Darin Zlatarev #261081234

Professor Roman Galperin

ORGB 672 – Exercise 2

21 March 2023

Part 1: Fake Book Bus Dataset

Node	Edges	Adjacency	Degree Centrality	Total Steps	Closeness Centrality	Cross Count	Betweenness Centrality
1	2;	1	0.111111111	30	0.3	0	0
2	1;A;	2	0.222222222	22	0.409090909	8	0.222222222
3	4;5;B;C;D;	5	0.555555556	16	0.5625	4.633333333	0.128703704
4	3;C;	2	0.222222222	20	0.45	0	0
5	3;D;6;	3	0.333333333	21	0.428571429	0.533333333	0.014814815
6	5;B;D;	3	0.333333333	19	0.473684211	0.933333333	0.025925926
A	2;B;C;	3	0.333333333	16	0.5625	14	0.388888889
B	3;6;A;B;D;	5	0.555555556	14	0.642857143	9.033333333	0.250925926
C	3;4;A;B;D;	5	0.555555556	14	0.642857143	8.7	0.241666667
D	3;5;6;B;C;	5	0.555555556	16	0.5625	3.266666667	0.090740741

Here is a table with the nodes and edges from the network along with the answers for part 2; it is also attached as a separate Excel file. In total the network has **10 nodes and 17 edges** in it.

Part 2: Centrality Calculations

Above in the table are the results for degree, closeness and betweenness centrality for all the seats in the bus. Below are the detailed calculations for each of the four empty seats:

Degree Centrality: $DC_A = \frac{\text{Adjacency}}{n-1} = \frac{3}{10-1} = \mathbf{0.333}$ $DC_B = DC_C = DC_D = \frac{5}{10-1} = \mathbf{0.556}$

Degree centrality is the proportion of nodes that each nodes has edges connected with. Hence, seats B, C and D all are adjacent to 5 other seats out of 9 possible neighbors, having the highest degree centrality among the available seats.

Closeness Centrality: $CC_A = \left(\frac{\sum \text{Total Steps}}{n-1}\right)^{-1} = \left(\frac{2+1+2+2+3+2+1+1+2}{10-1}\right)^{-1} = \left(\frac{16}{9}\right)^{-1} = \mathbf{0.5625}$

$CC_B = CC_C = \left(\frac{14}{9}\right)^{-1} = \mathbf{0.6429}$ $CC_D = \left(\frac{16}{9}\right)^{-1} = \mathbf{0.5625}$

Closeness centrality measures the distance between one node and all other nodes; the more nodes we have to go through to reach all other nodes, the lower the score. Here seats B and C have the greatest closeness centrality score meaning that they need the least amount of edges to reach all other nodes in the network; hence making them the closest to everyone in the bus. Adjacency (for degree centrality) and total steps (for closeness centrality) were counted by hand. Detailed calculations of cross count (for betweenness centrality) are on the Excel sheet.

Betweenness Centrality: $BC_A = \frac{\sum(\text{Cross Count} \times \text{Weight})}{(n-1)\times(n-2)\div 2} = \frac{\sum_1^{14} i*1}{9*8*0.5} = \frac{14}{36} = \mathbf{0.38889}$

$BC_B = \frac{9.0333}{36} = \mathbf{0.25093}$ $BC_C = \frac{8.7}{36} = \mathbf{0.24167}$ $BC_D = \frac{3.2667}{36} = \mathbf{0.09074}$

Betweenness centrality tells us how often a node heeds to be crossed if everyone needs to reach everyone; hence if two subsets in a network communicate mostly within each other, nodes with high betweenness centrality will act as a bridge between these two groups. Crossing can also be weighted by a decimal if there are several shortest paths available. For example, if there are two possible shortest paths, it will count as 0.5 crossings instead of one crossing. Here, seat A acts as a bridge from seats 1 and 2 towards the rest of the bus.

