Network & Graphs

Graph

Purpose of graphs: graph and graph theories are used to model and analyze networks

Components of a graph: nodes and edges.

Edges connect nodes

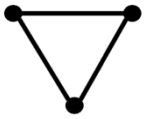
Edges can be directed & undirected

Definition of Graph: a graph G is an ordered pair of sets (V, E)

V is the set of all Nodes / Vertices

E is the set of all Edges

Ways to present a graph: Adjacency Matrix & Adjacency List



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

1 : {2, 3}
2 : {1, 3}
3 : {1, 2}

Charasteristics: degree & path & complete graph

degree: the degree of a node is the number of edges connected to it

path: a path between two nodes is a serie of edges that start with one node and end with the other one

complete graph: there is an edge between each pair of nodes

Problems of graph:

clique: vue en CS565

coloring: vue en CS630

TSP: vue en CS630

Shortest path: vue en CS630 & CS565

Vertex cover: vue en CS630

Network

Charasteristics:

Distribution of edges / nodes degrees

used for:

- anomaly detection
- ranking recommendation
- to desribe flow

Centrality of a node

used for:

- find influencer (vue en CS565, après le médian)
- find clusters/ groups
- how a node can affect connectivity and flow

Analysis:

How is a network or graph generated? By a process on its nodes and edges. Take Facebook for example, a new account is created == a new node is added; two people become friends == an edge linked these two nodes

How to present the state of network or graph? **At a given time**, the state is the **stochastic** result of these processes.

Compare graphs: metrics on graphs & metrics on nodes

Metrics on graphs: Diameter, Clustering Coefficient, Density

- Diameter: $Diam(G) = max_{ij}\ d_{ij}$, where d_{ij} is the shortest path between node i and node j
- Clustering Coefficient: amount of triangles / amount of triplets,
  triplet = node + edge + node + edge + node (3 nodes and 2 edges => not closed)
  triangle = a closed triplet
- Density: 2 * |edges| / |nodes|(|nodes| - 1)

Metrics on nodes: Degree Centrality, Closeness Centrality

- Degree Centrality: degree of the node (the more central a node is, the more connection with others it has) => $C_{deg}(v) = deg(v)$
- Closeness Centrality: the reciprocal of the sum of the length of the shortest paths between the node and **all other nodes** in the graph (the more central a node is, the closer it is to other nodes <=> the smaller the sum of the distance between this node and other nodes is) => $C_{close}(x) = 1/ (sum_y\ d(x, y))$
- Harmonic Centrality: a variant of Closeness Centrality (tout simplement, on sort le sum en dehors) => $C_h(x) = sum_y\ (1/d(x, y))$

Ranking aggregation: vue en CS565