Global network properties

Introduction to complex networks theory

ADA 2022/2023

Marija Mitrović Dankulov mitrovic@ipb.ac.rs
Aleksandra Alorić aleksandra.aloric@gmail.com

Recap centralities

- Degree centrality
  - The higher it is the more different nodes one can reach with in a single step
- Betweenness centrality
  - The higher it is, the more probable that the node will play a broker role between other nodes
- Closeness centrality
  - The higher it is, the faster information from the node can reach others

- The centralities are node measures, but they might contain some global information about the network - such as betweenness centrality

Network centrality measures

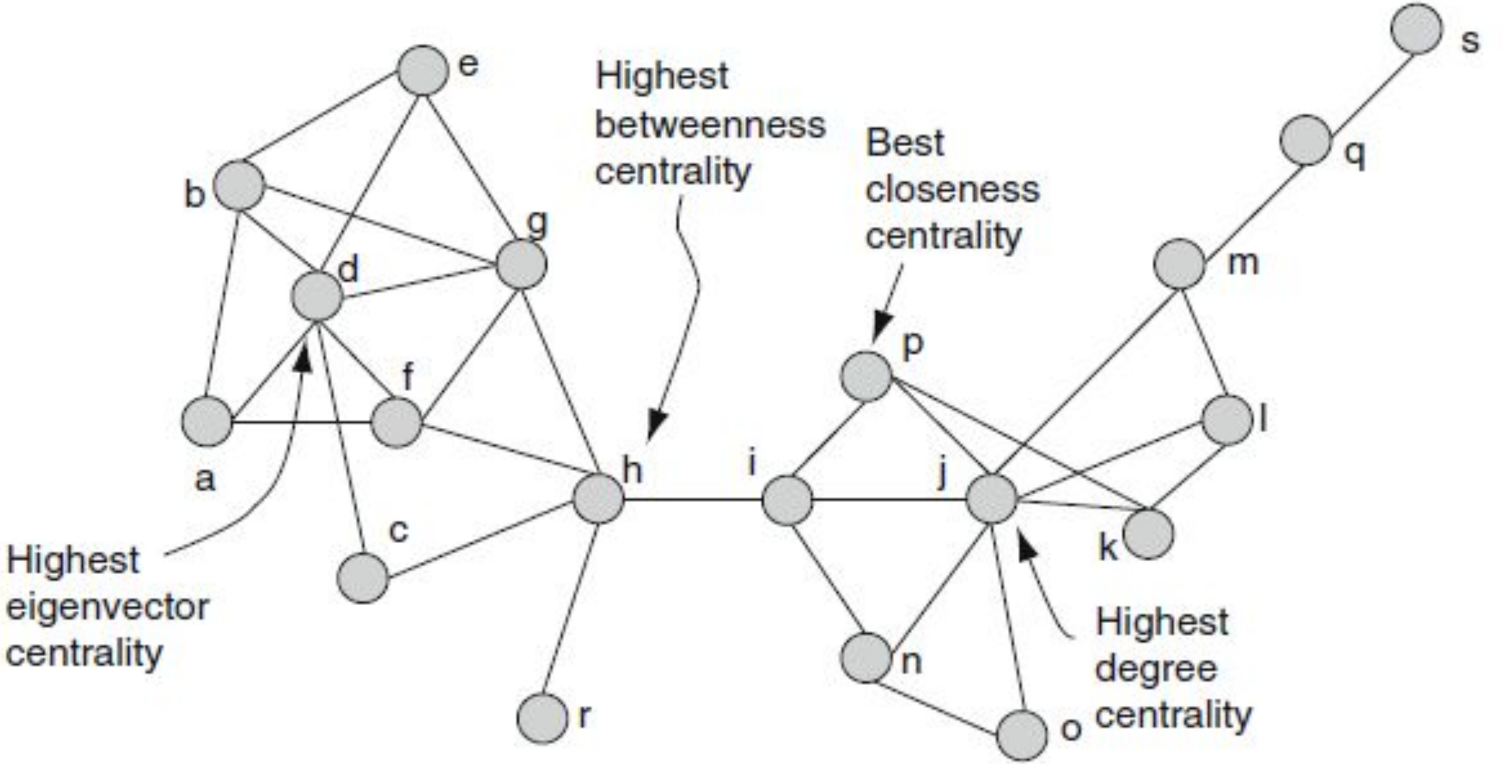


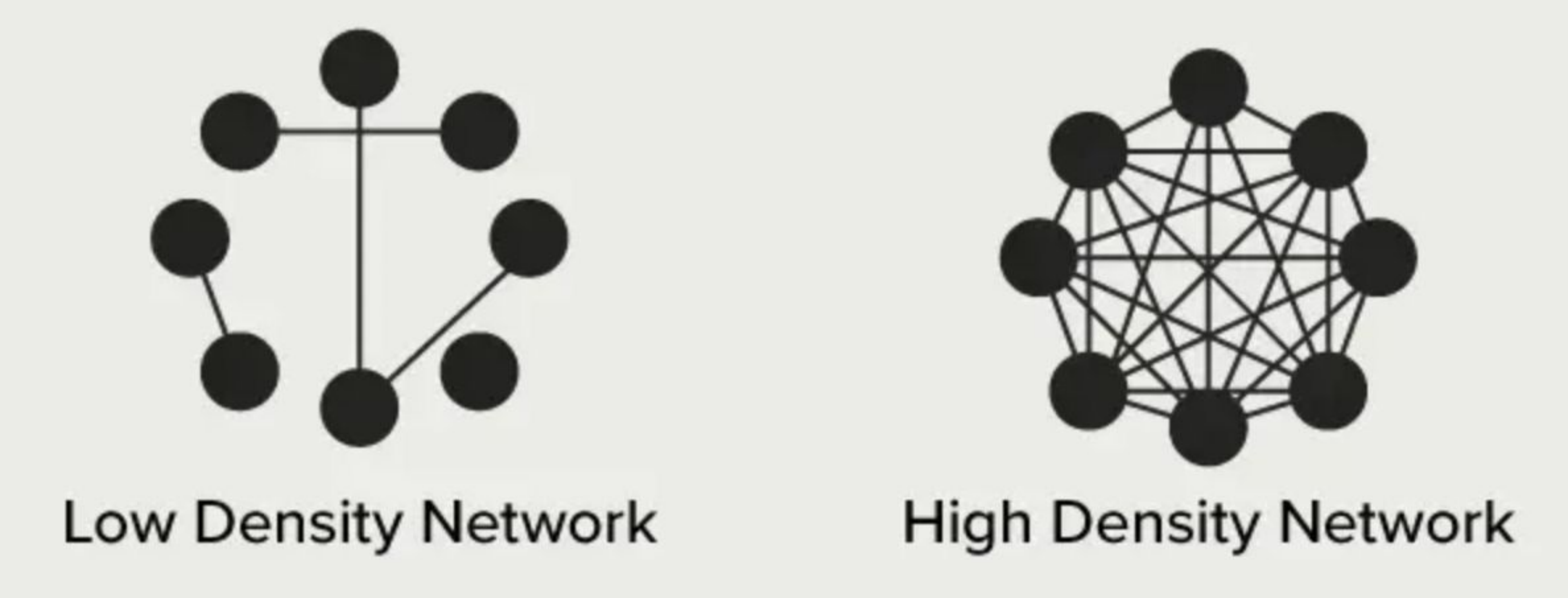
Image credits:
https://ultrabpm.wordpress.com/category/abpd/

Aim of this class

Is to continue characterising network structure, focusing on more global and local information:

- Global:
  - Network connectedness
  - Density
- Motifs:
  - Reciprocity
  - Transitivity
  - Cliques

Network density



Network density

- Tells us about how many connections between links are present in the network (out of all possible)
- Figure right shows effect of speed dating intervention on acquaintance network among scientists