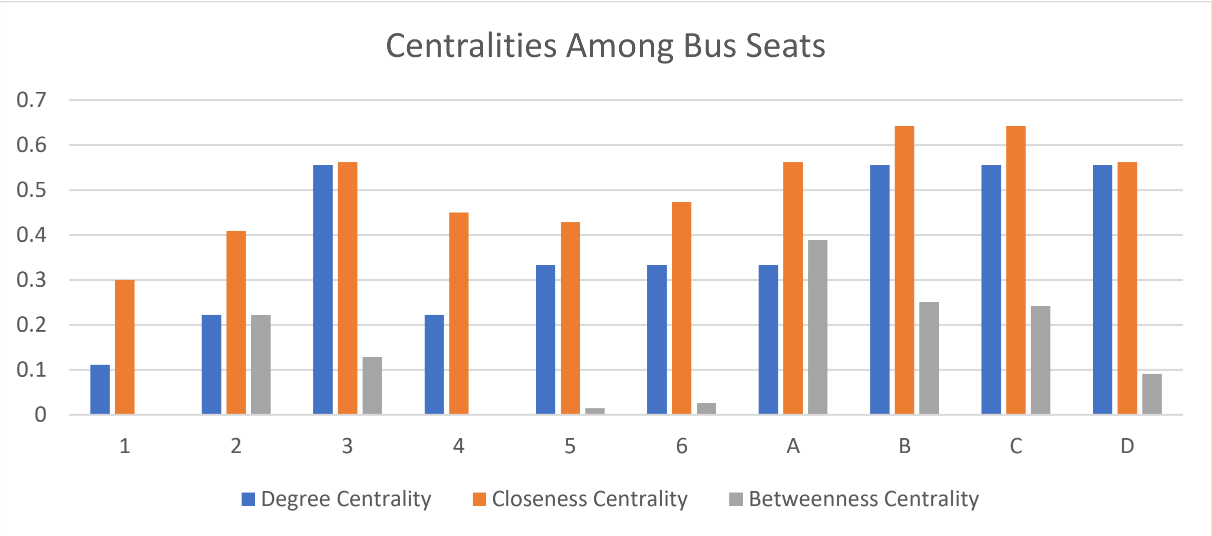
Part 3: Seat Choice Consequences

Each seat has its relative advantages and disadvantages. As the graph below shows, the worst seats in the bus are seat 1 and 4 as they have no betweenness centrality since they are at the very edge of the network. From the available seats, seat D is the worst since it has the lowest betweenness centrality. It also has the worst closeness centrality tied with seat A, as well as the same degree centrality as B and C; thus seats B and C are equally good or better among all aspects compared to seat D.

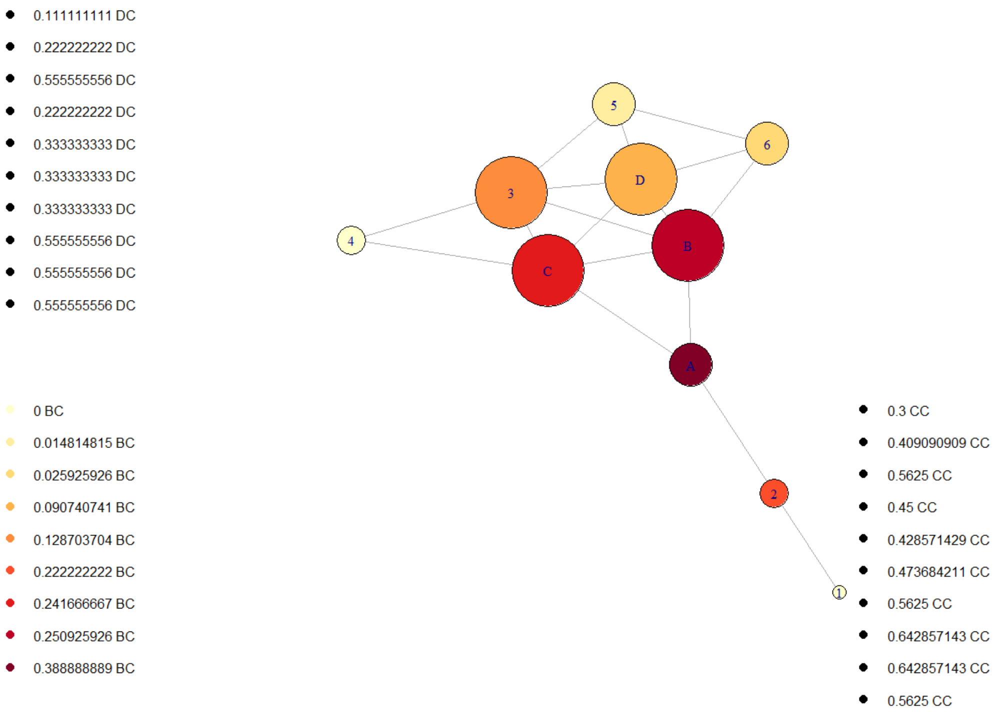
Seats B and C also have very similar statistics since their degree and closeness centrality is the same; however, B has a slightly higher betweenness centrality than C making it a better choice. Both B and C have the best closeness centrality in the bus at 0.6429. Thus, we should not choose C either as it is worse than B.

Lastly, seat A has the highest betweenness centrality overall at 0.38889 meaning that it acts as a bridge between the front and the back of the bus. However, it is worse off in all other centralities. Therefore, if somebody values bringing two separated groups together and acting as a mediator, they should chose seat A. Whereas, if they prefer being connected with the most people and easily reachable by everyone due to their central location they should chose seat B.

I personally would chose seat A since that would guarantee that my connections are diverse, instead of numerous. I do not need to know everyone in a particular group, having one key connection is enough in order to get things done if needed; quality over quantity!



Part 4: The Network Plotted



Here is the plotted version of the network. The size of the nodes represents the degree centrality, whose values are on the top left in alphanumeric seat order (i.e. from 1 to D). Node color represents betweenness centrality, with higher values being in darker colors; as we showed, node A is the darkest, with nodes 1 and 4 being the lightest; the legend is on the bottom left in increasing order of betweenness centrality. Lastly, the closeness centrality is on the bottom right also in alphanumerical order.