# Social Network Analysis

Heet Thakkar (202201431)
Sujal Manavadariya (202201084)
Sneh Joshi (202201048)
Kathan Khuman (202201091)
Stavan Ravisaheb (202201436)
Nirupam Nayak (202201437)

March 2023

# 1 INTRODUCTION

In this research, we'll investigate how discrete mathematics may be used to analyse social networks.The study of social networks has grown in significance across a number of disciplines, including business, computer science, and sociology.Insights into the structure, dynamics, and characteristics of social networks may be gained by utilising the power of discrete mathematics.

These insights can be used to better understand social behaviour, recognise important persons or communities, and forecast information dissemination patterns.

# 2 THE PROBLEM

This project depicts the Social Network Analysis(SNA) which we have solved by applying the algorithms of Discrete Mathematics.

In today's connected world, communication, information sharing, and collaboration depend heavily on social networks.In order to improve communication, recognise influential people, and promote the spread of ideas, it is essential to comprehend network dynamics.

The following issues will be dealt with by this project:

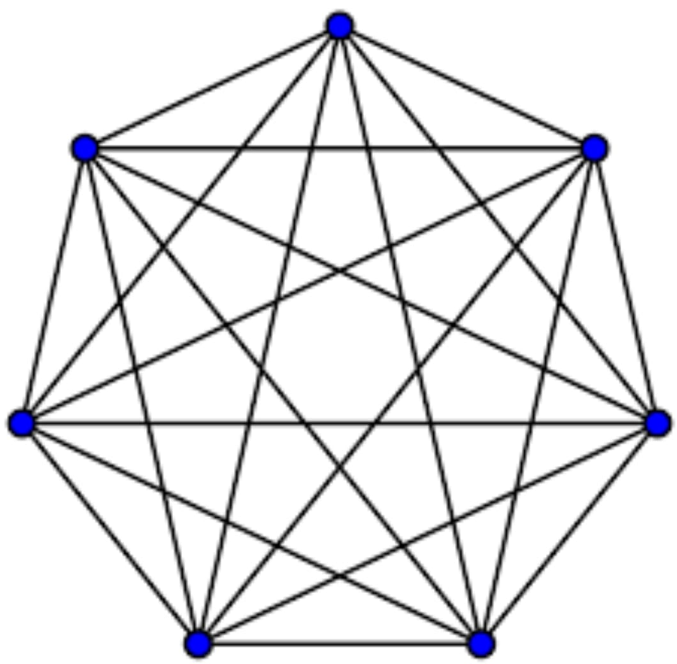
- Within a social network, who are the powerful people?
- What mechanisms does the network use to distribute information?
- How do influential people affect the spread of information?

We will use different algorithms to develop a centrality-based method for identifying influential people and analysing their influence. This method will be applied to discrete mathematics and social network analysis. We will be able to better understand social behaviour and network dynamics by using the new insights to inform our communication strategies.

# 3 PRE- REQUISITE KNOWLEDGE

## 3.1 COMPLETE GRAPHIC:

- An undirected graph in which every pair of distinct vertices is connected by a unique edge.
- The graph should just be symmetrical, and that's all we need to worry about.



## 3.2 Graph Notation:

- G(V,E) , in this term 'V' denotes the number of vertices and 'E' denotes the number of edges.
- Now, we are going to discuss about 'H' which is a sub-graph of 'G'.

### 3.3 Subgraph

It consists of a selection of edges and vertices from the main graph.

## 4 SOLUTION

There are various approaches and algorithms for Social Network Analysis(SNA).

### 4.1 Graph Theory

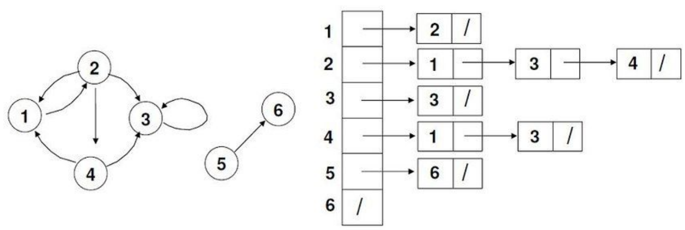
Graph Representation Algorithm:

