Community Structure in Networks

Social Networks Analysis and Graph Algorithms

Prof. Carlos Castillo — https://chato.cl/teach

upf. Universitat Pompeu Fabra Barcelona

Sources

- A. L. Barabási (2016). Network Science – Chapter 09
- D. Easly and J. Kleinberg (2010). Networks, Crowds, and Markets – Chapter 03
- F. Menczer, S. Fortunato, C. A. Davis (2020). A First Course in Network Science – Chapter 06
- URLs cited in the footer of slides

Typical community structures

- One dense sub-graph

  embedded somewhere within a larger graph

- Two groups (polarization)

  plus perhaps some ambiguous nodes

- Multiple communities

One dense sub-graph

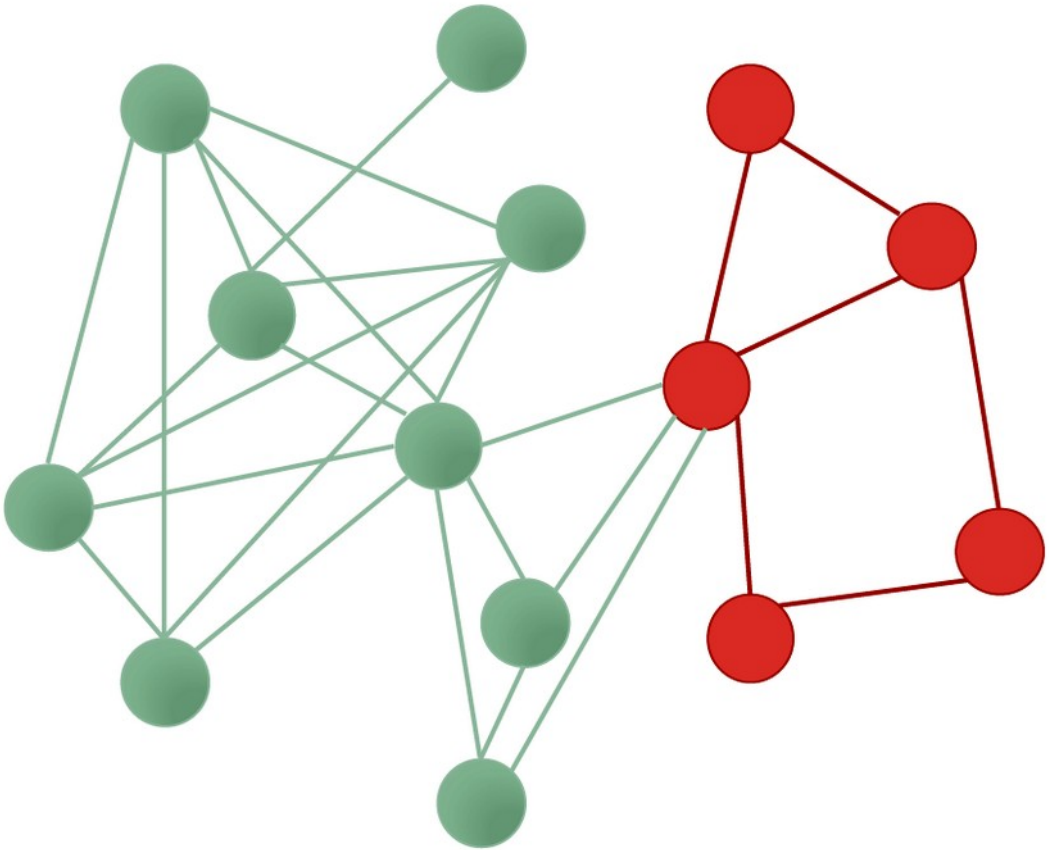
What is a sub-graph?

