Graph Theory

Semester Project

Topic:

Social Network Analysis

Submitted To:

Dr. Irfan Yousaf

Submitted By:

Shaheer Arshad (2020-CS-639)

Degree & Semester:

BSCS (6th Semester)

Department of Computer Science

University of Engineering and Technology, New Campus, Lahore

(1)

Different Measures

We have three graphs to consider they have the information of the friendship links or the edges present between different students or nodes. The txt files contain edge list of all the edges present between the nodes of that graph. Each tuple in the file represent the starting node of edge as first element and the second is the node on which the edge is incident. The basic details about these graphs is as below:

- Nodes in Graph A = 52
- Edges in Graph A = 167
- Nodes in Graph B = 44
- Edges in Graph B = 204
- Nodes in Graph AB = 95
- Edges in Graph AB = 397

The comparison of different measures of the three graphs is as follow:

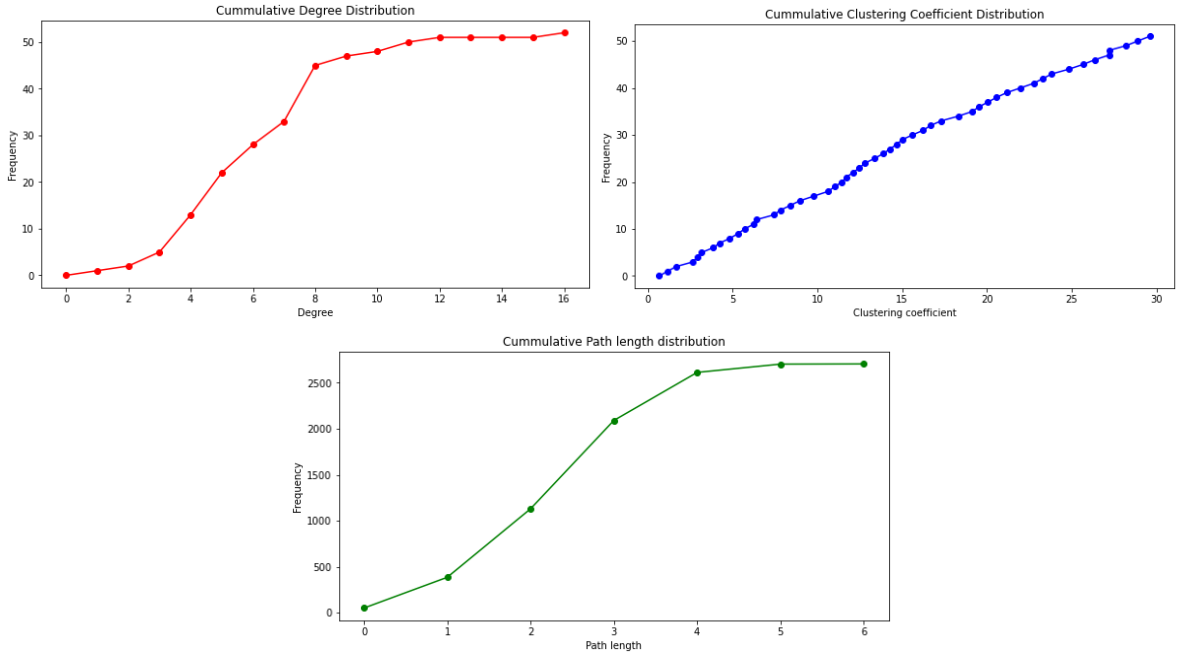
Measure	Graph A	Graph B	Graph AB
Avg Deg	6.423076923076923	9.272727272727273	6.423076923076923
Avg CC	0.5691239316239317	0.5966412931987798	0.5370493868017088
Avg Path Len	2.7345399698340875	2.184989429175476	2.9343784994400894
Diameter	6	4	6
Highest Deg	(623, 16)	(671, 17)	(671, 18)
Lowest Deg	(693, 1)	(694, 3)	(709, 2)

In this table, it is seen that the average degree for Graph B is the highest showing that Section B have more friendship relations as compared to Section A. When both sections combined, the average degree is same as Graph A. Similarly, the average clustering coefficient of Graph B is highest showing they have more tendency to make clusters. The average path length of the shortest paths in the Graph AB is highest showing that there is more variation in that as compared to Graph A and B. The average path length in Graph B is the shortest showing more compact relations. The diameter of Graph B is the lowest as they are friendlier. The highest degree node in Graph A is 623 as it is the most popular node of the graph with degree 16 and 693 is the least significant node with only 1 degree. In Graph B, the node 671 is most influential as it has the highest degree of 17 and 694 has the least degree of 3. In Graph AB, the node 671 is again most influencing and this time with degree 18 and node 709 is least with degree only 2.

