

Clique and Link Percolation

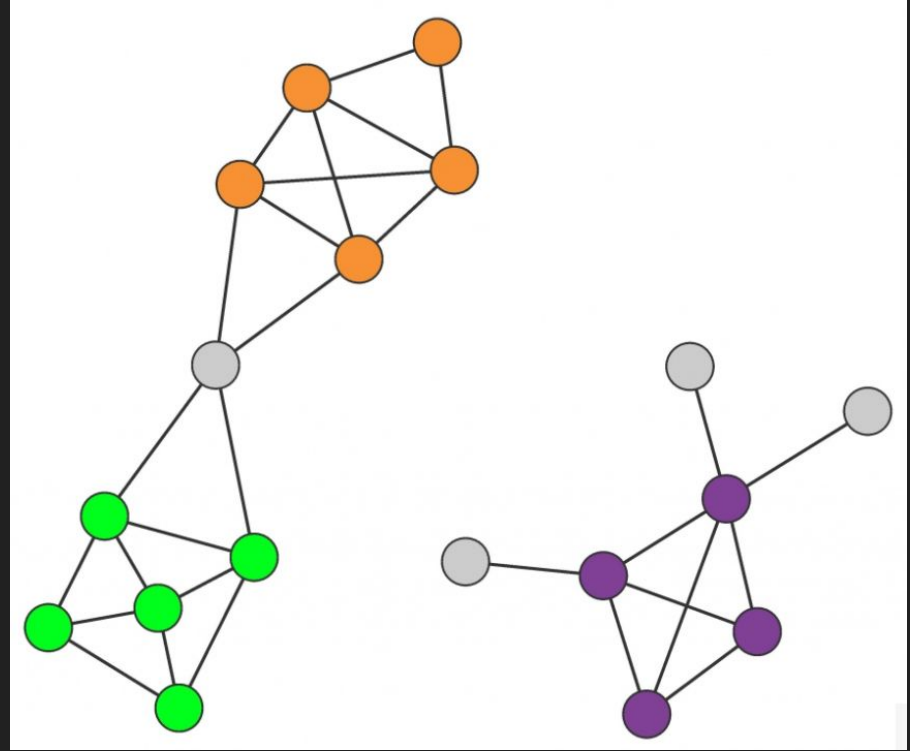
Andrew Celi

Overview

- Classifications of Communities
- Overlapping Communities
- Clique Percolation
- Link Clustering
- Applications

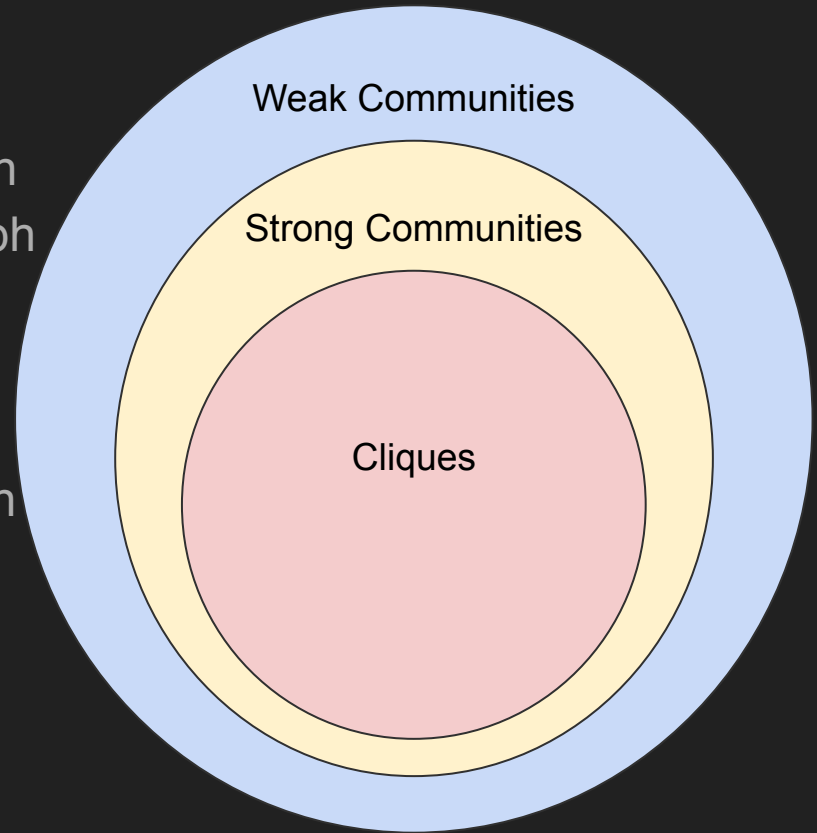
Definitions

- Connectedness Hypothesis: A community is a connected subgraph
- Density Hypothesis: Nodes in a community are more likely to connect to each other than to outside nodes



