Networks: Theory and Application
Assignment 2 (Due Friday, Aug. 3rd )

(Please submit as pdf attachment to the TA)

Please include an image in your assignment where you see the mark [*I*]. Otherwise there is no need to do so. Pajek can export the network visualization as EPS/Bitmap files (Export → 2D → ... ). You don’t have to solve the problems labeled with “[Extra Credit],” but by doing them you’ll get more experience of network analysis and also get extra credits.

1. **Network Centrality** [20pt]

![Graph](image1)

Find the centrality of vertex D and F in the left graph based on the following measures:

(a) Closeness
(b) Normalized Closeness
(c) Betweenness
(d) Normalized Betweenness
(e) Which network has a larger centralization? Left or right?

2. **Network Prestige** [20pt]

![Graph](image2)

(a) Label the influence range of the vertex C in the graph above (List the vertices in its input domain)
(b) Which vertex is in D’s input domain but not in C’s? Which vertices (or vertex) are in neither of their input domains?

c) Find the proximity prestige of C.

d) Find the normalized betweenness of E.

e) Is A a better hub (as opposite to authority) or E? Explain your answer.

3. Small World Networks [20pt]

(a) Which of the following graph has the highest clustering coefficient? (You don’t need to actually compute it!) [5pt]

(b) We introduced two ways of computing local clustering coefficient of a vertex. In one way, \( C(i) = \) density of the subgraph defined by i’s neighbors; in the other way, \( C(i) = \frac{\text{number of closed triplets centered on i}}{\text{number of connected triplets centered on i}} \). Will they always give you the same result? If yes, provide a proof. If not, give a counter example. (Assume that the network is undirected and there is at most one edge between any two vertices). [10pt]

(c) We introduced two variants of the Watts/Strogatz small world model (edge rewiring and edge addition). Which model leads to the larger clustering coefficient? The larger average path length? [5pt]

4. Centrality and Small World in Pajek [40pt]

Retrieve the Weibo social network from last assignment. Use a larger network (e.g., my facebook network) if your network has too few nodes.

(a) Compute the betweenness of vertices using Pajek. Visualize the values using Draw → Draw Vector [*I*]

(b) Compute the closeness of vertices. Visualize the values [*I*].

(c) Do these two measures agree with each other? Name the vertices with large closeness but small betweenness. Then name the vertices with small closeness but large betweenness. Then compute the correlation of the two measures.

(d) Which of the two centrality measures is closer to the degree centrality?

(e) Compute the average clustering coefficient (Net→Vector→Clustering Coefficients→CC1) and average shortest path (Net→Paths between two vertices→Distribution of distances→From all vertices, and look in the report window).
(f) Construct a random network with the same number of nodes and average degree with your Facebook network. (Net→Random Network→Erdos-Renyi→Undirected→General). Visualize it. [*I*] Compute the average clustering coefficient and the average path length of this random graph.

(g) Is your weibo network a small world? Explain your answer.

5. **[Extra Credit - 20pt]** Watts and Strogatz defined the average clustering coefficient as the un-weighted average of the local clustering coefficient of each vertex. People usually use this as a global measurement of clustering coefficient of the whole network. However, it is arguable whether treating every vertex equally is a good idea. Are some vertices more important than others? Would a weighted average make more sense? We remember that there is yet another measure for the clustering of the whole network - the global clustering coefficient. Unfortunately it is different from Watts and Strogatz’s average clustering coefficient. But good things are usually connected. By redesigning a *weighted* average clustering coefficient, can you establish an equivalence between the average clustering coefficient (based on density of neighbor graphs) and the global clustering coefficient (based on triangles and triplets)? Explain your answer. (Hint: your answer for the question 3(b) may be helpful).