

Social Network Analysis: Introduction to the Course

Even Semester of Academic Year 2022-23

Instructors: 1) Sankita Patel 2) Suhani Chauhan

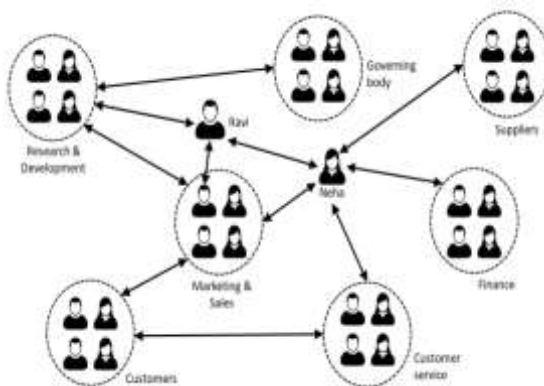
Slide Credits: Teaching Material on Social Network Analysis by Tanmoy
Chakraborty, Wiley, 2021

Motivation, Curriculum, Teaching Scheme and Course Outcome

Books to refer

- Social Network Analysis by Tanmoy Chakraborty, Wiley (Slide credits to this book)
- Network Science by Barabasi, Cambridge University Press
- Slide Credits: Teaching Material on Social Network Analysis by Tanmoy Chakraborty, Wiley, 2021

What is Social Network Analysis?



Network:

An abstract representation of relations among entities

Social Network:

A simplified representation of the social structure characterized by actors and ties

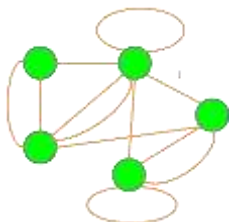
Social Network Analysis:

The application of networks and graph theory to analyze the relations present in a society

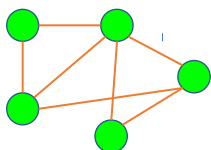
Network: Definition

- ❑ A network, also referred to as a *graph*, is defined as an ordered pair $G(V, E)$, where V is a set of nodes (also referred to as vertices or entities), and E is a set of edges (also referred to as links or relations) joining the nodes.
- ❑ Depending on the nature of application, the above definition may be revised or augmented, as follows:
 - ❖ the nature of edges may vary – *undirected* (also called symmetric, or reversible) edges, *directed* (also called asymmetric, or irreversible) edges, or *hyperedges*, etc.
 - ❖ both the nodes and/or the links are associated with one or more attributes/features like *weights*, *timestamps*, *textual features*, etc.
- ❑ An edge in a graph may have same node as end nodes. Such edges of a graph are called *self loops* (or, simply, loops).
- ❑ A graph may have more than one edge joining a pair of nodes. Such edges are called *parallel edges*.

Network: Definition (contd...)

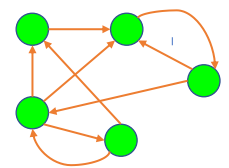


A graph with loops and parallel edges

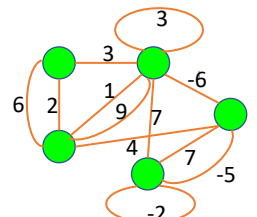


A simple graph

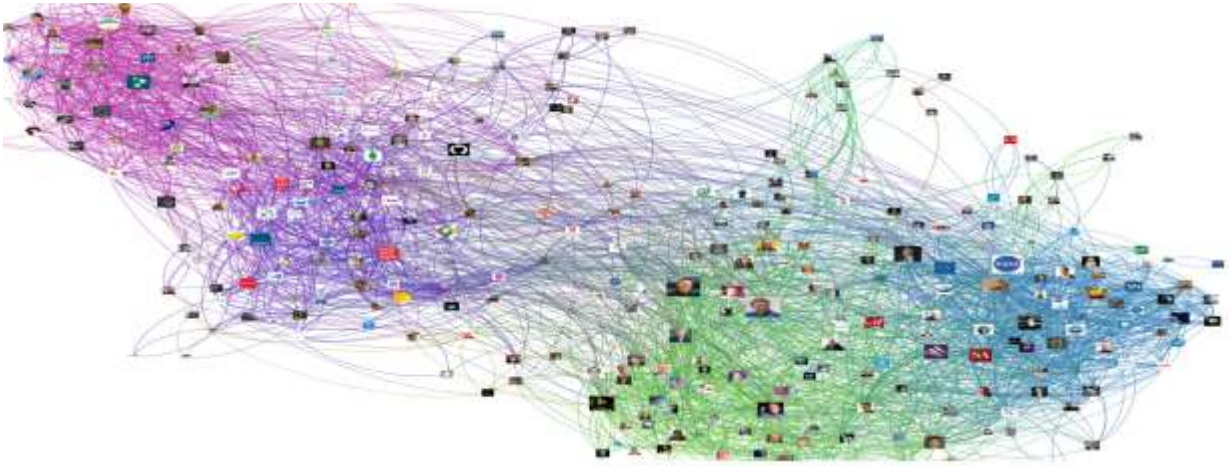
- ❑ A graph having neither self loops nor parallel edges are called a *simple graph*.
- ❑ A graph having directed edges (i.e. links having a direction) is called a *directed graph*. Directed edges are also referred to as *arcs*.
- ❑ A graph having no direction in its edges is called an *undirected graph*.
- ❑ A graph having weights associated with its edges are called a *weighted graph*. A weighted graph can be directed as well as undirected.
- ❑ A graph having its nodes and/or edges attributed with feature values is called an *attributed graph*.



