Graph Theory Project

PROJECT Report

Submitted to:

Dr. Irfan Yousaf

Submitted by:

Muhammad Ahmad Khan (2020-CS-651)

Department of Computer Science

UNIVERSITY OF ENGINEERING AND TECHNOLOGY LAHORE,
NEW CAMPUS

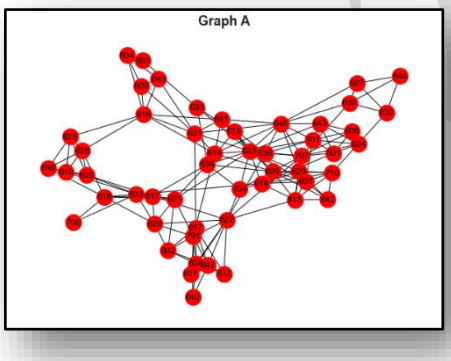
Date: 05-04-2023

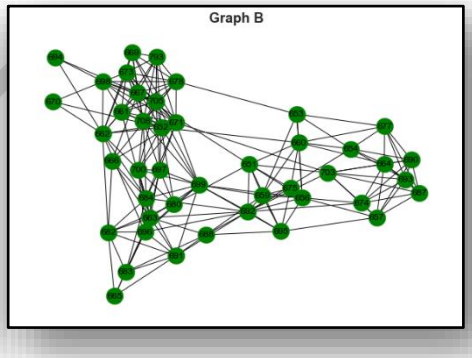
Social network analysis (SNA) is the process of investigating social structures using network and graph theories. It aims to understand and analyze the patterns of relationships and interactions among individuals, groups, or organizations in a social system. In the following report we can do some SNA on Graphs and exhibit some information.

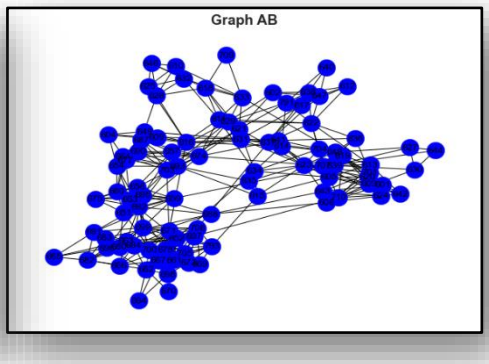
1. **Average degree, average clustering coefficient, average path length, diameter, highest degree node, lowest degree node etc. of the graphs.**

	Graph A	Graph B	Graph AB
Avg. Degree	6.423	9.27	8.358
Avg. Clustering Coefficient	0.569	0.596	0.537
Avg. Path Length	2.735	2.185	2.934
Diameter	6	4	6
Highest Degree Node	623 with degree 16	671 with degree 17	671 with degree 18
Lowest Degree Node	693 with degree 1	694 with degree 3	709 with degree 2

- **Graphs of Section A, B and AB:**





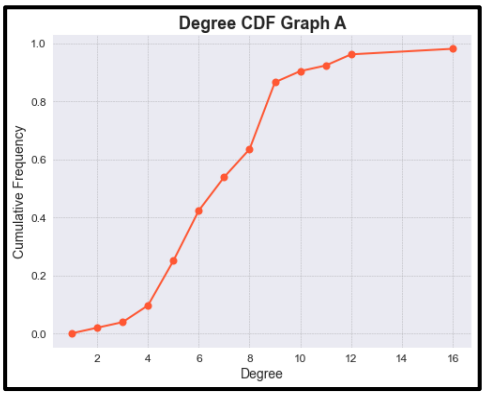


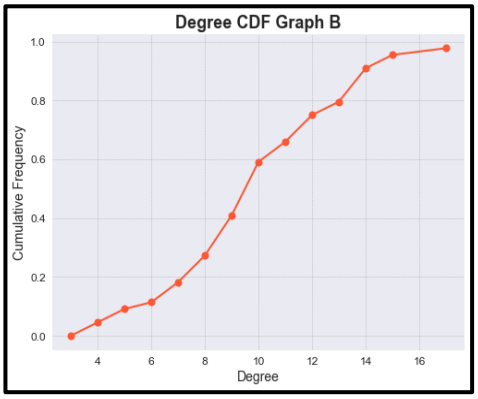
Graph A **Graph B** **Graph AB**

Graph A has 52 nodes and 167 edges connecting them while **Graph B** has 44 nodes and 204 edges connecting them. When these two graphs are combined, they form **Graph AB**, which has 95 nodes and 397 edges connecting them. This makes Graph AB very complex and connected network.

