

CSE 112 Networked Life 208

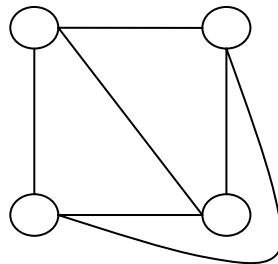
Homework 1

Grading Standard and Sample Solutions

1) 10 points

Full credit given for getting the correct answer 4 and drawing the correct diagram.

Sample Solution:



2a) 10 points

5 points for a clear description of what the network represents; 5 pts for correctly drawing the actual diagram. Your diagram should include a vertex representing yourself and all vertex should be clearly labeled, and you should not include vertices that are not directly connected to yourself, even though they might be connected to your neighbors.

2b) 10 points

Full mark if you get the correct answer, either in quotient form or decimal form. Otherwise, you receive 3 points if you get either the numerator (the number of edges between your neighbors, which does not include those edges that are incident to you) correct or the denominator (basically, you need to find the right k , which does not include yourself) correct.

3a-e) 5 points each

2 points for clearly stating what the vertices are; 2 points for clearly stating what the edges are and 1 point for the one sentence comment regarding its static structure, dynamics, and formation process.

Sample Solution:

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.

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.

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.

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.

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.

4a) 5 points

Full credit given unless answer missing.

Sample Solution:

Mark B. Cohen, whose profile I encountered while collecting data for question 4c, has 1009 friends on facebook. Surprisingly, someone who calls himself a fan of the TV show “Get Smart” can have this this many friends.

4b) 5 points

Full credit given unless answer missing.

Sample Solution:

“Glenjamin Franklin” has only one friend on facebook. I found him during a search for “Benjamin Franklin” on facebook.

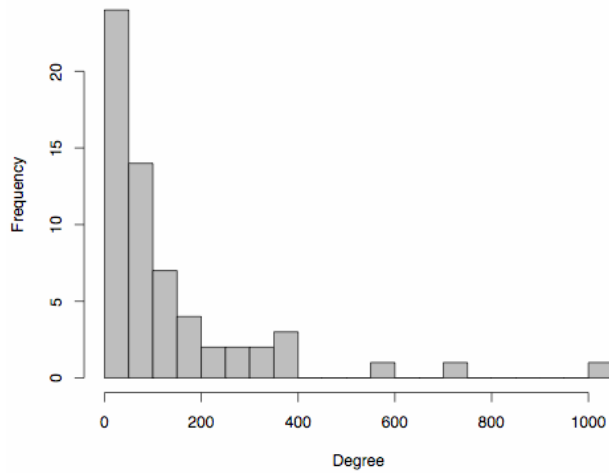
4c) 5 points

- minus 4 points for plotting something other than a degree distribution histogram.
- minus 1 point for not providing the average degree.

Sample Solution:

A search for individuals listing themselves as fans of the TV program “Get Smart” yielded 111 profiles within the Penn facebook. The maximum degree of any individual in this population was 1009. The minimum was 1. The average was 135. Below is the degree distribution histogram:

Degree Distribution for "Get Smart" Fans (TV series)



within our circle of friends:

Ted Sandler -> Jinsong Tan -> Jenn Wortmann -> John Blitzer -> Yun Mao -> Liang Huang -> Julia Deak -> Aviad Eliam -> Partha Talukdar -> Hannah Wallach -> Kilian Weinberger -> Yuan Ding -> Qian Liu -> Ben Ashpole -> Joao Graca -> Abhishek Gupta -> Daniel Fledger -> Ted Sandler

A total of 17 links, pretty small in comparison to some of the homework submissions...

5a) 10 points

- **minus 3 points for not mentioning local versus global links.**
- **minus 3 points for not mentioning something about clustering coefficient or cliques.**

Sample Solution:

All three networks share the same local, row-wise (horizontal) links.

If we assume that majors of the same subject are likely to be friends and hence to sit near each other, then the “long-distance” links of network A (the “same major” network) will actually be quite local. For network B, the “same state for high school” network, this effect should be diminished, if present at all, since states are large and we don’t expect people to be friends on the basis of this relationship. Thus, network B should exhibit more long distance connections than network A. Finally, network C should be similar to B in possessing more long distance connections since sharing the same last cell phone digit doesn’t increase likelihood of friendship.

The long distance links of all three networks, A, B and C, create cliques and hence all three networks will have quite very high clustering coefficients. The only difference between the networks in this respect is the distribution of clique sizes. For network C, we expect the distribution of clique sizes to be roughly uniform since the last digit of a cell phone number is (we imagine) equally likely to be any one of the digits 0-9. For networks A and B, we don’t expect clique sizes to be uniformly distributed since some majors are more popular and some states have larger populations; and since “Networked Life” is not equally attractive to students of all majors and Penn is not equally well attended by students from all states.

5b) 10 points

- **minus 4 points if described a strategy that the class typically did not, or was not allowed to use. E.g., saying that the class asked people what their majors were and passed it to those people with the same major as the target.**
- **minus 4 points if the alternative strategy proposed was to find out people's majors and pass the ball to people with the same major as the target. This wasn't allowed within the experimental constraints.**

Sample Solution:

The class tried to forward the ball as close as it could (proximally) to the racquet on each turn. As some students pointed, this strategy is suboptimal since a non-target (or

non-same-row-as-target) recipient has the same major/state/cell-phone-digit as the person forwarding the ball. Hence, passing the ball laterally more usefully provides a chance to uncover a long distance connection to the target directly or to her row.