(2)

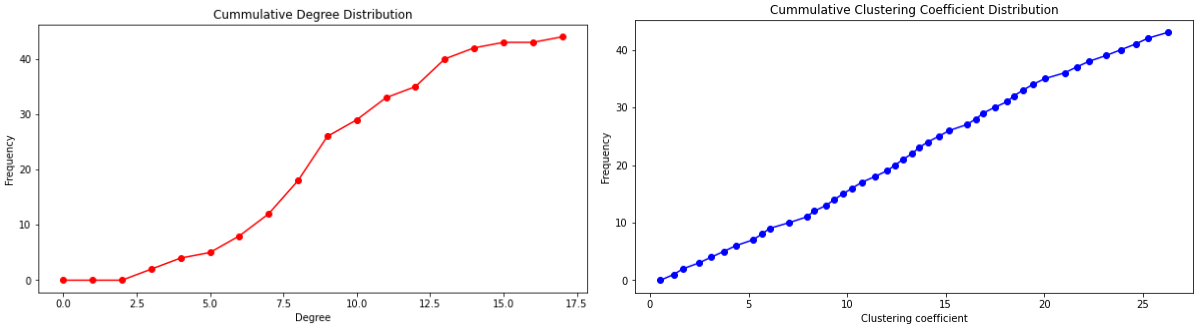
Cumulative Frequency Distributions

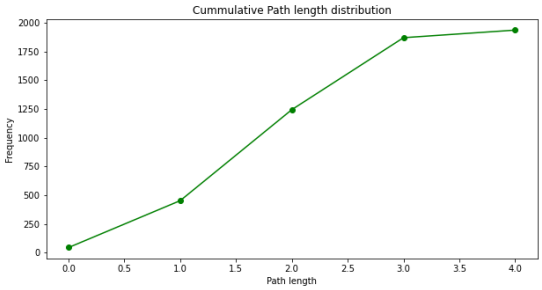
Graph A:



In Graph A, cumulative degree distribution graph as red above. It has observed that as degree increase the number of total nodes covering it are increased. There are most nodes with degree 8 as the curve is steep there, no nodes with degree 13, 14 and 15 as there is no change there, and there is no isolated node. The distribution of clustering coefficient as blue seems continuous, as there is slight change in the CC of nodes and it is increasing for cumulative distribution. The graph in green shows the path length distribution, there are more paths with length 2, 3 and 4 and a few with length 1 and even fewer with length 5 and 6.

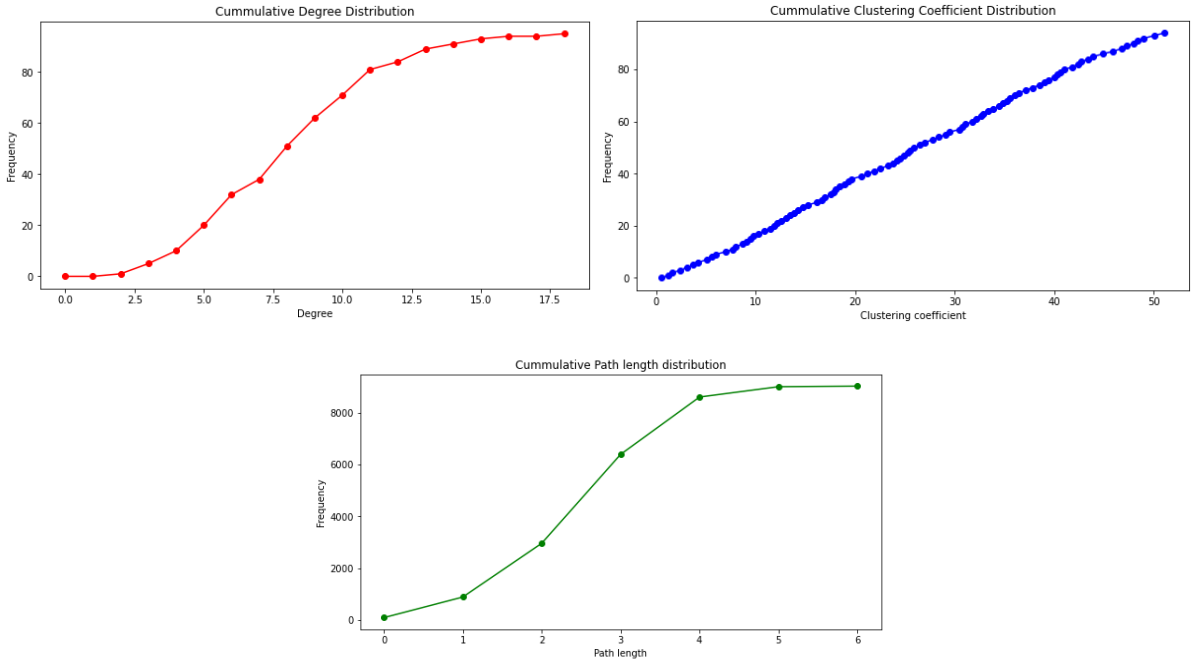
Graph B:





In Graph B, cumulative degree distribution graph as red above. It has observed that as degree increase the number of total nodes covering it are increased. There are most nodes with degree 6 to 11 as the curve is steep there, no nodes with degree 1, 2 and 16 as there is no change there, and there is no isolated node with degree 0. The distribution of clustering coefficient as blue seems continuous, as there is slight change in the CC of nodes and it is increasing for cumulative distribution. The graph in green shows the path length distribution, there are more paths with length 2 and 3 and a few with length 1 and even fewer with length 4.

Graph AB:

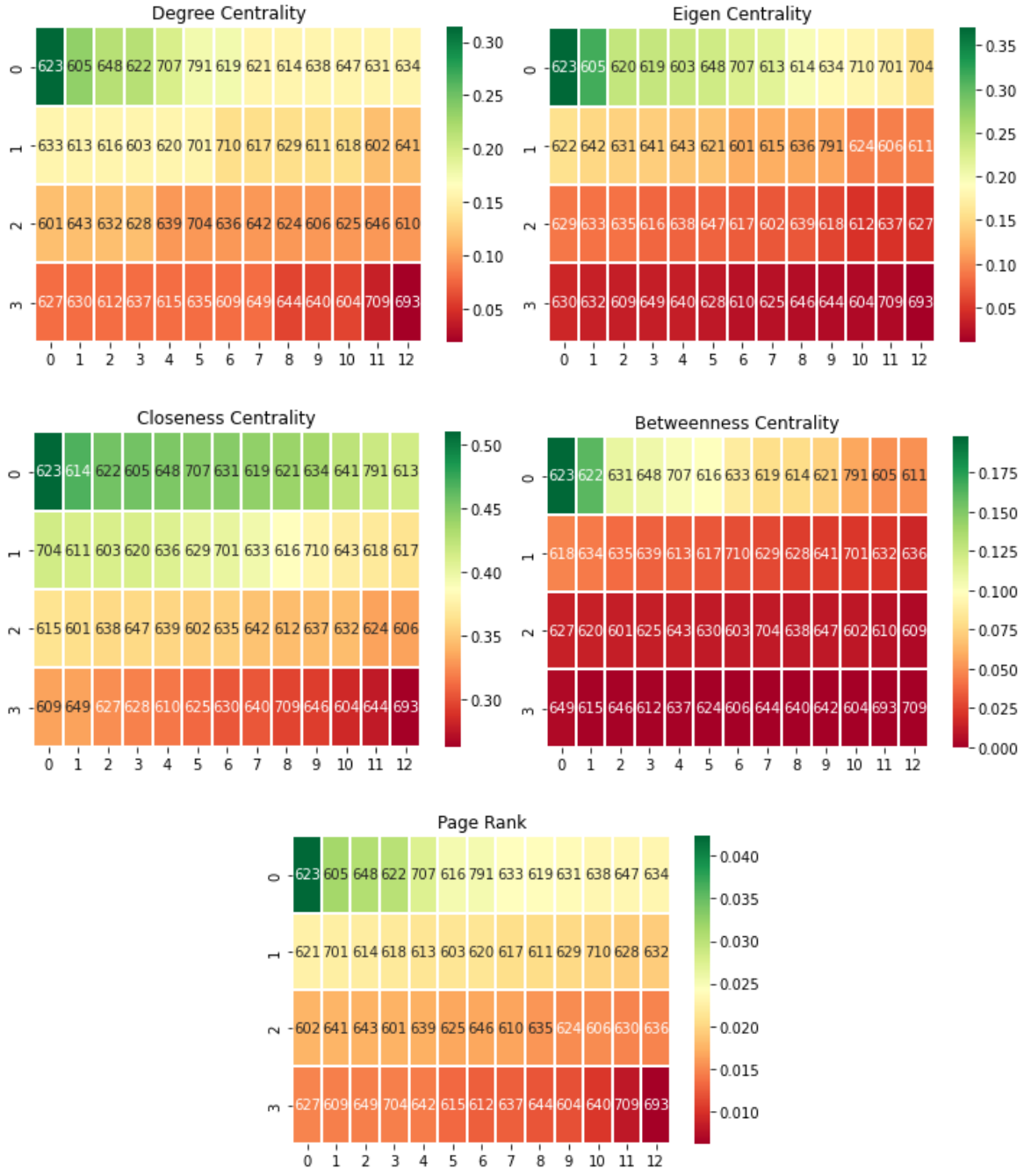


