Feedback — HW 5 small worlds

You submitted this homework on **Fri 16 Nov 2012 8:28 AM PST**. You got a score of **10.00** out of **10.00**.

Question 1

Download Lada's Facebook network. Load in Gephi (as undirected). Note the number of nodes and edges present. Calculate the clustering coefficient and average shortest path (this is OK to take as-is even though the network is not connected). Next close the project, and after you have blank slate, generate an Erdos Renyi random graph (File > Generate > Random graph...) with the same number of nodes and edges (you'll have to figure out the corresponding wiring probability to achieve this). It will produce a directed network. Calculate the clustering coefficient and average shortest path for this random network, making sure to treat the network as undirected. Which of the following observations are true (check all that apply).

Your Answer		Score	Explanation
☐ the average shortest path is longer than 3 hops	✓	0.33	The actual shortest path of 2.8 is slightly higher than the 2.3 of an equivalent random graph.
☐ the topology of Lada's network does not satisfy the definition of a small world network	✓	0.33	with clustering higher than that in the equivalent random network and average shortest path just about the same, Lada's network definitely has a small-world topology.
☑ the random graph has fewer connected components than Lada's actual	✓	0.33	At this density of edges, it is improbable that a node would be left out of the giant component (it occurs in Lada's network because some nodes' contexts are entirely separate and remote from the rest of her

network			network)
☐ the average clustering coefficient in Lada's egonetwork is lower than 0.4	✓	0.33	The average clustering coefficient is 0.497
☑ the average shortest path is slightly higher than that in the equivalent random graph	✓	0.33	True, the actual shortest path of 2.8 is slightly higher than the 2.3 of an equivalent random graph.
☑ the average clustering coefficient in Lada's egonetwork is higher than 0.3	✓	0.33	The average clustering coefficient is 0.497
Total		2.00 / 2.00	

Question Explanation

Lada's Facebook network has the characteristic structure of many ego-networks: presence of communities + occasional ties between communities)

Question 2

Download the a snapshot of the Gnutella peer-to-peer filesharing network (now over a decade old). Go through the same procedure: Load in Gephi (as undirected). Note the number of nodes and edges present. Calculate the clustering coefficient and average shortest path (this is OK to take as-is even though the network is not connected). Next close the project, and after you have blank slate, generate an Erdos Renyi random graph with the same number of nodes and edges (you'll have to figure out the corresponding wiring probability to achieve this). It will produce a directed network. Calculate the clustering coefficient and average

shortest path for this random network, making sure to treat the network as undirected. Which of the following observations are true (check all that apply).

Your Answer		Score	Explanation
☑ The clustering is low, although still higher than that of the equivalent random graph.	✓	0.40	The clustering of the gnutella network is 0.011 while the equivalent random graph has a clustering of ~0.002.
☐ The gnutella network is well modeled by the Watts-Strogatz small world model.	✓	0.40	No, the model does not capture the presence of hubs and the clustering coefficient for the Gnutella network is nowhere near as it is typically in the WS model (unless you turn up the rewiring probability to be extremely high).
☐ The clustering coefficient in the gnutella network is lower than that of the equivalent random graph.	✓	0.40	Not true. Although the clustering coefficient is low, at 0.011 it is about 5 times greater than the equivalent random graph.
☐ The average shortest path of the gnutella network is longer than that of the equivalent random graph.	✓	0.40	Its hubs make the gnutella network an ultra-small world network. This is a known property of scale-free networks. The average shortest path is near constant rather than growing with the number of nodes.
☑ The average shortest path of the gnutella network is shorter than that of the equivalent random graph.	✓	0.40	Its hubs make the gnutella network an ultra-small world network. This is a known property of scale-free networks. The average shortest path is near constant rather than growing with the number of nodes.
Total		2.00 / 2.00	

Question Explanation

The gnutella peer-to-peer filesharing network has a scale-free structure.

Question 3

Now consider Lada's clustering coefficient (the clustering coefficient of her node, where she is connected to each of her friends). What is it, approximately?

Your Answer		Score	Explanation
◉ 0.05	✓	2.00	clustering coefficient = e/n/(n-1) * 2 = 3598 / 388 / 387 * 2 = 0.048
Total		2.00 / 2.00	

Question Explanation

Lada's clustering coefficient is lower than the average clustering coefficient of nodes within her egonetwork. This is likely because each of those nodes is friends with others in the same cluster/community, while Lada knows these nodes through different contexts and so not every node is likely to know every other.

Question 4

Run a rewiring algorithm on a ring lattice. Do not apply the layout algorithm until the very end, because the position of the nodes on the circle is needed to compute the cost of wire. Read the description of the model, and experiment with different wire costs as well as trying to find the optimum (minimum energy) configuration for each wire cost setting. The model works best if you give it a long time to run. The 'find optimum' button will automatically first increase and then decrease the temperature to try to find a global optimal solution. You may want to follow this up with just a long period of 'rewire' to continue the process. Also, try increasing the speed with the slider to have the search occur faster.

Now for the question: As the cost of wire increases which of the following is true?

Your Answer	Score	Explanation

◉ edges become more localized	✓	2.00	localized edges are shorter and so more likely to occur when the cost of wire is higher.
Total		2.00 / 2.00	

Question Explanation

As the cost of wire decreases, long range links become more affordable and work to decrease the average shortest path.

Question 5

Skim through the paper by Liben-Nowell et al. on geographic routing in social networks inferred from LiveJournal. Based on the paper, which of the following is true (check all that apply):

Your Answer		Score	Explanation
☐ Applying a simple greedy strategy (pass the massage to your friend who is closest to the target geographically) would result in extremely long path lengths.	✓	0.50	such search is shown in the paper to be efficient in the LiveJournal network.
☑ What matters in navigation is that the probability of being acquainted with individual X depends on how many others live closer to you than than X.	✓	0.50	the paper demonstrates this by ranking each individual according to distance with respect to their friend
☑ Empirically, the probability that two individuals are acquainted falls off as 1/(distance) based on LiveJournal data.	✓	0.50	True, this is a discrepancy with a theoretical prediction valid for a regular 2 D lattice
☐ Population density has no bearing on the probability that two people living within a	✓	0.50	In rural areas, knowing everyone

certain distance know one another.		who is a mile away is very likely. In Manhattan it is impossible.
Total	2.00 / 2.00	

Question Explanation

There is a discrepancy between what the theory (i.e. Jon Kleinberg) predict would be the optimal probability of two individuals sharing an edge as a function of their geographical separation, and what is measured for the LiveJournal network. Yet the LiveJournal network is navigable. The explanation is that our locations are not well described by a regular 2D lattice.