



Network analysis with R

Shuning Lu (Ph.D.)
Assistant Professor
Department of Communication
North Dakota State University

Email: shuning.lu@ndsu.edu
Twitter: [@shuning_lu](https://twitter.com/shuning_lu)



Today's goals

Morning:

- Overview of network theory
- Network analysis tutorial

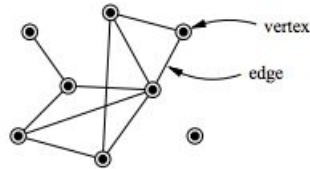
Afternoon:

- Group activity & report

What is network?

A structure that represents a set of objects and the relationships between them.

Two components: node and link





What can network analysis do?

Identify the most influential objects

Define the organizing structures of the group of objects

Capture network processes (formation, change, evolution)

Understand network outcomes

...



Node (vertex, actor, etc.)

Person, institution, country, website, message, word...



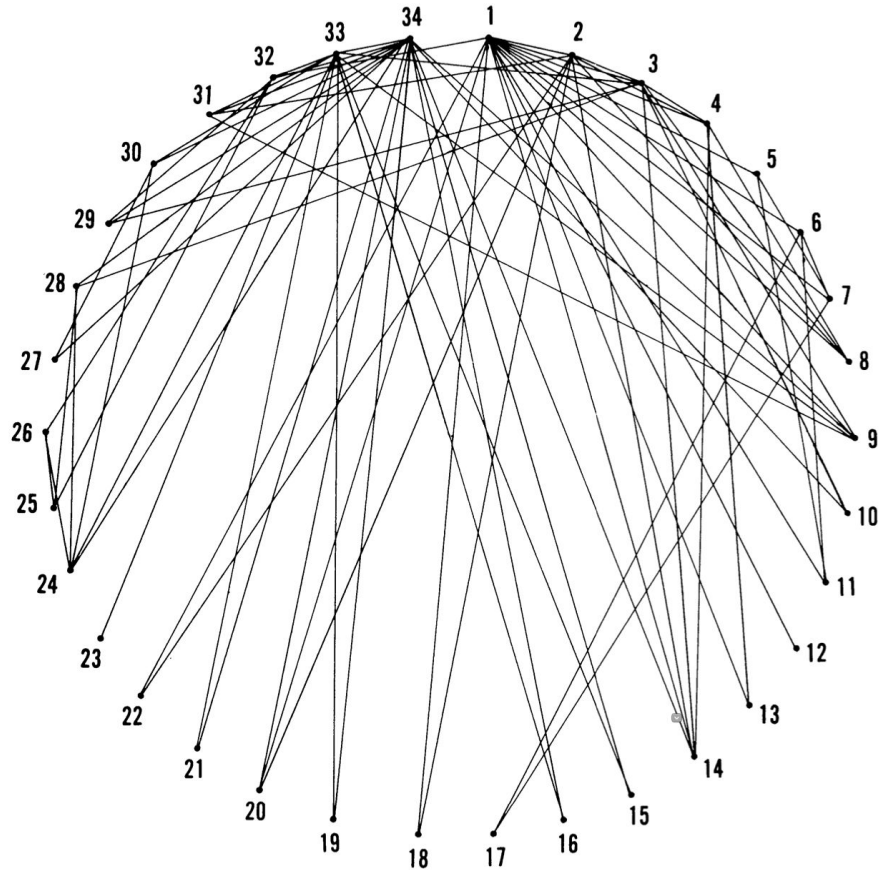
Edge (tie, link, relation, etc.)