2. **Distribution of degree, clustering coefficient and path length of nodes.**

- **Degree Distribution:**





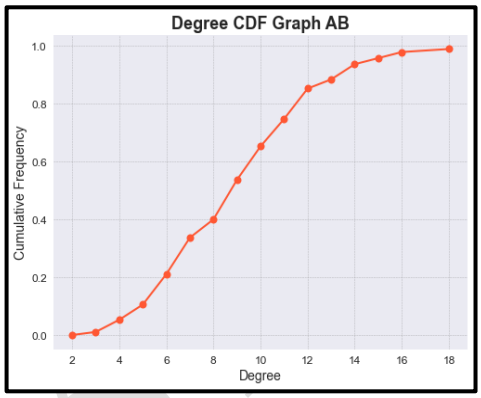


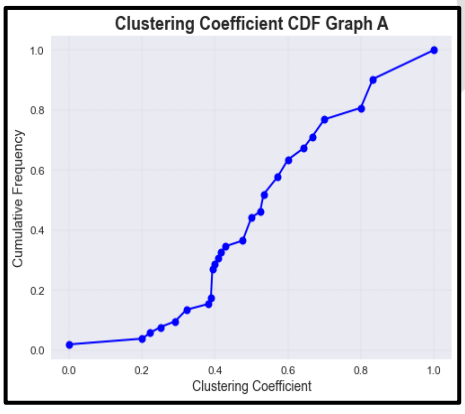
Fig.(a) **Fig.(b)** **Fig.(c)**

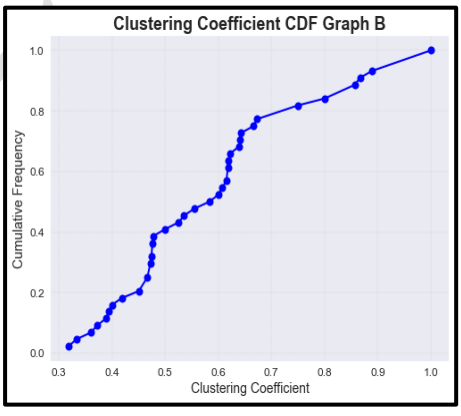
In fig.(a) that shows the CDF graph of Section A suggests that the phenomenon has a higher degree within the range of 2 to 9, and a certain degree the range of 10 to 16, with a gradual increase in certainty.

In fig.(b) CDF graph of Section B shows that the likelihood of the phenomenon having a degree between 4 to 6 is low, but it increases considerably from 7 to 10, implying that it's more likely to have a higher degree within this range. Additionally, the probability of the phenomenon having an even higher degree increases further between 11 to 16, almost certain implying that the phenomenon has a significant degree within this range.

In fig.(c) the CDF graph of AB it shows that the probability of having degree between 2 to 6 is low. As the degree increases beyond 6 up to 12 the probability of graph having higher degree. The graph also displays gradual increase in probability from degree 13 to 18.