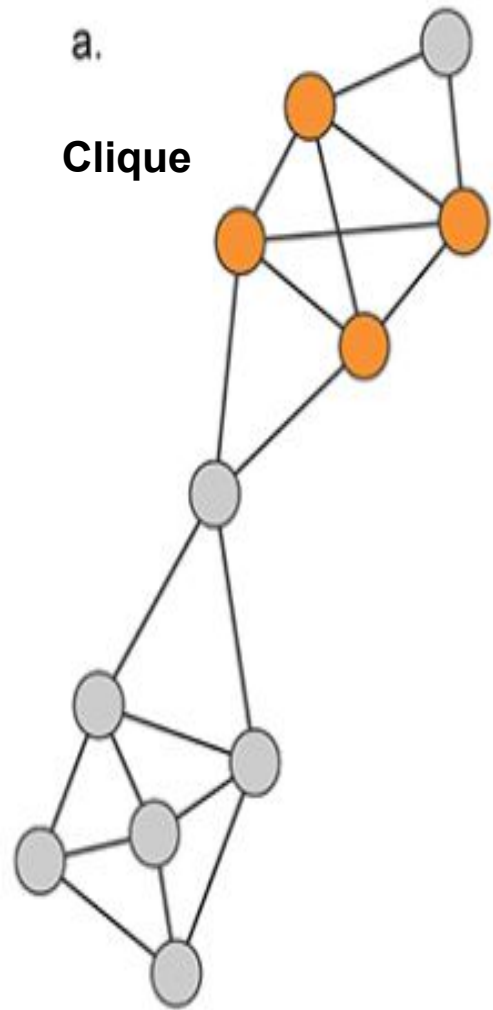
Classifications of Communities

- **Clique:** A completely connected subgraph
- **Strong Community:** A connected subgraph whose nodes have more links to nodes within the same community than nodes outside the community
- **Weak Community:** A connected subgraph whose total internal degree exceeds the total external degree



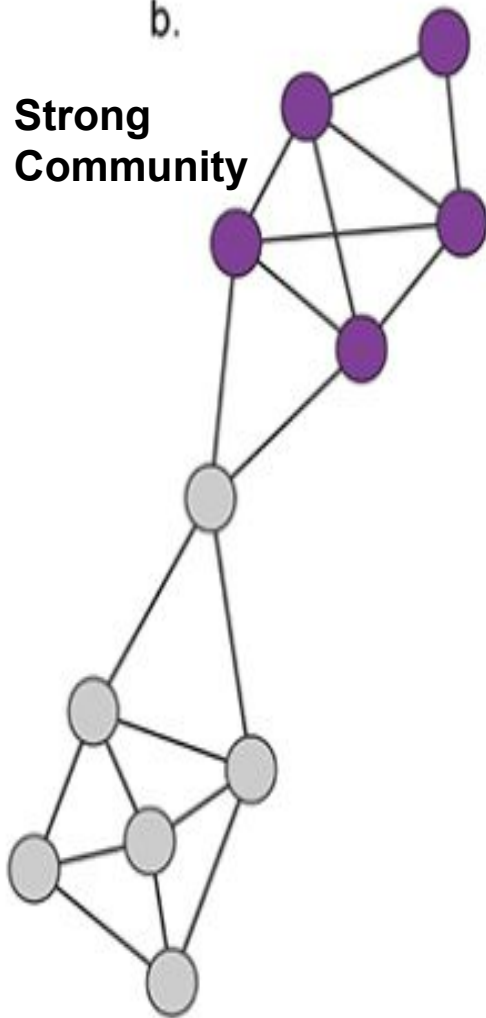
a.

Clique



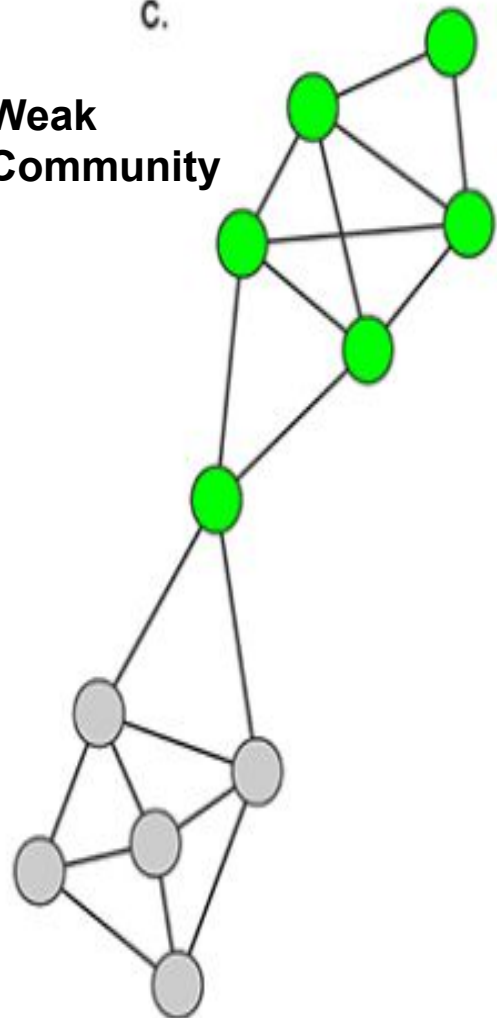
b.

**Strong
Community**



c.

