Introduction to Network Analysis using R

06/13/2022

Bomi Lee

(University of Kentucky)

What we need today…

https://github.com/bomibonnie/PolNetIntroNetworkAnalysis2022

- atop_sample2.csv
- edgeList.csv
- nodeList.csv
- pndata.xlsx
- rivdefme1980.csv

Outline

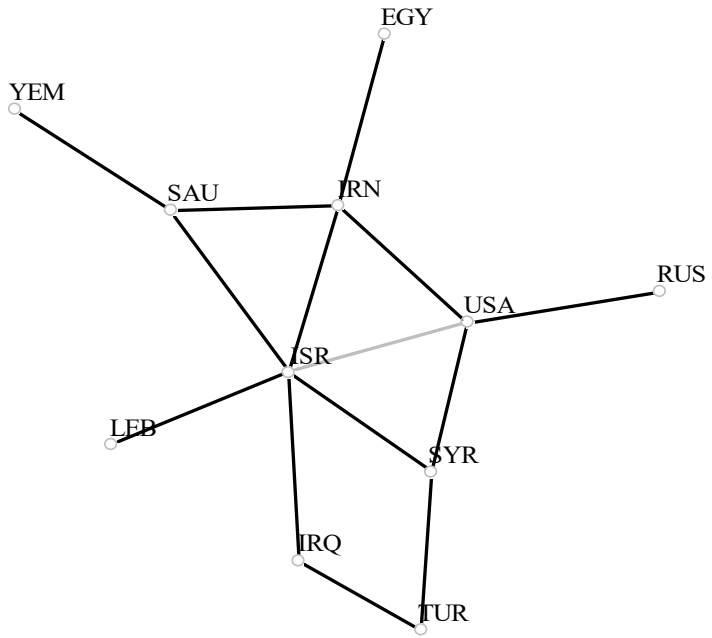
- Basic terms
- Directed/undirected, weighted/unweighted networks
- R code for igraph/network objects using *igraph* and *statnet* packages

------- Break: troubleshooting----------

- Network statistics and centrality
- R code for dyad/triad census, degree, eigenvector, betweenness, and closeness centrality
- Negative-tie networks
- R code for negative centrality using *signnet* package
- Q&A

Terms

A **network** is a collection of **nodes** joined by **edges.**



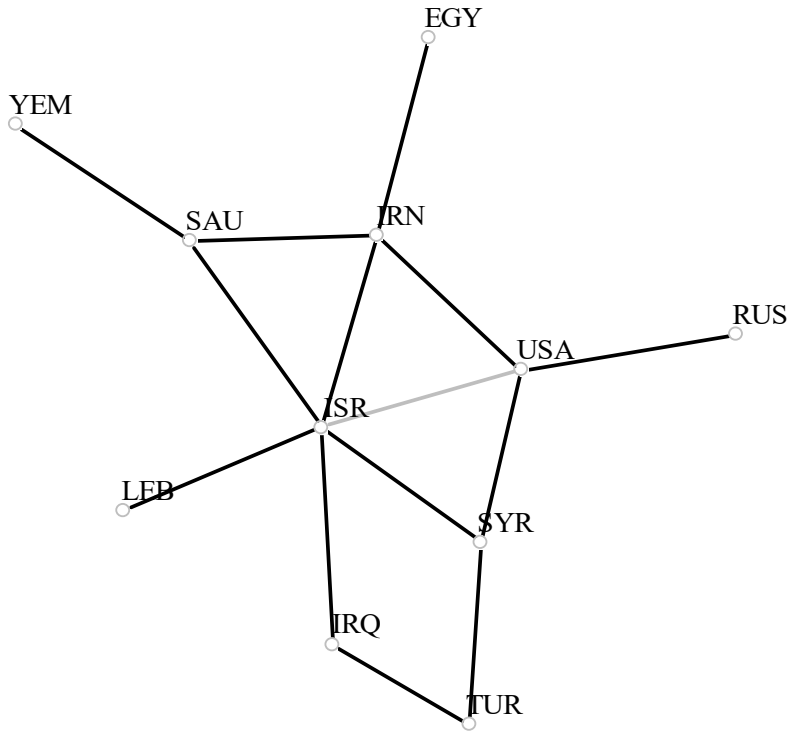
Social network perspectives

Wasserman and Faust (1994, 4)