- **Clustering Coefficient:**





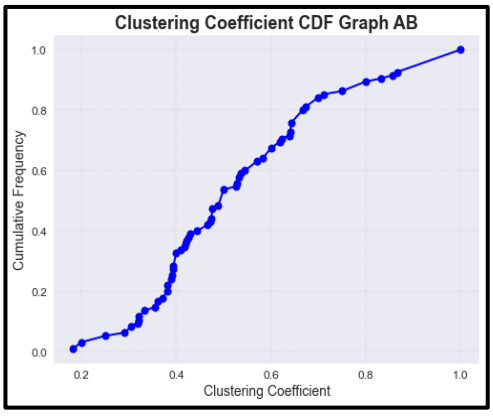


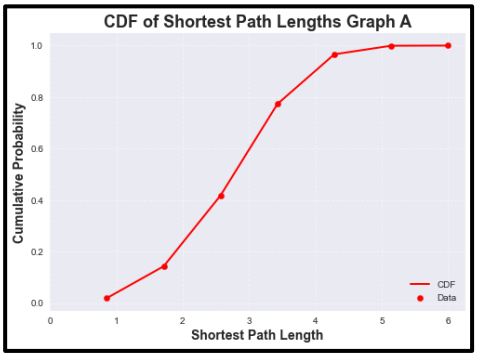
Fig.(a) **Fig.(b)** **Fig.(c)**

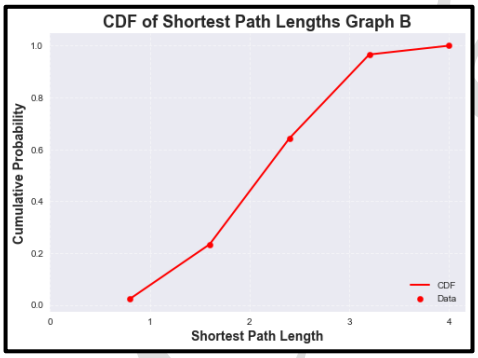
In Fig.(a) CDF of graph A shows the clustering coefficient that as the coefficient increases the probability of having higher coefficient also increases. The probability is low for coefficients between 0 to 0.4, moderate for coefficients between 0.4 to 0.8 and it has more clusters in it and high for coefficients greater than 0.8 up to 1 but in this range fewer clusters are there.

In Fig.(b) that illustrate the Graph B indicates that the probability of having a clustering coefficient between 0.3 to 0.7 is moderate, ranging from 0 to 0.8 but it contains more clusters. However, for clustering coefficients between 0.7 to 1, probability of having clustering coefficient is very high ranging from 0.8 to 1 but it contains less clusters.

In Fig.(c) that shows the Graph AB indicates that the probability of having more clusters is higher for clustering coefficients between 0.2 to 0.6 with probability ranging from 0 to 0.8. Conversely, for clustering coefficients between 0.7 to 1 the probability of having fewer cluster is higher with probability ranging from 0.8 to 1.

- **Path Length:**





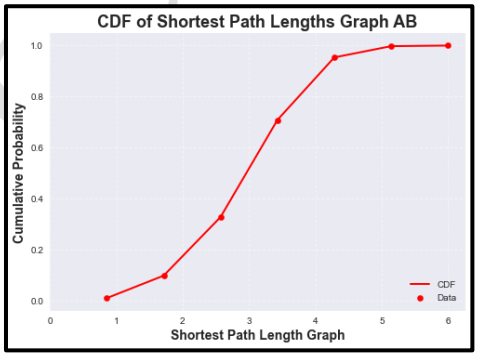


Fig.(a) **Fig.(b)** **Fig.(c)**

In Fig.(a) that illustrates the Graph A indicates that having a path length of 1 is impossible, while path lengths from 2 to 4 increasingly likely ranging from 0.1 to 0.9 probability. Path lengths greater than 5 are almost certain with a probability of 1 as shown by a sudden increase and then a flat line in the graph.

In Fig.(b) that represents Graph B shows that it's not possible to have a path length of 1. However, for path lengths ranging from 2 to 3 there is a considerable probability of having path of this length ranging from 0.2 to 0.9. For path lengths greater than 3 the probability of having path pf that length becomes very high with probability of 1.

In Fig.(c) that represents Graph AB shows that the probability of having a path length of 1 is 0. However, as the path length increases from 2 to 4, the probability of the phenomenon having a path length within this range increases significantly, ranging from 0.1 to 0.9. For path lengths greater than 4, the probability of having a longer path is very high, ranging from greater than 0.9 to 1. The graph shows a sharp increase in probability for path lengths 2-4, and then a gradual increase for longer path lengths until it reaches a probability of 1.

3. 5 types of centrality measures of all the nodes.

