# Motivation – Data Structures

Dr. Animesh Chaturvedi

Assistant Professor: Data Science and Artificial Intelligence, IIIT Dharwad

Post Doctorate: King's College London & The Alan Turing Institute

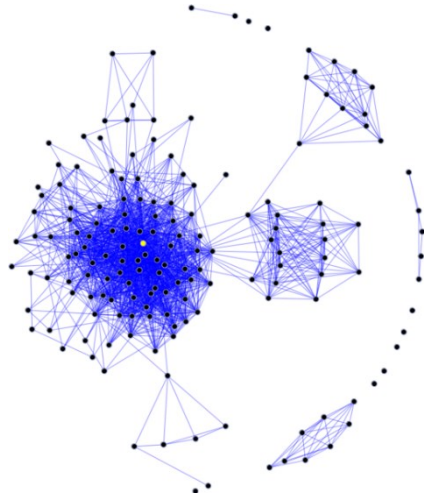
PhD: IIT Indore MTech: IIITDM Jabalpur

# Disciplines

- **Statistical physics:** methods of probability theory and statistics, and mathematical tools
  - deals with large populations and approximations,
- **Information theory:** quantification, storage, and communication of information (e.g. signals)
  - Subfield of Communication, Electronics, and Computer Science
- **Nonlinear dynamics:** the change of the output is not proportional to the change of the input.
  - Changes in variables over time, or, space etc.
  - Measures chaotic, unpredictable, or counterintuitive

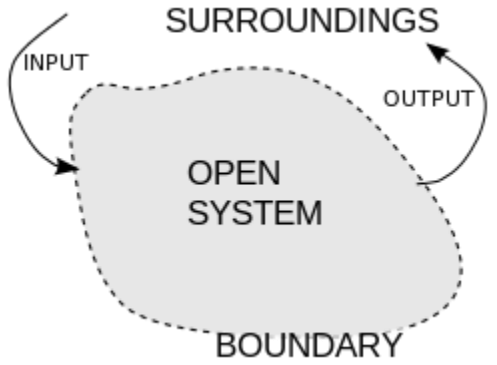
# Disciplines

- **Computer science:** programming creates complex systems, and application involving complex real-world data,
- **Sociology:** studies human behaviour, social behavior, society, patterns of social relationships, social interaction, and culture
- **Social network:** investigate social structures with networks and graph theory
- **Biology:** studies life and living organisms, physical structures of chemical, molecular interactions



# Systems

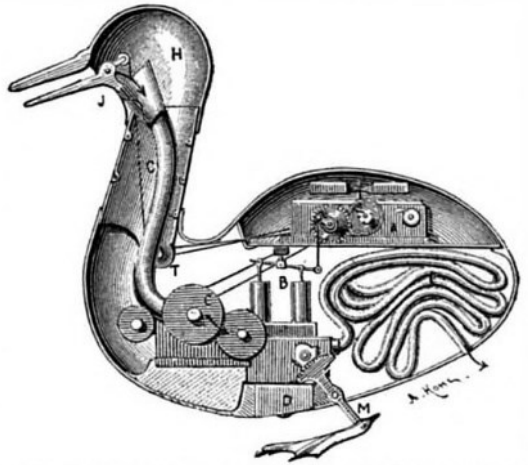
- An interdisciplinary domain
- Composed of many components (or entities) interacting with each other.
  - power grid, transportation or communication systems,
  - social and economic organizations (like cities),
  - organisms, a living cell, the human brain, and
  - an ecosystem, climate, entire universe.
- Behavior is hard to model with dependencies, relationships,
  - interactions between their components or
  - interactions between system and its environment




https://en.wikipedia.org/wiki/Complex_system

# Systems and Reductionism

- Reductionism is old domain since 16th century
- explains system in terms of parts and their interactions
  - Define a domain of possible parts
  - Generate inputs over the interaction between parts
  - Perform a deterministic computation on the input data
  - Aggregate the results
- interprets a complex system as the sum of its parts