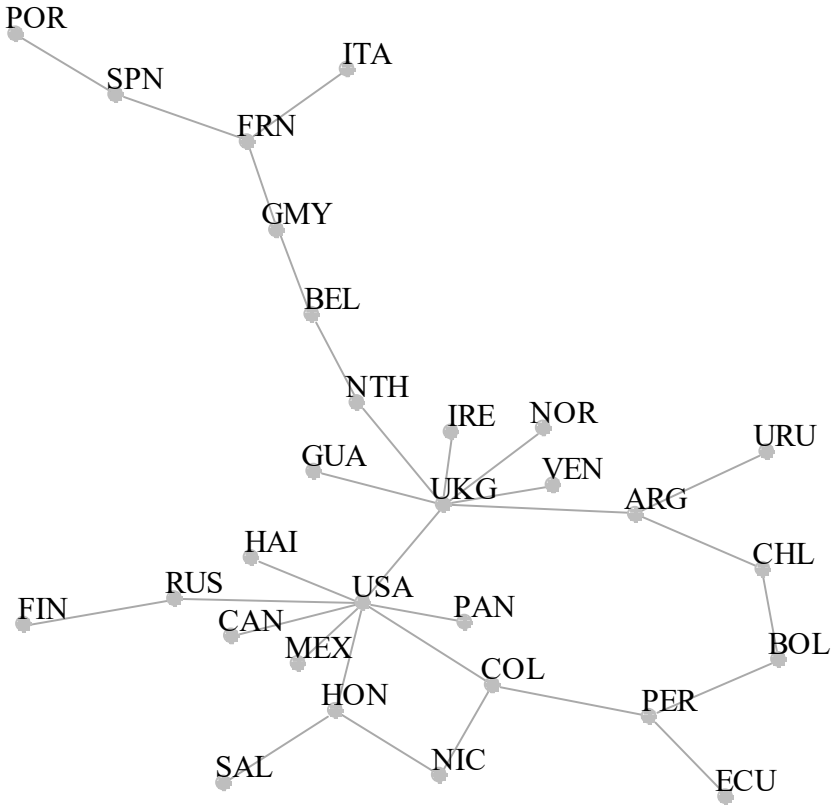
- Social network analysis is based on an assumption of the importance of relationships among interacting units.
- Relations, defined by linkages among units are a fundamental component of network theories.
  - Actors and their actions are viewed as interdependent rather than independent.
  - Relational ties (linkages) between actors are channels for transfer or "flow" of resources.
  - Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action.
  - Network models conceptualize structure as lasting patterns of relations among actors.