In Graph AB, cumulative degree distribution graph as red above. It has observed that as degree increase the number of total nodes covering it are increased. There are most nodes with degree 5 to 11 as the curve is steep there, no nodes with degree 16 and 17 as there is no change there, and there is no isolated node. The distribution of clustering coefficient as blue seems continuous, as there is slight change in the CC of nodes and it is increasing for cumulative distribution. The graph in green shows the path length distribution, there are more paths with length 2, 3 and 4 and a few with length 1 and even fewer with length 5 and no path with length 6.

(3)

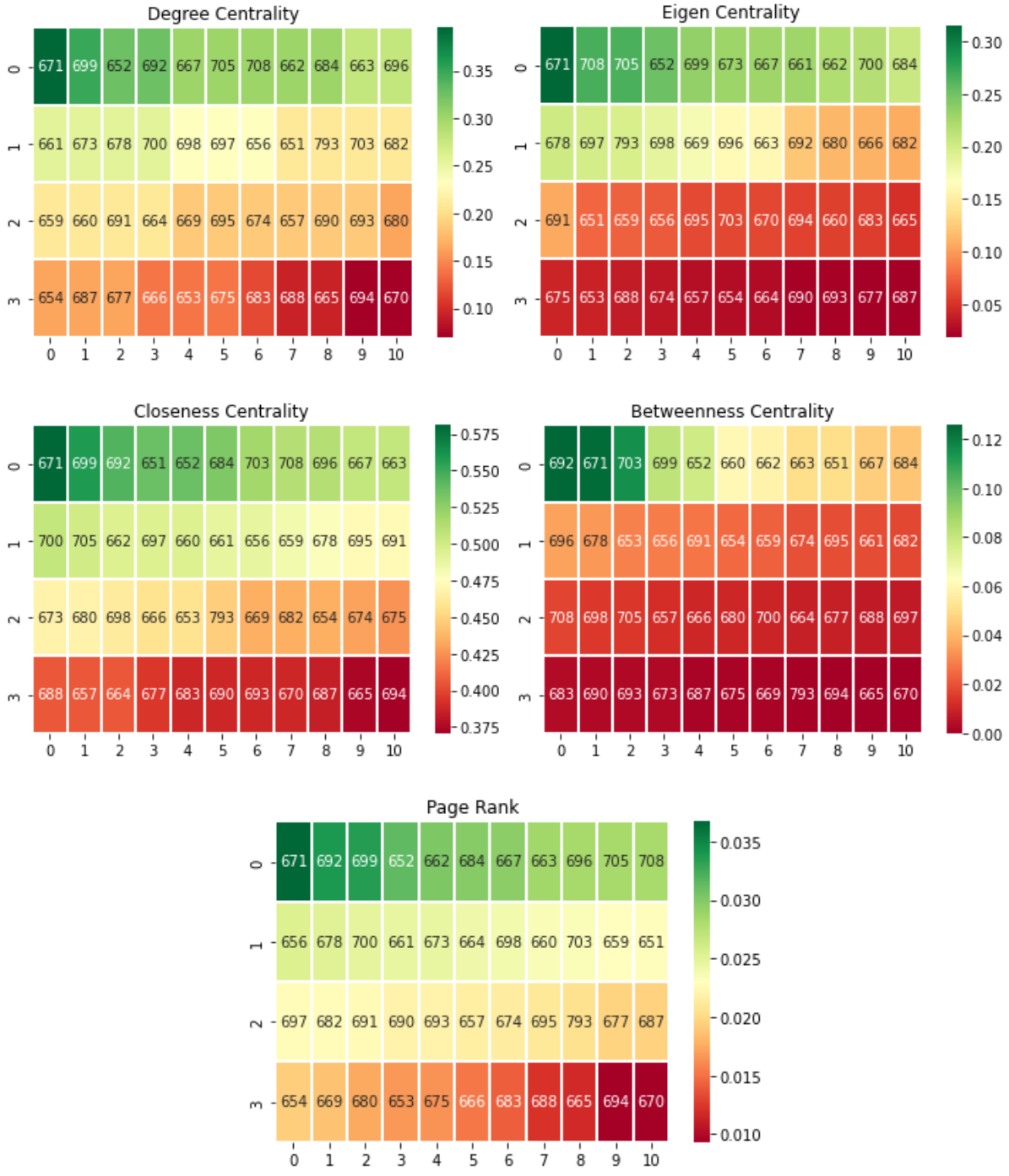
Centrality Measures

Graph A:



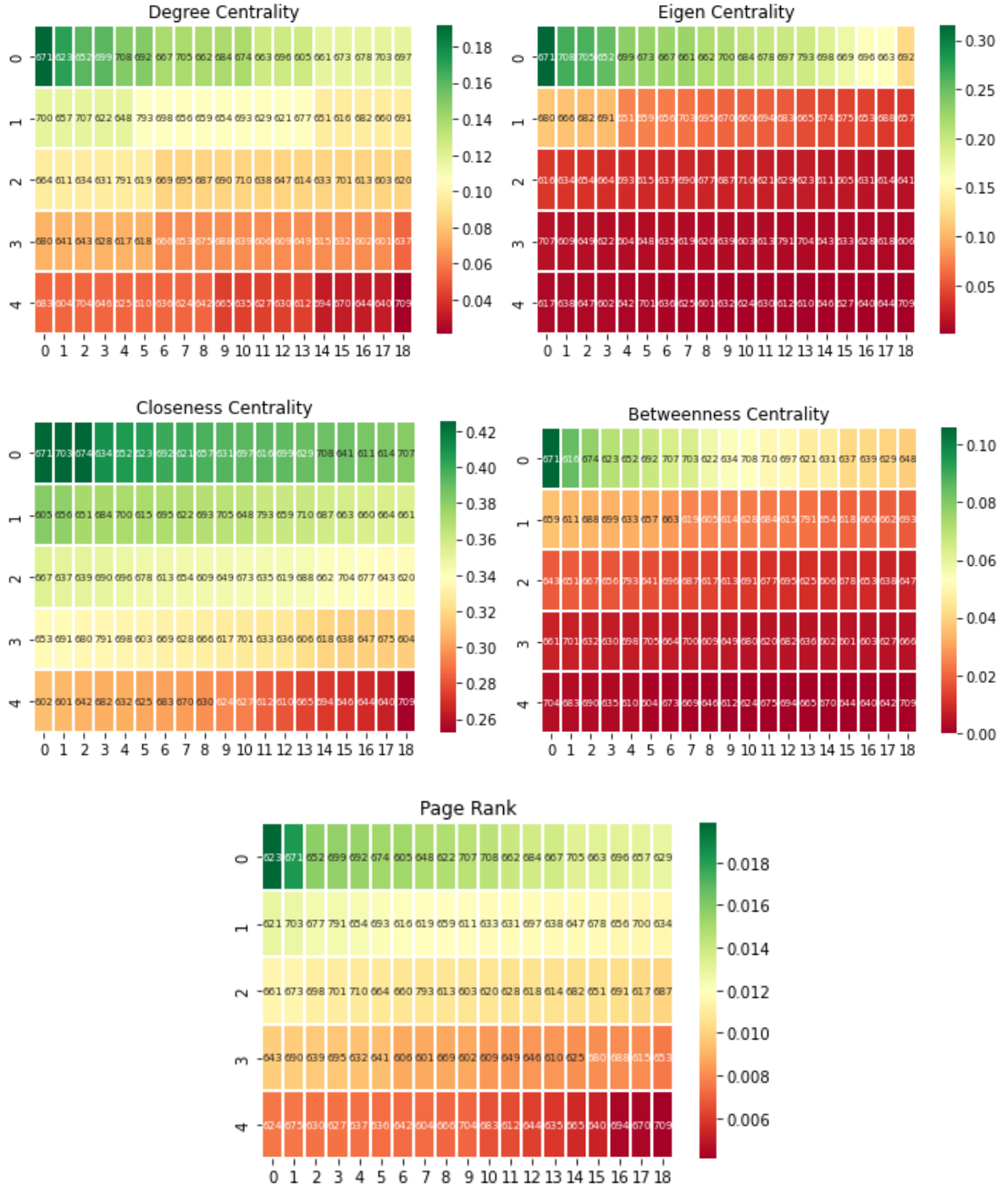
The above heat maps show the sorted maps for different centrality measures. The green coloured tiles show that they are the nodes with the highest values for the specific centrality or in other words, these nodes are most influential in the Graph A. Moreover, the red coloured tiles are the ones with least centrality measure. The ones with light green, yellow and orange shade are the ones with moderate value.

