CS584 Spring 2015, Problem Set 1
Due January 31, 11:59pm

Ymir Vigfusson

Problem 1

Consider the social network in Figure 1. The thickness of the line denotes tie strength, with thick lines denoting strong ties, and dotted ones denoting weak ties. Consider the strong triadic closure property from class and the book:

> We say that a node A violates the Strong Triadic Closure property if it has strong ties to two other nodes B and C, and there is no edge at all between B and C. We say that a node A satisfies the Strong Triadic Closure property if it does not violate it.

Which of the nodes in the network in the figure satisfy this property, and which ones do not?

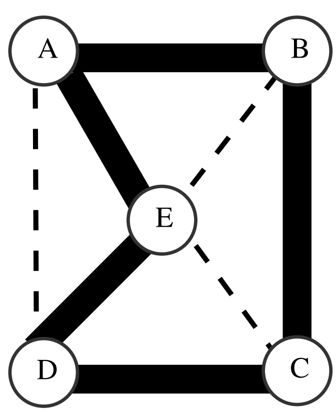

Figure 1: Social network for Problem 1. Thick edges are strong ties, dotted lines denote weak ties.

Problem 2

Atlanta is in turmoil, yet again. Zombies have taken over midtown, eradicating most of the population. A few surprisingly photogenic resistance fighters are unaffected and carry on with their lives outside the perimeter along with a conveniently situated camera crew. Compounding matters further, a swarm of aliens under the leadership of Sarah Kerrigan have now infested the remains of Buckhead and are setting up a base.

If we examine the social network of Atlanta, we will see that despite plot-driving differences, the resistence fighters and camera crew all get along; the aliens are all of the same brood and are mind-controlled to dedicate their life "for the swarm"; and the zombies are perfectly content with each other around. However, between each of these factions lies deep hatred: every alien hates humans and zombies, the zombies see the humans and aliens only as nutrition, and the humans wish most to eradicate both infestations and reclaim control of the Coca-Cola capital.

Is this social network strucurally balanced? Either show an example of a social network (including the number of people/zombies/aliens in each faction) that is structurally balanced and argue why it is balanced, or prove that no matter how many people/zombies/aliens are in the network, it can never be structurally balanced.

Problem 3

> "It takes considerable knowledge just to realize the extent of your own ignorance." (Thomas Sowell)

In a ficticious survey of 1000 people living uniformly alongside the 100-km I-285 perimeter, it was determined that people were on friendly terms with those living within 10 km of one another on the perimeter, but were highly suspicious and downright hostile towards those living outside those bounds. Assume that the perimeter is perfectly round and that the 1000 people all live same distance from each other alongside the circle.

Is this social network structurally balanced? Why or why not? Justify your answer (*i.e.*, provide a proof or a counterexample). What if the sample had contained 1001 people?

Problem 4

In this problem, we are going to investigate how our neighbors in an information network differ from us. There are various network properties that we could investigate — centrality, tie strengths, structural properties — but to start we will begin with some of the simplest: the node degrees.

To be concrete, let $G = (V, E)$ be a graph representing a network of high schoolers, where each node $v \in V$ represents a high-school student, and an edge $(u, v) \in E$ represents that students u and v are buddies. Let $n = |V|$ and $m = |E|$.

We let d_u denote the *degree* of node u. More specifically,

$$d_u = |\{v \in V : (u, v) \in E\}| .$$

(a) How many buddies do students have on *average* in the arbitrary network G? If D is a random variable denoting the degree of a random student from the high school population, you can think of the answer as being the expectation $\mathbb{E}[D]$. Show your work and return the simplest expression as you can find in terms of the variables we have defined.

(b) Let us look at how many buddies your buddies have on average in the graph. Denote by B the random variable of buddies-of-buddies. Prove that

$$\mathbb{E}[B] = \frac{\sum_{v \in V} d_v^2}{2m}.$$

Hint: Count the impact of each single edge in the graph on the total number of buddies-of-buddies. Think of paths of length exactly 2. Be careful to quantify exactly what is being averaged over.

(c) Let us try to interpret the expression from (b) by simplifying it in terms of the degree distribution. Recall from probability theory that for a random variable X, the *variance* is defined as

$$\text{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 .$$

and is always non-negative.

Given D as above, prove that

$$\mathbb{E}[B] = \mathbb{E}[D] + \frac{\text{Var}(d)}{\mathbb{E}[D]}.$$

(d) What does the result from (c) actually imply in terms of your popularity compared to that of your buddies, on average? How does the interpretation differ if we instead look at a social network such as LinkedIn connections? What about a romantic liaison network, where edges denote that two partners have had sexual intercourse? What about a paper co-authorship graph, where edges denote that two people have collaborated on a publication?