CPSC 619-600: Computer Networks
Homework 5 due April 13, 2017

1 Purpose

This homework extends the $G(n, p)$ model to exhibit characteristics of small-world graphs.

2 Description

2.1 GED

This problem studies how well GED achieves heavy-tailed degree and whether its clustering matches that of the Internet. Before you build the graph, you need to generate n weights w_i from the Pareto distribution with CCDF:

$$P(w_i > x) = cx^{-\alpha} = (x/\beta)^{-\alpha}. \quad (1)$$

Next, simulate several GED graphs with $n = 5,000$, in which the degree follows a Pareto distribution with $\alpha = 2$ and $c = 1.5$. Plot the tail of the degree distribution and use Excel's curvefit to understand how close the generated sequence follows the one you attempted to achieve. Examine 1) the probability that GED is connected; 2) the average size of the largest component (averaged over 100 iterations); and 3) the evolution of the clustering coefficient $\gamma(G)$ as $n \to \infty$ (vary n between 100 and 10,000). Use curvefit to figure out the exact rate of decrease.

Implementation notes: it is possible that p_{ij} in GED is above 1. In such cases, you need to cap it at 1.

2.2 BA

This problem analyzes the same metrics in the BA model. Generate several BA graphs and compare the distribution of observed degree to the one you attempted to build.[1] Using the procedure in the previous problem, examine what happens to clustering $\gamma(G)$ as $n \to \infty$. Examine two different cases with $m = 2$ and $m = 3$ (keep m_0 equal to 3).

[1] Even though the theoretical value of α is 2, your average α will be close to 1.8.

2.3 Extended BA

Observe that clustering can be increased if the existing nodes are allowed to form links between themselves without the addition of new nodes. Using this insight, create an algorithm that randomly decides at every step t whether to add a new node (like in BA) or to add some links to the existing graph. In other words, let q be the probability of adding a new node with m links and $1 - q$ be the probability of adding m links between the *existing* nodes.

Work out the details of this algorithm to the best of your ability, choose several values of q and generate the corresponding random graphs, and finally show their degree distribution and clustering $\gamma(G)$. Compare the obtained metrics to those in BA. Discuss $\gamma(G)$ as $n \to \infty$.