Affinity (alliance, friendship)

Transaction (information flow, financial exchange)

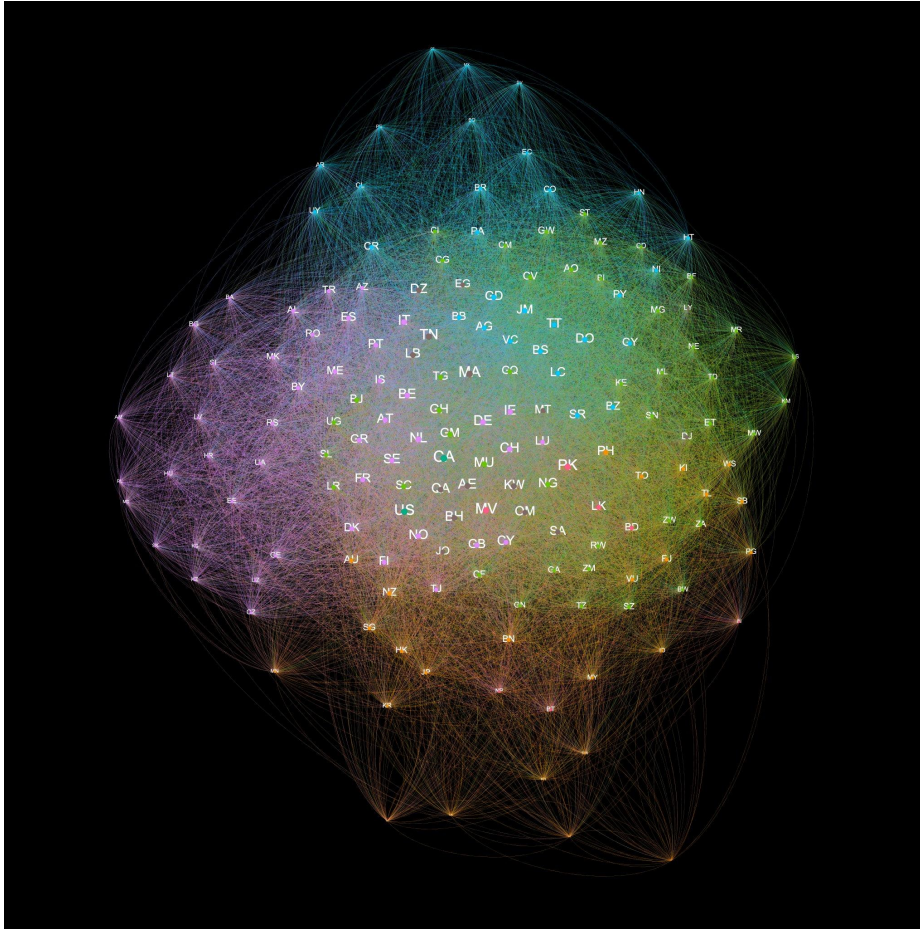
Semantic (word co-occurrence, cognitive association)

FIGURE 1
Social Network Model of Relationships in the Karate Club



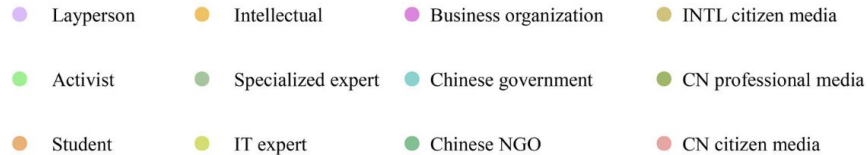
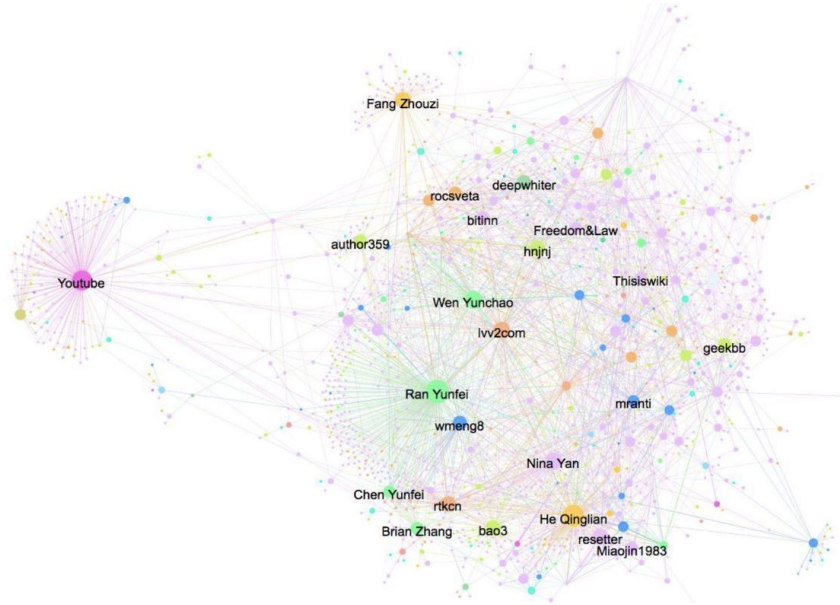
Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4), 452-473.

