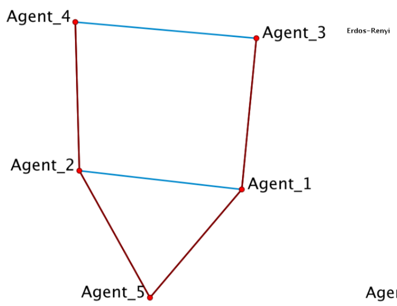
Social Media Analytics

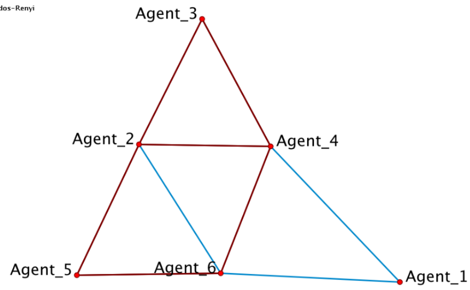
Module 3 Assignment

Paras Ahuja

September 20, 2020



Network A



Network B

Use networks A and B above to answer the following questions or perform the calculations as required. All work is to be done by hand without the use of software. You must show your work. No credit is awarded for the final answer.

1. Create the adjacency matrix for networks A and B above.

   For the given network graphs, we have the following adjacency matrices. An adjacency matrix shows if there is an edge between the two nodes. If an edge exists, we represent this with a 1, and 0 otherwise. Note that the agents are simply listed by the agent number.

$$A = \begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$B = \begin{array}{c c} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

2. Compute the degree centralities for each agent in networks A and B.

   Degree centrality of a node refers to the number of edges attached to the node. Following table shows the degree centrality of the Network A, and Network B.

Node	Sum	Degree
1	3	3/4
2	3	3/4
3	2	1/2
4	2	1/2
5	2	1/2

(a) Network A Degree Centrality

Node	Sum	Degree
1	2	2/5
2	4	4/5
3	2	2/5
4	4	4/5
5	2	2/5
6	4	4/5

(b) Network B Degree Centrality

3. Determine all possible geodesics in network A and B.

   Following tables outlines geodesics or shortest paths in Network A, and Network B.

From	To	Geodesic
1	2	(1, 2)
1	3	(1, 3)
1	4	(1, 3, 4) and (1, 2, 4)
1	5	(1, 5)
2	3	(2, 4, 3) and (2, 1, 3)
2	4	(2, 4)
2	5	(2, 5)
3	4	(3, 4)
3	5	(3, 1, 5)
4	5	(4, 2, 5)

(a) Network A Geodesics

From	To	Geodesic
1	2	(1, 6, 2) and (1, 4, 2)
1	3	(1, 4, 3)
1	4	(1, 4)
1	5	(1, 6, 5)
1	6	(1, 6)
2	3	(2, 3)
2	4	(2, 4)
2	5	(2, 5)
2	6	(2, 6)
3	4	(3, 4)
3	5	(3, 2, 5)
3	6	(3, 4, 6) and (3, 2, 6)
4	5	(4, 2, 5) and (4, 6, 5)
4	6	(4, 6)
5	6	(5, 6)

(b) Network B Geodesics

4. Compute the betweenness centrality score of all agents in networks A and B.

   Betweenness centrality is the measure of centrality in a graph based on shortest paths. It is calculated as:

$$C_B(n_i) = \sum_{uw} \frac{\sigma_{uw(n_i)}}{\sigma_{uw}}$$

   To oversimplify, what we are finding out here is how many times is a node interrupted or accessed to get to another node.

Table 3: Network A Betweenness Centrality

Path	**1**	**2**	**3**	**4**	**5**
(1, 2)	0	0	0	0	0
(1, 3)	0	0	0	0	0
(1, 2, 4) and (1, 3, 4)	0	1/2	1/2	0	0
(1, 5)	0	0	0	0	0
(2, 1, 3) and (2, 4, 3)	1/2	0	0	1/2	0
(2, 4)	0	0	0	0	0
(2, 5)	0	0	0	0	0
(3, 4)	0	0	0	0	0
(3, 1, 5)	1	0	0	0	0
(4, 2, 5)	0	1	0	0	0
Betweenness	1/4	1/4	1/12	1/12	0

Table 4: Network B Betweenness Centrality

Geodesic	**1**	**2**	**3**	**4**	**5**	**6**
(1, 6, 2) and (1, 4, 2)	0	0	0	1/2	0	1/2
(1, 4, 3)	0	0	0	1	0	0
(1, 4)	0	0	0	0	0	0
(1, 6, 5)	0	0	0	0	0	1
(1, 6)	0	0	0	0	0	0
(2, 3)	0	0	0	0	0	0
(2, 4)	0	0	0	0	0	0
(2, 5)	0	0	0	0	0	0
(2, 6)	0	0	0	0	0	0
(3, 4)	0	0	0	0	0	0
(3, 2, 5)	0	1	0	0	0	0
(3, 4, 6) and (3, 2, 6)	0	1/2	0	1/2	0	0
(4, 2, 5) and (4, 6, 5)	0	1/2	0	0	0	1/2
(4, 6)	0	0	0	0	0	0
(5, 6)	0	0	0	0	0	0
Betweeness	0	1/5	0	1/5	0	1/5

5. Compute the closeness centrality score of all agents in networks A and B. First we calculate the distance matrix. This matrix represents the number of edges from one node to the other.