**1)** Initialize an empty graph object.
**2)** Define the necessary data structures to represent the graph, such as nodes and edges.
**3)** Read the social network data and extract the information about nodes (individuals) and their relationships (edges).
**4)** For each node in the social network data: - Create a node object or data structure to store relevant information about the individual, such as their unique identifier, name, attributes, or any other relevant metadata. - Add the node to the graph object.
**5)** For each edge in the social network data: - Extract the source and target nodes of the edge. - Check if the nodes exist in the graph object. If not, create new node objects for them. - Create an edge object to represent the relationship between the source and target nodes. - Add the edge to the graph object, connecting the source and target nodes.
**6)** Once all nodes and edges have been added to the graph, the graph object now represents the social network data.
**7)** The graph can be further enhanced by adding attributes or properties to nodes and edges, such as weights, labels, or any other relevant information.
**8)** Additional graph operations or algorithms, such as centrality measures, community detection, or information diffusion models, can be applied using the graph representation.

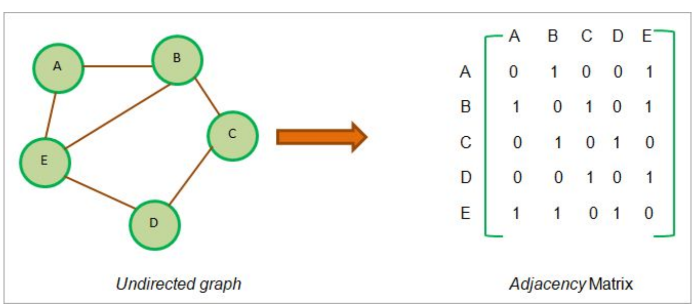
By following this algorithm, you can represent the social network data as a graph, where nodes represent individuals and edges represent their relationships. This graph representation forms the foundation for further analysis and exploration of the social network using discrete mathematics and social network analysis techniques.

In order to analyze a social network, it needs to be represented as a graph. There are two ways to represent a graph:

1) Adjacency lists: In this representation, each node in the network is stored as a vertex object, which contains a list of its adjacent vertices. This approach is memory-efficient for sparse graphs.



2) Adjacency matrices: In this representation, a matrix of size N x N is used, where N is the number of nodes in the network. The entry (i, j) in the matrix indicates whether there is an edge between node i and node j. This approach is suitable for dense graphs.



Undirected graph Adjacency Matrix

### 4.2 Centrality measures

Centrality measures help identify important nodes within a social network based on their connectivity and influence. Here are four common centrality measures:

**1) Degree Centrality**:

Definition: It quantifies the number of connections a node has in the network. The higher the degree centrality,the more central the node is. The algorithm involves counting the number of neighbors each node has.
When to use it: For finding very connected individuals, popular individuals, individuals who are likely to hold most information or individuals who can quickly connect with the wider network.

**2) Closeness Centrality**:

Definition: It measures how close a node is to all other nodes in the network. It is computed by calculating the average shortest path length from a node to all other nodes. Nodes with higher closeness centrality are more influential in terms of information flow.
When to use it: For finding the individuals who are best placed to influence the entire network most quickly.

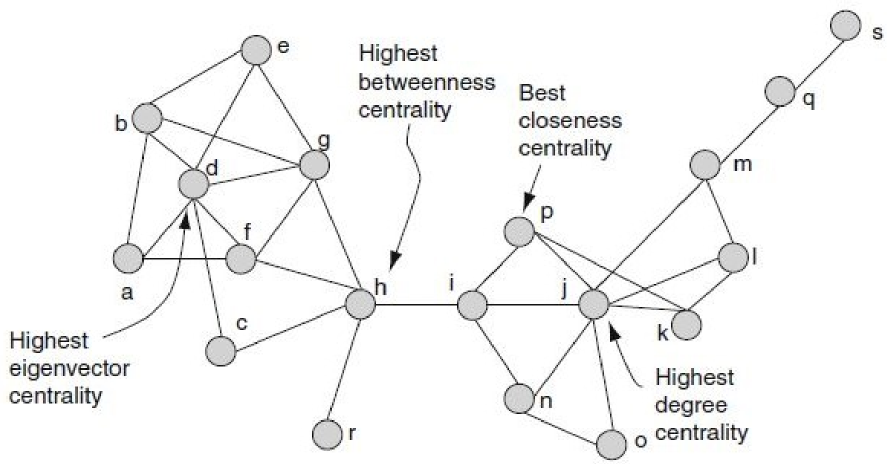
**3) Betweenness Centrality**:

Definition: It quantifies the importance of a node as a bridge between other nodes. It is computed by finding the shortest paths between all pairs of nodes and calculating the fraction of those paths that pass through a particular node. Nodes with higher betweenness centrality have more control over information flow.
When to use it: For finding the individuals who influence the flow around a system.

**4) Eigenvector Centrality**:

Definition: Based on the centrality of its surrounding nodes, eigenvector centrality calculates how central a node is. Nodes that are connected to other significant nodes receive higher scores. The eigenvector

linked to the largest eigenvalue of the network's adjacency matrix is used to calculate eigenvector centrality. The centrality scores for each network node are determined through an iterative process.
When to use it: Eigenvector Centrality is a good 'all-round' Social Network Analysis (SNA) score, handy for understanding human social networks, but also for understanding networks like malware propagation.

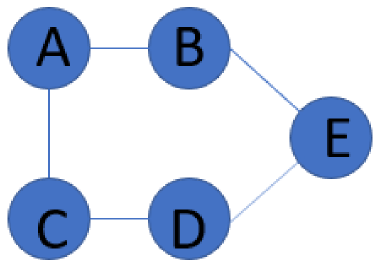



## 4.3 Community Detection

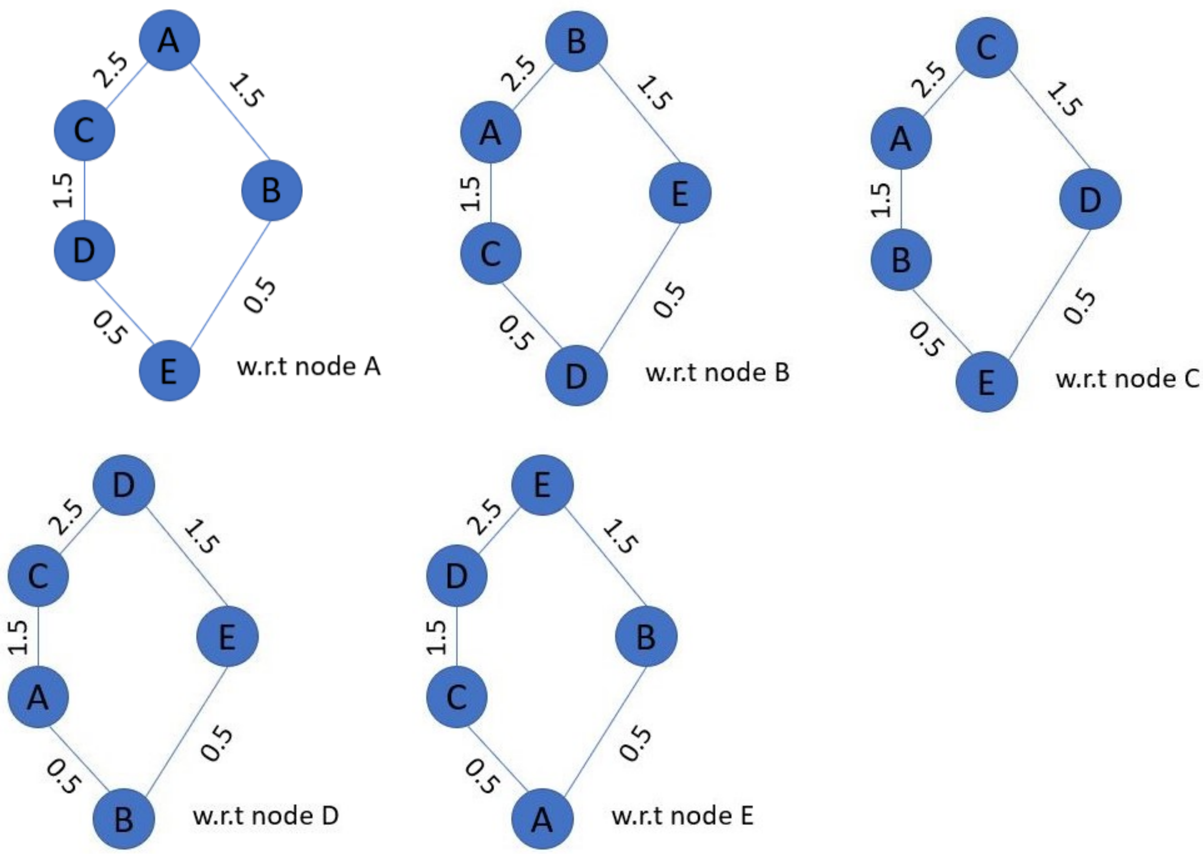
Community detection algorithms aim to identify groups or communities within a social network based on the patterns of connectivity. Here is one popular algorithm.