Facebook connectedness at country-level



Wang, R., **Lu, S.** & Prati, A. (2021). *Social butterflies: an exploration of the global Facebook connectedness index*. SICSS-Rutgers 2021 team project

Twitter discussion network on environmental crisis in China



Lu, S., Chen, W., Li, X., & Zheng, P. (2018). The Chinese smog crisis as media event: Examining Twitter discussion of the documentary Under the Dome. *Policy & Internet*, 10(4), 483-508.

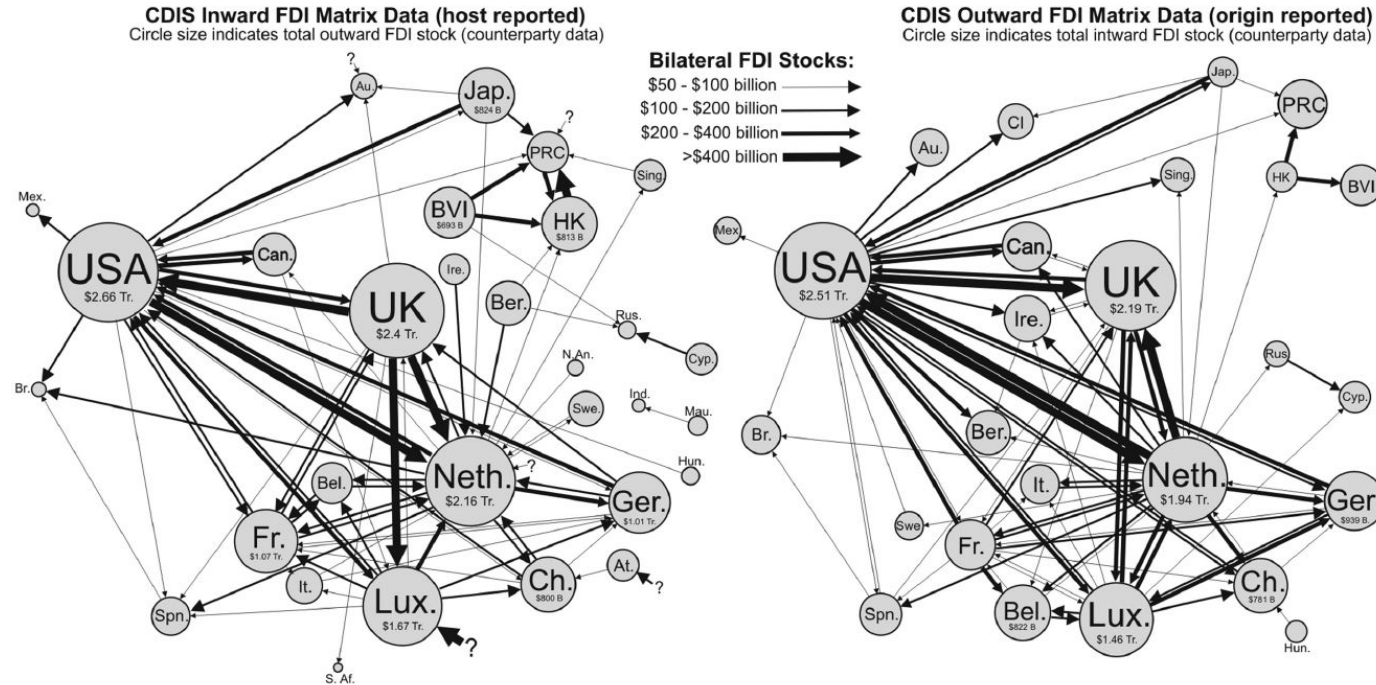
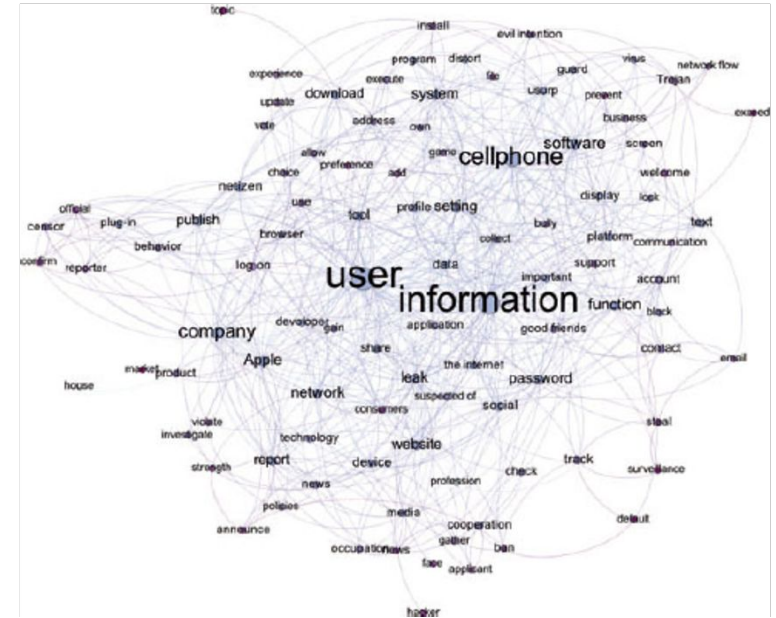
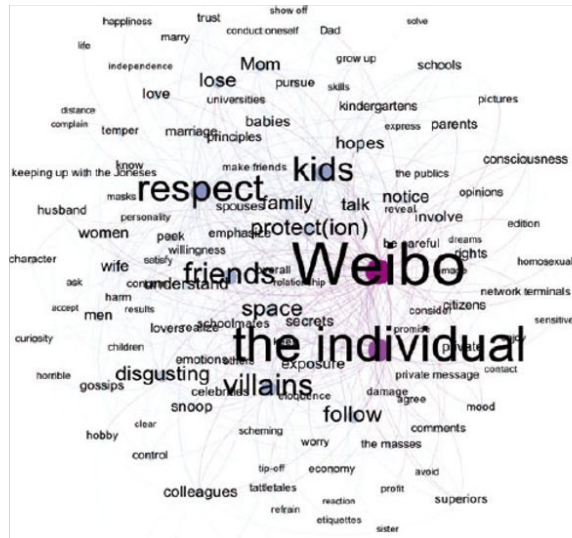


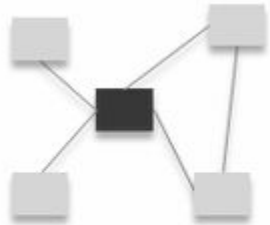
Figure 1. FDI Stocks > \$50 billion, YE-2010.
Source: Adapted from Haberly and Wójcik, 2014.

Figure 2 Group 1 the personal sphere of privacy on Sina Weibo. **Figure 3** Group 2 the technological and media sphere of privacy on Sina Weibo.