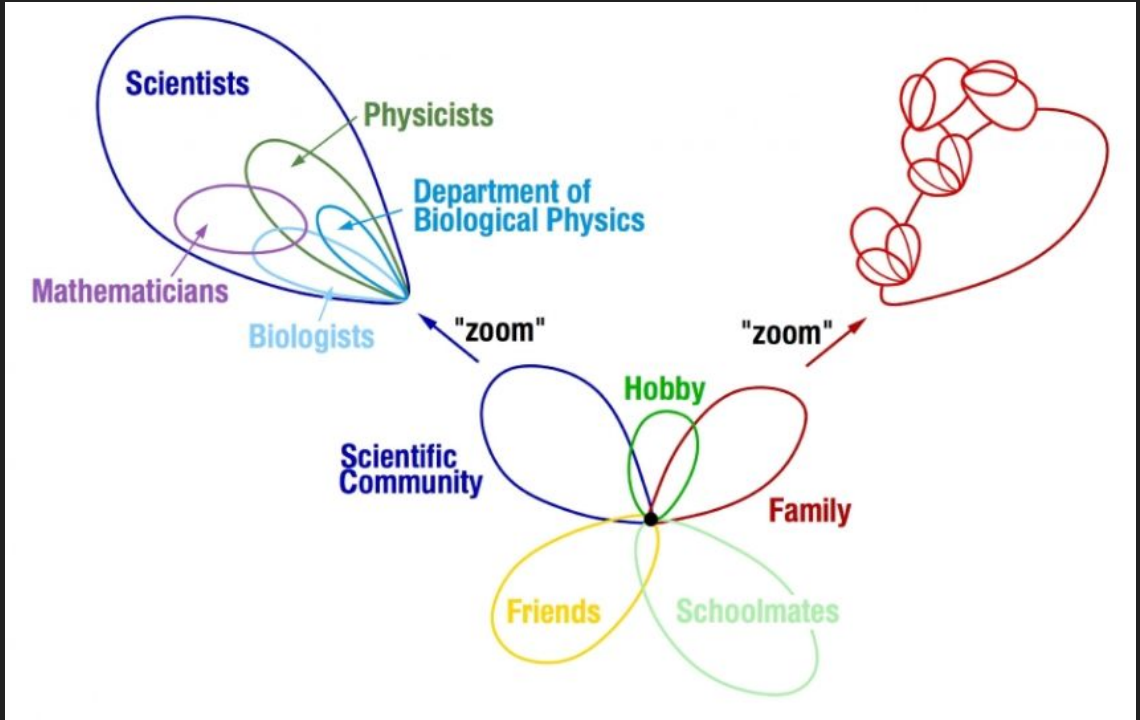
**Weak
Community**



Overlapping Communities

Overlapping Communities

- A node is rarely confined to a single community
- How can we detect overlapping communities?



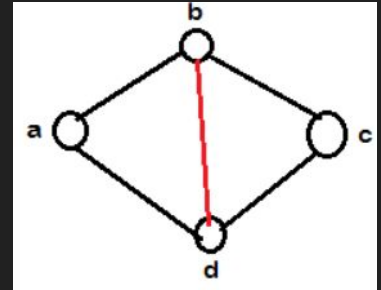
Clique Percolation (CFinder)

1) A community is the union of overlapping, or adjacent, cliques	

Clique Percolation (CFinder)

1) A community is the union of overlapping, or adjacent, cliques

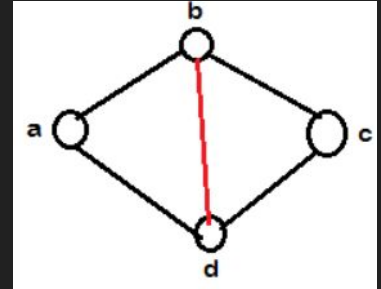
2) Two k -cliques are considered adjacent if they share $k-1$ nodes.



Clique Percolation (CFinder)

1) A community is the union of overlapping, or adjacent, cliques

2) Two k -cliques are considered adjacent if they share $k-1$ nodes.

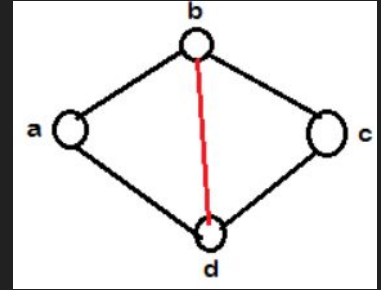


3) A k -clique community is the largest connected subgraph obtained by the union of all adjacent k -cliques

Clique Percolation (CFinder)

