**What is Social Network Analysis (SNA)?**

* Social Network Analysis (SNA) is the process of studying relationships between entities (nodes) in a network.
* It uses graph theory to study these relationships, helping measure interactions and flows between entities.
* SNA helps in identifying key influencers, detecting communities, and understanding how information flows through a network.
* It is widely used in various domains like social media analytics, business intelligence, epidemiology, fraud detection, and cybersecurity.

**Types of social network analysis:**

1. **Ego Network Analysis** – Focuses on relationships of a specific sample of individuals, analyzing factors like size and diversity. Traditional surveys are used to gather data on interactions, helping in information propagation, community detection, and resource access.
2. **Complete Network Analysis** – Examines relationships within an entire sample, using subgroup analysis, centrality measures, and equivalence analysis. This method aids organizations in decision-making by understanding network-wide connections.

**Key Terminologies in Social Network Analysis (SNA)**

**1. Basic Network Components**

* **Node (Vertex):** Represents an individual entity (person, organization, object).
* **Edge (Link):** Represents the connection or relationship between two nodes.
* **Graph:** A collection of nodes and edges that form a network.
* **Adjacency Matrix:** A matrix representation of connections between nodes.

**2. Types of Networks**

* **Directed Network:** Edges have direction (e.g., Twitter followers, email communication).
* **Undirected Network:** Connections are mutual (e.g., Facebook friends, co-authorships).
* **Weighted Network:** Edges have a weight representing the strength of the relationship (e.g., frequency of interaction).
* **Unweighted Network:** Edges either exist or do not exist (binary connection).
* **Bipartite Network:** A network with two distinct sets of nodes where connections only occur between different sets (e.g., users and movies in a recommendation system).

**3. Network Properties**

* **Degree (Degree Centrality):** The number of edges connected to a node.
  + **In-degree:** Number of incoming connections (for directed networks).
  + **Out-degree:** Number of outgoing connections (for directed networks).
* **Density:** The ratio of actual connections to possible connections in the network.
* **Clustering Coefficient:** Measures how likely a node's neighbors are to be connected.
* **Path:** A sequence of nodes connected by edges.
* **Shortest Path (Geodesic Distance):** The shortest route between two nodes.
* **Diameter:** The longest shortest path between any two nodes in the network.

**4. Centrality Measures (Influence in Networks)**

* **Degree Centrality:** Measures how connected a node is.
* **Betweenness Centrality:** Measures how often a node acts as a bridge in the shortest paths of other nodes.
* **Closeness Centrality:** Measures how close a node is to all other nodes in the network.
* **Eigenvector Centrality:** Assigns higher importance to nodes connected to other important nodes (used in Google's PageRank algorithm).

**Working:**

**1. Data Collection**

* Collect data on individuals (nodes) and their relationships (edges).
* Sources: Social media platforms, emails, co-authorships, citations, etc.
* Data types: Directed (one-way relationships like Twitter followers) or undirected (mutual relationships like Facebook friends).

**2. Network Representation**

* **Nodes (Vertices):** Represent entities (people, organizations, or objects).
* **Edges (Links):** Represent relationships (friendships, collaborations, interactions).
* **Edge Weight:** Represents the strength of relationships (e.g., number of messages exchanged).

**3. Graph Construction**

* **Adjacency Matrix:** A matrix representation of connections between nodes.
* **Edge List:** A list of pairs representing relationships.

**4. Network Metrics & Analysis**

* **Degree Centrality:** Number of direct connections a node has.
* **Betweenness Centrality:** How often a node acts as a bridge between other nodes.
* **Closeness Centrality:** How close a node is to all other nodes in the network.
* **Eigenvector Centrality:** Importance of a node based on its connections to other important nodes.
* **Community Detection:** Identifying groups within the network (e.g., clustering).

**5. Visualization**

* Networks are visualized using tools like Gephi, NetworkX (Python), Cytoscape, and Pajek.
* Graph layouts help in understanding the structure and influence within the network.

**✅ Advantages:**

1. **Finds Important People** – Helps find the most popular or influential person in a group.
2. **Improves Teamwork** – Helps companies see how employees work together.
3. **Detects Groups** – Finds smaller friend circles inside a big network.
4. **Helps in Marketing** – Businesses use it to find the best people to promote their products.
5. **Tracks Information Spread** – Shows how news, trends, or messages spread online.
6. **Catches Fraud & Crime** – Banks and police use it to find fraud and criminal networks.

**❌ Disadvantages:**

1. **Needs a Lot of Data** – Collecting information can be hard.
2. **Privacy Issues** – Analyzing personal connections can invade privacy.
3. **Takes Time & Power** – Big networks (like Facebook) need strong computers to analyze.
4. **Can Give Wrong Results** – Having many friends doesn’t always mean someone is important.
5. **Changes Over Time** – Networks keep changing, so analysis must be updated often.

**Applications:**

* **Social Media Analysis:** Identifying influencers, trends, and community structures.
* **Fraud Detection:** Detecting anomalies in financial transactions.
* **Epidemiology:** Tracking disease spread in social interactions.
* **Business & Marketing:** Customer segmentation and targeted advertising.

**Real-Time Example of Social Network Analysis (SNA):**

**Example: WhatsApp Friend Group Analysis**

Imagine you are in a **WhatsApp group** with 10 friends. Some of you chat frequently, while others are less active. You want to analyze **who is the most influential person in the group** based on interactions.

**How Social Network Analysis Works Here:**

1. **Nodes (People in the Group)**
   * Each member in the WhatsApp group is a **node** (e.g., You, Alex, Priya, Rahul, etc.).
2. **Edges (Connections Between Friends)**
   * If two people chat frequently, an **edge (link)** is created between them.
   * If Alex messages Priya often, there is a **strong connection** between them.
3. **Weight (Strength of Connections)**
   * If Rahul and Priya chat daily, their connection has **higher weight**.
   * If Alex and Priya chat once a month, their connection has **lower weight**.
4. **Centrality (Finding the Most Active Member)**
   * **Degree Centrality:** The person who chats with the most number of people is **the most connected**.
   * **Betweenness Centrality:** If one person (e.g., Rahul) is often **the middle link between friends**, they act as a **bridge**.
   * **Eigenvector Centrality:** The person connected to **other highly active members** is the most influential (like WhatsApp’s "most messaged" person).
5. **Community Detection (Finding Small Friend Groups)**
   * If 3-4 friends talk more among themselves, they form a **subgroup or clique** inside the WhatsApp group.
6. **Real-World Use Case**
   * Businesses use this technique on **social media** to find **influencers** who can spread messages widely.
   * WhatsApp or Facebook can analyze **who is most active** in a group and recommend people to connect.

**Final Takeaway**

Social Network Analysis helps in understanding **who is the most important, most connected, or most influential person** in a network, whether it's a **WhatsApp group, Facebook, Instagram, or even a business network**.