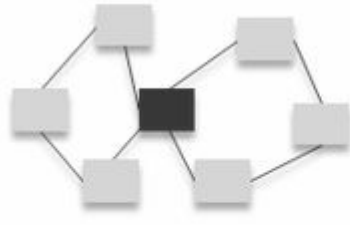
Yuan, E. J., Feng, M., & Danowski, J. A. (2013). "Privacy" in semantic networks on Chinese social media: The case of Sina Weibo. *Journal of Communication*, 63(6), 1011-1031.

Node centrality



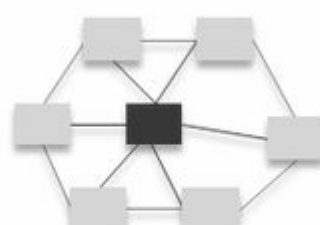
Degree

How many nodes can a given node reach directly?



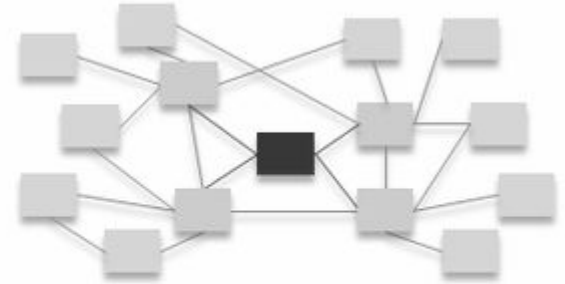
Betweenness

How likely is a node to be the most direct route between two nodes in the network?



Closeness

How fast can a node reach other nodes in the network?



Eigenvector

How well is a node connected to other well-connected nodes?



Why is node centrality important?

Degree centrality: Popular objects (influentials, key words)

Betweenness centrality: Authority over disparate networks (broker)

Closeness centrality: Efficient broadcaster (fast information flow)

Eigenvector centrality: Influentials over networks through central nodes of disparate networks



Edge characteristics

Link or no link

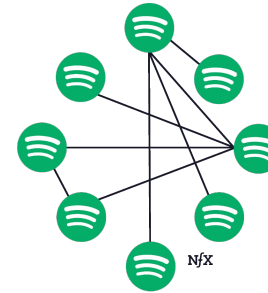
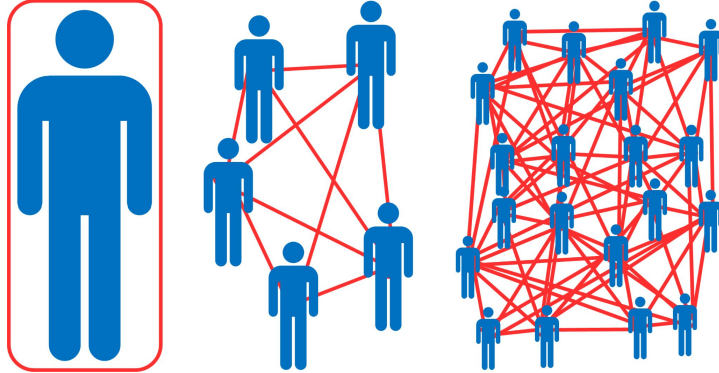
Directed (in, out) vs. non-directed

Valenced vs. non-valenced

Weighted edge (tie strength)

Multiplexity (more than one link type)