Subset of nodes,

and edges

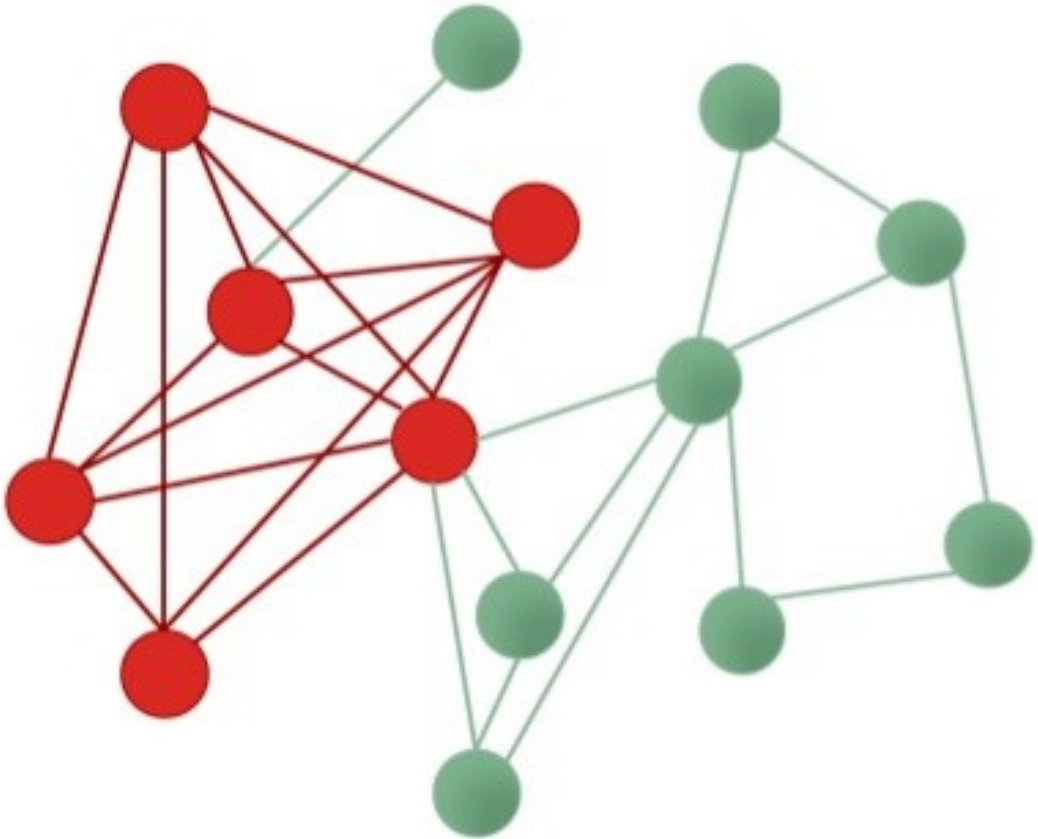
among those

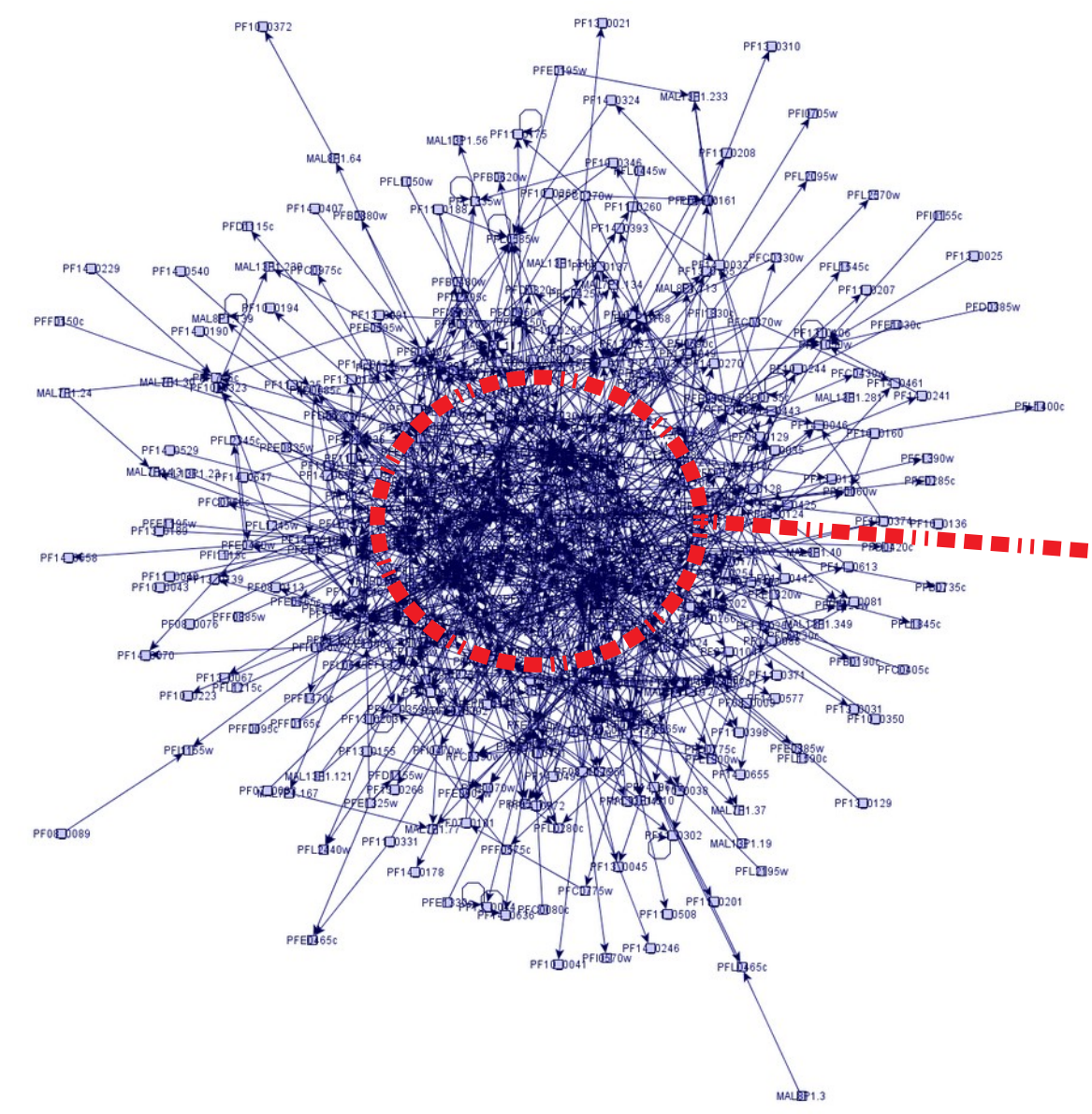
nodes



Densest sub-graph

Sub-graph having the maximum density

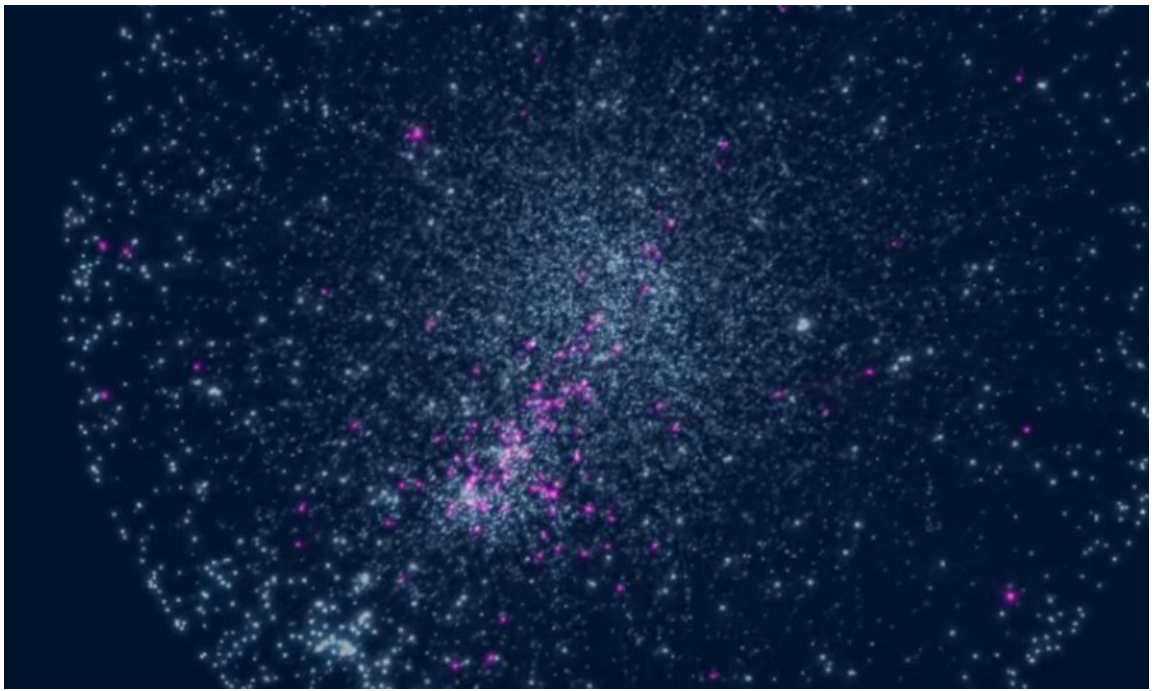




Many graphs look like “hairballs”

Sometimes, at the center these graphs may have an interesting dense sub-graph

Asthma-related genes

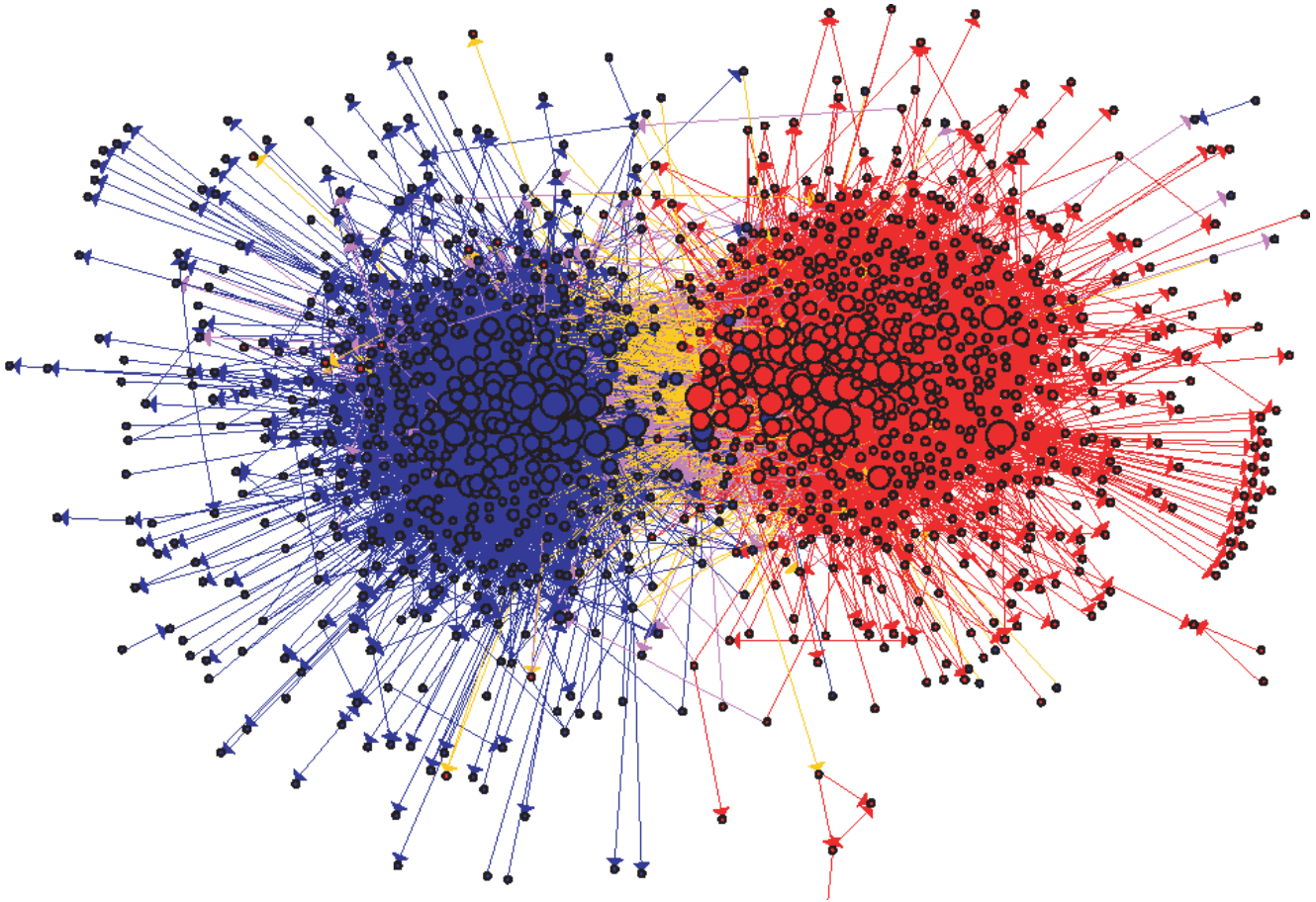


https://www.youtube.com/watch?v=VU_7FHAKMgA

Two groups (polarization)

US Political Blogs (2004)