A directed graph



A weighted graph



A sample of Twitter follower-followee network

(image source: <http://allthingsgraphed.com/2014/11/02/twitter-friends-network/>)

Social Network Analysis: Key Features

- **Required Knowledge Domains**

☐ Sociology ☐ Psychology ☐ Mathematics ☐ Statistics ☐ Computer Science

- **Study Benefits**

- ☐ To know the way social interactions influence a network
- ☐ To learn how the information flows inside a network
- ☐ To characterize roles of the individuals in a network
- ☐ To characterize communities inside a network
- ☐ To characterize the evolution of a network

Key Application Areas

Healthcare

- ❖ Combating Epidemics
- ❖ Mass Vaccination

Social Media & E-Commerce

- ❖ Friend & Follow Recommendation
- ❖ Know Your Customers
- ❖ Recommendation & viral marketing

Web & Cyberspace

- ❖ Search engine optimization
- ❖ Malware detection
- ❖ Spam detection

Police & Military

- ❖ Fighting cyber crimes
- ❖ Fighting terrorism
- ❖ Network-centric warfare

Scientific Research & Academic Collaboration

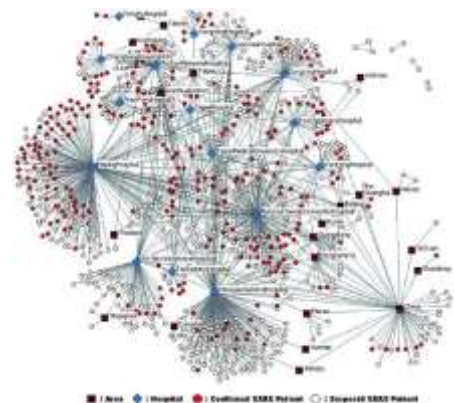
- ❖ Ranking scientific publications
- ❖ Ranking scientific authors
- ❖ Ranking publication venues

Miscellaneous

- ❖ Computer-supported collaborative learning
- ❖ Complex project management

SNA Applications: Healthcare

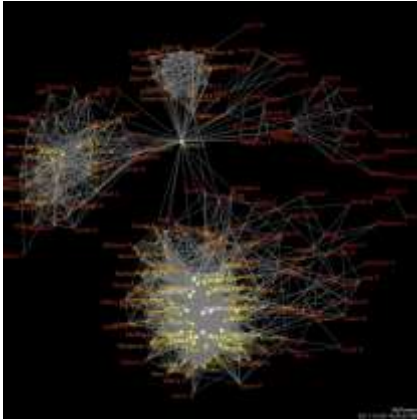
- ❑ Modeling the spread of infectious diseases
- ❑ Contact tracing during epidemic outbreak to identify possible patients
- ❑ Identify and isolate super-spreaders
- ❑ Planning lockdown schedule
- ❑ Identify vulnerable population during vaccination
- ❑ Planning vaccination schedule, etc.