Graph B:



The above heat maps show the sorted maps for different centrality measures. The green coloured tiles show that they are the nodes with the highest values for the specific centrality or in other words, these nodes are most influential in the Graph B. Moreover, the red coloured tiles are the ones with least centrality measure. The ones with light green, yellow and orange shade are the ones with moderate value.

Graph AB:

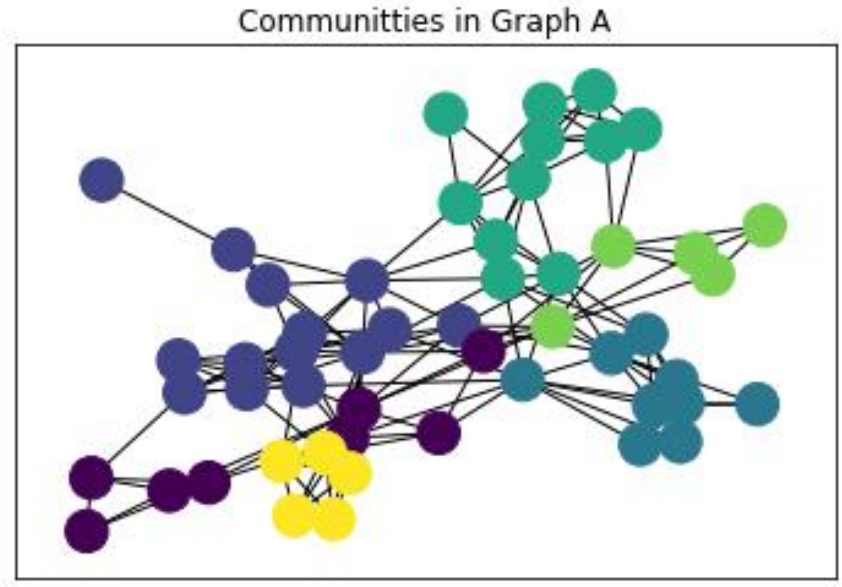


The above heat maps show the sorted maps for different centrality measures. The green coloured tiles show that they are the nodes with the highest values for the specific centrality or in other words, these nodes are most influential in the Graph AB. Moreover, the red coloured tiles are the ones with least centrality measure. The ones with light green, yellow and orange shade are the ones with moderate value.

(4)

Communities

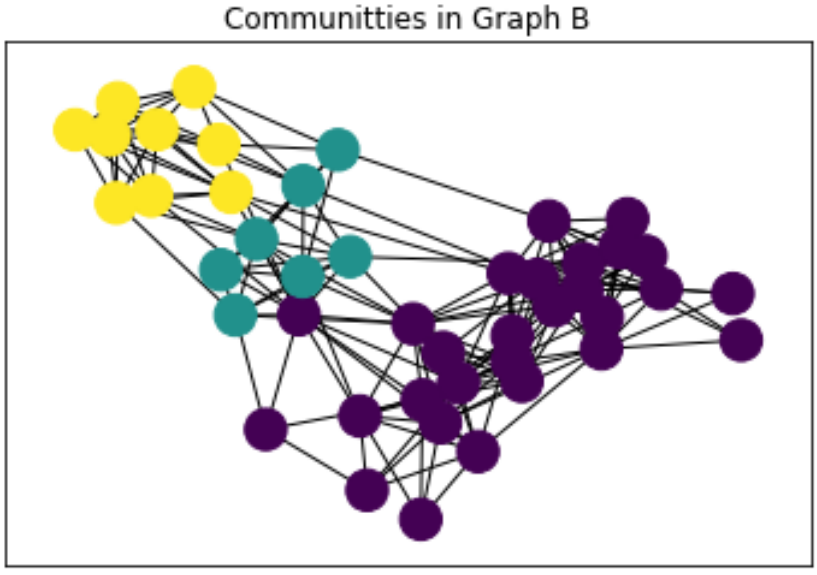
Graph A:



The Following are the different communities within Graph A:

- {627, 644, 630, 639}
- {704, 641, 707, 643, 614, 615, 648, 623, 624, 693, 631, 601, 634, 635, 636, 701, 606}
- {642, 613, 710, 619, 620, 603, 605}
- {640, 612, 647, 617, 622, 791, 602, 637, 638}
- {610, 611, 709, 646, 618, 621, 625, 628, 629, 632, 633}
- {616, 609, 649, 604}