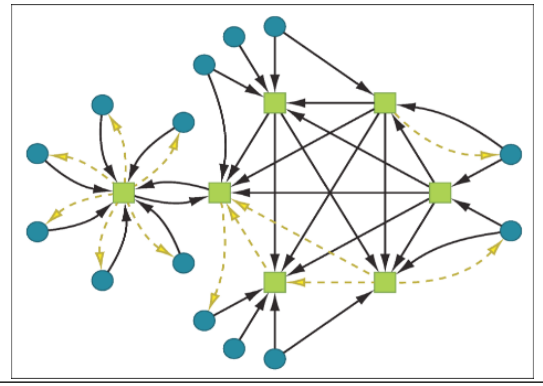



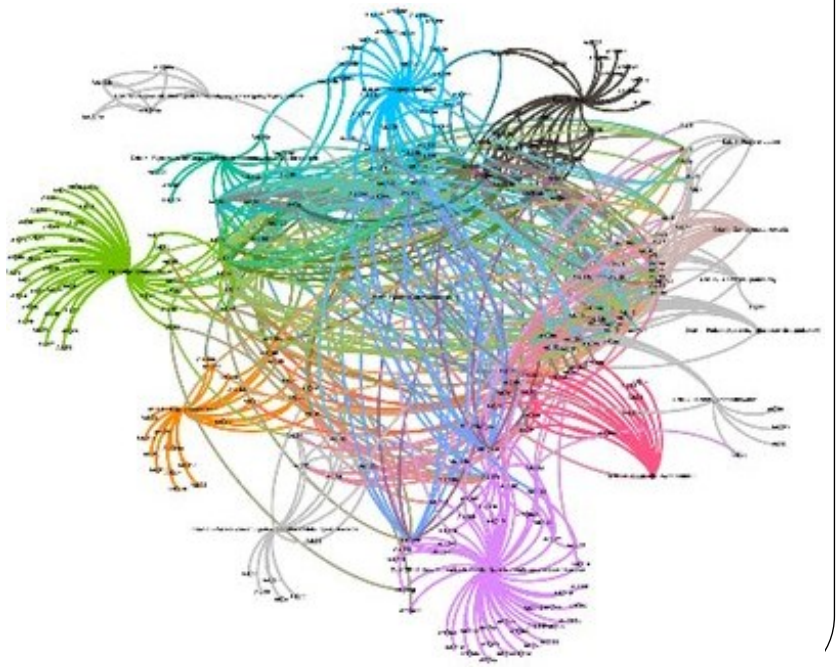
# System properties

- **Nonlinearity:** a set of simultaneous equations with the variables of a polynomial degree higher than one
- **Emergence:** entities have properties emerging only when the they interact together
- **Spontaneous order** arises from interactions between component of disordered system
- **Adaptation:** behavior of mutation and self-organization for the changes
- **Feedback loops:** outputs are routed back as inputs as part of a chain of cause-and-effect that forms a circuit or loop

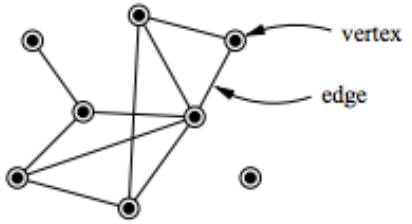
# System representation with Graph Theory

- Represented with a network (graph)
  - nodes represent the components (or entities) and
  - edges (or links) represents entities' interactions.
- Depicts collection of discrete objects and relationships between them
  - persons within an organization,
  - logic gates in a circuit,
  - genes in gene regulatory networks, or
  - between any other set of related entities