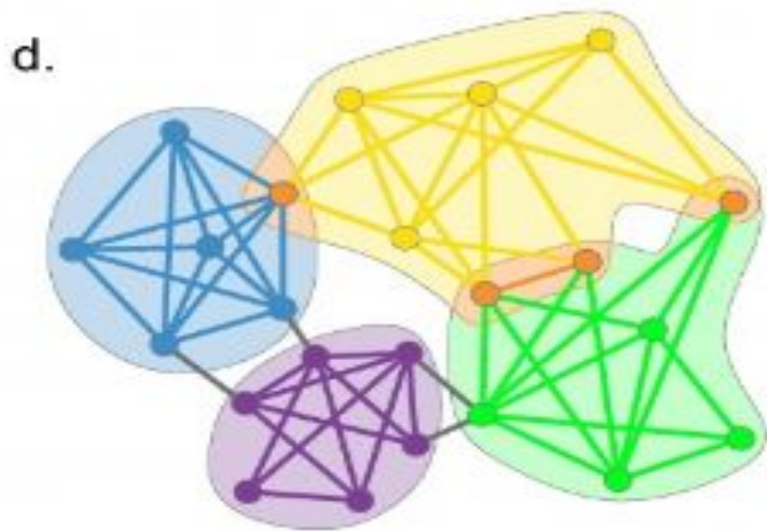
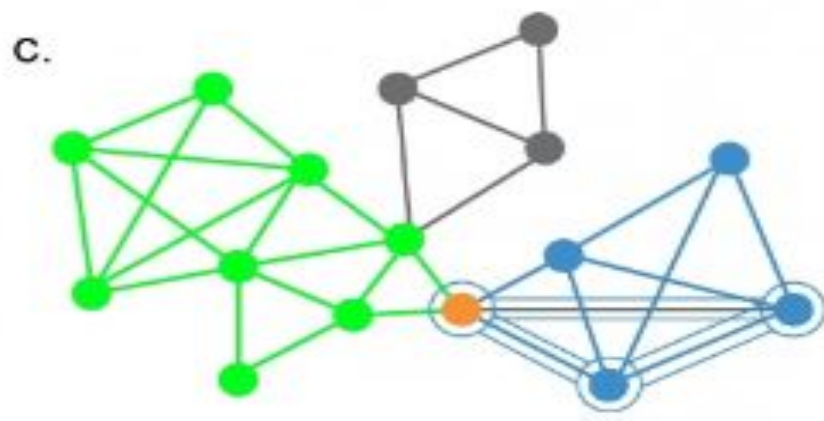
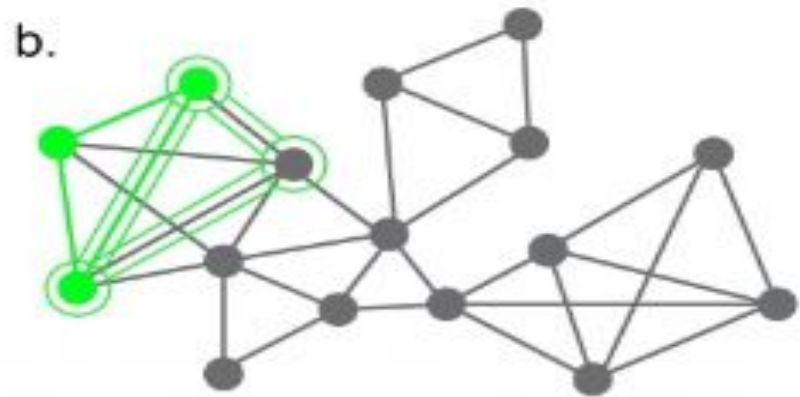
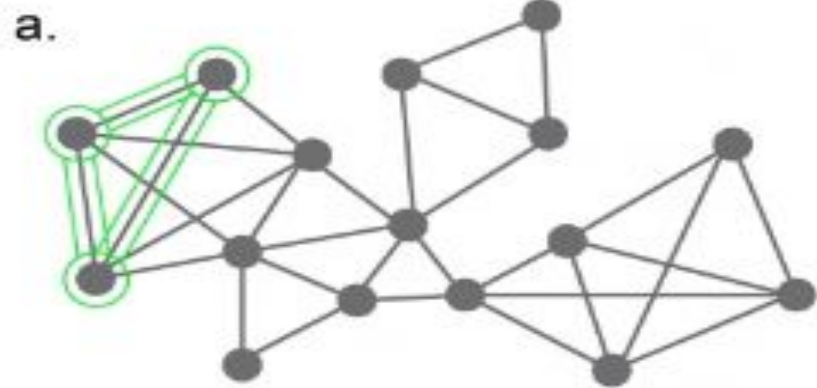
1) A community is the union of overlapping, or adjacent, cliques

