Networks and Markets

Homework 3

Due: 3/28/2024, 11:59pm

Submit solutions as a PDF file to Gradescope. On Gradescope, match pages with the corresponding problem (we will make a one-point deduction per problem if pages are not matched). Show work throughout, with legible handwriting. Clearly mark your answer by putting a box around it.

1 Congestion pricing (17 pts)

Consider the following network with two nodes and two edges connecting them:

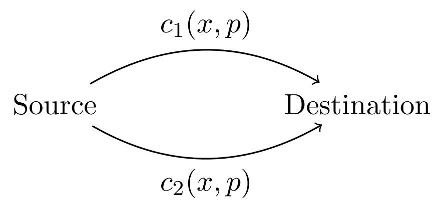

Figure 1: Network with 2 Nodes and 2 Edges

There are two cost functions, c_1 and c_2, that depend on the fraction x of traffic going through the edge and the price p charged by the government for that edge. We are going to consider the following cost functions $c_1(x,p) = x + \frac{1}{4} + p$ and $c_2(x,p) = 2x + p$.

First, we are going to consider the setting without any pricing, prices $p_1 = p_2 = 0$.

Part (a) (3 pts) Calculate the Nash equilibria for the given network, i.e., what are the equilibria x_1 and x_2? What is the social welfare at this equilibrium, i.e., what is the average cost faced by individuals on this network?

Part (b) (4 pts) Calculate the social welfare optimal flow for the network, i.e., what are the x_1 and x_2 that minimize the average cost? What is the social welfare at this point?

Part (c) (4 pts) Now suppose that we can price the second edge, i.e., we keep p_1 at 0 but now allow p_2 to be (positive) and non-zero. Calculate the price p_2 that leads to traffic optimality, i.e., where the new Nash Equilibria corresponds to the flow calculated in the previous part (b).

Part (d) (4 pts) Now, we make the make the setting closer to reality. Suppose that there are two groups of people (that we will call a and b), with group a corresponding to $\frac{3}{8}$ of the total population and group b corresponding to the remaining $\frac{5}{8}$ fraction of the total population.

Suppose that group a is extremely rich, and so doesn't care about the tolls that they pay. More concretely, we will suppose that groups a and b have cost functions $c_2^a(x, p) = 2x$ and $c_2^b(x, p) = 2x+p$ respectively for the second edge.

Suppose we set the prices $p_1 = 0$ and $p_2 = 2$. What are the fractions of traffic flow $x_1^a, x_1^b, x_2^a, x_2^b$ at equilibria? (Here, $x_1^a + x_2^a = \frac{3}{8}$ and $x_1^b + x_2^b = \frac{5}{8}$.)

Part (e) (2 pts) Consider the equilibria $x_1^a, x_1^b, x_2^a, x_2^b$ calculated in Part (d). Define the *traffic cost* to the component of the cost function that is just due to time, not due to prices, i.e., $t_1^a(x) = t_1^b(x) = x + \frac{1}{4}$, and $t_2^a(x) = t_2^b(x) = 2x$.

What is the average traffic cost faced by group a? What is the average traffic cost faced by group b?

2 Page Rank (15 points)

Recall that we briefly discussed in class the PageRank algorithm that was Google's initial magic – how do we algorithmically determine which websites are "best" when ranking results in a search engine? The idea is that we can use network information regarding which websites link to which other websites.

Consider the network of websites shown in Figure 2. We would like to apply a basic version of PageRank to determine which websites to show. The basic idea of PageRank is that the reputability of a website can be learned from the reputations of the websites that link to it. Indeed, a website is linked to by many high-reputation websites, we might expect that website to also be reputable. Conversely, a website not linked to by any other website, or only linked to by websites that are themselves not reputable, would not be reputable. The challenge, however, is that we do not know the reputation of any website beforehand.

Formally, we give each website i in a network a PageRank value v_i. Then, for a website with PageRank value $v_i = x$ and k outgoing neighbors, we assign each outgoing link a flow of x/k (i.e., each website splits up its PageRank among its links.) Then, a network is in equilibrium is each website's PageRank is equal to the sum of the flows of its incoming links. For example, in Figure 2, if E had PageRank value 0.4 and F had PageRank value 0.2, then the sum of the flows of the incoming links to H would be $0.4+0.2/2 = 0.5$. (The .2 is divided by 2 because F splits its reputation outflow between G and H).

Furthermore, we sum of the PageRank values across nodes is 1, i.e., we have $\sum_i v_i = 1$.

In this problem, we will compute the equilibrium PageRank values of the network in Figure 2.

Part (a) (7 points) First, let x denote the (unknown) PageRank value of node A. Write down the PageRank value for each other node in terms of x.

Part (b) (4 points) Using your answer to (a), determine the value of x, and from this find the PageRank values for all nodes (expressed as actual numbers).

Part (c) (2 points) Suppose you can add one edge to the network. What additional edge would you add in order to maximize the PageRank value of node F?

Part (d) (2 points) Suppose you can remove one edge from the network. What edge would you remove to minimize the PageRank of node G?