2003 SARS contact Network in Taiwan

https://doi.org/10.1007/978-1-4419-6892-0_15

SNA Applications: Social Media

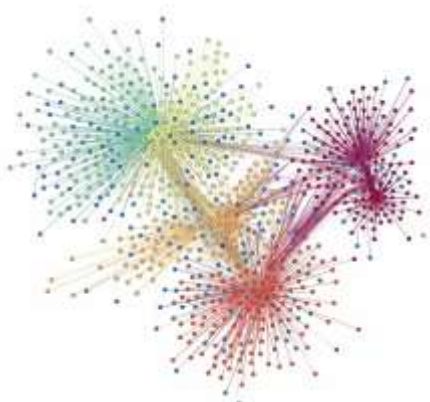


- ☐ Identifying friendship structures in online social media
- ☐ Recommending friends, and groups, or pages to follow
- ☐ Identifying information propagation patterns in social networks

An Example Map of a Facebook Friendship Network

<https://mathconceptions.wordpress.com/2012/01/16/application-snippet-friendship-and-influence-in-social-networks/>

SNA Applications: E-Commerce



- ☐ Customer profiling to know the customers
- ☐ Product/Service recommendation based on customer profile
- ☐ Instigating viral marketing by pinpointing influential players
- ☐ "People like you buy", "Frequently bought with this", or "Frequently browsed", "Trending" are common buzzwords



Influential Communities in Social Network

<https://towardsdatascience.com/influential-communities-in-social-network-simplified-fe5050dbe5a4>

SNA Applications: Network Centric Warfare

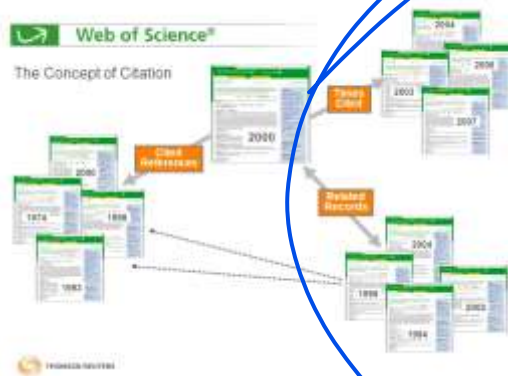


Network Centric Warfare

<http://www.indiandefencereview.com/news/network-centric-warfare/>

- ❑ Rising popularity of social network analysis influence the military doctrines
- ❑ Various military think-tanks proposed for network-centric warfare principle using social network analysis methods
- ❑ It is claimed that Saddam Hussein has been captured from his hideout exploiting network analysis techniques
- ❑ It is also claimed US Navy Seal Team Six assassinated Osama Bin Laden by tracking his secret hiding location in a similar manner

SNA Applications: Scientific Research & Academic Collaboration



<https://library.bu.edu/citedreferences>

- ❑ Scientific authors cite (refer) the works of other authors in their publications to authenticate their claims
- ❑ Finding the dynamics of these citations attracted social scientists
- ❑ Various networks of scholarly articles may be formed exploiting this relationship
 - ❑ Paper-paper citation network
 - ❑ Paper-paper co-citation network, etc.
- ❑ Various popular metrics are outcome of analysing these networks.
 - ❑ Publication related: H-index, i-10 index, g-index, etc.
 - ❑ Venue related: impact factor, CORE rank, etc.

SNA Applications: Scientific Research & Academic Collaboration



<https://scholarlykitchen.sspnet.org/2017/04/07/updated-figures-scale-nature-researchers-use-scholarly-collaboration-networks/>

- ❑ Scientific authors collaborate with one another to improve research quality
- ❑ Various scholar networks may be formed using these relationships
 - ❑ Author collaboration network
 - ❑ Author citation network
 - ❑ Author co-citation network, etc.
- ❑ Information retrieved from these networks may be used to measure authors' research quality

SNA Applications: Computer-supported Collaborative Learning

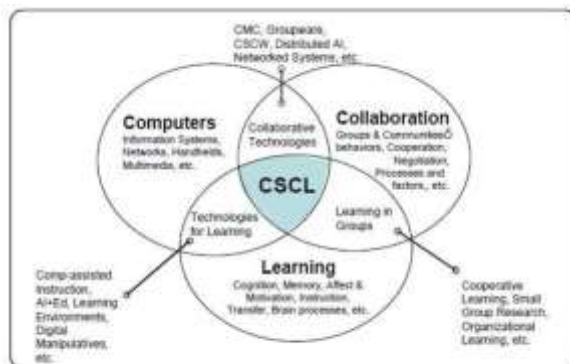
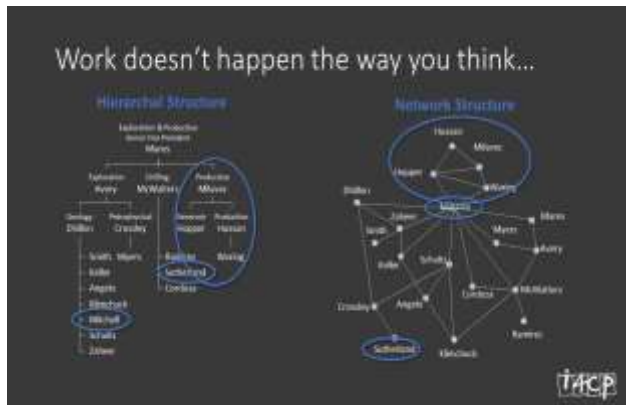