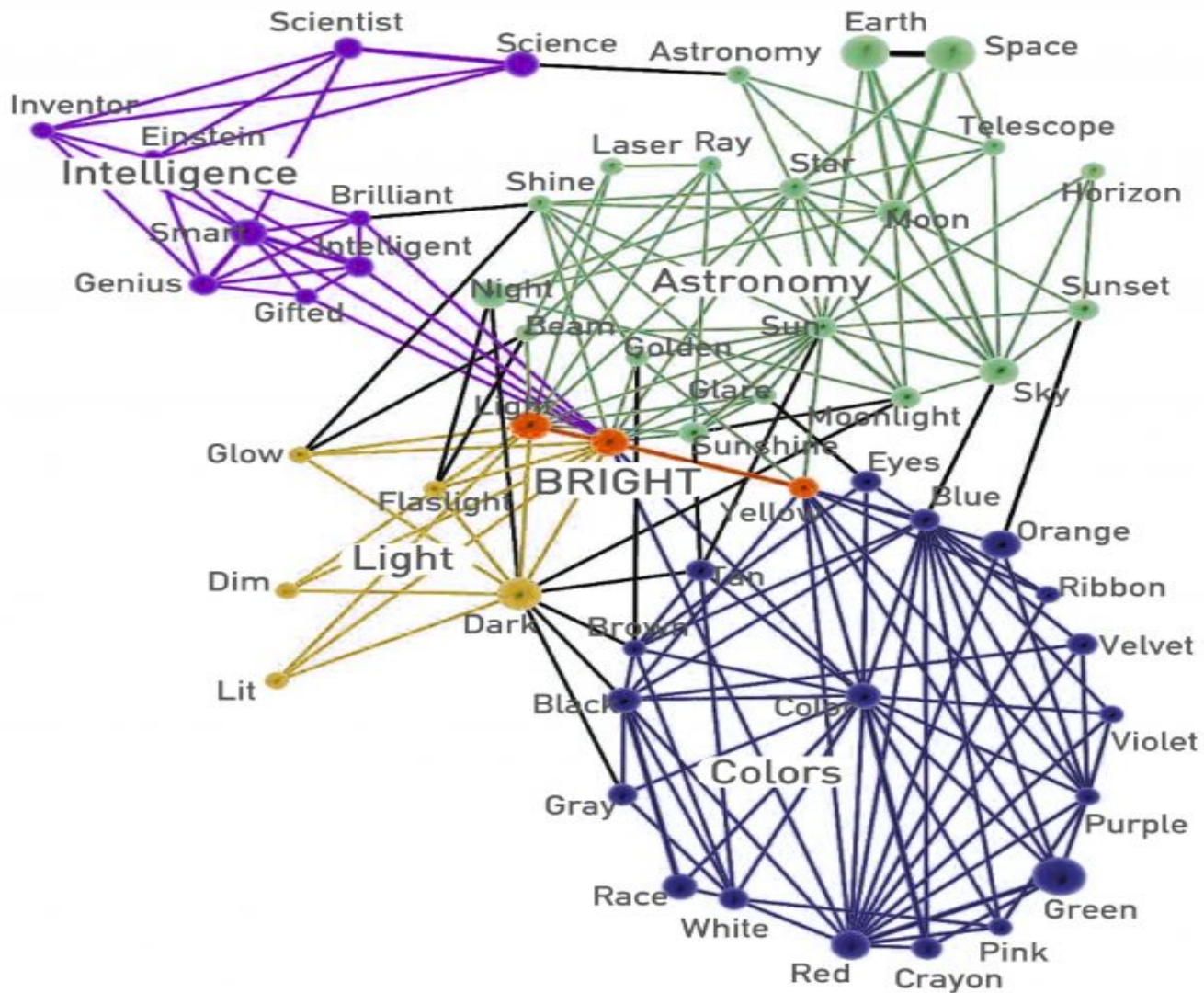
2) Two k -cliques are considered adjacent if they share $k-1$ nodes.



3) A k -clique community is the largest connected subgraph obtained by the union of all adjacent k -cliques

4) K -cliques that cannot be reached from a particular k -clique belong to another k -clique community





Can these emerge by chance?

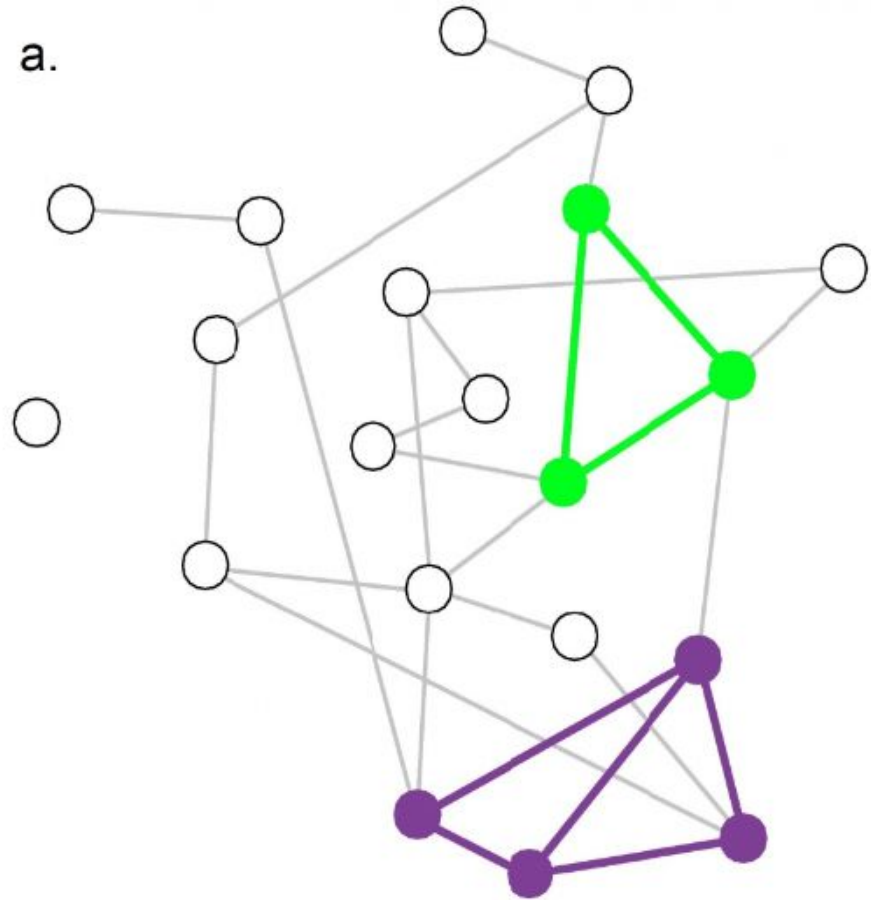
- Random networks can produce very large cliques if they have a very high density