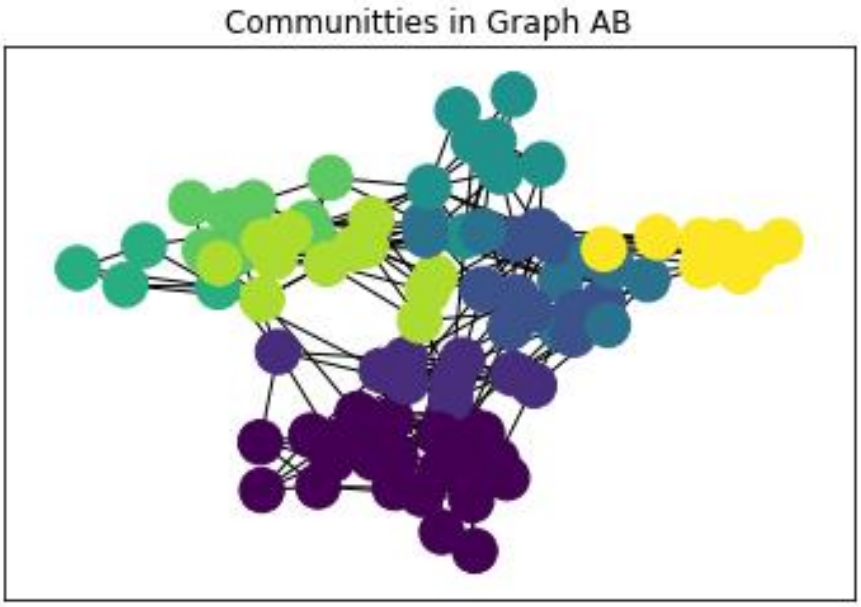
Graph B:



The Following are the different communities within Graph B:

- {652, 661, 662, 663, 793, 666, 667, 665, 669, 670, 671, 673, 678, 680, 682, 683, 684, 691, 694, 696, 697, 698, 699, 700, 705, 708}
- {675, 651, 653, 656, 688, 659, 692, 660, 695}
- {674, 677, 654, 687, 657, 690, 693, 664, 703}]

Graph AB:



The Following are the different communities within Graph A:

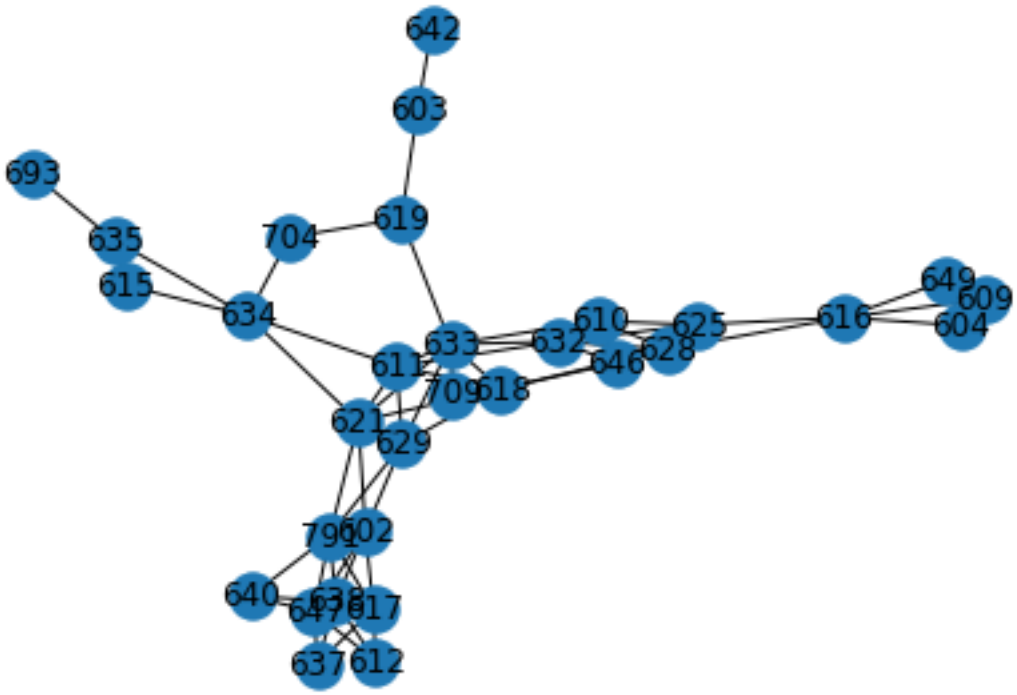
- {652, 661, 662, 663, 793, 666, 667, 665, 669, 670, 671, 673, 678, 680, 682, 683, 684, 691, 694, 696, 697, 698, 699, 700, 705, 708}
- {675, 651, 653, 656, 659, 692, 660, 695}
- {674, 611, 677, 621, 654, 687, 657, 690, 693, 629, 664, 703}
- {641, 610, 709, 614, 646, 616, 618, 625, 628, 631, 632, 633}
- {640, 612, 647, 617, 622, 791, 602, 637, 638}
- {707, 643, 648, 688, 624, 601, 636, 701, 606}
- {627, 644, 630, 639}
- {609, 649, 604}
- {704, 642, 603, 613, 710, 615, 619, 620, 623, 634, 635, 605}