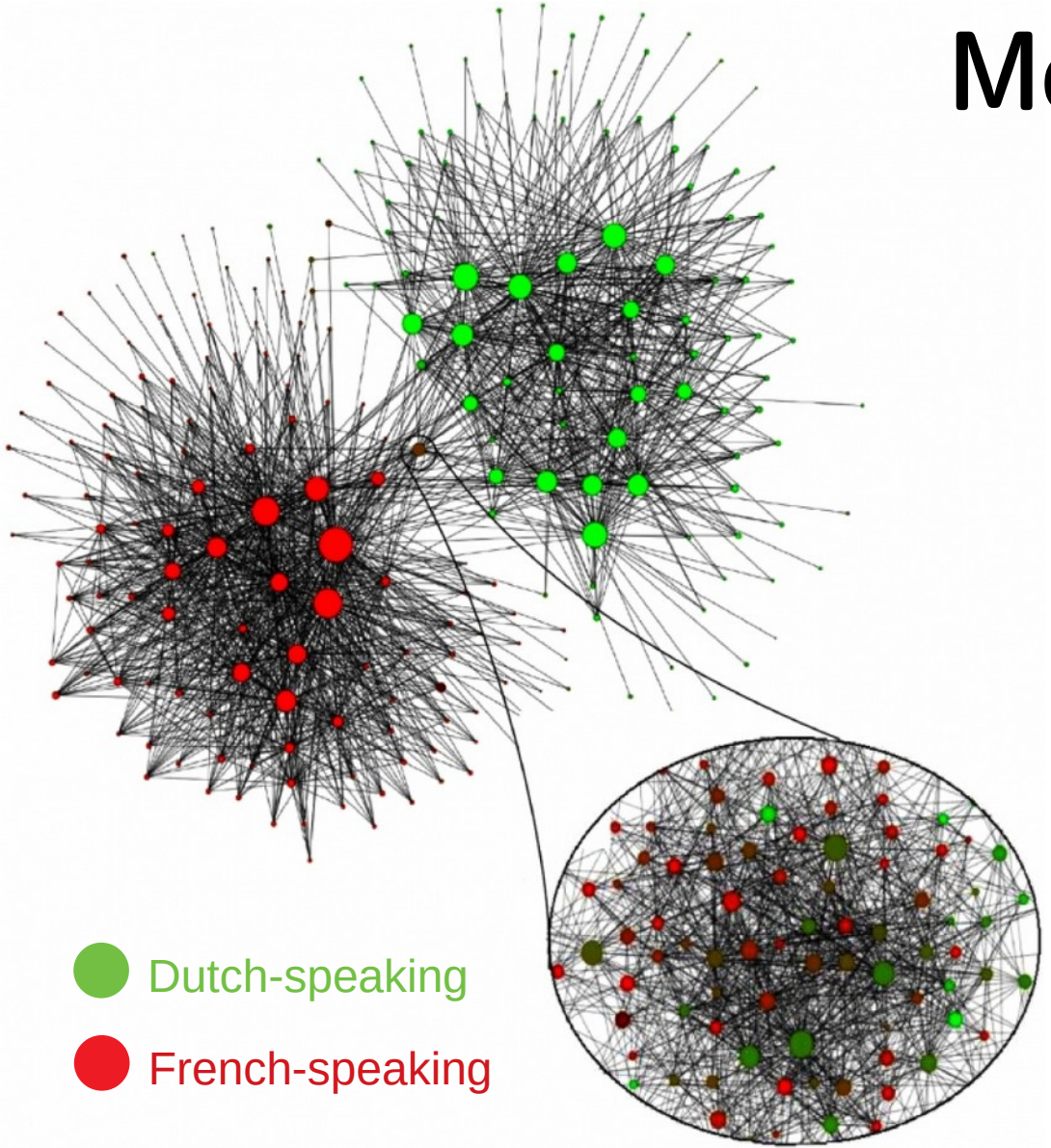
Liberal



Conservative

Adamic, L. A., & Glance, N. (2005, August). The political blogosphere and the 2004 US election: divided they blog. In Proc. LinkKDD (pp. 36-43). ACM.

Mobile phone users in Belgium (2008)

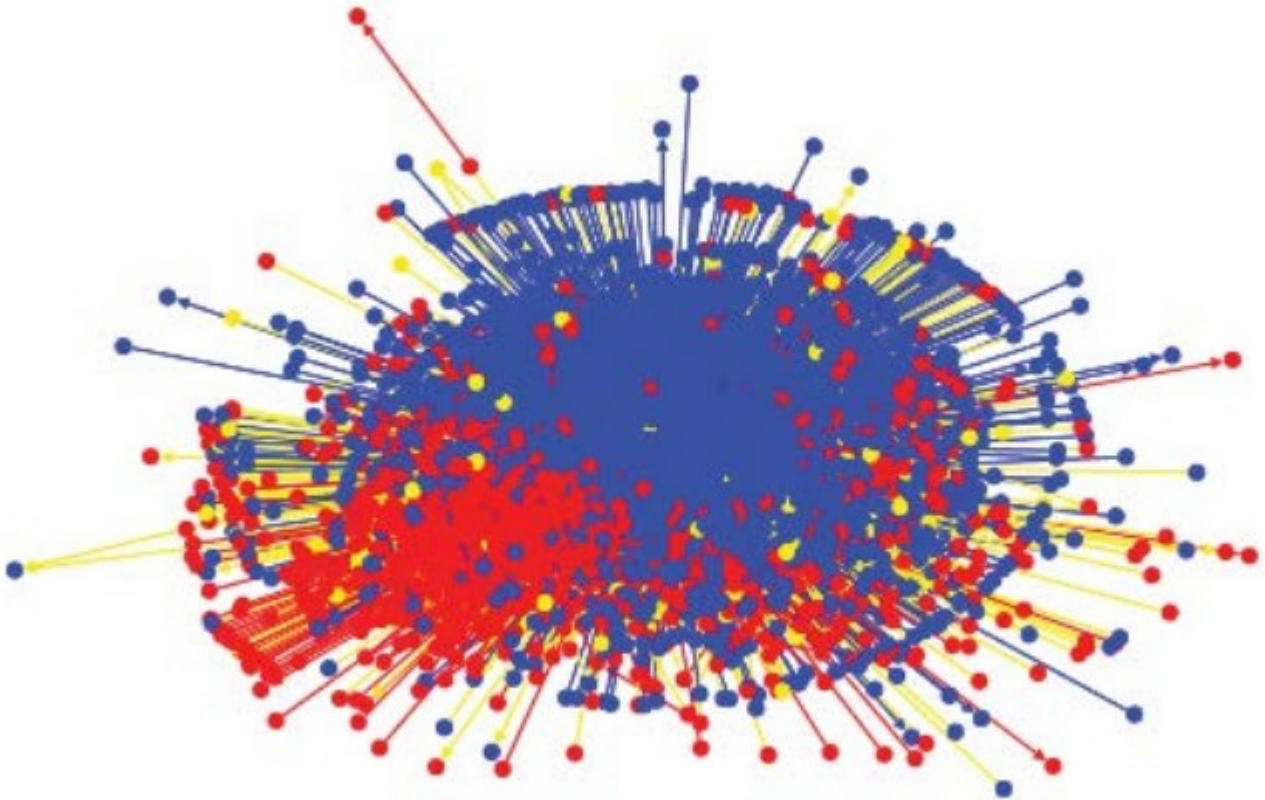


Each node is a community of 100 mobile users or more that tend to call each other

V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. J. Stat. Mech., 2008.

Egyptian Twitter Users (2013)

