2017

To start with this homework, first find a network that represents the structure of the Internet at the level of autonomous systems. You can use Network #29 with 22963 nodes in the paper used for Homework #2 or refer to the recent dataset from CAIDA. As usual, specify clearly where the dataset is from and provide some basic statistics of the network before proceeding with the following simulations.

1. (15%) To investigate the impact of router failures on the Internet, experiment with uniform removal of nodes (routers). You can refer to the implementation in graph-tool or the algorithm in this paper on how *site percolation* can be effectively simulated. Perform the simulation for several times (Monte Carlo simulation) and plot the (average) size of the largest component versus the number of nodes removed. What have you observed from the results?

2. (15%) Now perform non-uniform removal of routers on the Internet. Specifically, remove routers in descending order of their degrees (i.e. the router with the highest degree is removed first). Plot the size of the largest component as before and compare against the results obtained in Problem 1. Elaborate your observations on the such targeted attack.

3. (20%) To simulate the impact of other targeted attacks on the Internet, experiment with different types of non-uniform node removals. For example, you can find out the betweenness centrality of each router and then remove them in descending order of their betweenness centrality. Come up with ***two*** different non-uniform node removal approaches, and then compare the percolation results. What is the most effective approach (that can decrease the size of the largest component by removing the smallest number of routers) that you can come up with to attack the Internet based on your experiment?

4. (25%) In addition to percolation, propagation of epidemics is also one interesting process to investigate for complex networks. Assume that a computer worm spreads from one randomly chosen router on the Internet, and routers are defenseless against the worm yet. Use the SI epidemic model with parameter β (refer to Slide #7 in the lecture slides) to simulate the spread of the worm on the Internet. (You can use existing tools for this task by specifying which package you use; if you write the simulation code by yourself, explain in details how you implement the SI model.) Perform the simulation for several times and plot the (average) proportion of infected routers as time progresses for one β value. Plot different curves in one figure, where each curve corresponds to one β value chosen. Elaborate your observations from the results. How would you relate "bond percolation" with "epidemics" in this case?

   Bonus (20%): Compare your simulation results against first-order and second-order analytical results (refer to Slide #20 in the lecture slides). Explain in details how you obtain the analytical results from the given network.

5. (25%) As a computer genius who has designed the computer worm, you aim to upgrade the

worm that is “intelligent” to infect its neighbors with different infection rates depending on the “significance” of each neighbor. Come up with an infection strategy for the worm such that it can spread faster than the uniform case in Problem 4. In other words, each link now has its own β value and you aim to come up with a distribution of β for your chosen metrics (e.g. node degree). For fair comparison, the average infection rate should be the same for the cases of uniform and non-uniform infections. Simulate your new worm on the same network and compare the results with those obtained in Problem 4. How good is your worm in spreading itself to the whole Internet compared to an ordinary worm?