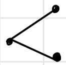
4/6

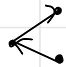
Networks and Graphs

Graphs:

use graph theory to analyse Networks



undirected edge



directed edge

graph can be stored as adjacency matrix or adjacency list

Characteristics: degree of a Node = # of edges connected to that node

path = sequence of edges that joins 2 Nodes

complete = an edge between every pair of Nodes

Network Characteristic

distribution of edges / nodes degrees:

- Anomaly detection
- Ranking / Recommendation
- network flow

Centrality of a node:

- discover clusters

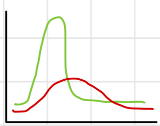
Network Analysis

state of a Network / Graph is Stochastic

Random Graph Model: distribution of the degree of Nodes

$$P(deg(v) = k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Power Law: Random Graph



Power Law

Metrics on Graph

1) diameter : $diam(G) = max_{ij} d_{ij}$

2) Clustering coefficient $= C = \dfrac{\#\text{ of triangle}}{\#\text{ of triplets}}$

3) density : N = nodes, M = edges. Density $= \dfrac{2M}{N(N-1)}$

Metrics on Nodes

1) Degree Centrality : $C_{deg}(V) = Deg(V)$

2) Closeness Centrality : $C_{close}(V) = \dfrac{1}{\sum_u d(u,v)}$

3) Harmonic Centrality : $C_h(v) = \sum_{u \neq v} \dfrac{1}{d(u,v)}$

4) Betweeness Centrality : $C_h(v) = \sum_{s \neq v \neq t \in V} \dfrac{\sigma_{st}(v)}{\sigma_{st}}$

Ranking Aggregation

kendall τ distance : $d_\tau(w_1, w_2) = \dfrac{\#\text{ pairs ranked in different order}}{\#\text{ pairs}}$

Given m Rankings $w_1, \cdots, w_m$, the aggregate ranking w^*

$$w^* = \arg\max_{w} \sum_{i=1}^{m} d_\tau(w, w_i)$$