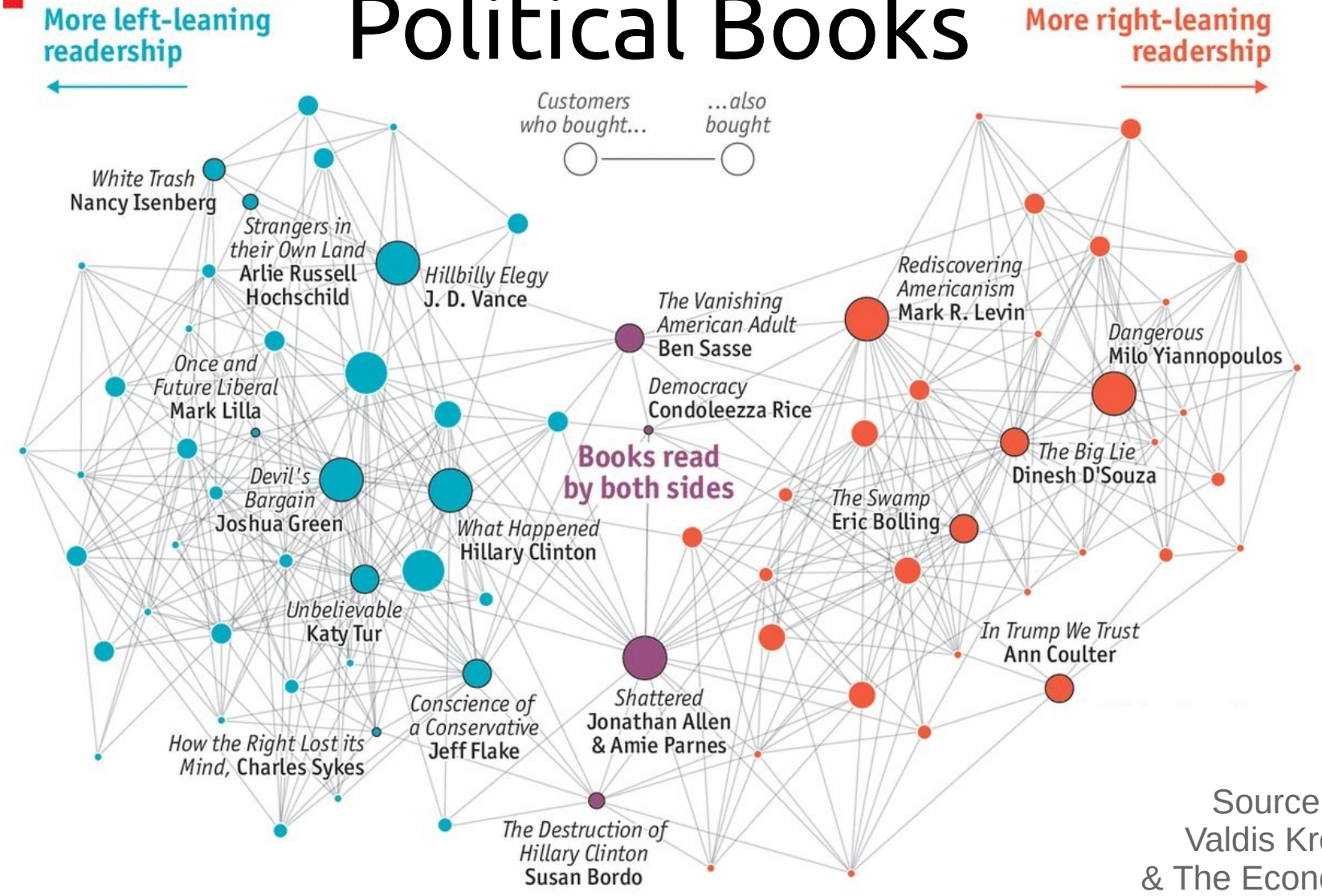
Islamist



Secularist

Weber, I., Garimella, V. R. K., & Batayneh, A. (2013, August). Secular vs. Islamist polarization in egypt on twitter. In Proc. ASONAM (pp. 290-297). ACM.

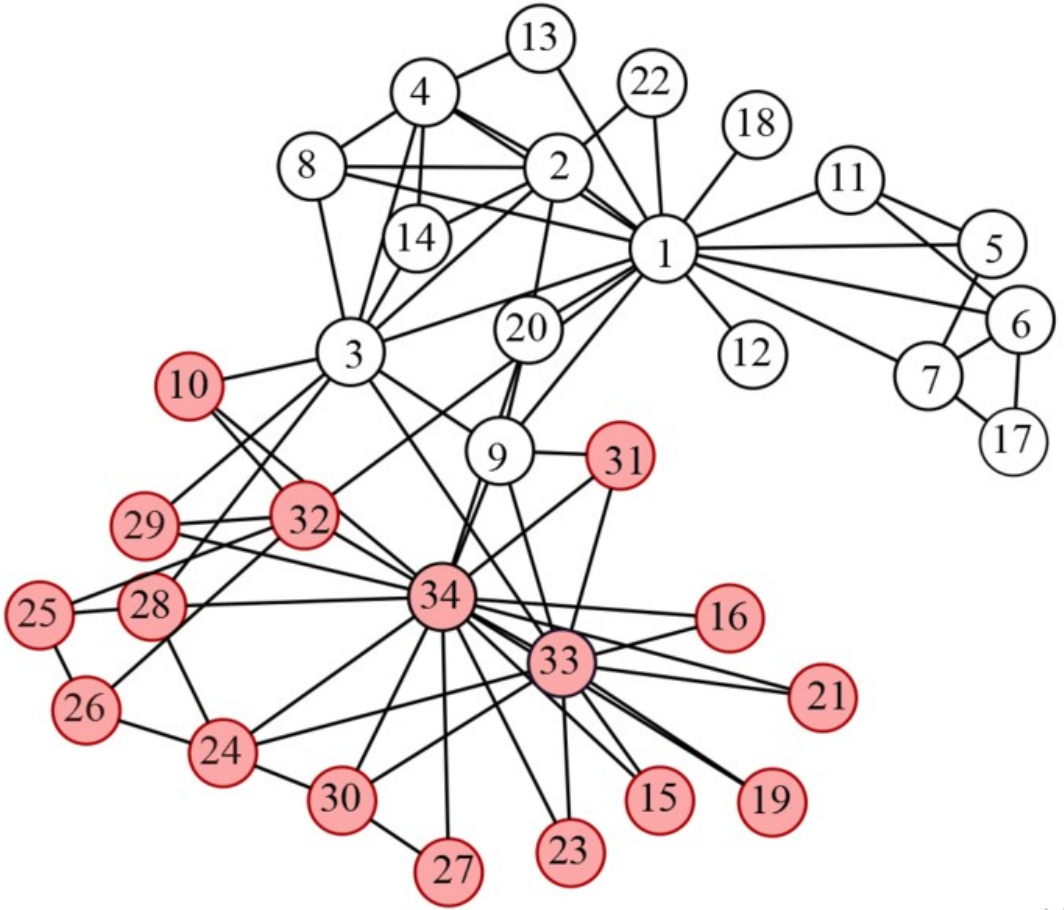
Political Books



Source:
Valdis Kreb
& The Economist

Wayne Zachary's PhD Thesis (1972)

- Studied 34 members of a karate club
- Found 78 links between members who regularly interacted outside the club
- The club splitted in two during the study
- 1=sensei, 34=president



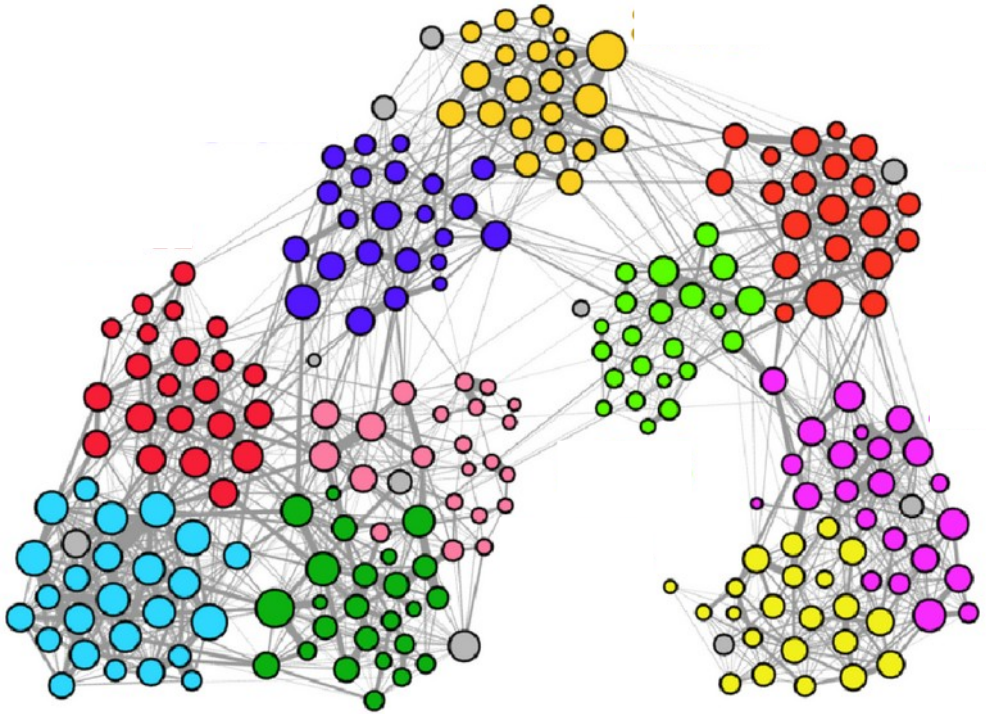
Multiple communities

Primary school contacts

Links connect students who spent more than two minutes face to face

Students wore RF-ID badges hanging on their chest, which have a range of about 1.0-1.5 meters

What do you think the colors represent in this visualization?