Figure 1: Multidisciplinary of CSCL

<https://www.semanticscholar.org/paper/The-Characteristics-of-the-Computer-Supported-a-on-Hashim-Ismael/42176e6bf76dd15a2c9874e6fa8696e153a3f554>

- ❑ Pedagogical process of observation where students learn progressively through active group interaction using ICT
- ❑ SNA techniques used to extract relationship between various actors (human and non-human) of CSCL
- ❑ Study insights are used to improve the students' learning outcome and user experience

SNA Applications: Organizational Network Analysis

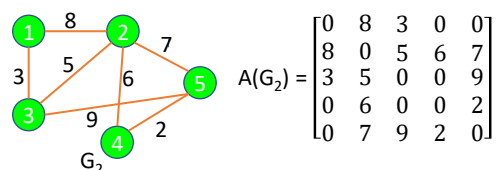
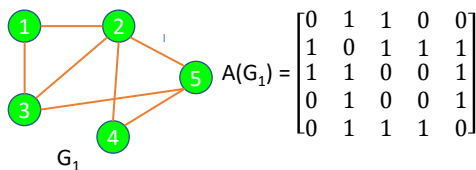


<https://www.i4cp.com/productivity-blog/what-organizational-network-analysis-is-and-how-it-benefits-companies>

- ❑ Classic (linear) organization charts do not effectively describe the real social network of an organization
- ❑ The informal networks, historically considered an annoying problem and an indicator of undisciplined attitude of workers, carry **huge potential** if applied suitably
- ❑ ONA provides information on how to improve performance in the organization
- ❑ ONA represents the complete set of real relationships between the players
 - ❑ who is in touch with whom
 - ❑ specific features of each player
 - ❑ Type and intensity of relationship, etc.

Network Representation: Adjacency Matrix

- ❑ An adjacency matrix $A = (a_{ij})$ for a graph $G(V, E)$ is a square matrix of dimension $|V| \times |V|$ such that each element a_{ij} of A indicates the existence of an edge between the node v_i and node v_j (also the weight of the corresponding edge in case of a weighted graph) in G .



Network Representation: Adjacency Matrix (Cont...)

advantages

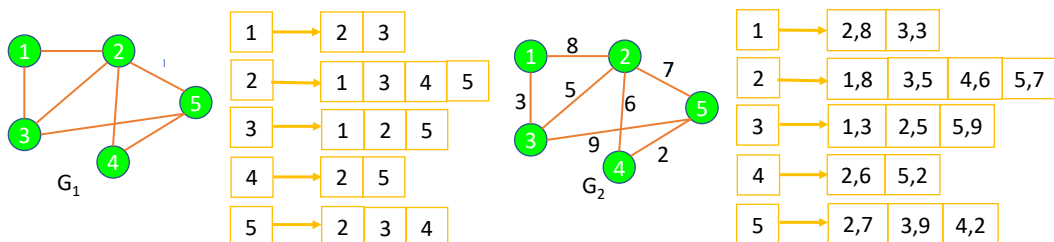
- ❑ Easy to implement and follow
- ❑ Addition, and removal of an edge require $O(1)$ time
- ❑ Query to an edge require $O(1)$ time

disadvantages

- ❑ Consumes $O(|V|^2)$ storage space, even if the graph is sparse
- ❑ Addition or removal of a node require $O(|V|^2)$ time

Network Representation: Adjacency List

- ❑ An adjacency list for a graph $G(V, E)$ is a collection of unordered lists such that each node correspond to a list from the collection that indicates the set of neighbours of the node. Every entry in an adjacency list A_i for node v_i in the graph is a node adjacent to node v_i



Network Representation: Adjacency List (Cont...)