$$p_c(k) = \frac{1}{[(k-1)N]^{1/(k-1)}}$$

k = clique size, N = number of nodes

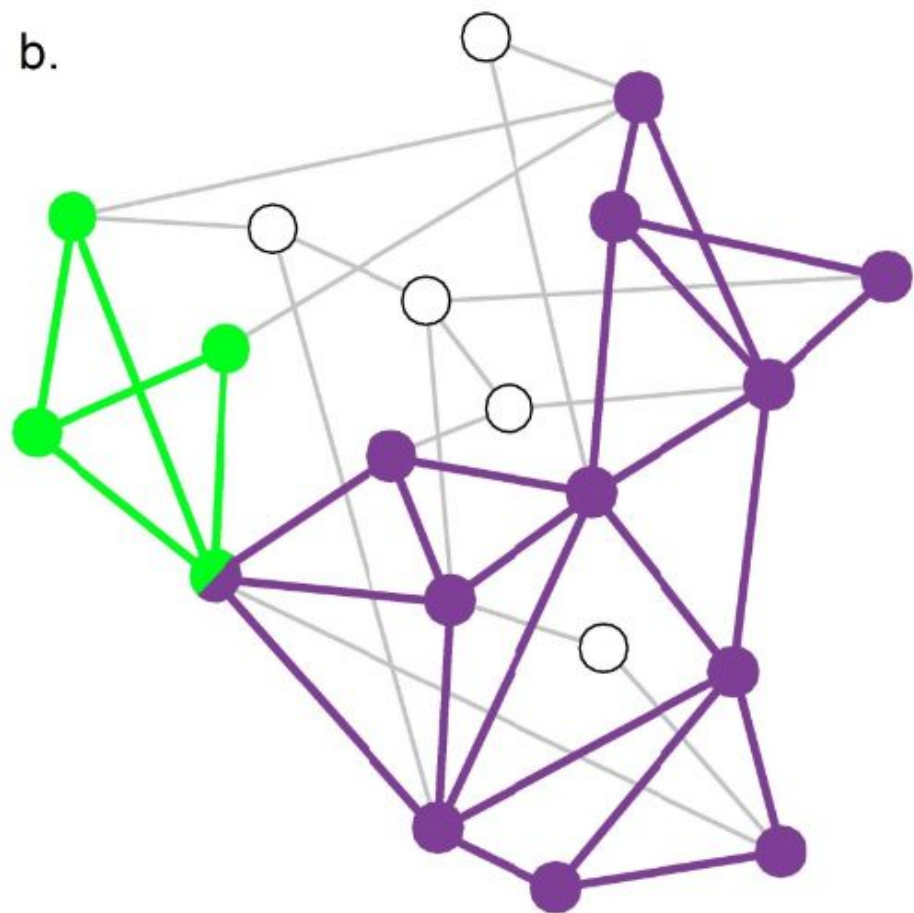
- Subcritical Communities: $p < p_c(k)$
- Supercritical Communities: $p > p_c(k)$

a.



$$p = 0.13, p_c(3) = 0.16$$

b.



$$p = 0.22, p_c(3) = 0.16$$

Link Clustering

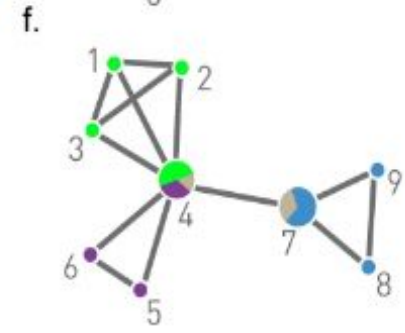
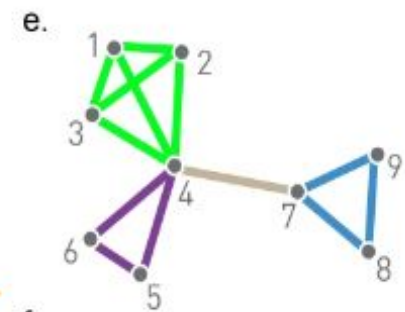
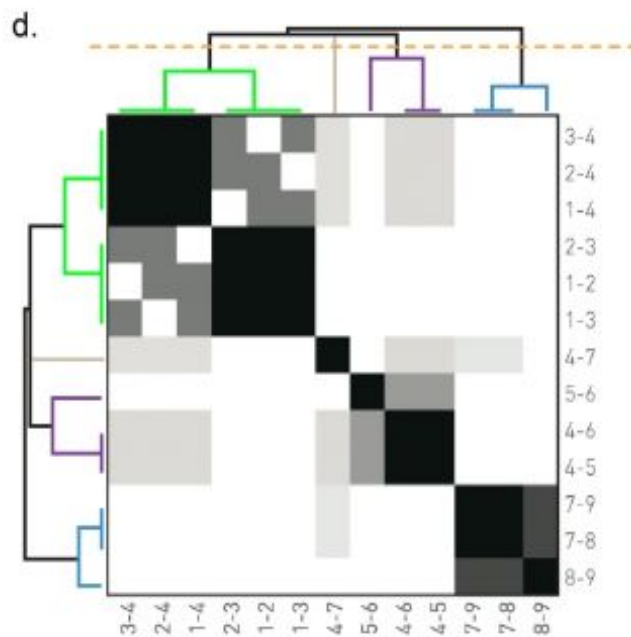
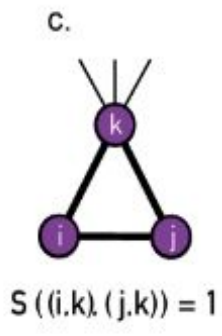
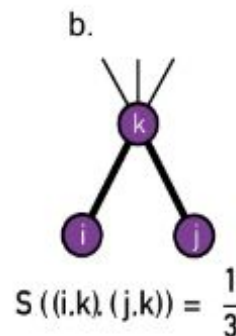
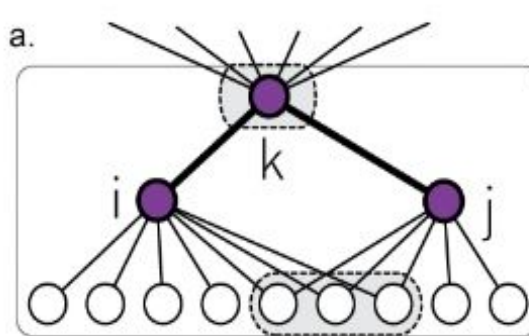
- Nodes can belong to many communities but links typically are community specific
- Algorithm finds link communities by exploring the similarity between their neighbors and themselves