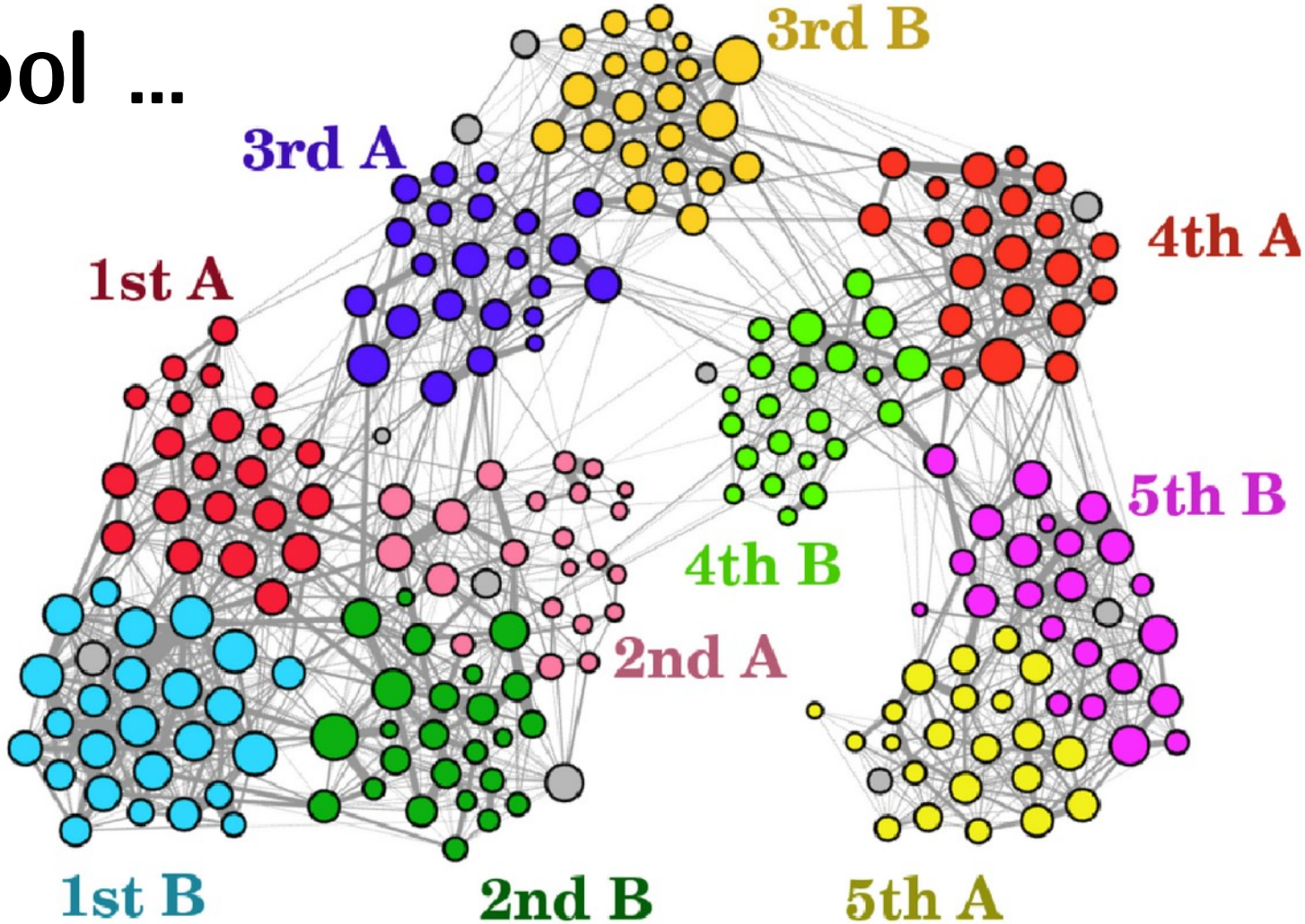


Stehlé, J., et al. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. PloS one, 6(8), e23176.

Primary school ...

Colors represent classes. Note the **hierarchical** clusters

Teachers are shown in gray color. Node sizes are number of connections.



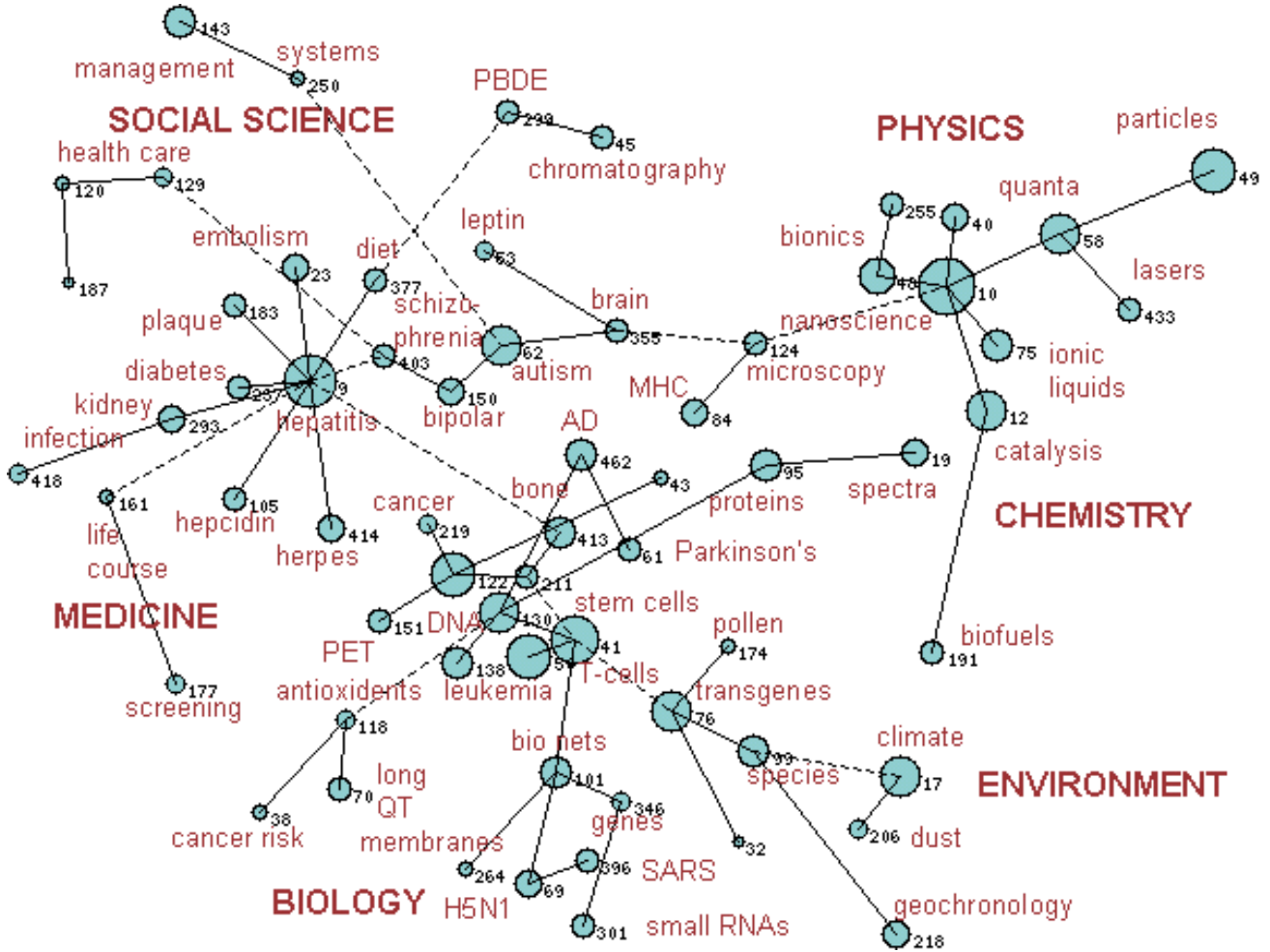
Stehlé, J., et al. (2011).
High-resolution measurements of face-to-face contact patterns in a primary school.
PloS one, 6(8), e23176.

Science

Two topics T1, T2, are connected if there is at least one paper that cites:

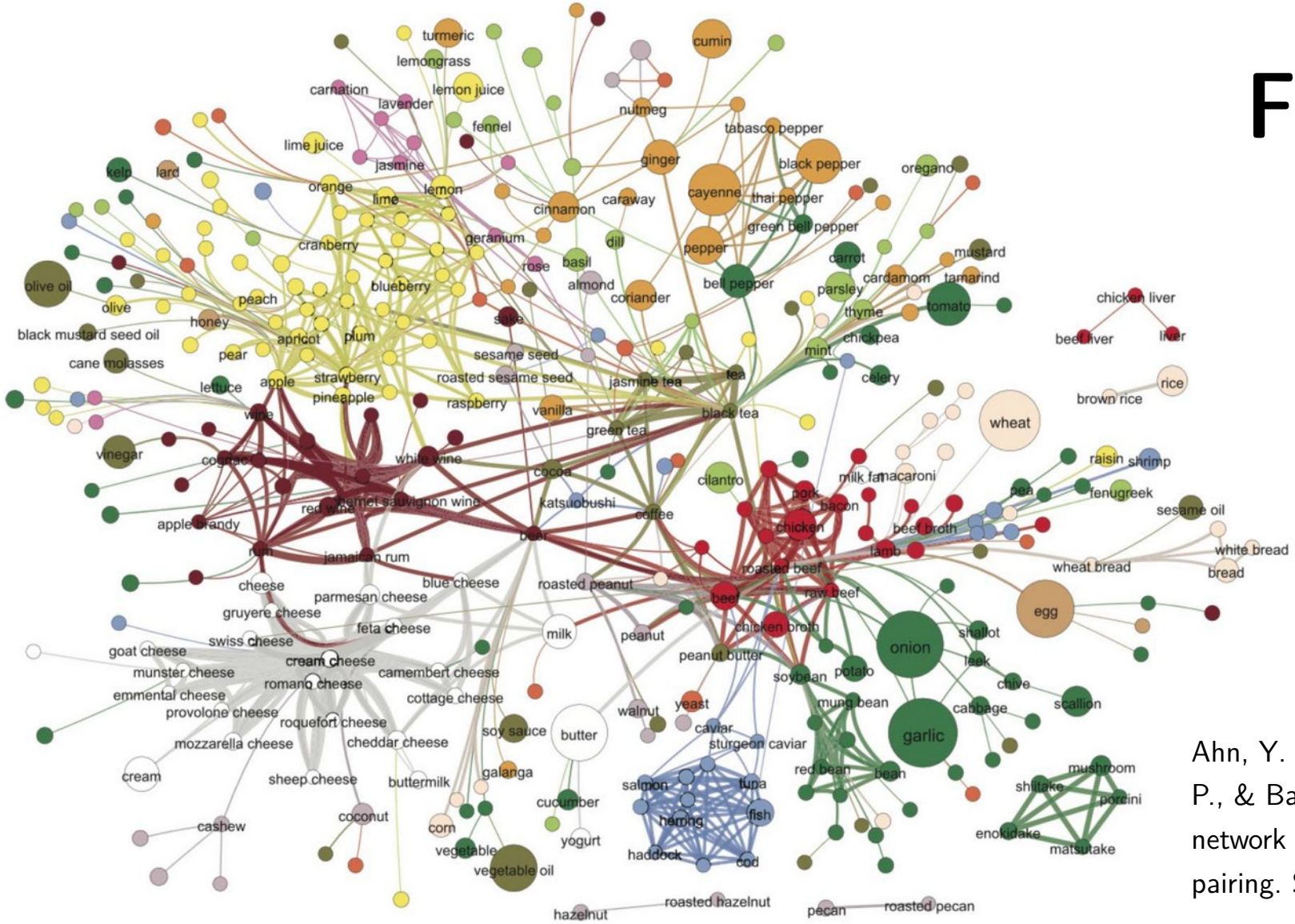
a paper u in T1 and

a paper v in T2.



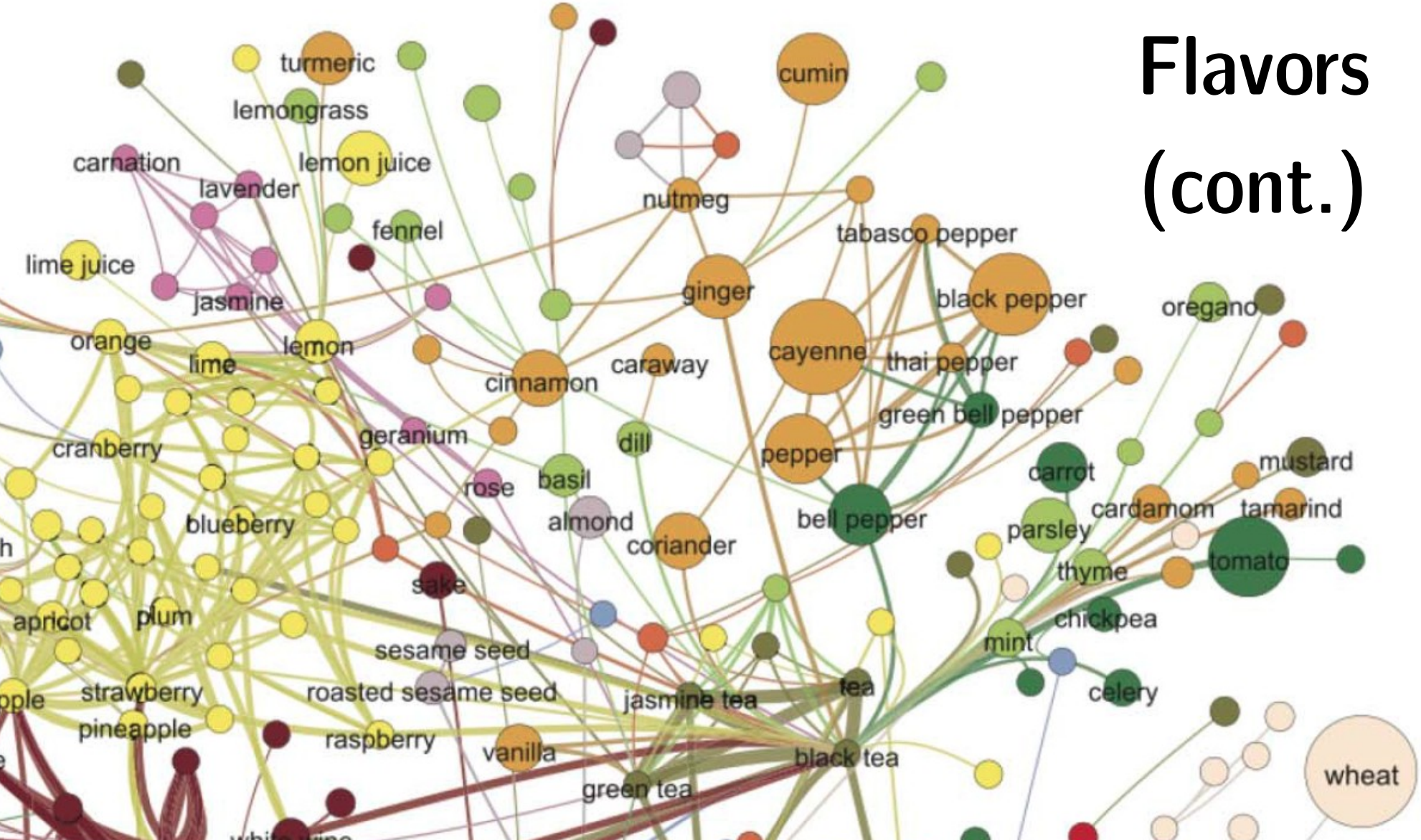
http://archive.sciencewatch.com/dr/rfm/mos/10maymosGLOBAL/