**Girvan-Newman Algorithm**: This algorithm works by iteratively removing edges with high betweenness centrality, causing the network to split into smaller communities. It measures the edge betweenness centrality using concepts from graph theory.

The Following is an example explaining how Girvan-Newman Algoritm works



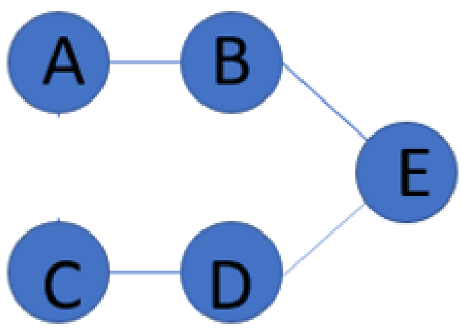

**Step 1** : In the first step we will find Edge betweenness of each edge with respect to every node.

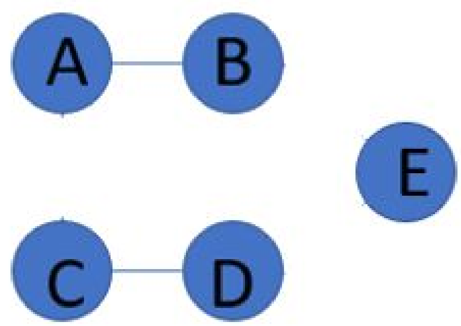


**Step 2** : Now total Edge betweenness of each edge will be the sum of Edge betweenness w.r.t every node.

| | A | B | C | D | E | Total |
|---|---|---|---|---|---|---|
| AB | 1.5 | 2.5 | 1.5 | 0.5 | 0.5 | 6.5 |
| AC | 2.5 | 1.5 | 2.5 | 1.5 | 0.5 | 8.5 |
| CD | 1.5 | 0.5 | 1.5 | 2.5 | 1.5 | 7.5 |
| DE | 0.5 | 0.5 | 0.5 | 1.5 | 2.5 | 5.5 |
| BE | 0.5 | 1.5 | 0.5 | 0.5 | 1.5 | 4.5 |

**Step 3** : Remove the edge having highest Edge betweenness. Repeat the same process until we find a proper communities.




At last we got 3 communities i.e. **(A,B), (C,B), (E)**

## 5 Practical Interpretation of Application

Social network analysis has been applied to social media as a tool to understand behavior between individuals or organizations through their linkages on social media websites such as Twitter and Facebook.

SNA has applications in various fields such as sociology, marketing, healthcare, security, and politics. It can be used to study the spread of diseases, identify key players in a market, or detect terrorist networks.

## 6 CONCLUSION

SNA is a powerful tool to uncover hidden connections and understand the structure and function of social networks. It can provide valuable insights for decision-making and problem-solving in various fields. However, it requires careful data collection and analysis.

# Bibliography

## References


[1] DISCRETE MATHEMATICS AND ITS APPLICATIONS, SEVENTH EDITION by Kenneth H. Rosen

[2] Complete graph
https://en.wikipedia.org/wiki/Complete_graph

[3] Articles
https://www.sciencedirect.com/topics/social-sciences/social-network-analysis
https://www.sciencedirect.com/science/article/abs/pii/S095741741930394X