Link Clustering Algorithm

- Link Similarity can be defined as S :

$$S((i,k), (j,k)) = \frac{|n_+(i) \cap n_+(j)|}{|n_+(i) \cup n_+(j)|}$$

- Identify link clusters using a matrix

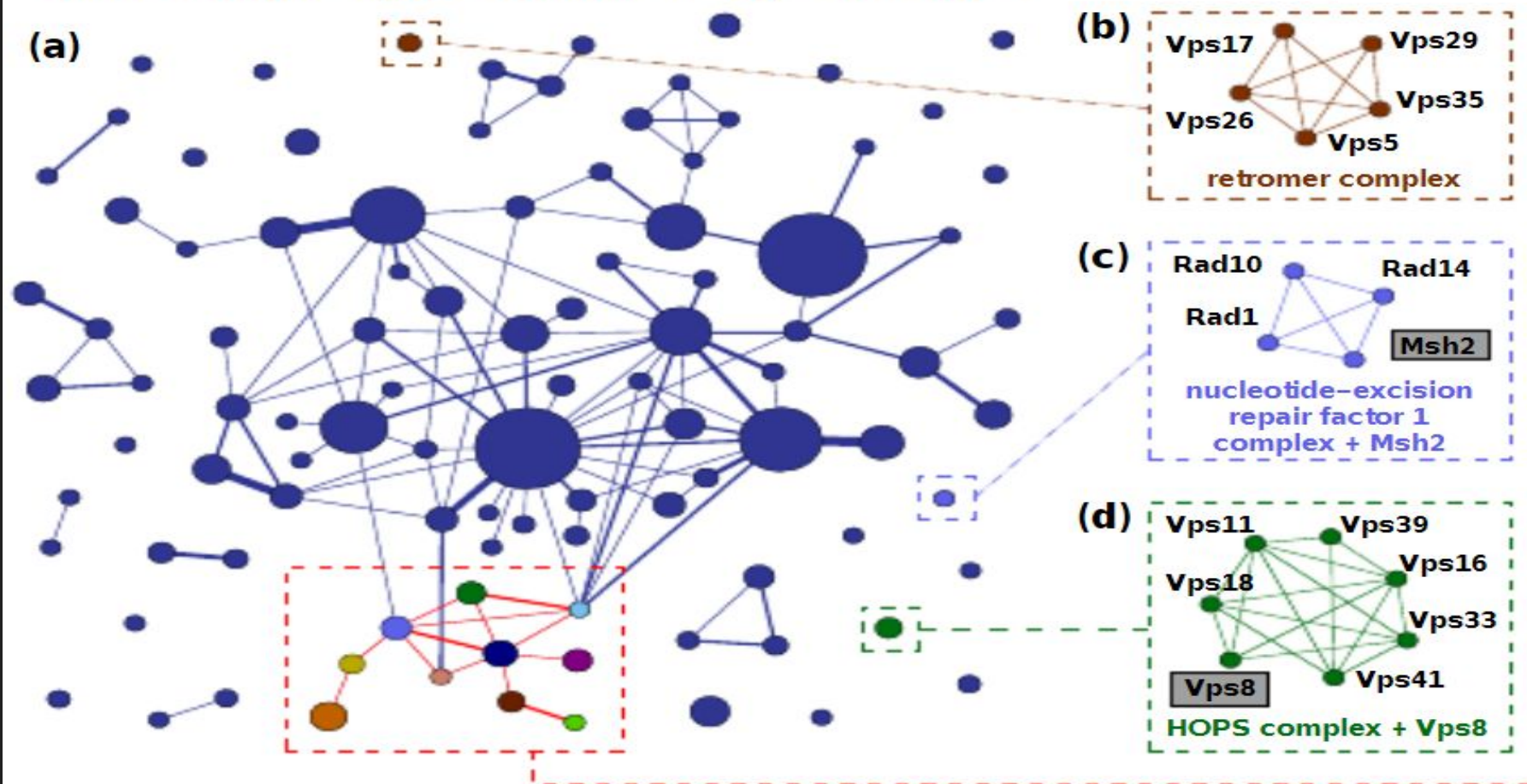


Clique Percolation in Biological Networks

- CFinder: Locating cliques and overlapping modules in biological networks by Balázs Adamcsek, Gergely Palla, Illés J. Farkas, Imre Derényi, and Tamás Vicsek
- Identify which groups of proteins interact with each other

