MS&E 135 Networks – Problem Set 1 – Due 4/7

Stanford University, Spring 2016, Prof. Johan Ugander
Change history: v1.2. If errors are spotted they will be corrected and announced on Canvas.

Instructions Please answer the following questions concisely, drawing illustrations where appropriate. The assignments should be submitted at the start of class on Thursday 4/7, or before the start of class in the submission box in the Huang basement.

Discussing the assignment with classmates is permitted, but each student must submit their own solution set, and you should be writing your solutions independently. Please list the names of people you have discussed the problem set with (if any) at the top of your submission. "E&K" refers to Easley & Kleinberg, the course text, which can be accessed in full as a searchable PDF from the course homepage.

Question 1 (10 pts) E&K, Chapter 2, Problem 1

Question 2 (10 pts) E&K, Chapter 2, Problem 2

Question 3 (10 pts) E&K, Chapter 2, Problem 3

Question 4 (5 pts) E&K, Chapter 3, Problem 2

Question 5 (5 pts) E&K, Chapter 3, Problem 4

Question 6 (10 pts) E&K, Chapter 4, Problem 2

Question 7 (10 pts) E&K, Chapter 4, Problem 3

Question 8 (10 pts) E&K, Chapter 4, Problem 4

Question 9 (15 pts) There's been a lot of discussion about "degrees of separation" growing smaller as online platforms increase connectivity (in the shortest path sense, not the navigability sense that we will be discussing later in the course). In 2008, researchers at Microsoft analyzed the communication network of MSN Messenger and found an average distance between users of 6.6. A similar effort at Facebook in 2011 to analyze their social graph found an average of 4.74 degrees of separation between Facebook users. A new analysis in 2016 showed that this number had been reduced to 4.57. In this problem, we ask: how low can it go?

Facebook users can form a maximum of 5000 reciprocal friends (we're not counting directed "subscription" or "follow" edges pointing to celebrity accounts). It turns out that this maximum points to an illustrative lower bound on how small the degrees of separation can get on Facebook.

(a) If everyone has 5000 friends, what's the maximum number of people you can have as a friend of a friend? What's the maximum number of people you can have as a *friend of a friend of a friend*?

Consider a social network with 1 billion people, and you're in the middle of that network, with 5000 friends who all have 5000 friends, and the remainder of the people are 3 degrees removed from you. Convince yourself (no answer needed) that this arrangement is the closest that 1 billion people could possibly be arranged to you, given the constraint that everyone has at most 5000 friends.

(b) In the above network, what's the probability that a randomly selected person is your friend? What's the probability that a randomly selected person is your friend-of-friend? What's the probability that a randomly selected person is exactly 3 degrees removed from you?

(c) What's the average distance to a random node from you?

Since this arrangement is the closest that 1 billion people can get to you under the constraint of a maximum of 5000 friends, the answer to (c) serves as a lower bound of how low the "degrees of separation" can go: you're in the middle and have the smallest average shortest distance, and everyone must have at least a large average distance as you. Thus, it also serves as a lower bound for the average shortest path between two random people, given the maximum degree constraint.

A more sophisticated lower bound (≈ 3.6) will be mentioned briefly in the lecture on 4/7.

Question 10 (3 pts, extra credit) How does the above analysis change if the maximum degree changes? Derive a correct lower bound on the average distance between users as above, for a network of 1 billion people, but as a function a generic maximum degree D. To simplify matters, assume $D > 70$.

Question 11 (15 pts) In this problem we're going to explore the friendship paradox.

(a) Visit `http://explorecourses.stanford.edu/` to look up the number of "students enrolled" in each of your courses this quarter. Do not count independent studies. List these class sizes. If you pick one of your courses uniformly at random, what is the expected class size?

(b) Across all these courses, you can think of each student in each class as forming a student-class pair representing a "classroom experience." If you pick a student-class pair uniformly at random, what is the expected size of the classroom for that "average classroom experience"?

(c) Consider the following two social networks.

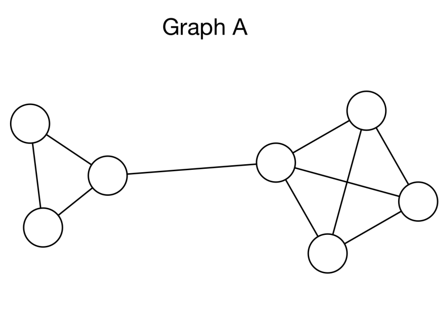

Figure 1

For each of these two connected graphs, separately compute (1) the expected degree of a node selected uniformly at random, and (2) the expected degree of a randomly selected friend (selecting friendships uniformly, and then uniformly selecting a random member of that friendship). Next, (3) what fraction of the nodes in each graph have less friends than the average friend count of their friends?

Notice that the "friendship paradox" only makes a formal statement about the relationship between these first two quantities: (2) is greater than or equal to (1). It does not require that the third quantity be more than 50%. In empirical social networks, it is however often observed that the fraction of nodes that have less friends than the average of their friends is quite high: 92.7% on Facebook in 2011.