CSC 485B/578B: Assignment 1

Due: 23:55 pm, May 26, 2014

Remarks

- Each question has two weights, with the first one for CSC 485B students and the second for CSC 578B students.
- You need to submit your solution in a **pdf** file to connex dropbox by the due time. Other format will not be accepted.
- For Question 2, you also need to submit your java source code to connex dropbox.

Question 1

There are two quantities to measure the distance between nodes in a network. One is the diameter of the network, which is defined as the maximum distance between any pair of nodes in the network. The other is the average distance, which is defined as the average distance over all pair of nodes in the graph. The two measures in most networks are close to each other. But in some situation they are quite different.

1. (15%, 15%) Describe an example of a network where the diameter is more than three times as large as the average distance.
2. (10%, 10%) Describe how you could extend your construction to produce networks in which the diameter exceeds the average distance by as large a factor as you would like, i.e., for every positive number c, can you produce a network in which the diameter is more than c times as large as the average distance?

Warning: Please carefully check the definitions in the course lecture notes before you work on this problem.

Problem 2

Implement a naive algorithm in Java to find the diameter of a network based on the Floyd-Warshall algorithm. Use your code to calculate the diameter of an anonymized personal Facebook network described with *Personal-May8-Anonymous.gephi*, which could be downloaded from Connex → Resource.

1. (20%, 15%) Submit your java source code into connex dropbox.

2. (15%, 10%) How many isolated nodes in the network? An isolated node means there is no link between this edge and other nodes.

3. (10%, 10%) Clearly, when a network is disconnected, its diameter is infinity. Nevertheless, when you use Gephi to calculate the diameter of *Personal-May8-Anonymous.gephi*, you will get a number which is obviously not infinity. Find out how the network diameter is calculated in Gephi. If your code returns a different result compared to Gephi, modify your code accordingly.

4. (0%, 15%) Assume that you are allowed to introduce four new edges in the network. Which four new edges would you introduce into the network so that clustering coefficient of your network represented by *Personal-May8-Anonymous.gephi* is minimized?. In practice, if you want your personal network to be a closely-knit group, the new edges captures the "problems" among your friends that you might want to put efforts to fix.

Note: (1) Please treat the network as a directed network. (2) You can build your code over the gephi platform or simply implement your standalone java program. The structure of *Personal-May8-Anonymous.gephi* is easy to understand if you open the file with a text editor.

Problem 3

(10%, 10%) Consider the social network represented in the following figure. Suppose that this social network was obtained by observing a group of people at a particular point in time and recording all their friendship relations. Now suppose that we come back at some point in the future and observe it again. According to the theories based on empirical studies of triadic closure in networks, which **new** edge is most likely to be present? Provide a brief explanation for your answer.

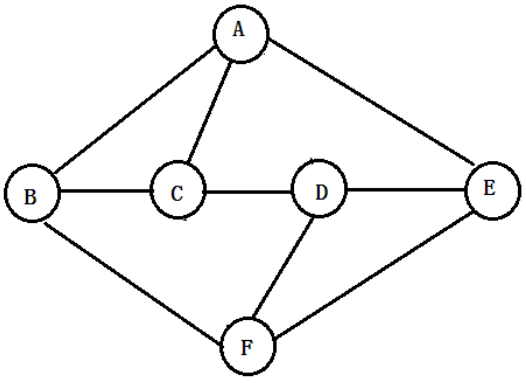

Figure 1: The network in Problem 3

Problem 4

(20%, 15%) Together with some anthropologists, you are studying a sparely populated region of a rain forest, where 50 farmers live along a 50-mile-long stretch of river. Each farmer lives on a tract

of land that occupies a 1-mile of the river bank, so their tracts exactly dived up the 50 miles of river bank that they collectively cover. The farmers all know each other, and after interviewing them, you have found that each farmer is friends with all the other farmers that live at most 20 miles from him or her, and is enemies with all the farmers that live more than 20 miles from him or her.

Build the signed complete graph corresponding to this social network. If the network structurally balanced or not? Explain your answer.