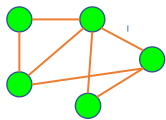
advantages

- ☐ Require $O(|V| + |E|)$ space
- ☐ Insertion of vertex and edge require $O(1)$ time
- ☐ Removal of vertex require $O(|V| + |E|)$ time

disadvantages

- ☐ Removal of edge require $O(|E|)$ time
- ☐ Query to an edge require $O(|V|)$ time

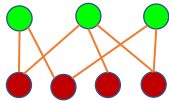
Network Types: Link-centric View Unipartite Network



- ☐ Consists of a vertex set V and an edge set E . There is no restriction on the formation of edges between nodes of the network
- ☐ Example: An organizational LAN, where nodes are the devices, and edges are the local area links.
- ☐ Used to model the situation when links can join any pair of nodes of the network

Network Types: Link-centric View

Bipartite Network



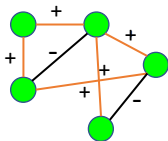
❑ Consists of a vertex set V that is divided into two sets V_1 and V_2 that are disjoint and independent. Each edge of the network connects a vertex in V_1 to another vertex in V_2

❑ Example: An e-commerce user-product network. One part consists of the users, the other part consists of the products, the links are based on the basis of who bought what.

❑ Generalization of Bipartite network is n-partite networks, where the vertex set is partitioned into n number of part, an edges join a node from one part with a node from another part.

Network Types: Link-centric View

Signed Networks

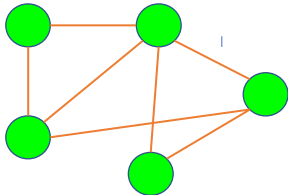


❑ Consists of a set of nodes V , a set of edges E , and a function $f: E \rightarrow \{+, -\}$ that assigns each edge a positive or a negative sign

❑ Example: Consider a social media website that allows users to tag other users as friends or foes. The positive edges are friendship links and negative links are between foes.

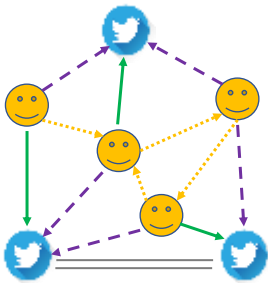
❑ studied specifically in the context of balance and status theory which determines the stability or existence of certain types of structural patterns in a network.

Network Types: Node and Link-centric View: Homogeneous Network



- Consists of a set of nodes V , all of which are of same type, and a set of edges E , all of which are of same type
- Example: Follower-Followee network of any kind.
 - ❖ Nodes are the users of the platform (all nodes are of same type),
 - ❖ Links are the follower-followee link between these nodes (all links are of same type)

Network Types: Node and Link-centric View: Heterogeneous Network



- Consists of a set of nodes V and a set of edges E , and two associated mapping functions, f_v and f_e , for nodes and edges, respectively. f_v maps a node to a node type and f_e maps an edge to an edge type
- Example: We consider a specific instance of Twitter network like the figure.
 - Two types of nodes:
 - Twitter Users
 - Tweet Posts
 - Four types of edges representing four types of relations between these nodes:
 - .Posted-by (User – Post: Directed links)
 - Followed-by (User – User: Directed links)
 - Similar (Post – Post: Undirected links)
 - Retweet (User – Post: Directed links)

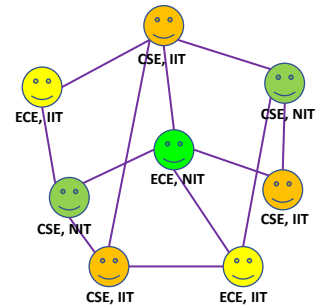
Network Types: Node and Link-centric View: Attributed Network

- Consists of a set of nodes V and a set of edges E , and two associated mapping functions, f_v and f_e , for nodes and edges, respectively. f_v maps a node to a node attribute vector and f_e maps an edge to an edge attribute vector

- Example: We consider a specific instance of Facebook network like the figure

- Nodes are some Facebook users
- Edges are given by Facebook friendship relationship between these users
- Node attributes are the users' academic affiliations
- There is no edge attribute in this network

- The example is a node-attributed network



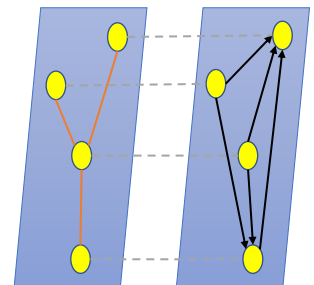
Network Types: Node and Link-centric View: Multidimensional Network