# Graph Theory

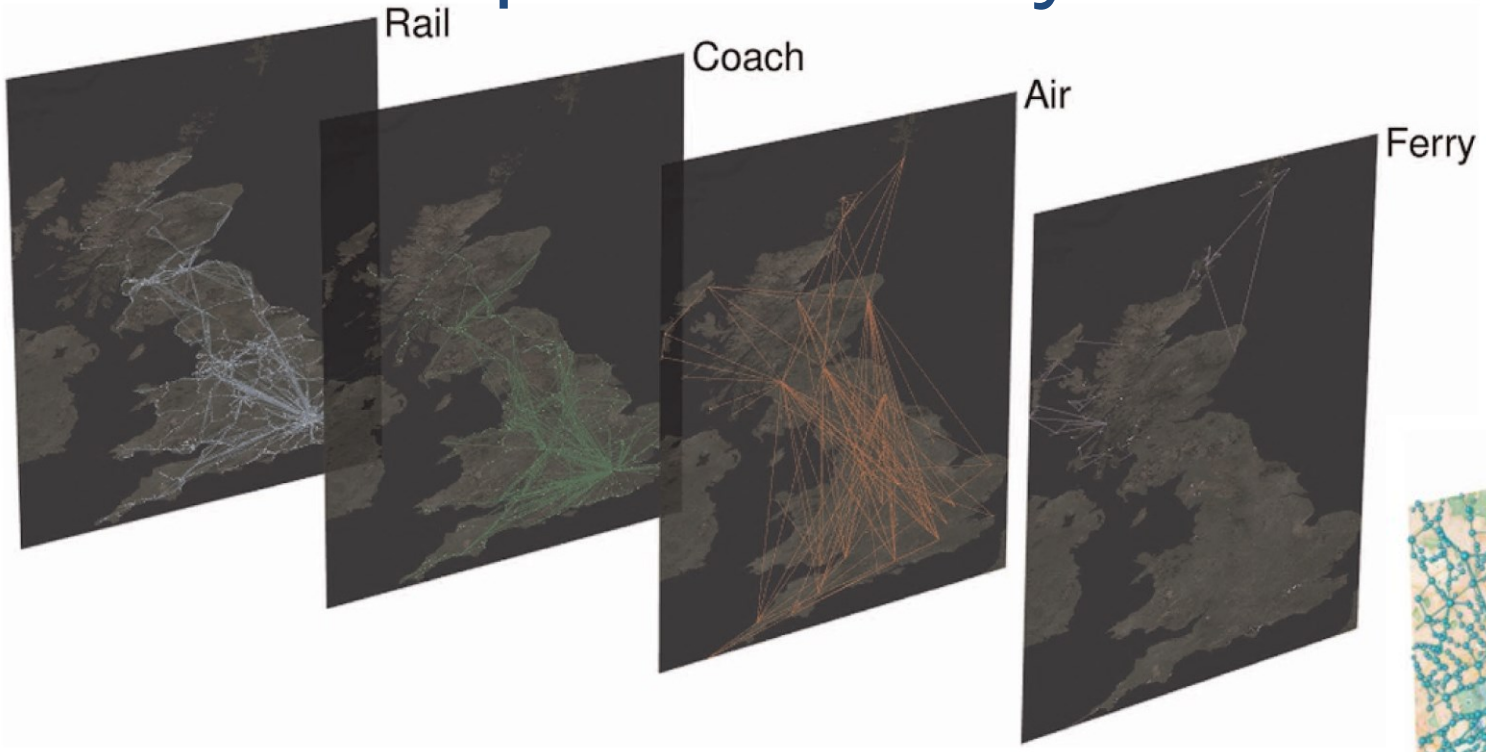



- Study of mathematical structures used to model relationship between discrete elements of a set
- Vertices (or nodes) are connected by edges (or links)
- Undirected graphs, where edges link two vertices symmetrically
- Directed graphs, where edges link two vertices asymmetrically
- Graph is an ordered triple $G = (V, E, \emptyset)$
  - V is a set of vertices (or nodes)
  - E is a set of edges (or links)
  - $\emptyset: E \rightarrow \{\{x, y\} \mid x, y \in V \text{ and } x \neq y\}$ an incidence function mapping every edge to an unordered pair of vertices

# Network Theory

- Graph representing symmetric or asymmetric relations between discrete objects
- Graph with nodes and/or edges have attributes (e.g. names)
- Network theory is a subdomain of Graph theory
- Network theory is applied graph theory