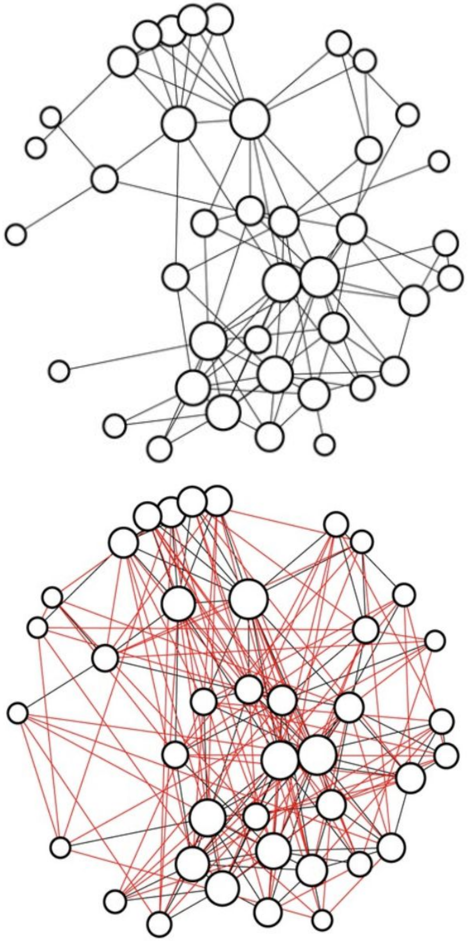
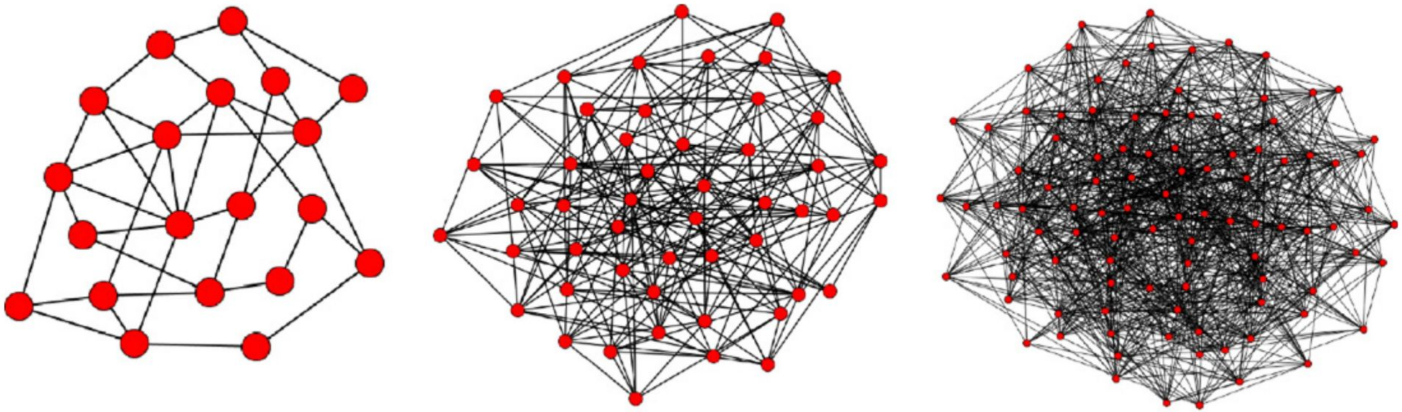


Image credits:
https://elifesciences.org/articles/02273

Network densities

- Pay attention when comparing networks of different sizes, although for larger network the density might be lower, it could still have higher average degree which could be more important for your question



20 nodes, 38 links
Density = 0.20
Links per node = 1.9

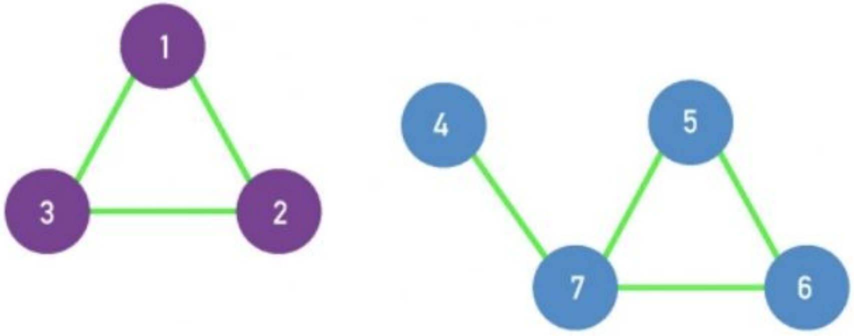
50 nodes, 245 links
Density = 0.20
Links per node = 4.9

100 nodes, 990 links
Density = 0.20
Links per node = 9.9

Image credits:
https://leadershiplearning.org/system/files/SNA%20and%20Leadership%20Networks%20-%20LQ.pdf

Connectedness

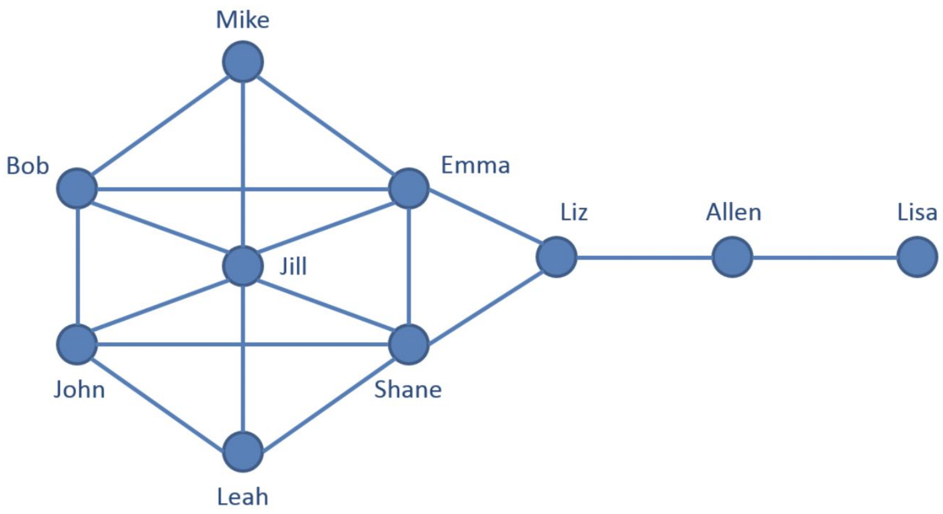
- Two nodes i, j are connected if there's at least one path between them
- If there is a pair of nodes in a network for which the path does not exist, the network is **disconnected**
  - Remember what is network diameter? What's diameter of disconnected network?
- Small networks we can visually classify as (dis)connected, but for large scale networks we rely on algorithms