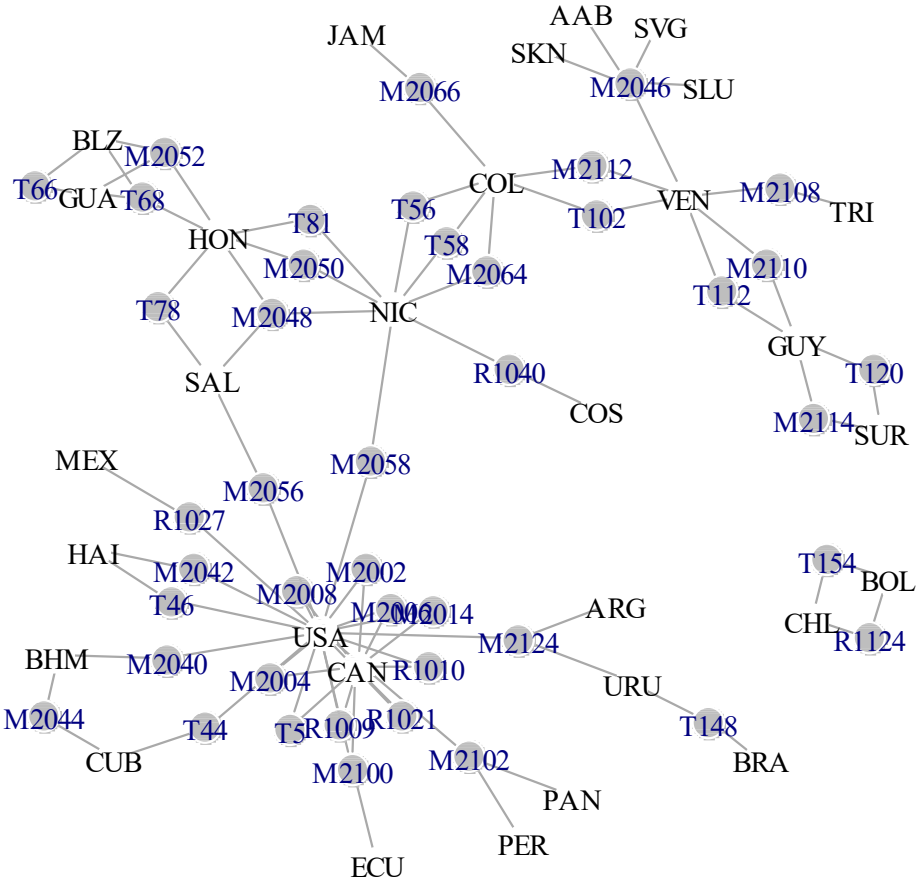
Terms

- Node: unit we are examining
  - Could be countries, people, cities, etc.
  - Synonyms: vertex, actors
- Edge: connection between two nodes



Look at the nodes!