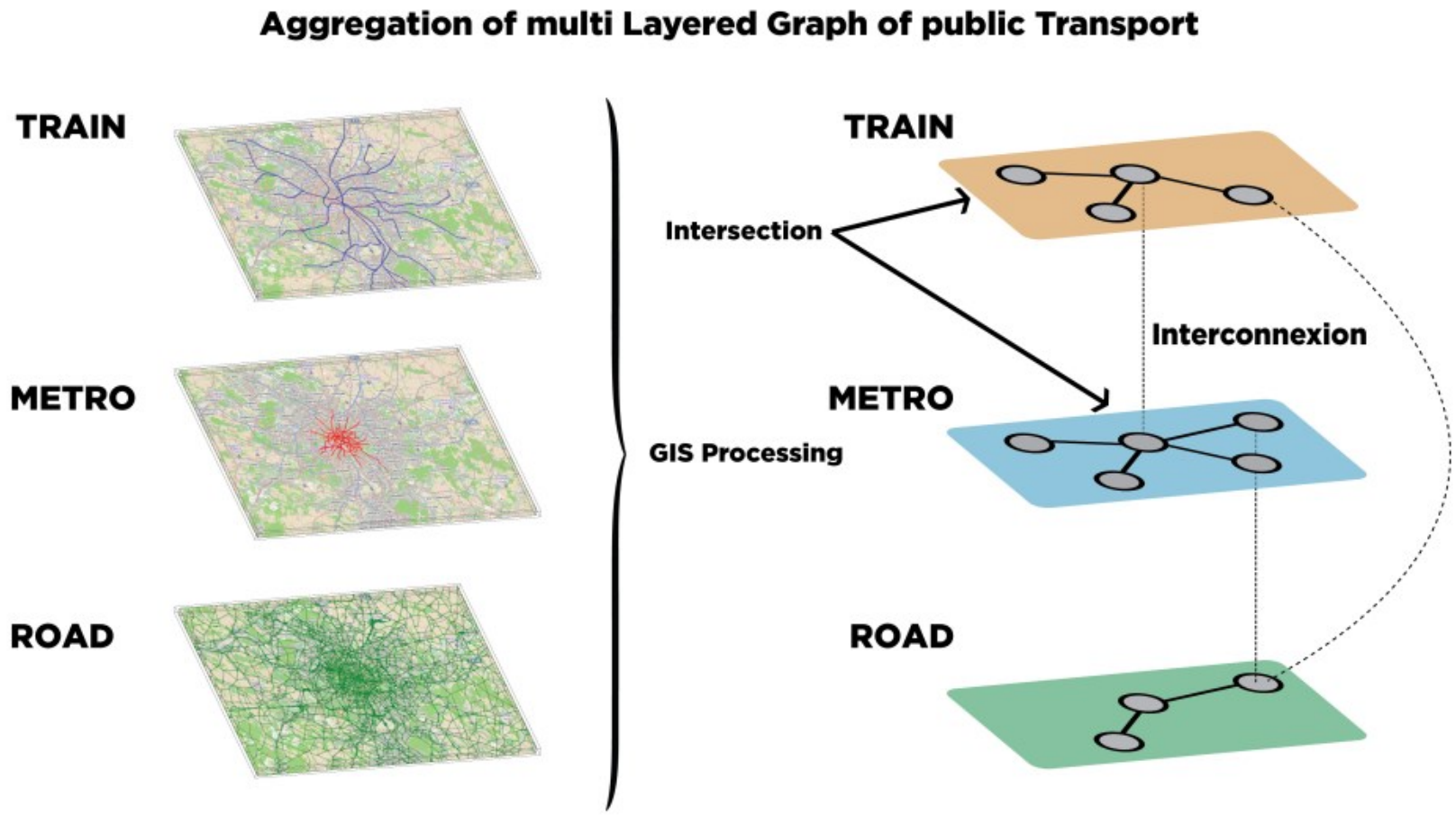
# Transport multi-layer network





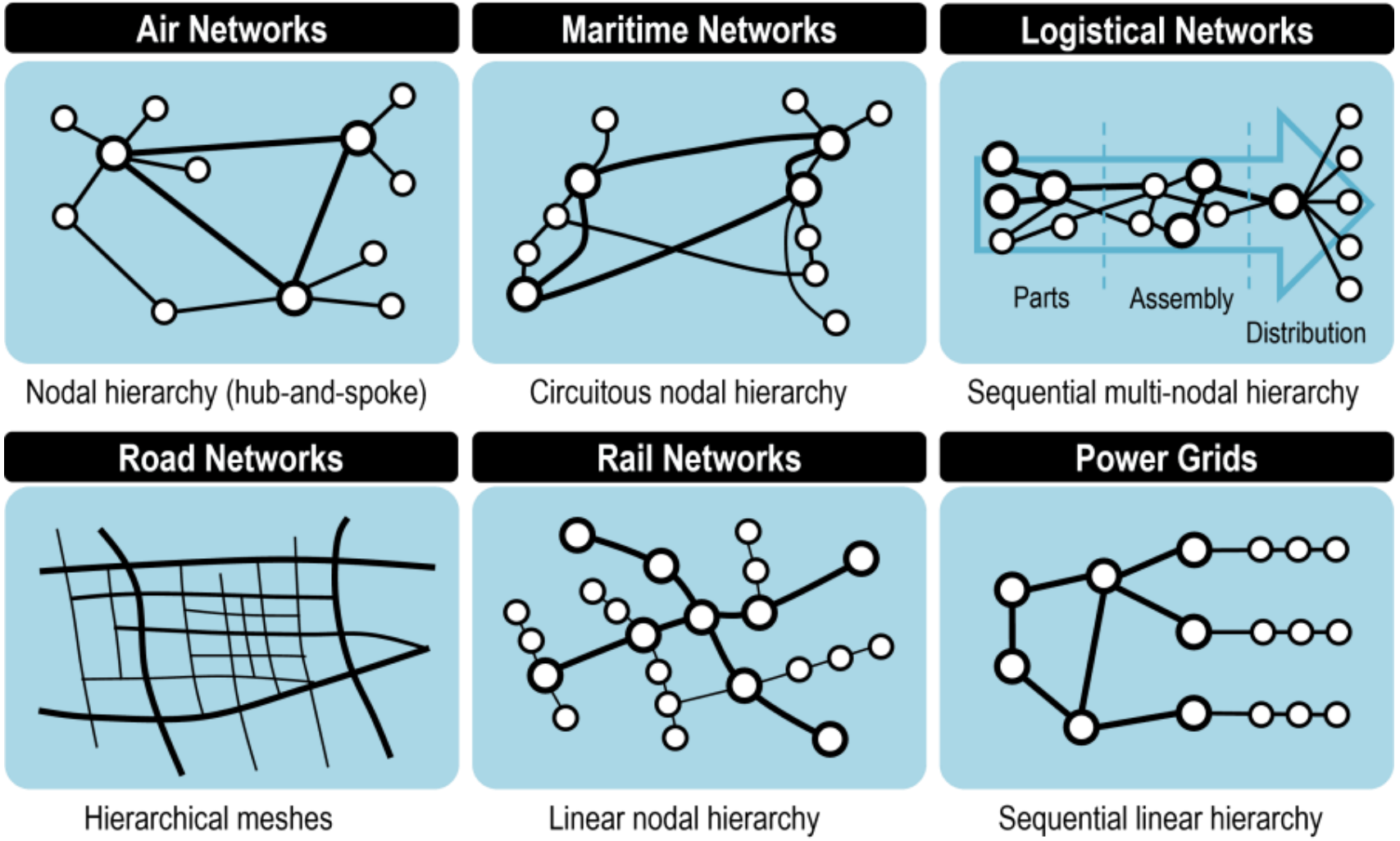
Gallotti, Riccardo, and Marc Barthelemy. "The multilayer temporal network of public transport in Great Britain." Scientific data 2.1 (2015): 1-8.

# Transport multi-layer network




Asgari, Fereshteh, et al. "CT-Mapper: Mapping sparse multimodal cellular trajectories using a multilayer transportation network." Computer Communications 95 (2016): 69-81.

# Types of Transportation Networks



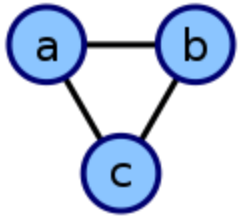

https://transportgeography.org/contents/chapter9/transportation-and-disasters/transportation-networks-vulnerabilities/

# System Representation

- Systems can be represented by Networks
- Networks can represent connections between component (or entities)
- Network can be represented as a graph with
  - vertices (or nodes) representing entities and
  - edges (or links, or connections) representing the relationship between components.

# System Representation

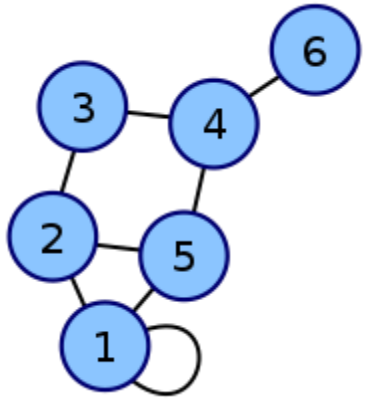
- Adjacency list is a collection of unordered lists used to represent a finite graph.
- List contains the set of neighbors of a vertex in the graph.
- Undirected graphs:
  - data structure two different linked list nodes for each edge




| Adjacency list representation | |
|---|---|
| a | b |
| a | c |
| b | a |
| b | c |
| c | a |
| c | b |

# System Representation

- Adjacency matrix is a square matrix used to represent a finite graph.
- Matrix with rows and columns are indexed by vertices
- an adjacency list is more space-efficient than an adjacency matrix (stored as a two-dimensional array)
- adjacency list is as simple
- Matrix are 2-Dimensional representation.

