Networks: Fall 2021	Homework 1
David Easley and Yian Yin	**Due 3:30pm, Thursday, August 31, 2023**

Homework solutions should be submitted by upload to Gradescope, which can be accessed via Canvas. The file you upload **must be typed and submitted in PDF format**. Handwritten assignments will not be graded. However, you can draw graphs and insert them into your pdf. You can create a separate file with the solutions (you don't need to repeat the questions); it is fine to create the homework in any format provided it's typed and handed in as a single PDF file. When you upload your pdf to Gradescope be sure to assign your answers to the correct question.

To be eligible for full credit, your homework must come in by 3:30pm Thursday. We will also accept late homeworks after 3:30pm Thursday until 3:30pm Friday for a deduction of 10% of the total number of points available. Gradescope will stop accepting homework uploads after 3:30pm Friday; after this point we can only accept late homework when it is accompanied by a University-approved reason that is conveyed to and approved by the TA in charge of this homework prior to the due date of the homework. (These include illness, family emergencies, and travel associated with university activities.)

The TA in charge of this homework is Miles Ma, hm387@cornell.edu

Reading: The questions below are primarily based on the material in Chapters 2, 3 and 5. (The material from Chapter 3 will be covered in class on Friday August 25. The material from Chapter 5 will be covered in class on Monday August 28.)

---

(1) [6 points]

A group of economists are studying the interactions between 10 banks. They drew the graph in Figure 1 to describe these interactions with nodes representing the banks and edges between two nodes representing the idea that these two banks interact with each other. So for example, banks A and D interact, but banks A and B do not interact. These "interactions" are meant to describe economic transactions between the banks.

These economists are interested in measuring how important individual banks are purely as a function of their location in this network. One way they are considering is to count the degree of each bank. The *degree* of a node in an undirected graph such as Figure 1 is the number of edges attached to the node.

(a) For each node in Figure 1, what is it's degree?

An alternative way to ask about importance is to ask how many connected components the graph would have after removing one of the banks.

(b) For each node in Figure 1 how many connected components would the graph have if this bank was removed from the graph?

Another way to measure the importance of a bank would be to consider how close it is to all the other banks. Recall that the length of a path in a graph is the number of edges

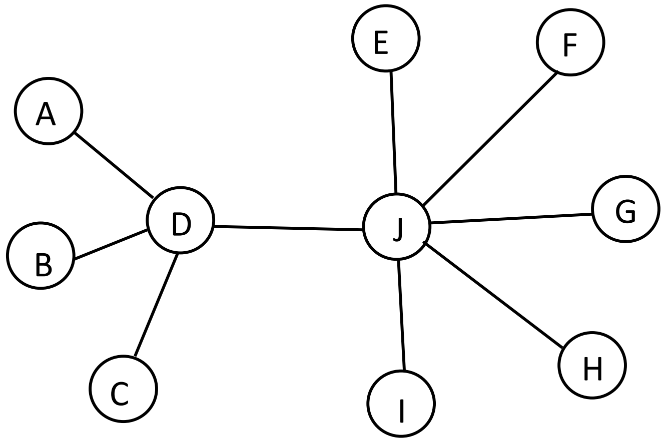

Figure 1: Q1 Banks

it contains, and the distance between two nodes is the length of the shortest path between them. (For example, the distance between A and J is 2, since there are two edges on the path from A though D to J.) Using the notion of distance, we'd like to consider what it means for a node to be "centrally located" in the bank network. One option for defining this notion of centrality is that the number of hops that are necessary to get to any other bank should be as small as possible.

(c) Is there a bank (or banks) in Figure 1 with the property that the maximum distance from this bank to any other bank at most 2? If there is, name all such banks and say why they have this property. If there isn't, give a brief explanation why not.

(d) In the example given in Figure 1 there is a close connection between a node have a high degree and being centrally located (using the definition of "centrally located" given above). Is this always true? That is, are there graphs such that a node with the " highest degree" is not centrally located? Either give an example in which "highest degree" and "centrally located" are different or give an argument for whey they agree for all graphs.

---

(2) [6 points]

In a directed graph as in Figure 2 links are from one node to another node. So for example there is a link from node A to node B but not one from B to A. In a directed graph a path from one node to another must follow the direction of the edges that it uses. Finally, a directed graph is strongly connected if there is a path from every node to every other node.

(a) Is the directed graph in Figure 2 strongly connected? If so, explain why; if not explain why not.

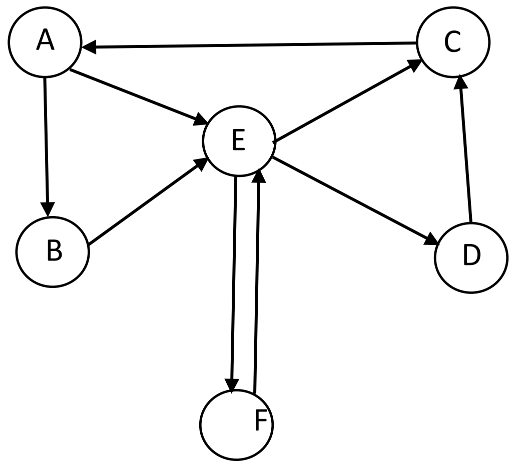

Figure 2: Q2 Directed Graph

(b) For the graph in Figure 2, if you think that it is not connected what is the minimum number of directed edges that you would need to add to make it strongly connected? If you think that it is strongly connected what is the maximum number of edges that you can remove such that the resulting graph is strongly connected.