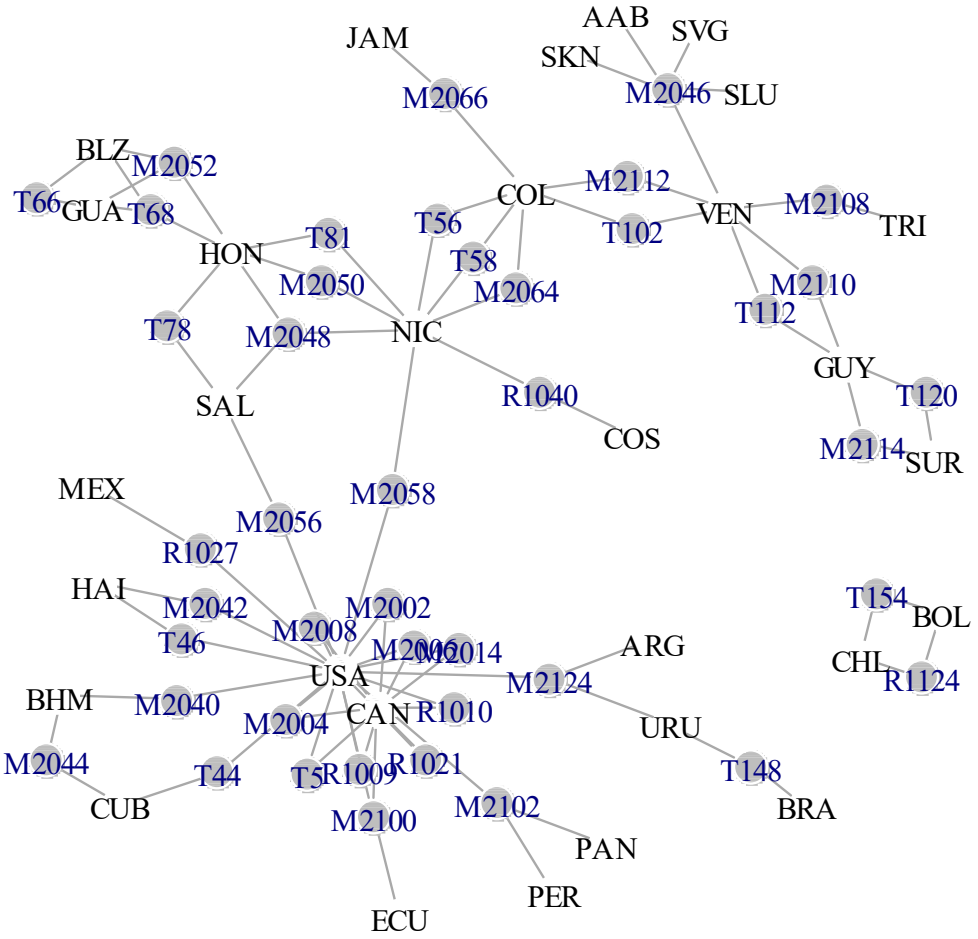




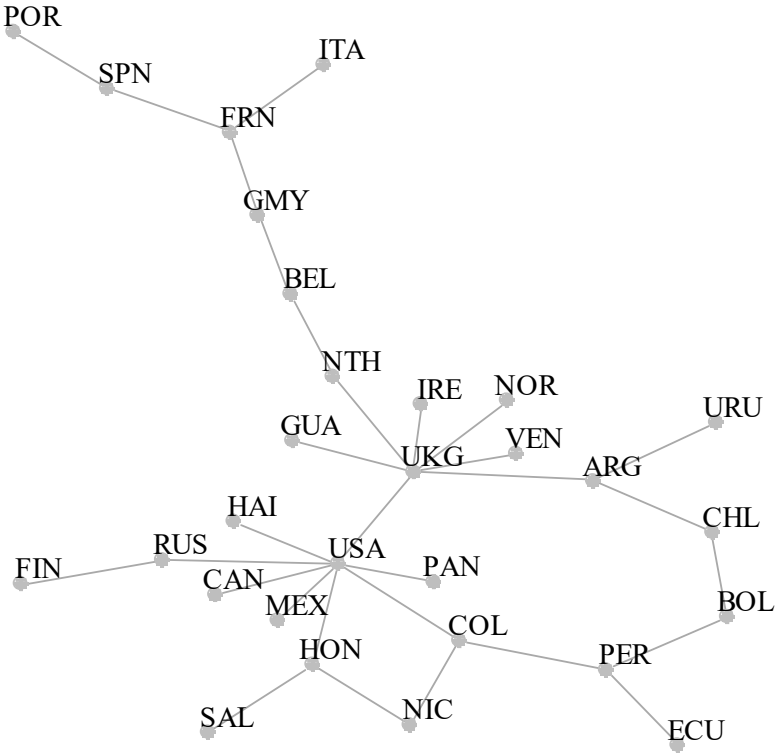
Terms

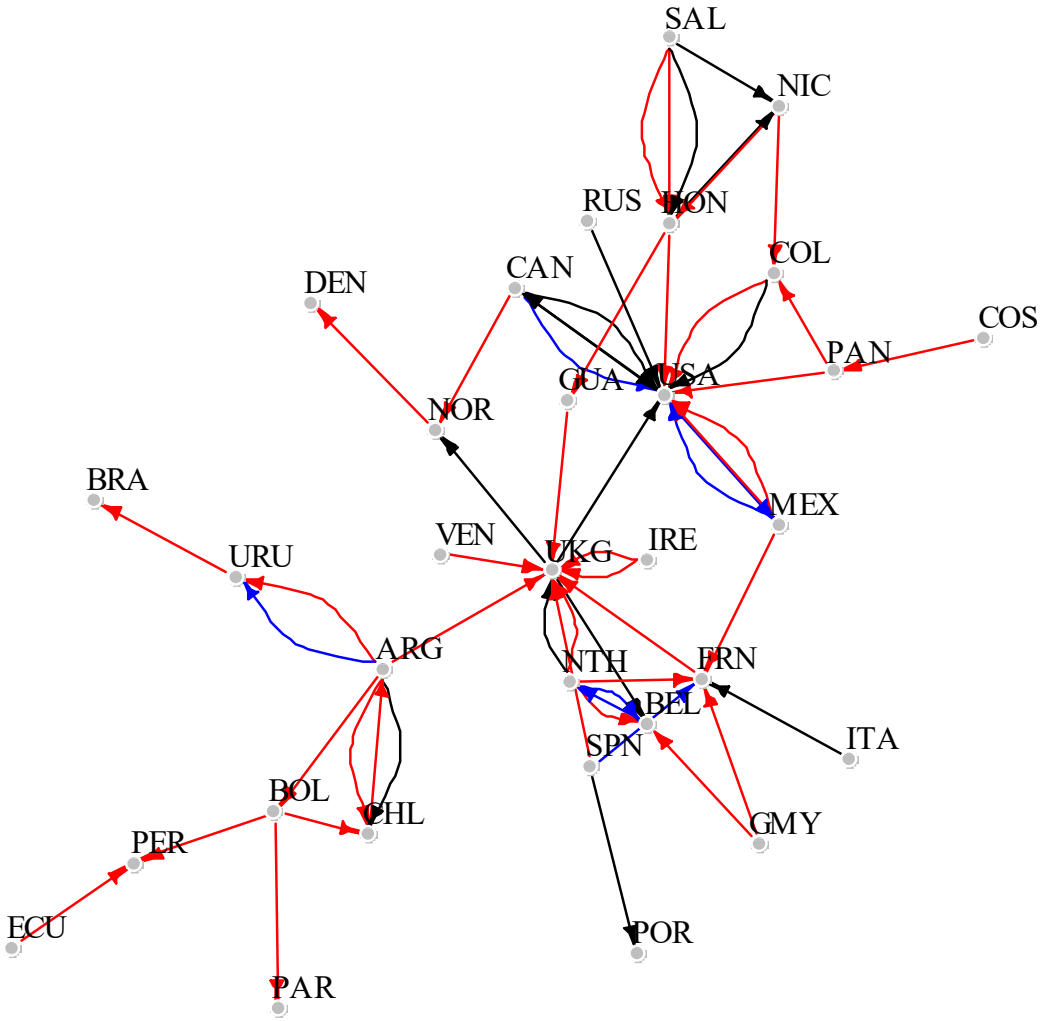
- One-mode networks: all actors are tied to one another according to one relation (e.g., friendship)