Network characteristics: size and density



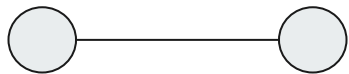
Low Density



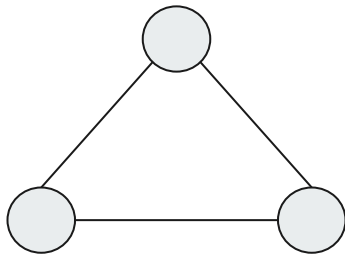
High Density



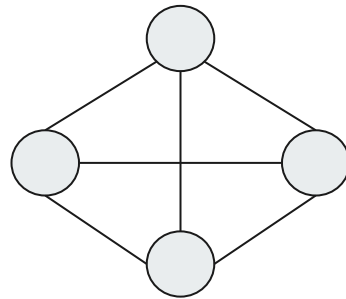
Clique



2-clique

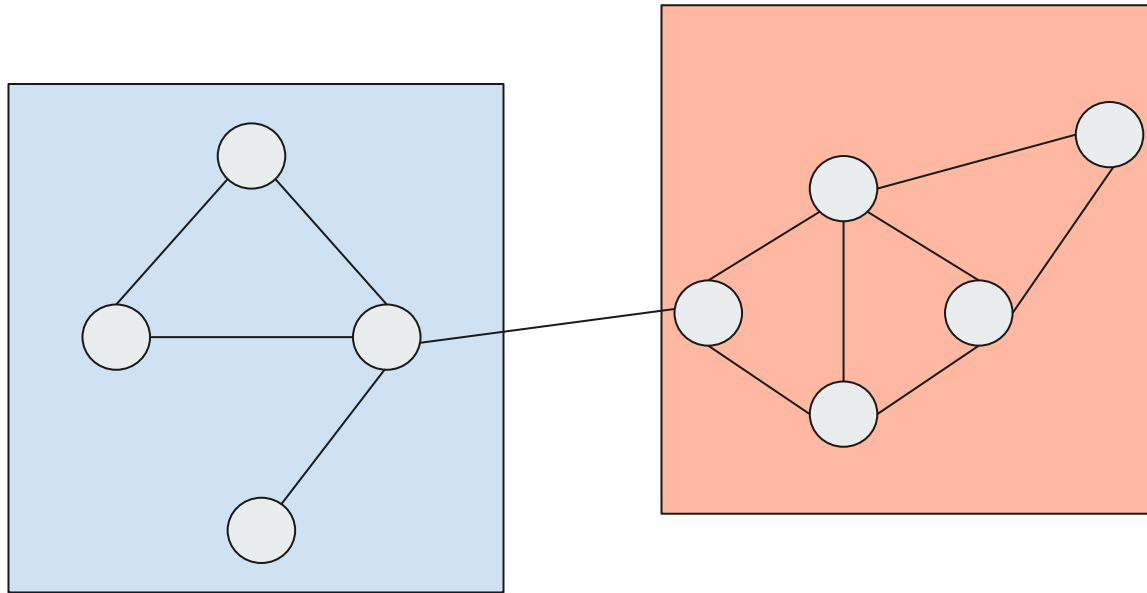


3-clique