---

(3) [6 points]

We've looked at examples from large online social platforms in our lectures. However, due to privacy settings, these platforms might not show all friendships. Only some of the friendships are publicly visible.

(a) Suppose you are interested in the social network structure on a popular website named Y. Figure 3 shows such a friendship network based on publicly available information from Y, where solid lines represent strong ties and dashed lines represent weak ties. Identify all bridges and local bridges in the network.

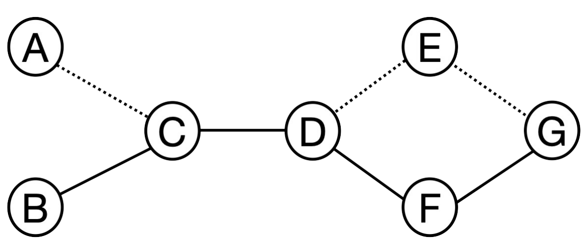

Figure 3: Q3 Social Network

(b) Use the Strong Triadic Closure Property to find the potential 'hidden' friendships that aren't shown in Figure 3 but likely exist in the real world.

(c) Add all hidden ties identified in **(b)** to the original network (treat them as weak ties). Revisit the bridges and local bridges you listed in **(a)**: are they still bridges or local bridges?

---

(4) [6 points]

Geographic proximity between cities plays a crucial role in shaping economic, social, and infrastructural interactions. Cities that are closer tend to have stronger trade relationships and more knowledge exchanges. Let us look at a region with five cities (A, B, C, D, and E), where the roads connecting them are shown in Table 1. People can drive both ways on all these roads. For example, it's 120 miles from City A to City B (or from City B to City A).

City 1	**City 2**	**Driving Distance (miles)**
A	B	120
A	C	90
A	D	160
B	C	130
B	D	105
C	E	180
D	E	135

Table 1: Roads between cities

(a) Construct a network of the five cities. Here's how we'll decide to connect them: (i) If you can drive between two cities in 150 miles or less, add a strong tie between them. (ii) If you can't drive between two cities within 150 miles, but can do so in 200 miles or less, add a weak tie between them. You can travel on either a direct route (a single road between cities) or an indirect route (a combination of roads going through other cities). Once you've created the network, write a list of all the strong ties and another list of all the weak ties.

(b) In this network, which cities satisfy the Strong Triadic Closure Property, and which violate it? Provide an explanation for your answer.

(c) Can you change the '200 miles' threshold in our definition of weak ties, so that the Strong Triadic Closure Property will always be satisfied of all cities? Explain why.

---

(5) [6 points]

In our class, we discussed the Balance Theorem: *If a labeled complete graph is balanced, then either all pairs of nodes are friends, or else the nodes can be divided into two groups, X and Y, such that each pair of people in X likes each other, each pair of people in Y likes* *each other, and everyone in X is the enemy of everyone in Y.* Here we will explore how this property can help us to better characterize the structure of balanced networks.

(a) Let us start from a simple example. We can, consider a balanced complete graph with 3 nodes.

From the theorem we know the three nodes either (i) like each other, or (ii) can be divided into two groups (with sizes 1 and 2 respectively). Hence there are only 2 possible structures, corresponding to 0 or 2 negative edges, as illustrated in Fig. 4. + denotes positive edges between friends and - denotes negative edges between enemies. For simplicity, graphs with the same number of negative edges are considered as the same structure.

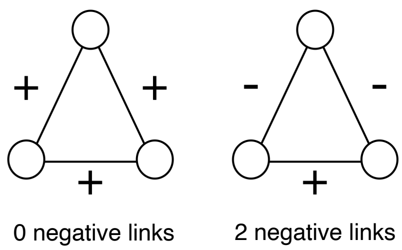

Figure 4: Q5 Balanced Networks

What are the possible structures of a balanced complete graph with 4 nodes? For each possible scenario, draw an example and calculate the number of negative edges in it.

(b) What are the possible structures of a balanced complete graph with 5 nodes? For each possible scenario, draw an example and calculate the number of negative edges in it.

(c) These exercises show how balanced theorem simplifies the counting of different balanced structures. Indeed, if one takes a naive approach by simply enumerating all possible signed complete graphs of 5 nodes (10 edges) and checking which of them are balanced, they will need to examine $2^{10} = 1,024$ different configurations!

Now let's try to generalize the results in **(a)** and **(b)**: For an integer $n \geq 3$, what are the possible number of negative edges of a balanced complete graph with n nodes?