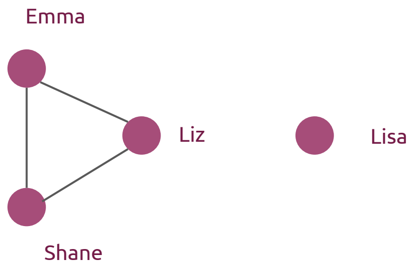
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Subgraph

- Choosing any set of nodes, you can create a subgraph from the original graph
- The subgraph contains all the edges between chosen nodes that existed in the original graph



Subgraph example:



Connected component

- Maximal connected subgraph
- Connected graph has only one connected component
- In disconnected graph, the largest connected component usually contains substantial fraction of nodes, see example on the right

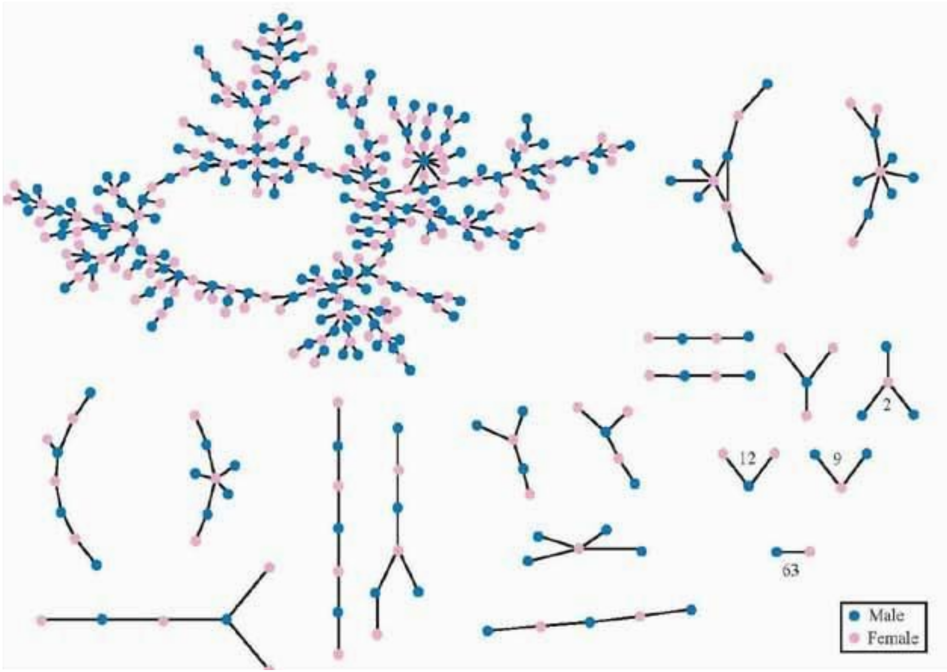
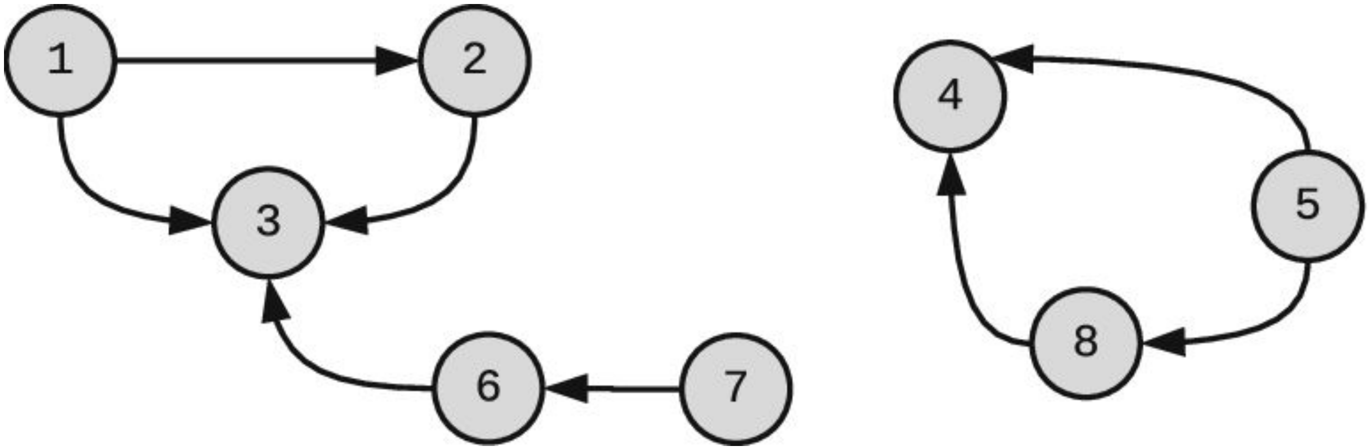