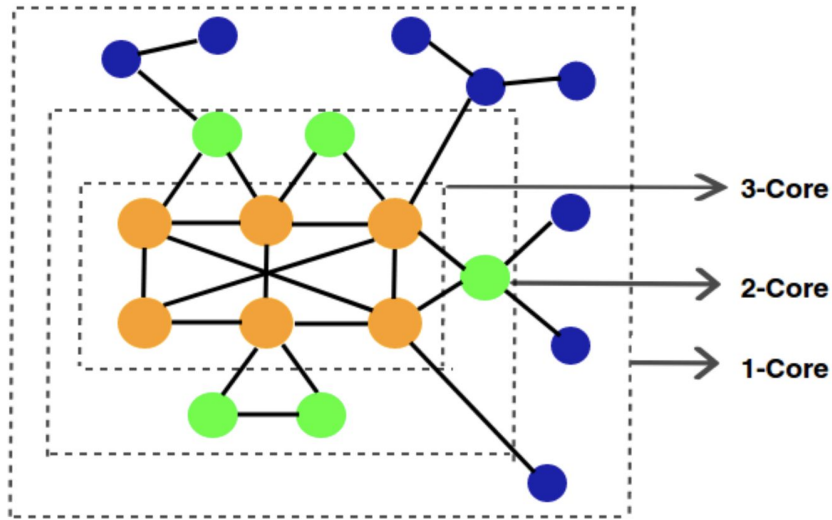


4-clique

Component/community detection

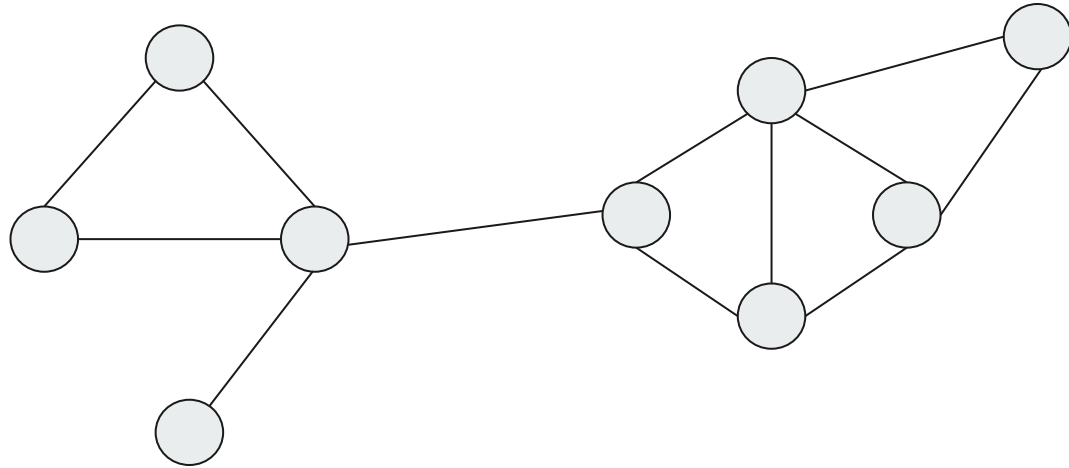
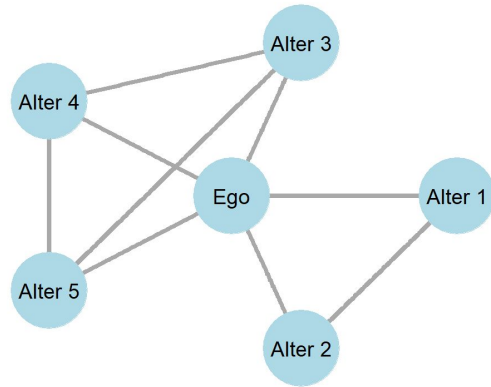


K-core decomposition

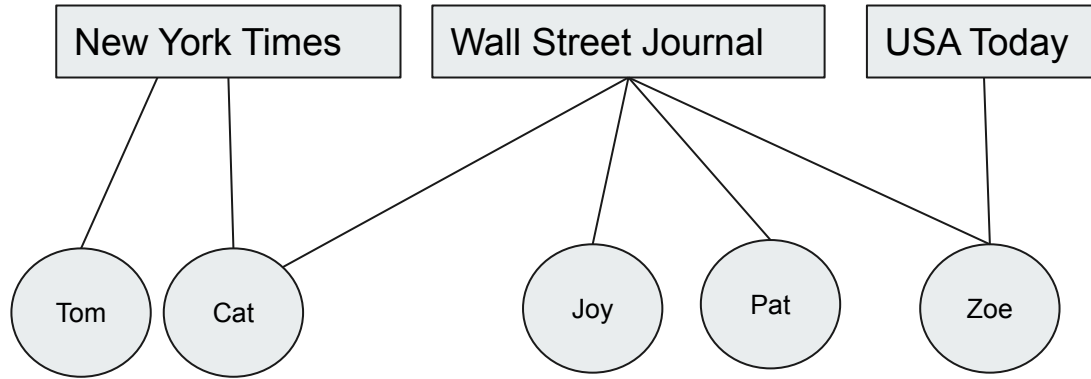


Illustrative example of the K-Core algorithm.

