Social Network Analysis MATH-517: Homework 3
(Total 10 points - Covers Chapters 8 and 11)

1 [**1 point**] What (roughly) is the time complexity of:

a) Finding a word in a (paper) dictionary if the size of the input is the number n of words in the dictionary?

2 [**1.5 points**] Consider the following situations:

a) You are asked to calculate the closeness centrality of a single node in an undirected network with m edges and n nodes. What algorithm would you use to do this, and what would be the time complexity of the operation in terms of m and n?
b) You are given a road map and told the average driving time along each road segment, then you are asked to find the route from A to B with the shortest average driving time. What algorithm would you use to do this, and what would be the time complexity of the calculation?
c) What algorithm would you use to find all the components in an undirected network, and what would be the time complexity of the operation?

3 [**2 points**] What is the time complexity, as a function of the number n of nodes and m of edges, of the following network operations if the network in question is stored in adjacency list format?

a) Calculating the mean degree.
b) Calculating the median degree.
c) Calculating the air-travel route between two airports that has the shortest total flying time, assuming the flying time of each individual flight is known.
d) Calculating the minimum number of routers that would have to fail to disconnect two given nodes on the Internet.

Hint: While statement 1 should be straightforward to prove, statement 2 is a little more involved. It may be useful to consider a path from s to a node at distance $d+1$, and then consider the shortest distance from s to the penultimate node along that path.

4 [**2 points**] For an undirected network of n nodes and m edges stored in adjacency list format show that:

a) It takes time $O(n(m+n))$ to find the diameter of the network.

b) It takes time $O(\langle k \rangle)$ on average to list the neighbors of a node, where $\langle k \rangle$ is the average degree in the network, but time O(· k2·) to list the second neighbors.

5 [**1.5 points**]For a directed network in which in- and out-degrees are uncorrelated, show that it

takes time $O(m^2/n)$ to calculate the reciprocity of the network. Why is the restriction to uncorrelated degrees necessary? What could happen if they were correlated?

6 [**2 points**] Write a computer program in the language of your choice that generates a random graph drawn from the model $G(n, m)$ for given values of n and the average degree $c = 2m/n$, then calculates the size of its largest component. Use your program to find the size of the largest component in a random graph with $n = 1\,000\,000$ and $c = 2 \ln 2 = 1.3863\ldots$ and compare your answer to the analytic prediction for the giant component of $G(n, p)$ with the same size and average degree. You should find good agreement, even though the models are not identical.