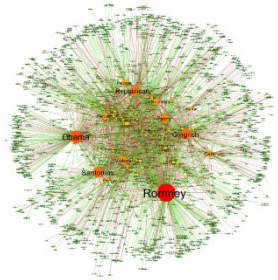



$$\begin{pmatrix} 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

# System Representation

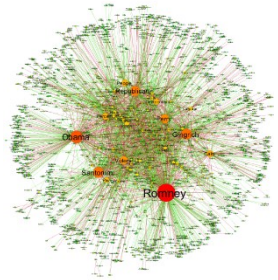
- Multilayer Network representation:
- Systems can be described by representing many interacting entities in
  - a Single network or
  - Multilayer networks
- Examples social networks, neural networks
- Co-occurring network structure made by links, activity as the nodes (e.g. community structure between multi-links).
- Dynamic entity relationships can be represented with multilayer networks, multiplex networks, and network of networks.

# Graph or Network Theories




- **Graph colouring:** e.g. coloring a graph so that
  - no two adjacent vertices have the same color, or
  - no two coincident edges are the same color
- **Matching** is a set of edges without common vertices
- **Route problems:** e.g. shortest path problem
  - minimum spanning tree (MST): a subset of the connected edges that connects all the vertices together, without any cycles, and with minimum possible total edge weight.
  - traveling salesman problem (TSP): "Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?"

# Graph or Network Theories




- **Network flow:** "Directed graph where each edge has a capacity and each edge receives a flow, where the amount of flow on an edge cannot exceed the capacity of the edge"
- **Transport problem:** study of optimal transportation and allocation of resources.
- **Trans-shipment problem:** a subgroup of transportation problems, where, transportation may or must go through intermediate nodes, possibly changing modes of transport
- **Critical path analysis:** identifying the longest dependent activities and measuring the time required to complete a project
- **PERT** to analyze and represent the tasks involved in completing a given project

# System Network analysis

- Non-trivial topological features in networks representing real systems
- e.g. theories: Network motifs are small subgraphs that are over-represented in the network
- Examples of Complex networks are
  - computer networks,
  - biological networks,
  - technological networks,
  - brain networks,
  - climate networks and
  - social networks.

https://en.wikipedia.org/wiki/Complex_network

# System Network analysis

- **Electric power systems analysis:**
  - a graph consists from representing electric power aspects (e.g., transmission line impedances)
- **Biological network analysis:**
  - analysis of molecular networks
  - visualize the nature and strength of interactions between species
  - Gene Regulatory Networks (GRN), Metabolic networks, Protein-Protein Interaction, molecular interactions
- **Operations research:** logistical networks, social networks, epistemological networks

# System Network analysis

- **Computer science:** graphs are used to represent networks of communication, data organization, computational devices, the flow of computation, etc.
  - **Link analysis:** the link structure of a World Wide Web, Internet, Computer Network
  - website can be represented by a directed graph,
  - the vertices represent web pages and directed edges represent links from one page to another
- **Social network analysis:**
  - graph with the structure of relationships between social entities.
  - entities are persons, groups, organizations, nation states, web sites, or scholarly publications


ขอบคุณ
Thai
Grazie
Italian
תודה רבה
Hebrew
धन्यवादः
Sanskrit
ধন্যবাদ
Bangla
Ευχαριστώ
Greek
Thank You
English
ಧನ್ಯವಾದಗಳು
Kannada
Спасибо
Russian
Gracias
Spanish
شكراً
Arabic
https://sites.google.com/site/animeshchaturvedi07
Obrigado
Portuguese
Merci
French
多謝
Traditional
Chinese
धन्यवाद
Hindi
Danke
German
多谢
Simplified
Chinese
நன்றி
Tamil
Tamil
ありがとうございました
Japanese
감사합니다
Korean