Assignment of ET 4389

February 24, 2016

Consider the network whose index is the same as the last digit of your student number. All the networks in the attachments are given in the following format: each row ”a b” denotes a link from a to b. By default, we assume each network is undirected, thus, if a is connected to b, b is also connected to a. Let G be the network you get. Compute the following network properties of your network.

1) What is the number of nodes N, the number of links L, the link density p, the average degree $E[D]$ and the degree variance $Var[D]$? Plot the degree distribution. Does the degree distribution follow a power law distribution $\Pr[D = k] \sim ck^{\gamma}$? If so, what is the power exponent γ (plot the fitting curve)?

2) What is the degree correlation (assortativity) ρ_D? What is its physical meaning?

3) What is the clustering coefficient C?

4) What is the average hopcount $E[H]$ of the shortest paths between all node pairs? What is the diameter $H_{\max}$?

5) What is the largest eigenvalue (spectral radius) λ_1 of the adjacency matrix?

6) What is the second smallest eigenvalue μ_{N-1} of the Laplacian matrix (algebraic connectivity)?

Consider the real-network NetSci G_N (with index 10).

7) Compute all the metrics asked in the previous questions (excluding the degree distribution) for this NetSci network G_N.

8) Assume that your network G assigned earlier and the real-world network G_N are the two possible designs for a communication network. Discuss each metric ($C, E[H], H_{\max}$,λ_1, μ_{N-1}) individually: what is its physical meaning with respect to a communication network, which design is the better and why. Taking all the metrics into account, which design would you recommend?

9) Generate 100 instances of Erdős-Rényi (ER) random graphs with the same number N of nodes and the same link density p as the network G assigned to you . Plot the degree distribution of these 100 network instances and the theoretical degree distribution. [Hint: the probability that two nodes are connected in an ER random network equals the link density of G]

10) Computer all the metrics asked in 2)-6) for these 100 ER instances and present in your report only the average of the each metric over the 100 ER random networks.

11) We are going to evaluate which of the three types of networks (the network assigned to you, Netsci and ER random networks) may allow information to propagate to a larger fraction of the network. A simple information propagation process on a network can be described by a SIR model: initially 20 nodes are chosen uniformly at random to be infected (I), i.e. possessing the information, whereas the rest nodes are susceptible. Each infected node could infect each susceptible neighbor independently in the subsequent time step with probability $\beta = 0.3$ and afterwards becomes resistant (R). Hence, each newly infected node has only one time step to attempt to infect its susceptible neighbors. The process stops or equivalently the system reaches the steady state when the number of R nodes does not increase over time any more. The number of R nodes in the steady states n_{R_∞} is the number of nodes that have ever been infected, implying the prevalence of the information propagation. a) simulate the SIR process for 100 iterations on the network assigned to you. What is the average number $E[n_{R_\infty}]$ of resistant nodes in the steady state? b) Simulate the SIR process for 100 iterations on the Netsci network. What is the average number $E[n_{R_\infty}]$ of resistant nodes in the steady state? c) Simulate the SIR process on ER

random networks for 100 iterations. Note that within each iteration, an ER random network is generated, upon which the SIR process is further deployed. What is the average number $E[n_{R_\infty}]$ of resistant nodes in the steady state? Compare the $E[n_{R_\infty}]$ of these three types of networks. What is your observation? What are the possible reasons leading to such an observation? Can we conclude which type of networks better facilitate information propagation based on the simulation results obtained so far? If so, what can you conclude? If not, what can be done further to conclude?

Hint: A table with all metrics computed for the network assigned to you, the Netsci and ER networks (the average over 100 ER random networks) is recommended.