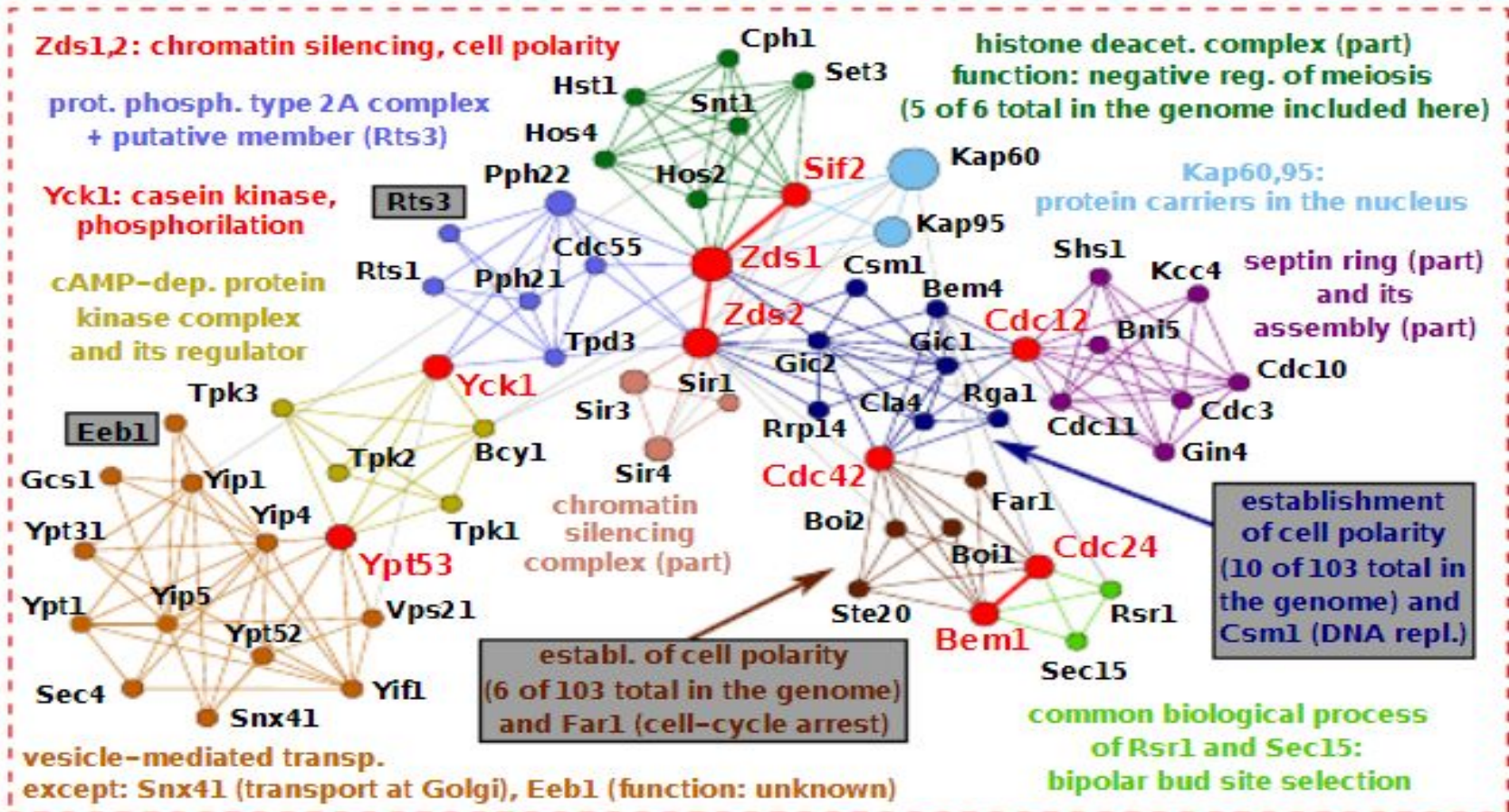
network of yeast PPI modules

node: module of proteins, link: overlap of modules



enlarged portions of the network of modules

(e)



References

Adamcsek, Balazs, et al. *CFinder: Locating Cliques and Overlapping Modules in Biological Networks*.

https://www.researchgate.net/publication/7302566_CFinder_Locating_cliques_and_overlapping_modules_in_biological_networks.

Barabási Albert-László. *Network Science*. Cambridge University Press, 2016.