Flavors

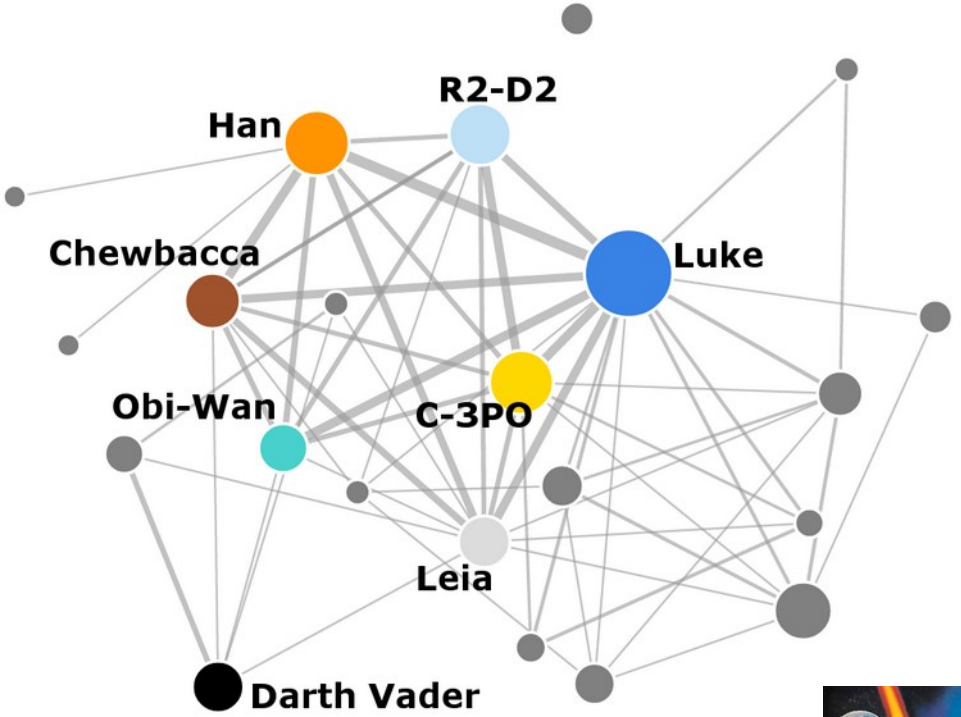


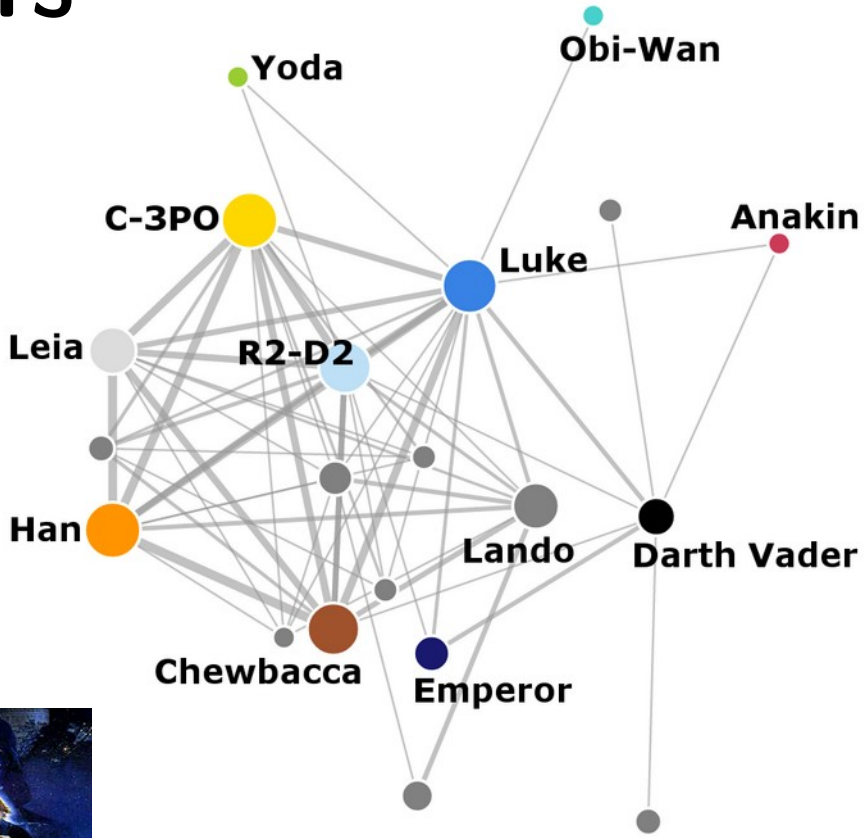
Ahn, Y. Y., Ahnert, S. E., Bagrow, J. P., & Barabási, A. L. (2011). Flavor network and the principles of food pairing. Scientific reports, 1, 196.

Flavors (cont.)



Star Wars

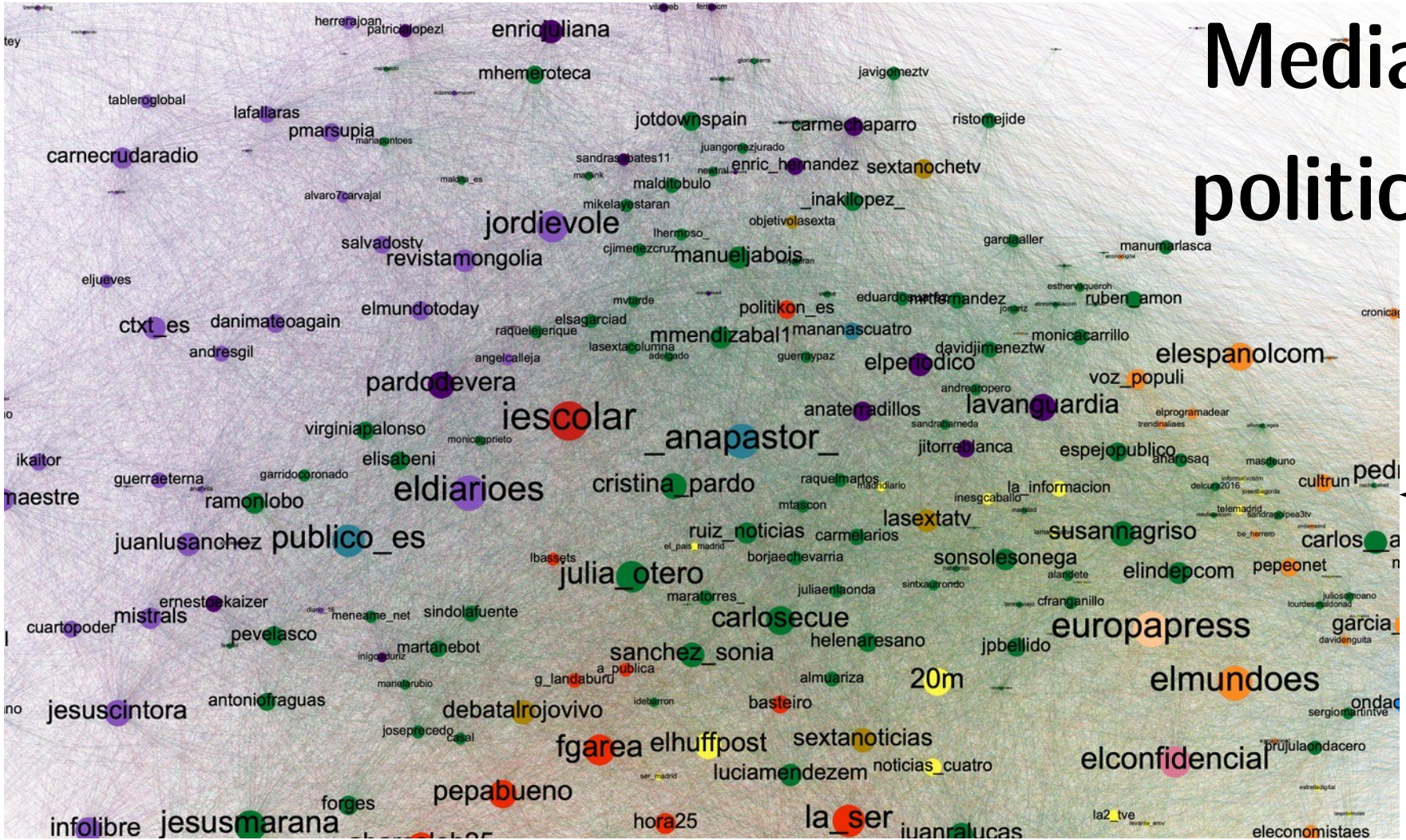

R2-D2
Han
Chewbacca
Luke
Obi-Wan
C-3PO
Leia
Darth Vader


Yoda
Obi-Wan
C-3PO
Anakin
Luke
Leia
R2-D2
Han
Lando
Darth Vader
Chewbacca
Emperor


STAR WARS
EPISODE IV


RETURN OF THE JEDI

Media and politicians?