- Two-mode networks: actors are tied to particular events



Look at the edges!





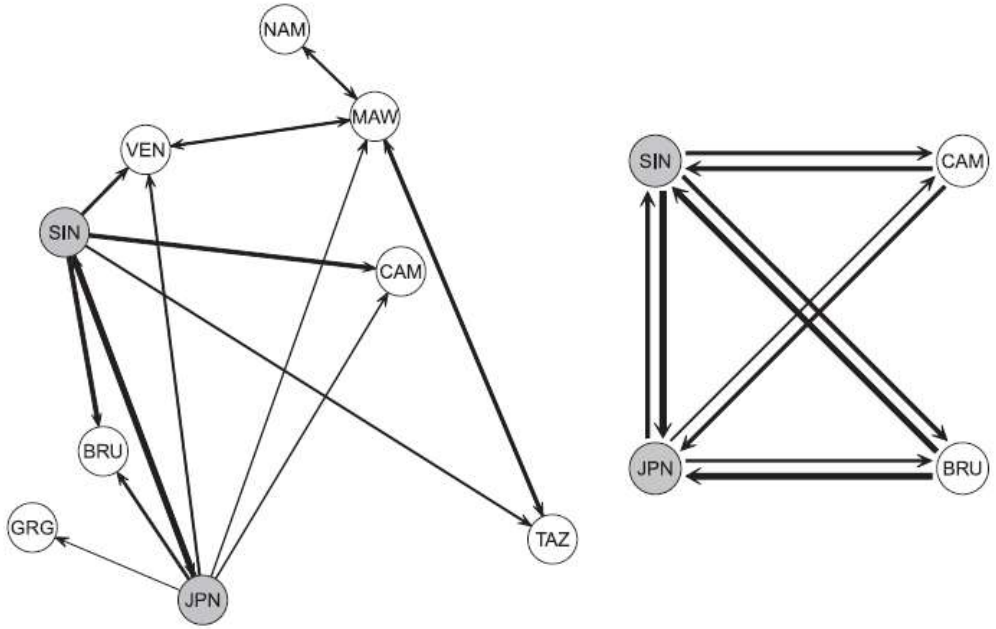
Terms

Edge

- Binary or valued?

Kinne (2012)