- A special type of multilayer network where each layer represents a particular type of relationship among nodes

- Example 1: A special instance of Twitter network:

- Nodes are Twitter users in both layer
- Layer 1 edges: user – user similarity links (based on mutual interests) – Undirected links
- Layer 2 edges: user – user follower-followee links – Directed links

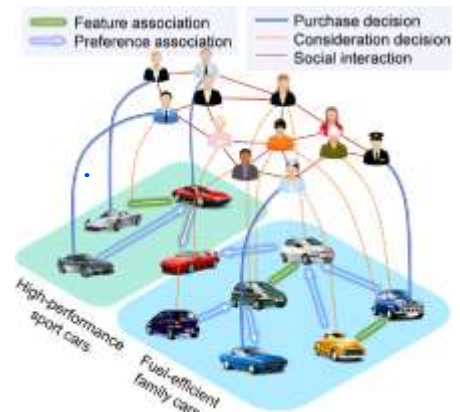
- Note: In this example, each layer is node homogeneous.



Network Types: Node and Link-centric View: Multidimensional Network (Cont...)

❑ Example 2: Customer – product relationships as a multidimensional complex network system

- ❑ Layer 1: One type of nodes, one type of edges:
 - ❑ Nodes: Customers
 - ❑ Edges: Customer Social Interaction – Undirected links
- ❑ Layer 2: Two types of nodes, two types of edges
 - ❑ Nodes:
 - ❑ High performance Cars,
 - ❑ Fuel efficient cars
 - ❑ Edges:
 - ❑ Car feature association – Undirected links
 - ❑ Car preference association – Directed links
- ❑ Inter-layer edges: two types of edges
 - ❑ Customer – Car: Purchase decision – Undirected links
 - ❑ Customer – Car: Consideration decision – Undirected links

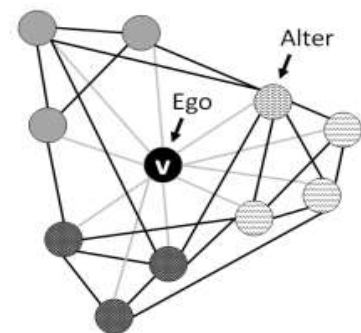


<https://www.semanticscholar.org/paper/Modeling-customer-preferences-using-network-in-Wang-Chen/539c7f0632041903521b8cbc42eabd27b8844673>

Network Types: Local View: Ego-centric Network

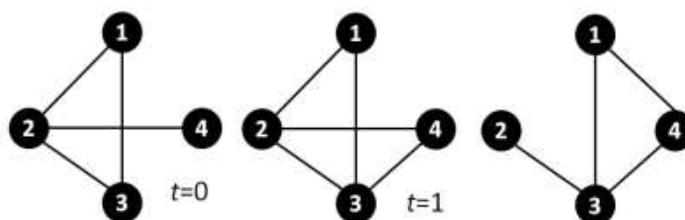
❑ A network of the form $G(V, E, u)$, that corresponds to a node $u \in V$, usually known as the 'ego', and consists of the node u as the central node, the nodes that are connected directly to the node u , usually known as the 'alters', and the induced subgraph for the same.

- ❑ Example: A subgraph of a Facebook Friendship Network:
 - ❑ Ego node corresponds to a user,
 - ❑ Alter nodes are his Facebook friends from different capacities and affiliations.



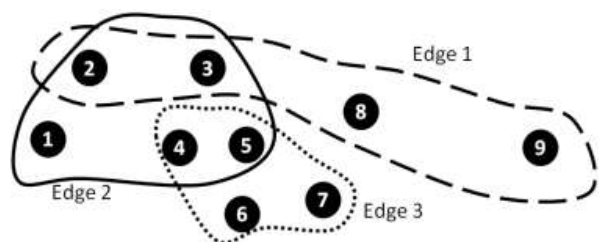
Network Types: Temporal View: Time-varying Network

- Consists of a set of nodes V and a set of edges E where each edge $e_{ij} \in E$ is represented by a three- tuple $e_{ij} = \{v_i, v_j, t_{ij}\}$. Here, v_i and v_j are two end-points, and t_{ij} indicates the persistence duration of the edge e_{ij}
- Example: Person-to-person communication network over a span of time. The visible components are snapshots of the network at different time instances.