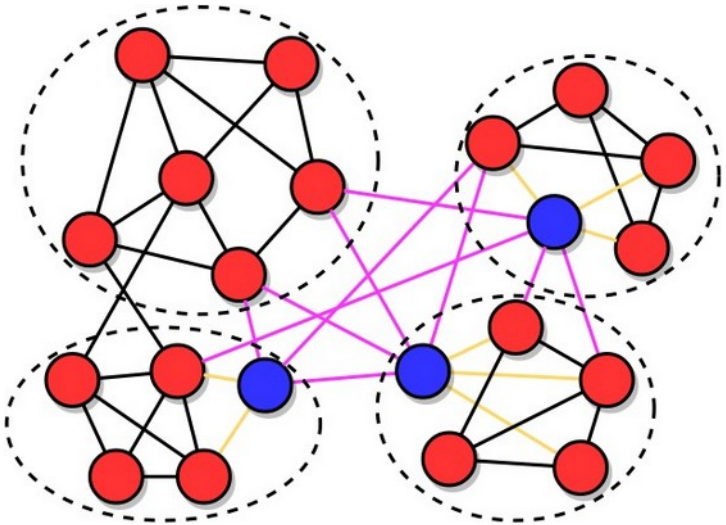


https://twitter.com/jbo/status/1120444347772821504/photo/1

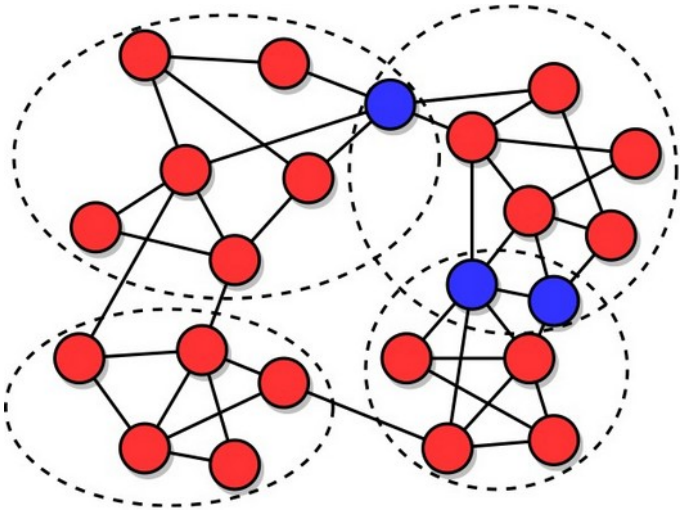
Partitions vs Overlapping communities

Hierarchical communities

Partition vs Overlapping communities



Partition, or *hard* clusters



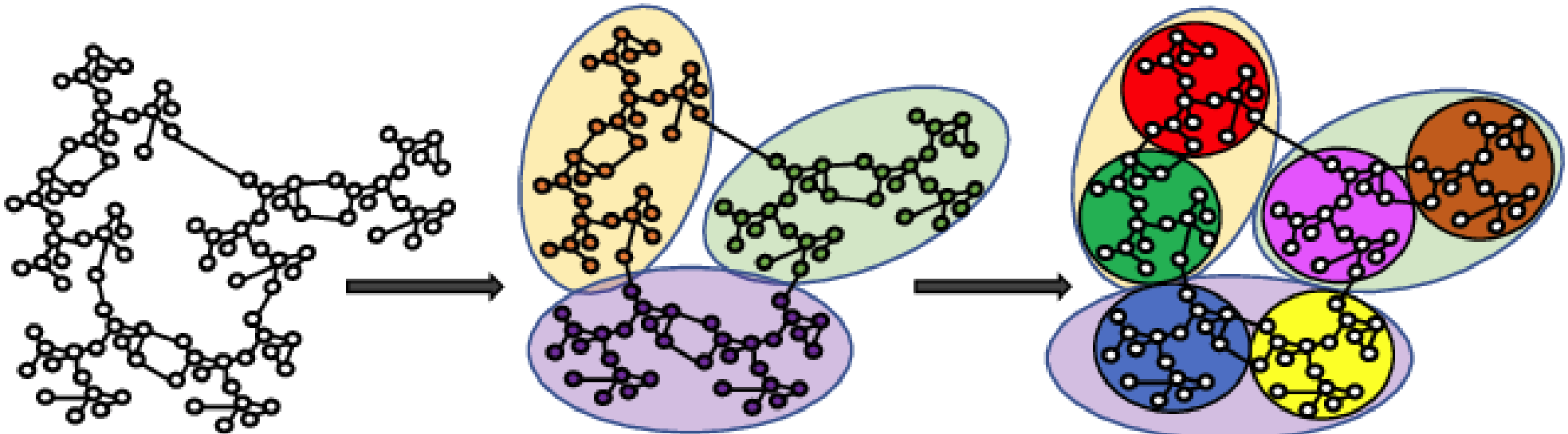
Overlapping communities, or *soft* clusters

What's special about blue nodes?

Blue nodes are in more than one community

Source: Menczer, Fortunato, Davis: A First Course on Networks Science. Cambridge, 2020.

Hierarchical communities



https://www.catalyzex.com/paper/arxiv:2012.05980

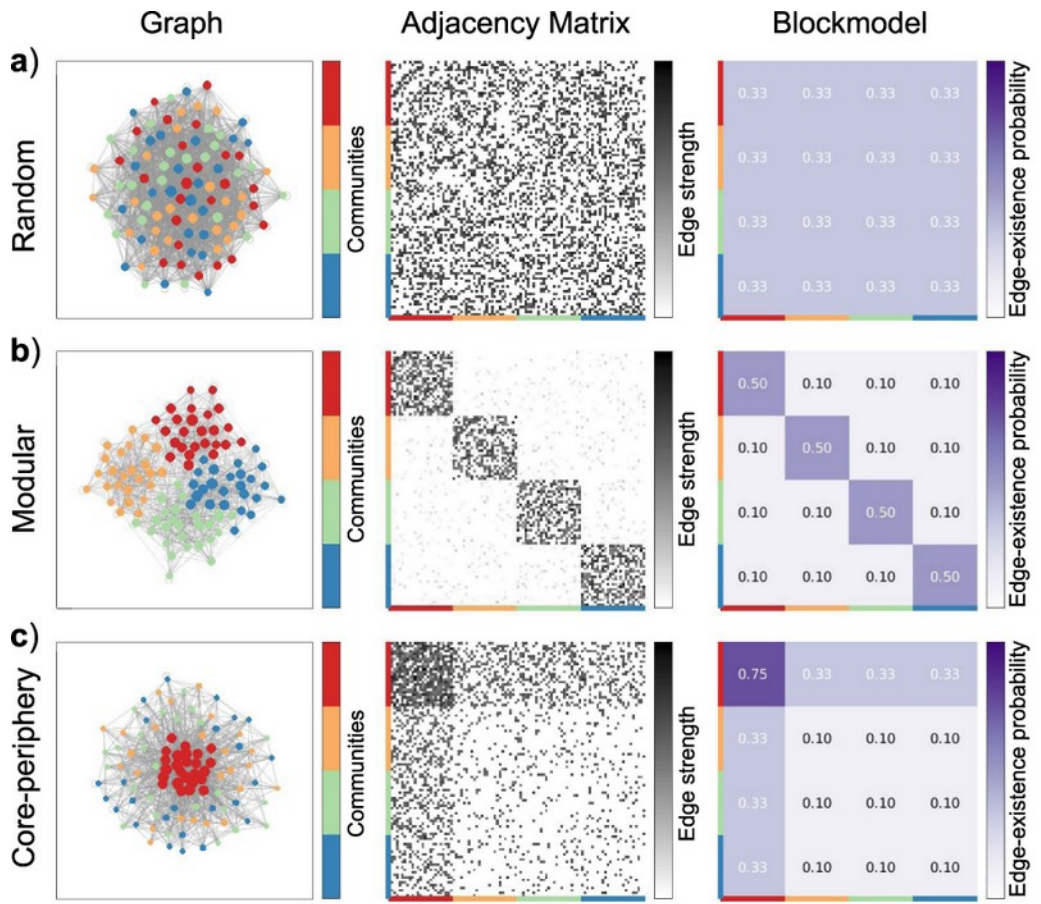
Stochastic block model

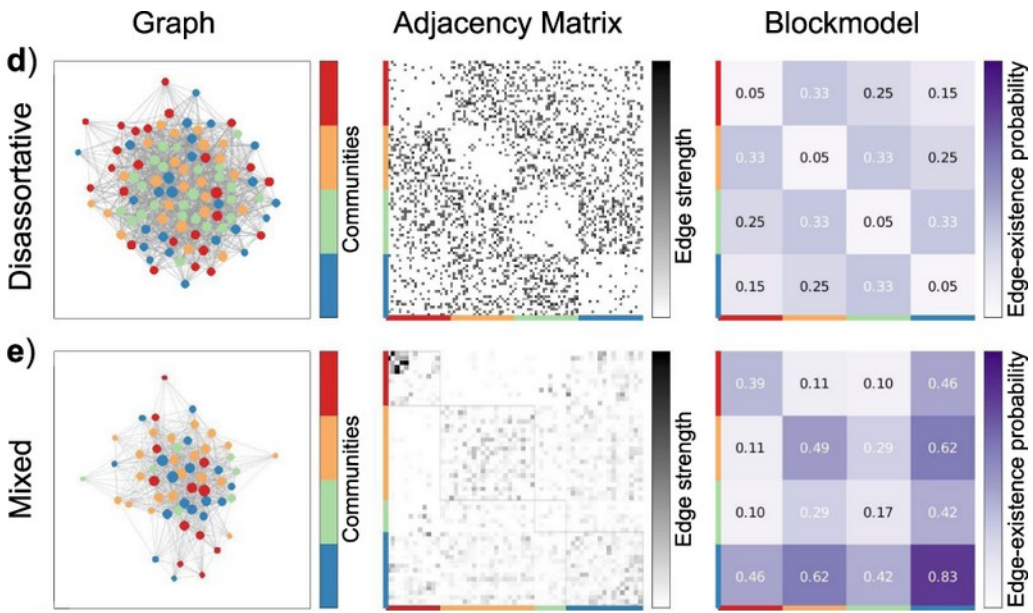
How to generate community structure?

- The **stochastic block model** generates graphs with community structure
  - It can also be used for inference, but we will not see that in this course
- Can be described as a variation of the ER model in which:
  - There are m groups
  - Link probability scalar p becomes an m x m matrix P that contains in position (i,j) the probability of a link between a node in group i and a node in group j

Summary

Examples of stochastic block model





$$Pr((i,j) \in E) = P_{z_i,z_j}$$

z_i and z_j are the communities to which nodes i and j belong to

Faskowitz, J., Yan, X., Zuo, X. N., & Sporns, O. (2018). Weighted stochastic block models of the human connectome across the life span. Scientific reports, 8(1), 1-16.

Things to remember

- Many networks have community structure
- Sometimes it's:
  - One dense sub-graph
  - Two communities (polarization)
  - Multiple communities
- Partitions vs overlapping communities
- Hierarchical communities