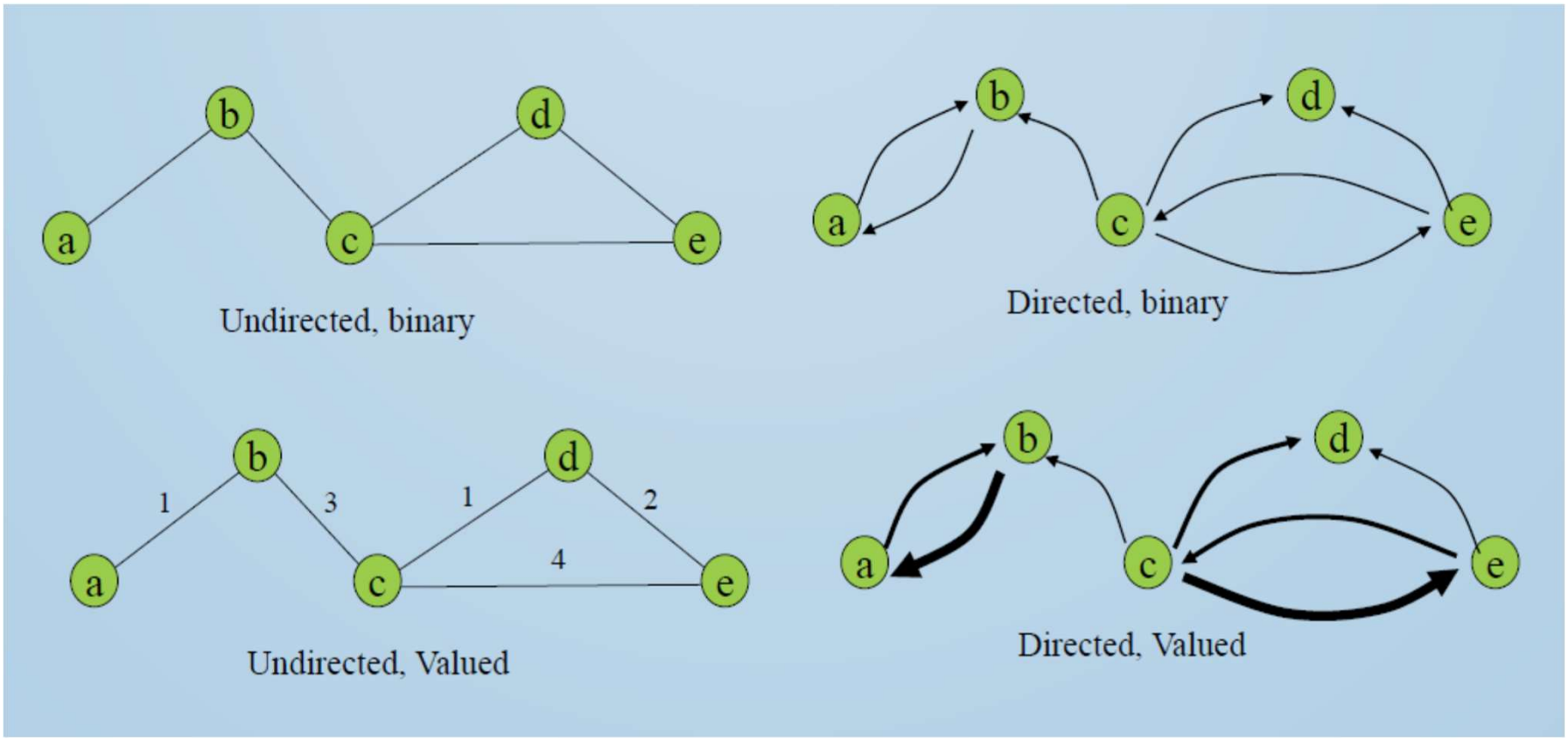
FIGURE 1 A Trade Network with Directed and Weighted Ties



(a) Subgraph of global trade network (b) Directed ties in detail

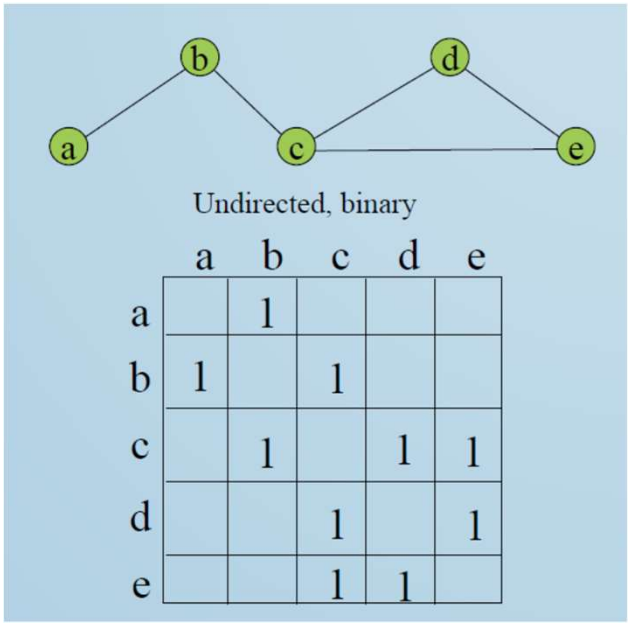
Note: Line widths indicate strength of trade tie. Node positions are arbitrary. In (a), Singapore and Japan's incoming ties are suppressed for clarity of presentation. Graph (b) shows directed ties of four countries in detail. Graph based on trade data from year 2000. See the online appendix for adjacency matrix and exact tie weights.

Relationships?