Image credits:
https://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/teensex.pdf

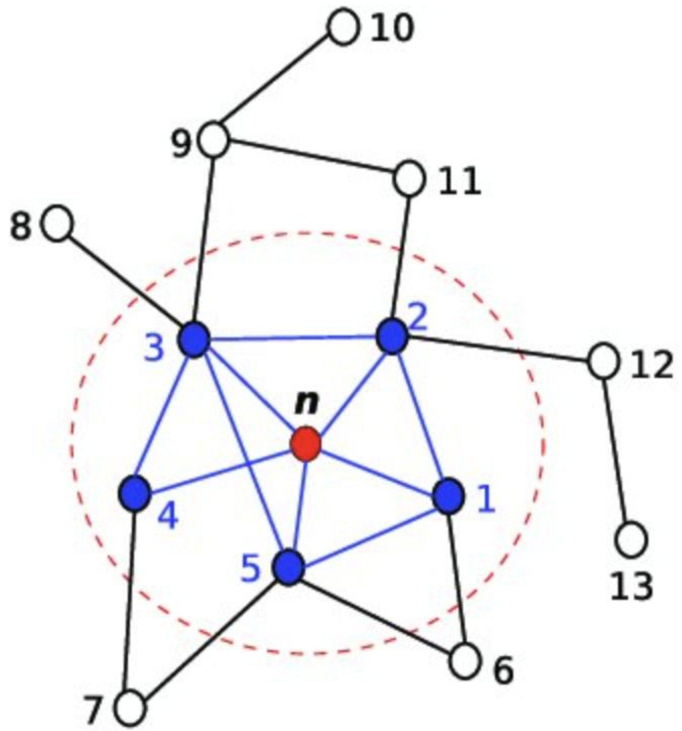
Connectedness in directed network

- Network is weakly connected if the underlying undirected network is connected
- Network is strongly connected if there is directed path between any two nodes
- Assess connectedness of this network:



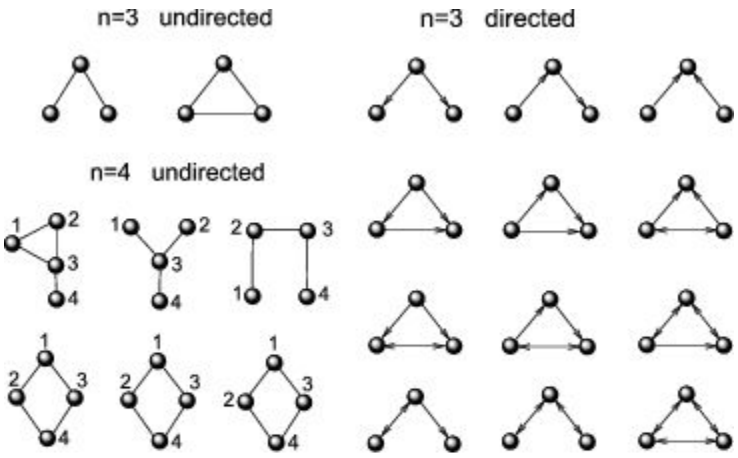
Ego network

- Special example of subgraph
- Starting from a node of interest, we create a subgraph containing all the neighbours of the studied node, together with all the connections between them

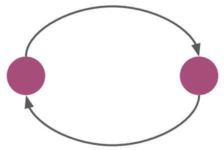


Motifs

- Particular connected subgraphs that are of importance, usually small
- Count frequency of motifs in a network
- Computational cost of finding and counting all motifs in a graph quickly increases with motif size, that is why we mostly focus on small motifs



Reciprocity



- Smallest motif that is of importance
- Think of directed social networks such as Instagram, Twitter, the metrics of interest is how many of the directed links are reciprocated
- We can ask this question locally
  - How many of the outgoing links of a node in question are reciprocated (e.g. how many of the people I follow, follow me back?)