$$D_A = \begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 0 & 1 & 1 & 2 & 1 \\ 1 & 0 & 2 & 1 & 1 \\ 1 & 2 & 0 & 1 & 2 \\ 2 & 1 & 1 & 0 & 2 \\ 1 & 1 & 2 & 2 & 0 \end{bmatrix} \end{array}$$

$$D_B = \begin{array}{c c} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 0 & 2 & 2 & 1 & 2 & 1 \\ 2 & 0 & 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 & 2 & 2 \\ 1 & 1 & 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Next we sum up individual rows of the matrices and obtain the following tables:

Table 5: Network A Closeness Centrality

Row	Sum	Closeness Centrality
1	5	4/5
2	5	4/5
3	6	2/3
4	6	2/3
5	6	2/3

Table 6: Network B Closeness Centrality

Row	Sum	Closeness Centrality
1	8	5/8
2	6	5/6
3	8	5/8
4	6	5/6
5	8	5/8
6	6	5/6

6. Look at the two networks. Which agents may have greater influence? Why?

   Neo4j document states that:

   > *Betweenness centrality is a way of detecting the amount of influence a node has over the flow of information in a graph. It is often used to find nodes that serve as a bridge from one part of a graph to another. The closeness centrality of a node measures its average farness (inverse distance) to all other nodes. Nodes with a high closeness score have the shortest distances to all other nodes.*

   Network A: Based on the above specified criteria, we see that for Network A agents 1 and 2 have greater influence because they have a higher betweenness centrality, and closeness centrality.

   Network B: For Network B agents 2, 6 and 6 have greater influence because they have a higher betweenness centrality, and closeness centrality.

7. Calculate the network density of networks A and B.

   For an undirected network, the density of a network is the fraction between 0 and 1 that tells us what portion of all possible edges are actually realized in the network. For a network G made of n nodes and m edges, the density $\rho(G)$ is given by:

   $$\rho(G) = \frac{2m}{n(n-1)}$$

   Then, we have the following:

   $$\rho(G)_A = \frac{2 \times 6}{5(5-1)} = \frac{12}{20} = \frac{3}{5} = 0.6$$
   $$\rho(G)_B = \frac{2 \times 9}{6(6-1)} = \frac{18}{30} = \frac{3}{5} = 0.6$$

   As a percentage we can say that the density of both networks is 60%.

8. Calculate the diameter of networks A and B.

   For both of these networks the longest shortest path requires traversing of at least 3 edges, therefore we can conclude that the diameter of Network A, and B is 3.

Network

September 20, 2020

[1]:
```
# import libraries
import igraph as g
import cairo as c
```

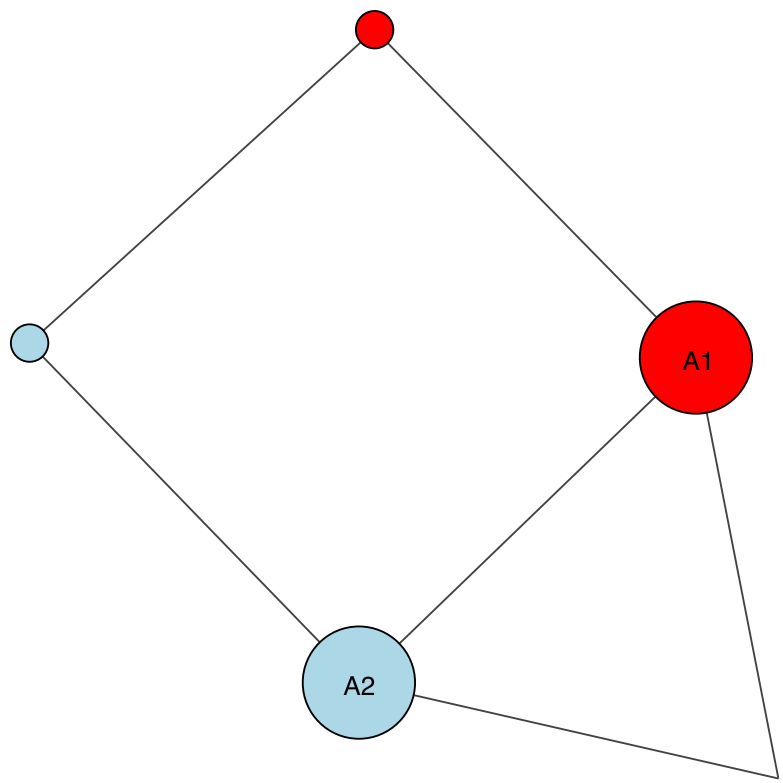
[2]:
```
# add edges to the network
# igraph starts edges/vertices from 0
network_a = g.Graph([(0,1), (0,2), (2,3), (1,3), (0,4), (1,4)])
```

[3]:
```
# create layout
layout = network_a.layout_kamada_kawai()
layout = network_a.layout("kamada_kawai")
layout = network_a.layout_reingold_tilford(root=[2])
layout = network_a.layout("rt", 2)
layout = network_a.layout("kk")
```

[4]:
```
betweenness = network_a.betweenness()
save_as = "network_a.pdf"
labels = ["A2", "A1", "", "", ""]
colors = ["lightblue", "red", "lightblue", "red", "lightblue"]
size = [40 * s for s in betweenness]
```

[5]:
```
# plot graph
g.drawing.plot(network_a, save_as, layout = layout, vertex_label = labels,␣
 ↪vertex_color = colors, vertex_size = size, margin=100)
```

[5]:


A1
A2