Unit II - Measures of cohesion

▼ Triadic closure

p - probability of a and b becoming friends when they have a common friend

(1-p) - probability of a and b not becoming friends

$(1-p)^k$ - probability of a and b not becoming friends when they have k common friends

$1 - (1-p)^k$ - probability of a and b becoming common friends when they have k common friends.

▼ Measuring cohesion

- The adjacency matrix multiplied by itself, (raised to the 2nd power) counts the number of pathways between two nodes that are of length two when the matrix is binary.

- Measuring the density of a network gives us an index of the degree of dyadic connection in a population.

- Measure the average geodesic as an indicator to how close the actors are to one another.

▼ Degree distribution

- Entropy - If entropy is high, distribution is relatively random

- Assortativity - $-1 < r < 1$

▼ Network linking behaviour

▼ Transitive linking

Transitive linking needs at least three nodes.

Higher transitivity leads to a denser graph.

- Perfect transitivity occurs in networks where each component triple is a fully connected subgraph or clique.

▼ Reciprocity

Dyadic relationship

▼ Social balance theory

multiplication of edge values in a triangle must be positive for the cycle to be socially balanced.

▼ Clustering Coefficient - Watts and Strogatz

Local clustering coeff - measures transitivity at the node level.

Global

▼ Newman - Strogatz - Watts

▼ Bowtie structure of the web

▼ Cliques

A clique is a complete subgraph such that every member of the set is connected by an edge to any other vertex

A clique is an induced subgraph that is a complete graph

- Maximal clique - one where no more nodes can be added without disturbing the clique property
  NP hard problem

▼ N - Cliques and N - clans

N-Clique

Any subset S of nodes in which the geodesic path between every pair of nodes is {u, v} belonging to S is ≤ n

N-Clan

Restricting the total path distance between any two members. All ties must occur through other members of the group.

An n-clique in which the subgraph induced by set S has a diameter ≤ n will be an n clan.

▼ K-plex and K-core

K-plex

K-plex of size n is a maximal subset of n vertices within a network such that each vertex is connected to at least n-k of the other vertices of the induced sub graph

<u>K-Core</u>

A maximal subgraph of nodes such that each node in the subgraph has at least a degree k

K-crust - What is left of the original network after removing the core. The periphery of the network.

K-corona - the subgraph of the k-core in which all nodes have exactly k neighbours

▼ Community detection

    ▼ Clique percolation method

        Finding overlapped communities - Two cliques of size k are adjacent when they share k-1 nodes.

        Algorithm -

- Construct a Cnxn matrix

- Fill the diagonals with the k-clique values

- Delete any values < k

- Fill the non diagonal elements, delete values k - 1.

- Convert all non zero elements to 1

- Combine adjacent cliques to form a community

    ▼ Girvan Newman

    ▼ Louvain

▼ Modularity

A measure that defines how likely the community structure found is created at random.

(-1, 1) with -1 being non modular and 1 being fully modular

▼ Purity

The fraction of instances that have labels equal to the community's majority label