$$r_i = \frac{1}{k_i^{out}} \sum_{j=1} A_{ij} A_{ji}$$

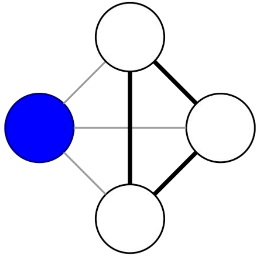
- ...or globally
  - How many links are reciprocated within a network

$$r_i = \frac{1}{L} \sum_{i,j} A_{ij} A_{ji}$$

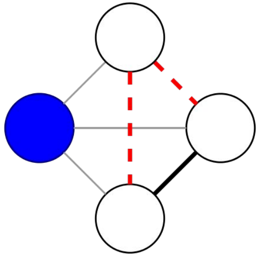
Clustering coefficient

- Triadic closure
- How often are my friends friends to each other
- For the node you're trying to calculate clustering, think about its **ego** network
- Clustering is ratio between the number of links without the node (L_i) and total number of links between the friends there could have been created $C_i = \frac{2L_i}{k_i(k_i-1)}$
- Using adjacency matrix:

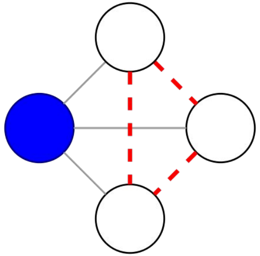
$$C_i = \frac{\sum_{jk} A_{ij} A_{jk} A_{ki}}{k_i(k_i-1)}$$



c = 1



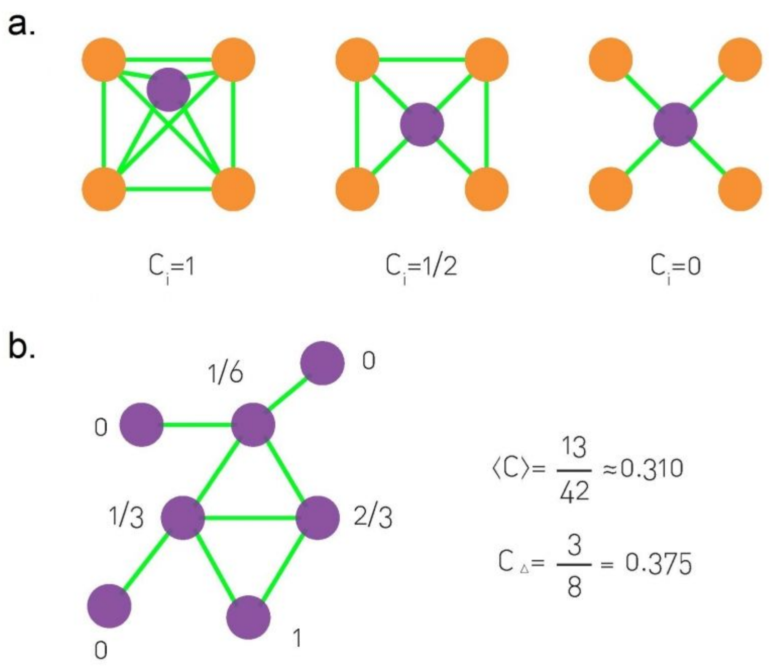
c = 1/3



c = 0

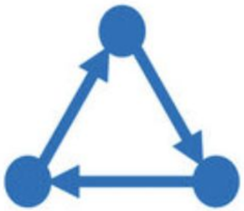
Clustering coefficient

- Local link density, the more interconnected the neighbourhood of node i, the higher the local clustering coefficient
- For the full network, we can either calculate average clustering coefficient, or global clustering coefficient

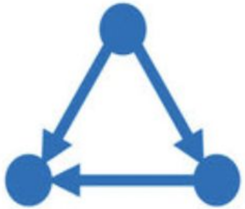


Feed forward and feedback loops

- 3 node motifs in directed network
- Appear in gene regulatory networks and other data driven networks more frequently than it would be expected by chance
- Example - biological circuitry for governing the flow of activation across molecular networks
    - Feed forward loop - speed up the response time of the target gene expression following stimulus steps in one direction