Sr.	Max. 10 Degree Centrality						Min. 10 Degree Centrality					
	Graph A		Graph B		Graph AB		Graph A		Graph B		Graph AB	
	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value
1	**603**	0.156	**696**	0.279	**705**	0.138	**693**	0.019	**694**	0.069	**709**	0.021
2	**620**	0.156	**667**	0.302	**662**	0.138	**709**	0.039	**670**	0.069	**694**	0.031
3	**701**	0.156	**705**	0.302	**684**	0.138	**644**	0.058	**688**	0.093	**670**	0.031
4	**791**	0.176	**708**	0.302	**674**	0.138	**640**	0.058	**665**	0.093	**644**	0.031
5	**619**	0.176	**662**	0.302	**708**	0.149	**604**	0.058	**683**	0.116	**640**	0.031
6	**707**	0.196	**684**	0.302	**692**	0.149	**627**	0.078	**666**	0.139	**665**	0.042
7	**648**	0.215	**652**	0.325	**652**	0.159	**630**	0.078	**653**	0.139	**635**	0.042
8	**622**	0.215	**692**	0.325	**699**	0.159	**612**	0.078	**675**	0.139	**627**	0.042
9	**605**	0.235	**699**	0.348	**623**	0.17	**637**	0.078	**680**	0.162	**630**	0.042
10	**623**	0.313	**671**	0.395	**671**	0.191	**615**	0.078	**654**	0.162	**612**	0.042

Table: 1 Degree Centrality of Graph A, B and AB

Sr.	Max. 10 Betweenness Centrality						Min. 10 Betweenness Centrality					
	Graph A		Graph B		Graph AB		Graph A		Graph B		Graph AB	
	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value
1	**621**	0.074	**667**	0.045	**634**	0.053	**644**	0	**694**	0	**694**	0
2	**614**	0.078	**651**	0.05	**622**	0.057	**640**	0	**665**	0	**665**	0
3	**619**	0.079	**663**	0.05	**703**	0.063	**642**	0	**670**	0	**670**	0
4	**633**	0.086	**662**	0.05	**707**	0.063	**604**	0	**793**	0.0009	**644**	0
5	**616**	0.099	**660**	0.06	**692**	0.066	**693**	0	**669**	0.001	**640**	0
6	**707**	0.102	**652**	0.077	**652**	0.069	**709**	0	**675**	0.001	**642**	0
7	**648**	0.106	**699**	0.082	**623**	0.07	**624**	0.0007	**687**	0.001	**709**	0
8	**631**	0.109	**703**	0.115	**674**	0.076	**606**	0.0007	**673**	0.002	**675**	0.0003
9	**622**	0.161	**671**	0.124	**616**	0.084	**612**	0.0013	**690**	0.003	**624**	0.0003
10	**623**	0.197	**692**	0.125	**671**	0.105	**637**	0.0013	**693**	0.003	**612**	0.0003

Table: 2 Betweenness Centrality of Graph A, B and AB

Sr.	Max. 10 Closeness Centrality						Min. 10 Closeness Centrality					
	Graph A		Graph B		Graph AB		Graph A		Graph B		Graph AB	
	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value
1	**634**	0.435	**700**	0.505	**631**	0.394	**693**	0.263	**694**	0.37	**709**	0.252
2	**621**	0.439	**708**	511	**657**	0.396	**644**	0.277	**665**	0.373	**640**	0.264
3	**619**	0.443	**696**	0.511	**692**	0.4	**604**	0.283	**687**	0.383	**644**	0.267
4	**707**	0.447	**703**	0.518	**621**	0.4	**646**	0.293	**683**	0.387	**646**	0.27
5	**631**	0.447	**684**	0.53	**623**	0.405	**709**	0.297	**690**	0.387	**694**	0.272
6	**648**	0.451	**651**	0.538	**652**	0.406	**630**	0.304	**693**	0.387	**665**	0.275
7	**622**	0.455	**652**	0.538	**634**	0.41	**640**	0.304	**670**	0.387	**610**	0.28
8	**605**	0.455	**692**	0.544	**671**	0.42	**625**	0.311	**677**	0.394	**612**	0.283
9	**614**	0.467	**699**	0.558	**703**	0.42	**610**	0.316	**688**	0.405	**627**	0.291
10	**623**	0.51	**671**	0.581	**674**	0.42	**628**	0.321	**657**	0.405	**624**	0.294