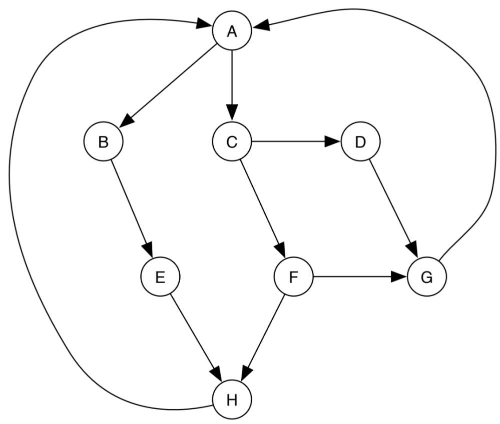

Figure 2: Network for Problem 2

3 Network Models (26 points)

To better theoretically analyze network structures, it is useful to create *models* of networks. We would like these models to capture relevant features of networks that appear in the real world. In this problem, we will explore some basic network models, and try to incorporate features of social networks that we have already encountered. In particular, we will focus on two properties: *small worlds* (everyone in the network is connected by a short path), and *traidic closure* (a connection of a connection is often also a connection).

The Erdős–Rényi Model

We begin with perhaps the most basic model of networks, the *Erdős–Rényi* model. In this model, there are n nodes, and between any two nodes, an edge forms independently at random with probability p. This random network is denoted $G(n, p)$. Parts (a)-(d) will focus on the network $G = G(100, 0.1)$, i.e., there are 100 nodes and any pair of nodes as an edge with probability 0.1.

Part (a) (2 points). Consider a single node in G. What is its expected number of neighbors (two nodes are neighbors if they have an edge between them)?

Part (b) (3 points). Now consider two nodes in G, labeled a and b. Consider a third node, c. What is the probability that *both* a and b are neighbors with c?

Part (c) (4 points). We now analyze to what extent G resembles a "small world," meaning that nodes are often connected by short paths. Again consider two nodes in G, labeled a and b. What is the probability that a and b are connected by a path of length 2 or less? Write your answer as a decimal to the nearest thousandth. (You may use WolframAlpha or other computational tools to compute.)

(Hint: Consider both paths of length 1 *and* 2*, and you will use your answers to the above two parts.)*

Part (d) (2 points) We now consider the presence of triadic closure in G. In particular, consider three nodes labeled a, b, and c. Suppose that a is neighbors with both b and c. What is the probability that b and c are also neighbors?

Part (e) (3 points) We say that a node a exhibits the *strong triadic closure property* if every two of a's neighbors are also neighbors. Suppose that the node a has 10 neighbors in G. What is the probability that it exhibits the strong triadic closure property? Express your answer in scientific notation.

A More Sophisticated Network

One limitation of the Erdős–Rényi model, as suggested by the previous problem, is that despite exhibiting small world behavior, it does not obey triadic closure well. We will now consider a somewhat more sophisticated model, that loosely resembles popular models of social networks.

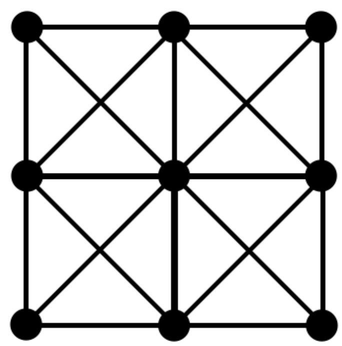

Figure 3: Stage 1 Network

Stage 1 of the network is shown in Figure 3. Call it G_1. It consists of nine points arranged in a grid-like manner, where each point is connected to its horizontal, vertical, and diagonal neighbors.

Part (f) (2 points) What proportion of the nodes in G_1 exhibit the strong triadic closure property?

Part (g) (2 points) What is the diameter of G_1 (i.e., the maximum distance between any two nodes)?

Stage 2 of the network, shown in Figure 4, consists of 9 copies of the Stage 1 network. Such that the "central nodes" of the Stage 1 network are connected in the same pattern as before. Call the Stage 2 network G_2.

Part (h) (2 points) What proportion of the nodes in G_2 exhibit the strong triadic closure property?

Part (i) (2 points) What is the diameter of G_2?

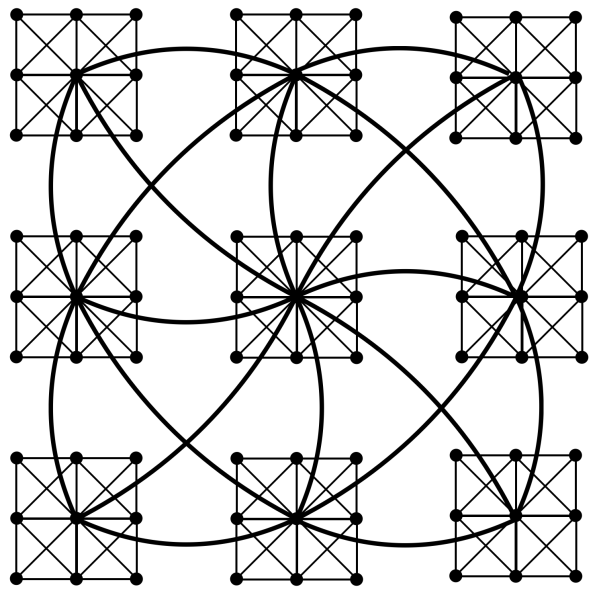

Figure 4: Stage 2 Network

We may further extend this process, such that the Stage k network consists of making 9 copies of the Stage $k-1$ network, such that the "central nodes" of each of the 9 Stage $k-1$ networks are connected in the original pattern. Call the Stage k network G_k.

Part (j) (2 points) Come up with a social network that might plausibly resemble the network structure G_3. Explain briefly in at most 3-4 sentences.

Part (k) (4 points) Let n_k be the total number of nodes in G_k. Let d_k be the diameter of G_k. Write d_k as a simplified function of n_k.