(5)

Checking Influence

Graph A:

After removing node 639, its friends and friends of friends the Graph A is as follow:



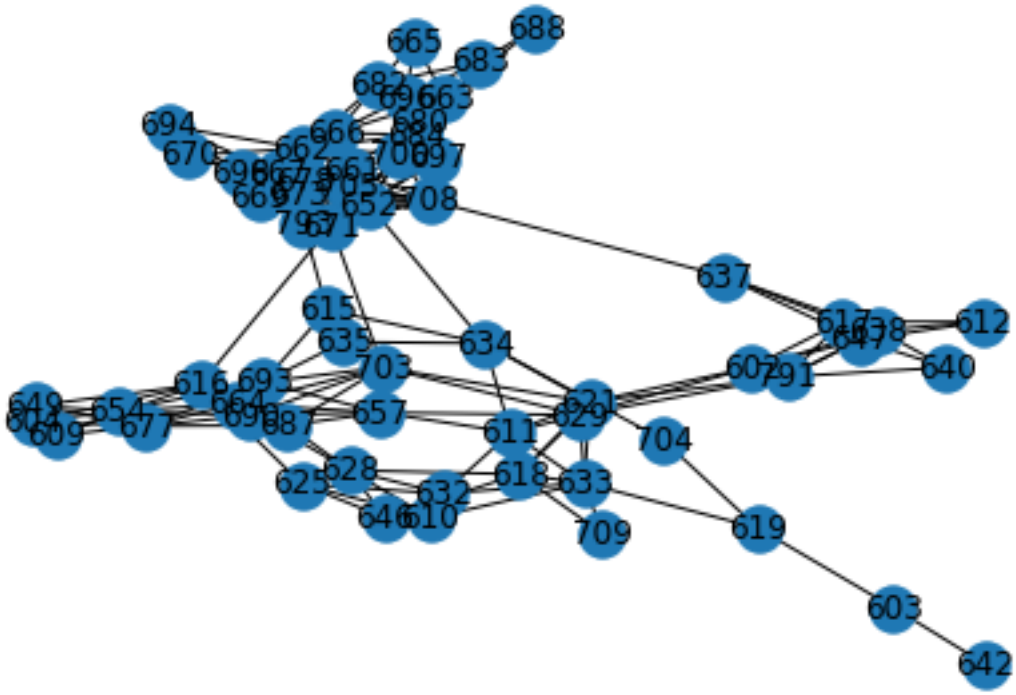
Cut-vertex set: {603, 616, 619, 634, 635}

After removing my node, my friends and friend of friends. The Graph A remains connected, as my node is not that influential in the graph. Removing my node just removes 4 to 5 edges and the neighbouring nodes are only 4 or 5. Further removing the friends of my friends there is no such big change in the graph and the graph remains connected.

In this case, we have to find the cut vertex set given above after removing me, my friends and friends of friends we further have to remove 603, 616, 619, 634 and 635 to make the graph disconnected.

Graph AB:

After removing node 639, its friends and friends of friends the Graph AB is as follow:



Cut-vertex set: {619, 603}

After removing my node, my friends and friend of friends. The Graph AB remains connected, as my node is not that influential in the graph. Moreover, I have not much links in Graph B and combining the both graphs i.e. Graph AB, removing my node just removes 4 to 5 edges and the neighbouring nodes. Further removing the friends of my friends there is no such big change in the graph and the graph remains connected.

In this case, we have to find the cut vertex set given above after removing me, my friends and friends of friends we further have to remove 603, 616, 619, 634 and 635 to make the graph disconnected.