Table: 3 Closeness Centrality of Graph A, B and AB

Rank	Eigen Centrality					
	Graph A		Graph B		Graph AB	
	Node	Value	Node	Value	Node	Value
1	**623**	0.369	**671**	0.314	**671**	0.315
2	**605**	0.314	**708**	0.266	**708**	0.266
3	**620**	0.242	**705**	0.265	**705**	0.262
4	**619**	0.24	**652**	0.252	**652**	0.252
5	**603**	0.235	**699**	0.234	**699**	0.232
6	**648**	0.232	**673**	0.229	**673**	0.227
7	**707**	0.222	**667**	0.228	**667**	0.226
8	**613**	0.217	**661**	0.228	**661**	0.225
9	**614**	0.197	**662**	0.219	**662**	0.216
10	**634**	0.183	**700**	0.216	**700**	0.214

Table: 4 Eigen Centrality of Graph A, B and AB

Rank	PageRank Centrality					
	Graph A		Graph B		Graph AB	
	Node	Value	Node	Value	Node	Value
1	**623**	0.042	**671**	0.037	**623**	0.02
2	**605**	0.031	**692**	0.034	**671**	0.018
3	**648**	0.031	**699**	0.034	**652**	0.016
4	**622**	0.03	**652**	0.031	**699**	0.016
5	**707**	0.027	**662**	0.03	**692**	0.015
6	**616**	0.025	**684**	0.03	**674**	0.015
7	**791**	0.025	**667**	0.03	**605**	0.015
8	**633**	0.024	**663**	0.029	**648**	0.015
9	**619**	0.024	**696**	0.028	**622**	0.015
10	**631**	0.024	**705**	0.028	**707**	0.015

Table: 5 PageRank Centrality of Graph A, B and AB

4. Find Communities in the graph and show them in different colours in the graph.

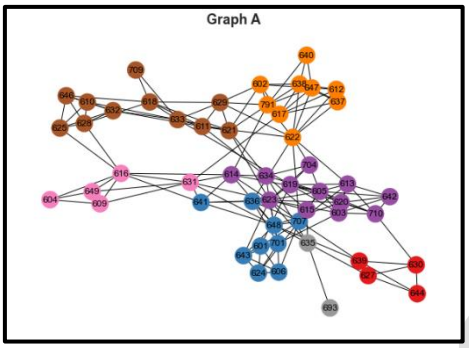


Fig.(a)

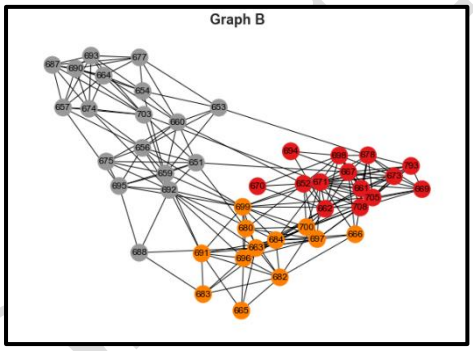


Fig.(b)

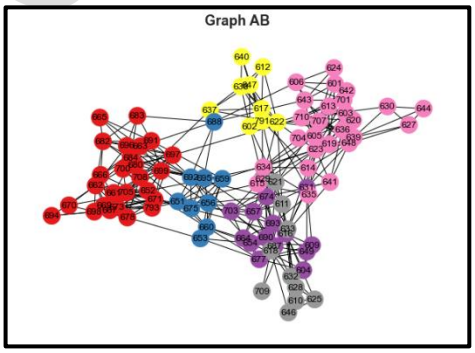


Fig.(c)

In Fig.(a) The graph shows 7 communities within a network. These communities consist of nodes that are grouped together based on similarities or interactions. Each community has a higher number of connections within the community compared to outside the community. The nodes within a community tend to be more densely connected with each other, creating a substructure within the overall network. The identification of communities within the graph can provide insights into the structure and function of the network, as well as the relationships between the nodes.

