

CSE 112: Networked Life Problem Set 1 Solution

1. (15 points)

For each section, 1 point for stating name of the network that was *not* discussed in class and for its brief description in terms of its statics, dynamics and formation; 1 point for clearly stating vertices and 1 point for clearly identifying the relationship that determines an edge.

Social Network: High School Sex Network. Vertices are students in a high school and physical contact is the relationship that connects two vertices. Large connected component is possible as in the “Jefferson High School” case. Romantic attraction, spread of STD are examples of things happening on such a network. (3 points)

Content Network: iTunes music database where songs are vertices and a song is connected to another if the two are similar. Similarity can be defined based on genre, artist (or even signal attributes). The network grows as new songs are composed. Similarity based search, categorization etc. are things happening on this network. (3 points)

Business or Economic Network: Network of trade among countries. Vertices are countries and there is an edge between two vertices (countries) if there was trade between those two countries in the past year. Large industrialized nations are likely to have high degrees as they are expected to have export to many countries. The number of vertices in the network relatively remains constant while new edges are added when new trade relations are established between two countries. Dynamics of this network captures trade among countries. (3 points)

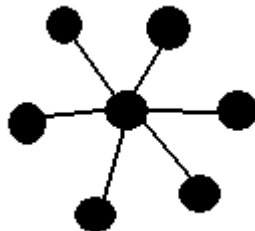
Physical Network: Telephone network (excluding cell phones). Vertices are telephones and switches. There is an edge between two vertices if there is a physical link between them in the telephone network. Vertices corresponding to switches are going to have heavy degree. New nodes are added as new telephone connections are started. Communication among people and control signals flow through such a network. (3 points)

Biological Network: Protein interaction network. Vertices are known proteins in the human body and there is an edge between two vertices (proteins) if there is interaction between them. Dynamics of such a network includes biochemical cascade reactions where protein-protein interactions are involved. (3 points)

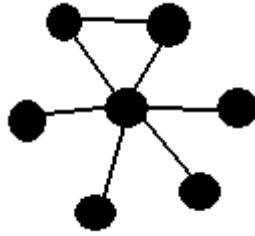
2. (10 points)

The worst-case diameter for all three cases is 2.

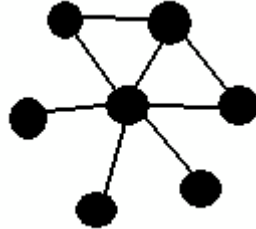
(a) 4 points: 3 points for answer of 3. 1 point if vertices put in a straight line with answer of 6.



(b) 3 points; 2 points for answer of 3.



(c) 3 points; 2 points for answer of 3.



3. (15 points)

Graphs generated using random connectivity models are known to lack structural properties that show up in most real-life social networks such as clustering and heavy-tailed distributions of degree. They do not take into account the idea that two nodes are more likely to be connected if they have many common neighbors. (10 points for this part, you need to mention at least two specific properties, e.g. clustering, heavy-tailed distribution or any of your own insights that makes sense, to receive full credit.)

One type of model that might better capture properties of real social networks is preferential attachment. Generating models using preferential attachment will yield a heavy-tailed degree distribution, but will not lead to clustering. (5 points for this part, full credit is given for any *clearly stated* model designed to solve at least one of the problems with random connectivity graphs.)

4. (15 points)

(a) Answers for this problem will, of course, vary from student to student. (1 point)

(b) 2 points each for exhibiting understanding of connectors, mavens, and salesmen, and describing how each of the three did or did not play a role in the spread of this trend. (6 points)

(c) Full credit for exhibiting understanding of Gladwell's definition of stickiness. (4 points)

(d) Full credit for any reasonable answer displaying understanding of tipping phenomena. (4 points)

5. (15 points)

(a) The largest number of people in this network is $1+10+10*9+90*9=911$, and the network is a *tree*. (7 points for this part, you receive partial credit if your answer is partially correct, e.g. stating

correctly that the network is tree but making mistake in figuring out the total number of people)

(b) The smallest number of people in this network is 11, and the network is a *clique*. (8 points for this part, you receive partial credit if your answer is partially correct)

6. (15 points)

(a) 5 points: 2 points for finding the individual and giving his or her degree; 3 points for the method.

(b) 2 points: 1 point for finding the individual and giving his or her degree; 1 point for the method

(c) 5 points: 1 point for choosing the phrase that returns 30-100 individuals. 1 point for histogram plot. 1 point each for reporting the minimum, maximum and average degree of the population.

(d) 3 points: received full credit if you discussed more than 1 aspect covered in class.

(e) Extra credit; 3 points: 2 points for finding a valid cycle of length ≥ 3 . 1 point as extra credit for briefly describing the method.

7. (15 points)

(a) Minimum diameter is 2 (node G) and maximum diameter is 5 (node D)

(4 points (2 for minimum and 2 for maximum) – Received 1 point for each if you were 1 off (ex: you gave maximum diameter of 4). Received no credit for impossible solutions (ex: maximum diameter of 6).)

(b) The worst-case diameter is 4, from F to B, F to A, F to C and F to E.

(4 points possible – received 1 point for each correctly identified pair.)

(c) The largest clique is size 4, ABED and BECD.

(4 points possible – Received 2 points for each correctly identified clique. Received 1 point for an answer of 5 (ABECD).)

(d) Minimum number of edges that can be removed is 1 (between D and E)

(3 points.)