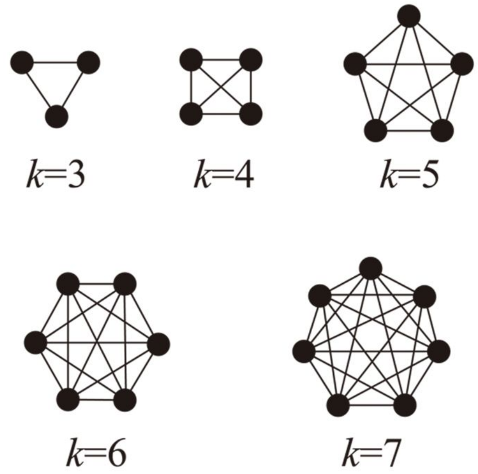
Feedback loop

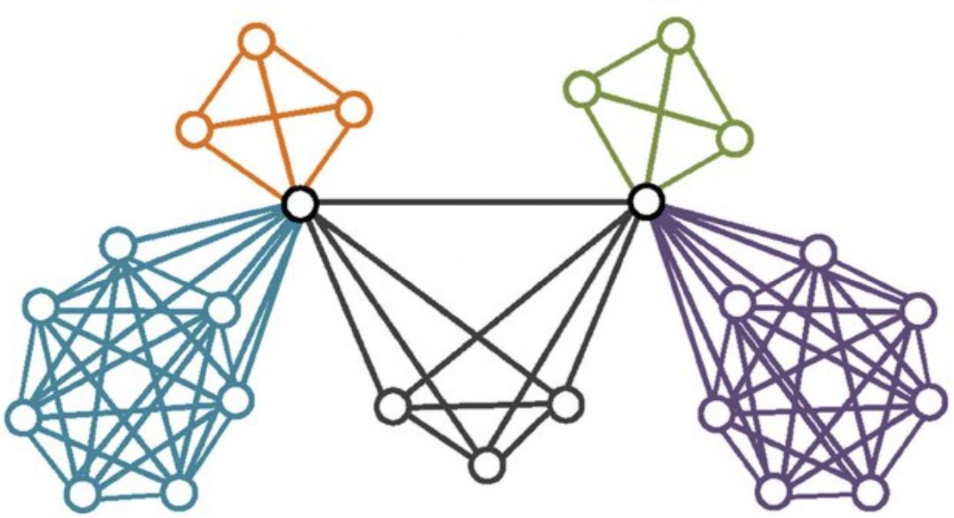


Feed-forward loop

Cliques

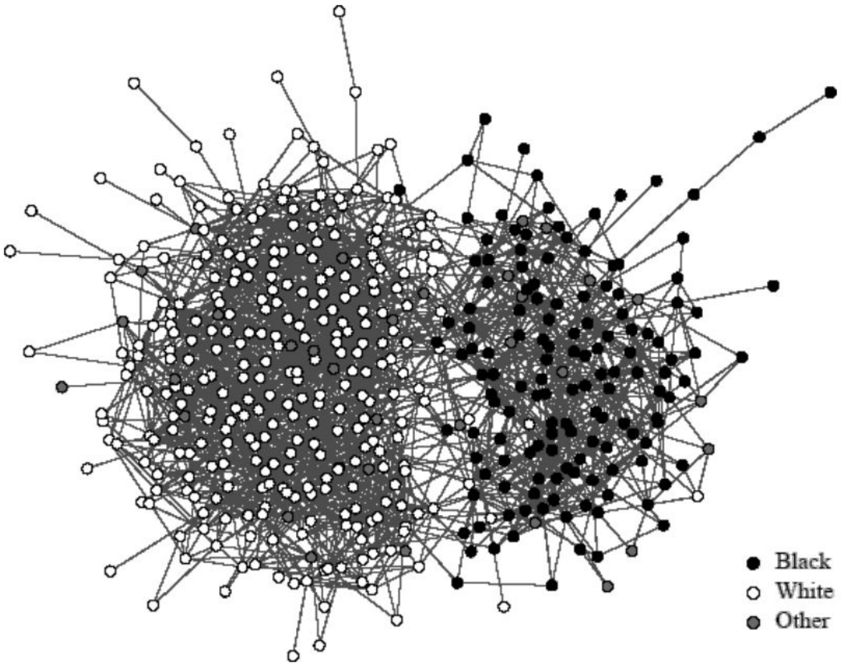
- Fully connected subgraphs
- Tightly knit subgrouop of people, of interest when studying social dynamics
- Large cliques are rare in real networks





