Patrick Austin

CS 491: Social Networks: Lab 3

7 March, 2018

**Similarity**

3.7.10 (a)

I assume a node is not considered to be in its own neighborhood for purposes of calculating similarity. This approach was discussed in class but I have opted not to use it here. (If used, then the similarity by both metrics is 1, as v4 and v5 will have identical neighborhoods of {v3, v4, v5, v6}.)

Jaccard similarity =

|{v3, v5, v6} ∩ {v3, v4, v6}| / |{v3, v5, v6} ∪ {v3, v4, v6}| = 2 / 4 = 0.5

Cosine similarity =

|{v3, v5, v6} ∩ {v3, v4, v6}| / ( |{v3, v5, v6}| \* |{v3, v4, v6}|)0.5 = 2 / (3\*3)0.5 = 2/3 = 0.6666

**Network Models**

4.7.4

Phase transition occurs in random graphs when average node degree c equals 1. At that point the giant component, which began to appear as c approached 1, begins to grow. The diameter of the graph, which has grown to its maximum value as c approached 1, begins to decrease.

4.7.7

Random graphs are poorly suited to modeling real-world graphs due to having several characteristics that do not reflect the characteristics of real-world graphs. While random graphs have similar average path lengths to real-world graphs, proportional to log(n), the degree distribution follows a binomial distribution, not a power law distribution as in real-world graphs. Random graphs also underestimate local clustering compared to real-world graphs.

Random graphs, regular lattices, and small-world models are three different types of graphs generated according to different rules.

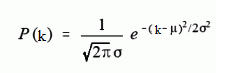
A random graph assumes that edges between nodes are formed at random. In the G(n,p) model, each of the possible edges among the n nodes is formed with probability p. In the G(n,m) model, a graph with n nodes and m edges is selected at random from all the possible graphs with n nodes and m edges. Random graphs have the properties described above.

A regular lattice is a regular network with a certain pattern of ordered connections between nodes. For example, in a ring, a type of regular lattice, each node might connect to its next-nearest neighbors on either side. This type of regular lattice has equal degree for all nodes, a high, fixed clustering coefficient, and high average path lengths.

Small world models begin with a regular lattice, but randomly rewire edges at a specified probability p to create a graph with different properties. If p=1 the graph behaves like a random graph, and if p=0 the regular lattice remains unchanged. At “sweet spot” p values of approximately .01-.1 a small world graph emerges with a high clustering coefficient and low average path lengths. However, the degree distribution remains binomial and not a power-law distribution, so they still serve as imperfect models of real-world graphs.

4.7.9

To get the fraction of pages with k inlinks assuming inlinks are governed by a normal distribution, we expect a probability density function along the lines of



Where μ is the mean inlink value and σ is the standard deviation around that mean.

If inlinks are governed by a power law instead, we expect a probability density function along the lines of

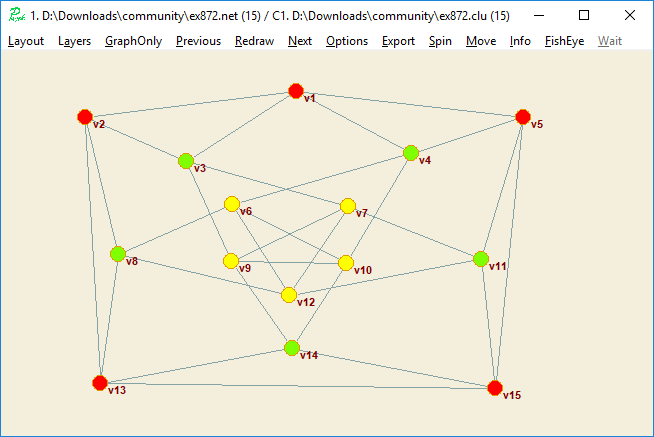
P(k) = ak-b

Where a is a constant value and b is the power-law exponent, typically in the range [2,3].

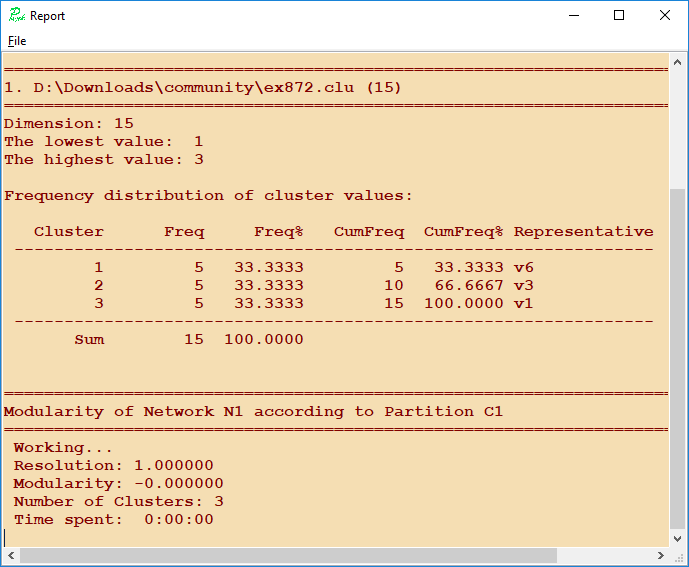
**Assortativity**

8.7.2 (c)

I opted to use the Pajek graph software linked on the course website for this problem. Input files are available upon request. Here is a visualization of the data after readin, with nodes partitioned according to the specified categories:



And here is the calculated result for modularity of that graph, using a resolution value of 1:



Therefore the modularity of the graph for the provided category labels is 0.

This result makes intuitive sense: all the nodes have 4 edges and there are 3 classes, so in a non assortative graph we would expect ⅓ \* 4 = 4/3 of the average node’s edges to go to nodes of the same class. When averaged over all nodes in the graph this is exactly what we see, since the ten a class and c class nodes all have edges to two like nodes, while the five b class nodes have no edges to other b class nodes. Thus there are 20 homogenous edges averaged over 15 nodes, so we see the 20/15 = 4/3 homogenous ratio we would expect from a non assortative graph with modularity of 0.