Network Types: Generalized View: Hypergraph

- Defined by a set of nodes V and a set of edge or hyperedges E , where each hyperedge e connects multiple nodes of the hypergraph
- Example: A special representation of Coauthorship Network:
 - Nodes are authors
 - Papers are hyperedges connecting the coauthors of the paper



Popular Real-world Networks

❑ Social Network

- ❖ Telephone call network
- ❖ Email message network
- ❖ Film actor collaboration network
- ❖ Academic co-authorship network

❑ Biological Network

- ❖ Protein-protein interaction networks
- ❖ Genetic regulatory networks
- ❖ Neural networks
- ❖ Metabolic networks
- ❖ Food Web
- ❖ Cell signalling networks

❑ Information Network

- ❖ World Wide Web (WWW)
- ❖ Citation network

❑ Technological Network

- ❖ Electric power grids
- ❖ Networks of airline routes
- ❖ Network of Railway Routes
- ❖ Electronic circuits
- ❖ Delivery networks of post-office/Courier
- ❖ The Internet

❑ Language Network

- ❖ Network formed by using the persons speaking a particular language

Levels of Social Network Analysis: Microscopic Level

❑ We begin by analyzing how a pair of nodes interacts and gradually trace the interactions at the group level or subgraph level.

❑ **Dyadic level** → Interaction patterns among two nodes

- ❑ Examined properties: homophily, reciprocity, social equality, mutuality, etc.
- ❑ Derived global statistics: assortativity, mixing coefficient, etc.

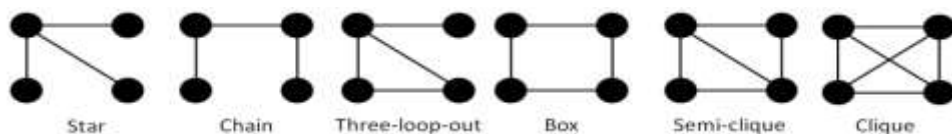
❑ **Triadic level** → Interaction patterns among three nodes

- ❑ Examined properties: triadic closure
- ❑ Derived network properties: Clustering Coefficient, local bridges, etc.

❑ **Ego-centric circles** → Interaction pattern between ego node with its alters

Levels of Social Network Analysis: Mesoscopic Level

- ❑ Mesoscopic analysis is an intermediary between microscopic and macroscopic analyses, which mostly deals with a subset of the entire population.
- ❑ **Communities** → Formed due to frequent interactions among homogeneous nodes in a network
 - ❖ Within a community, the nodes exhibit a particular kind of dynamicity
 - ❖ Across communities, the dynamic behaviour differs
- ❑ **Network Motifs** → Subgraphs that repeat themselves frequently within or across a network
 - ❑ Highly effective in capturing functional properties in a network



Undirected motifs with size 4 and their names

Levels of Social Network Analysis: Macroscopic Level

- ❑ At macroscopic level, we deal with the entire network as a whole and try to understand the micro-level dynamics by exploring the overall graph property.
- ❑ Features of Interest :
 - ❑ Connectedness,
 - ❑ Diameter or Average path length,
 - ❑ Degree Distribution,
 - ❑ Edge Density, etc.
- ❑ Example:
 - ❑ We find that the diameter of a network is too small ⇒ network may look like a star, or a clique
 - ❑ We further find that overall edge density is too high ⇒ network looks like a clique

Graph Visualization Tools

❏ Web-based tools

- ❖ **Pollinode:** <https://www.polinode.com> (Non-open source application)
- ❖ **NodeGoat:** <https://nodegoat.net> (Non-open source application)
- ❖ **Linkage:** <https://linkage.fr> (Open source application)
- ❖ **EchoDemo:** <https://osome.iuni.iu.edu/demos/echo> (Non-open source application)
- ❖ **Palladio:** <https://hdlab.stanford.edu/palladio> (Open source application)

❏ Standalone tools

- ❖ **NDlib-Viz:** <https://ndlib.readthedocs.io/en/latest/viz/ndlib-viz.html> (Open source application)
- ❖ **CytoScape:** <https://cytoscape.org> (Open source application)
- ❖ **Gephi:** <https://gephi.org> (Open source application)
- ❖ **Vizster:** <http://vis.stanford.edu/jheer/projects/vizster> (Open source application)
- ❖ **SparklingGraph:** <https://sparkling-graph.github.io> (Open source application)