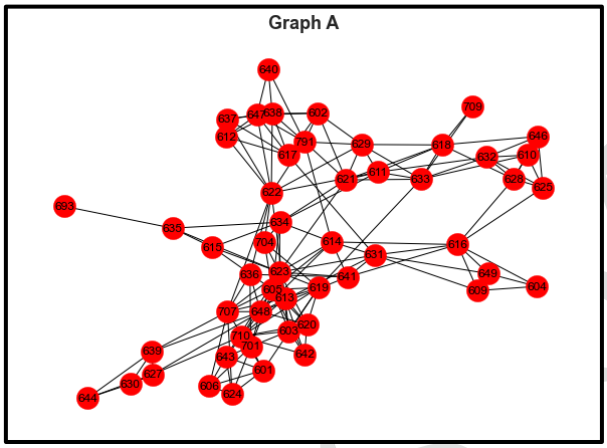
In Fig.(b) The graph depicts a network that has been partitioned into three distinct communities. These communities are formed by groups of nodes that exhibit a higher degree of interconnectedness amongst themselves compared to their connections with the nodes in other communities. The nodes within a community tend to have a more dense and cohesive structure, resulting in a distinct substructure within the overall network.

In Fig.(c) The graph displays a complex network that has been segmented into six distinct communities. Each community is comprised of a group of nodes that are densely

interconnected with one another and exhibit a higher degree of cohesion compared to their connections with nodes outside of the community.

5. **Remove yourself, your friends, and friends of friends and then check if the graph is connected or not?**

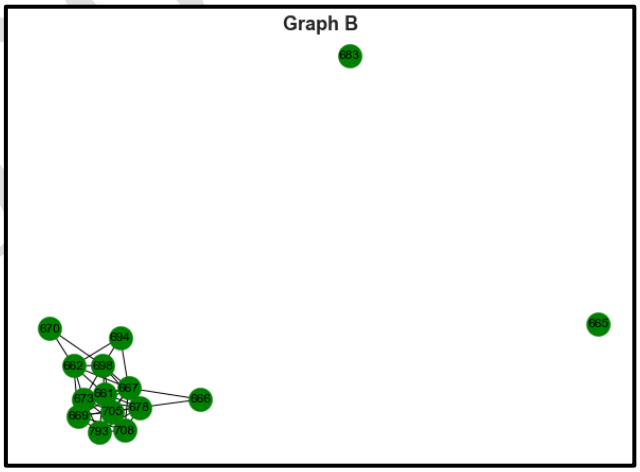
- **Graph A:**



The above Graph A is connected.

But it has **Cut-Vertex Set {635}**. By removing this node, the above graph becomes disconnected.

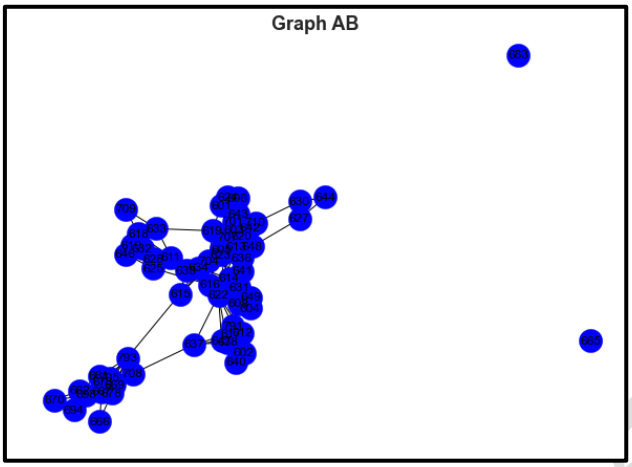
- **Graph B:**



The above graph is disconnected by removing myself, my friends and my friends of friends.

The graph has **3 connected components: [{673, 705, 708, 678, 666, 661, 694, 662, 793, 698, 667, 669, 670}, {665}, {683}].**

- **Graph AB:**



The above graph is disconnected by removing me, my friends, and my friends of friends.

The graph has **3 connected components: [{640, 641, 642, 643, 644, 646, 647, 648, 649, 661, 662, 791, 793, 666, 667, 669, 670, 673, 678, 694, 698, 701, 704, 705, 707, 708, 709, 710, 601, 602, 603, 604, 605, 606, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 622 , 623, 624, 625, 627, 628, 630, 631, 632, 633, 634, 635, 636, 637, 638}, {665}, {683}].**