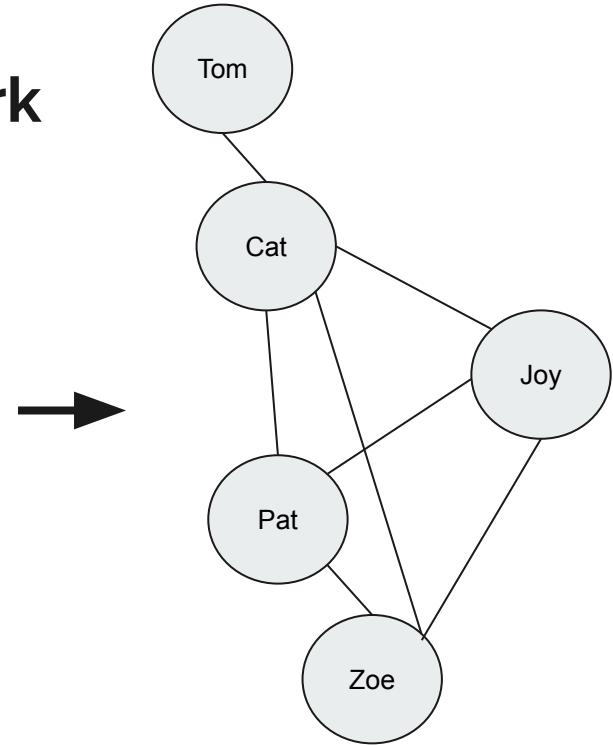
Egocentric network vs. complete network



One and two-mode/bipartite network



a fictitious bipartite network of freelancers and news organizations



One mode network of freelancers



What questions you want to solve?

Based on your disciplinary background and research, identify:

- Actors of interest to you?
- What kinds of ties that bind them?
- What are some attributes we could collect about the actors?
- What are the characteristics of the ties?

Discuss this with your tablemates

Data formats

adjacency_matrix

A	.	1	1	1	.	1
B	1	.	.	1	1	.	1	.	.	.
C	1	.	.	1	.	1
D	1	1	1	.	1	1	1	.	.	.
E	.	1	.	1	.	.	1	.	.	.
F	1	.	1	1	.	.	1	1	.	.
G	.	1	.	1	1	1	.	1	.	.
H	1	1	.	1	.
I	1	.	1
J	1	.

Edge list

A--B

A--C

A--D

A--F

B--D

B--E

B--G

C--D

C--F

D--E

D--F

D--G

E--G

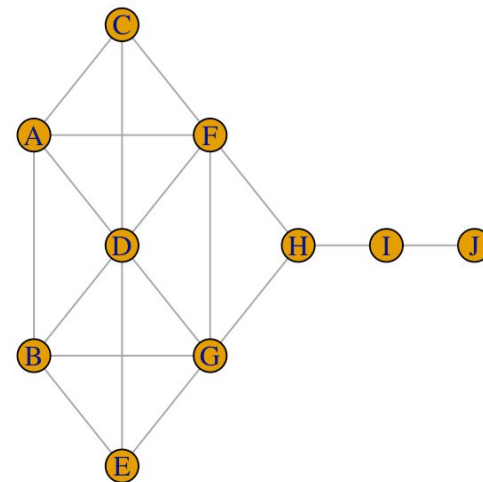
F--G

F--H

G--H

H--I

I--J



a fictitious social network with ten actors

Source: igraphdata library, data(kite)



Network analysis workflow

Data collection <-> Boundary setting

Set up R environment

Import data

Prepare network object

Run descriptive statistics

Plot data

Fine-tuning

*advanced: statistical modeling of network processes



Tool time: Network analysis with R

Code along!



Group activity Day 4

Summary

An open-ended group exercise to create a network graph, report descriptive statistics, and suggest a direction for future research.

Activity

- Split into small groups and select person(s) to take notes and report group process/results.
- 13:40-13:50: brainstorm potential research ideas and select one to pursue
- 13:50-14:00: discuss data collection strategy (collect by yourself or use/combine existing network data)
- 14:00-15:20: (collect) and analyze network data to answer 1-2 key research questions
- 15:20-15:30: reflect on the strengths/limitations of what you have completed and ways to address
- 15:30-16:00: come back together as a large group and discuss projects at the end of the day