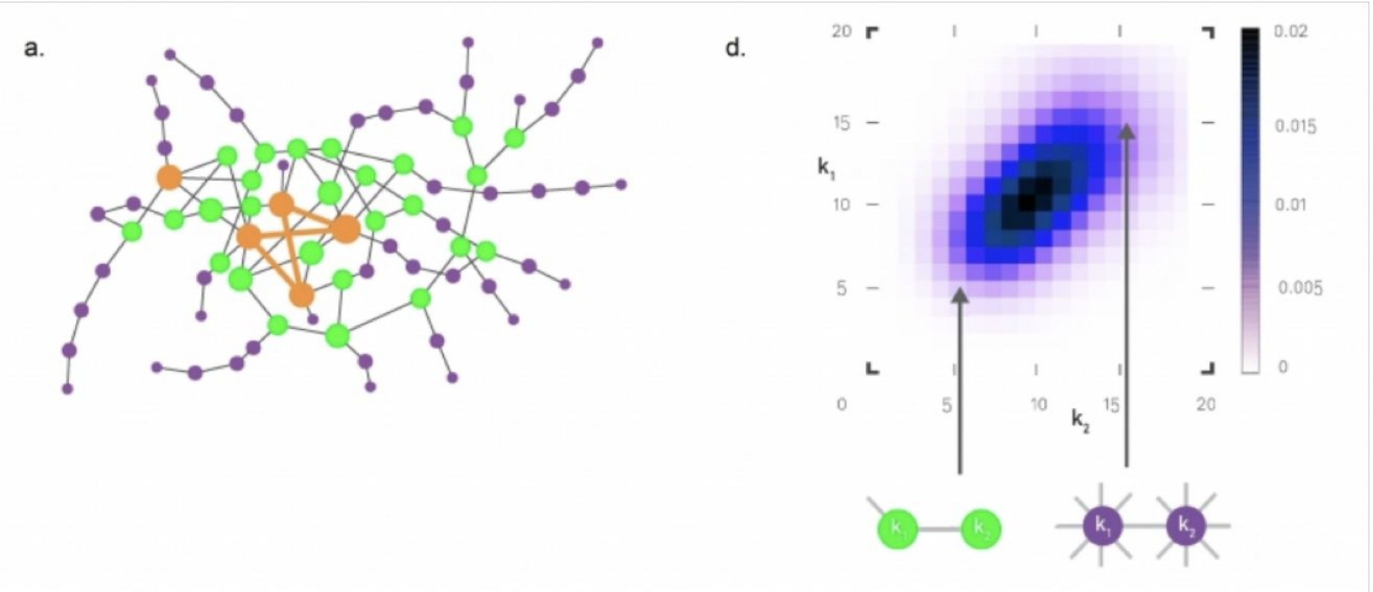
Assortativity

- Birds of a feather flock together
- Tendency to connect with nodes that are similar
- Example: friendship network in a highschool in US where race assortativity is high
- Example disassortative network - sexual relationships among heterosexual people



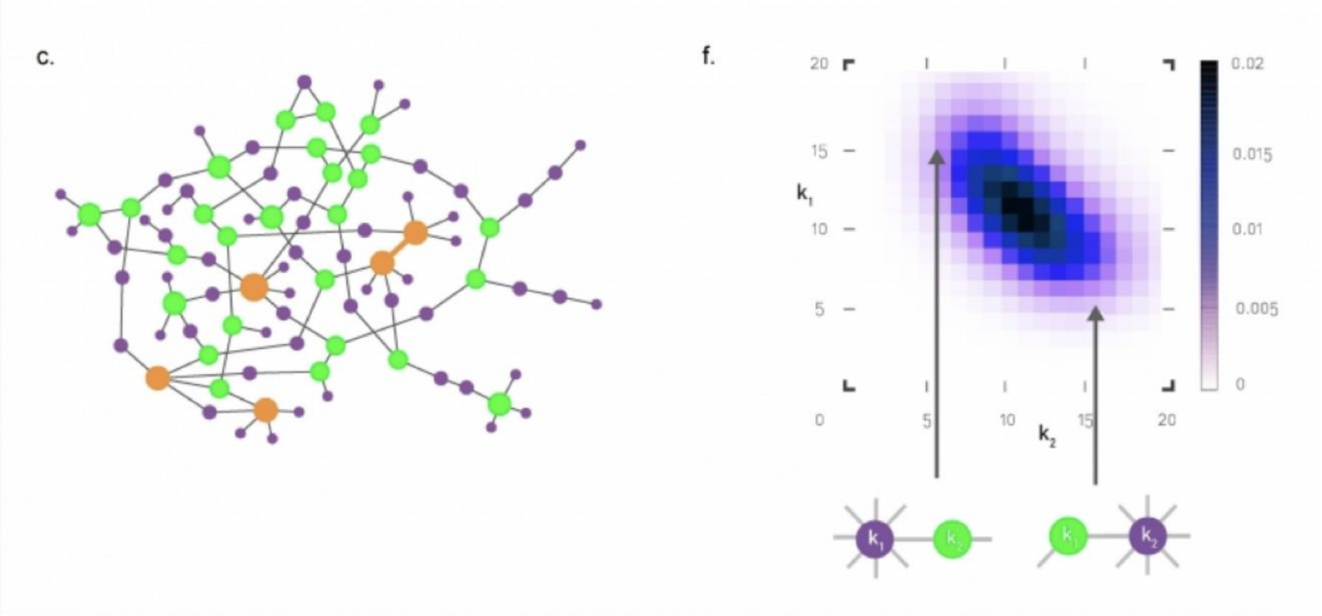
Degree assortativity

- We investigate degrees of nodes on each side of the edges.
- Social networks are often assortative



Degree assortativity

- We investigate degrees of nodes on each side of the edges.
- Social networks are often assortative, while biological and technical are often disassortative



Further reading

- Network science book
  - Sections 2.9 i 2.10
  - Check also summary section 2.11 and 2.A on global clustering coefficient
  - Assortativity - section 7.2
- Lecture notes from A. Clauset

Homework

We return to this simple graph and find clustering coefficients, cliques and subgraphs in it.

The questions are in the submission [link]