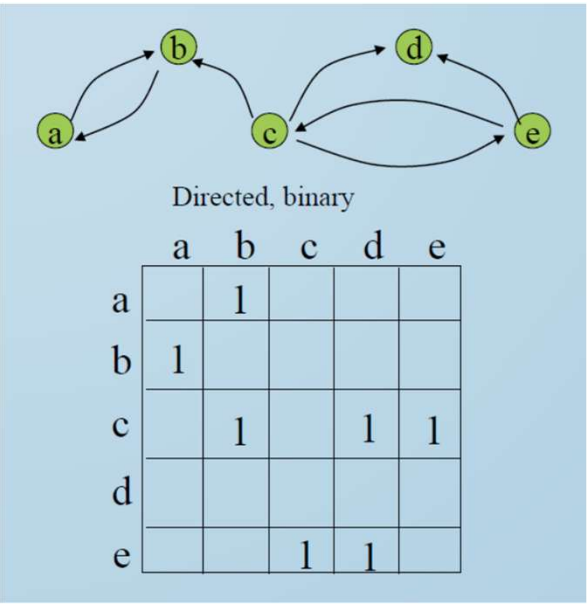
Data: adjacency matrix vs edge list

- Adjacency matrix- values show connection between nodes
  - Often omit 0, just care about 1 if binary



Undirected, binary

	a	b	c	d	e
a		1			
b	1		1		
c		1		1	1
d			1		1
e			1	1	

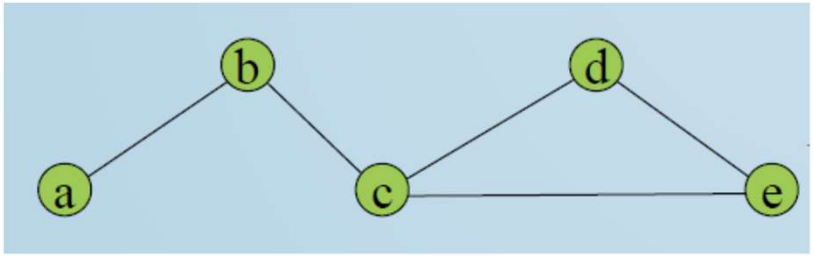


Directed, binary

	a	b	c	d	e
a		1			
b	1				
c		1		1	1
d					
e			1	1	

Edgelist

- 2 columns- node 1, node 2
  - Sometimes 3rd column for weight if not binary
  - Missing edges not included (usually)



Edge List

a b
b a
b c
c b
c d
c e
d c
d e
e c
e d

R code

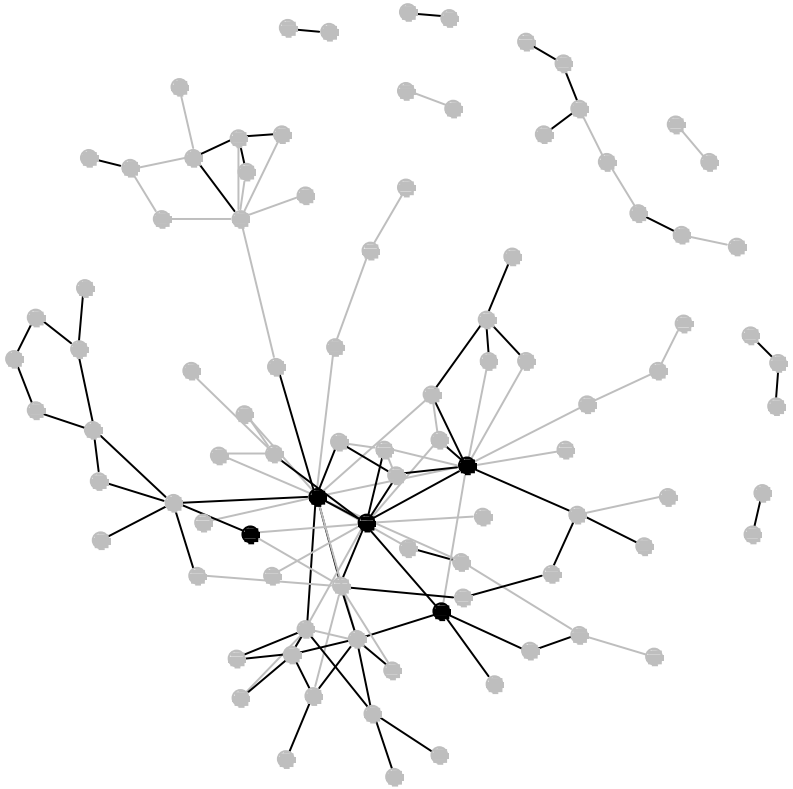
Social network perspectives

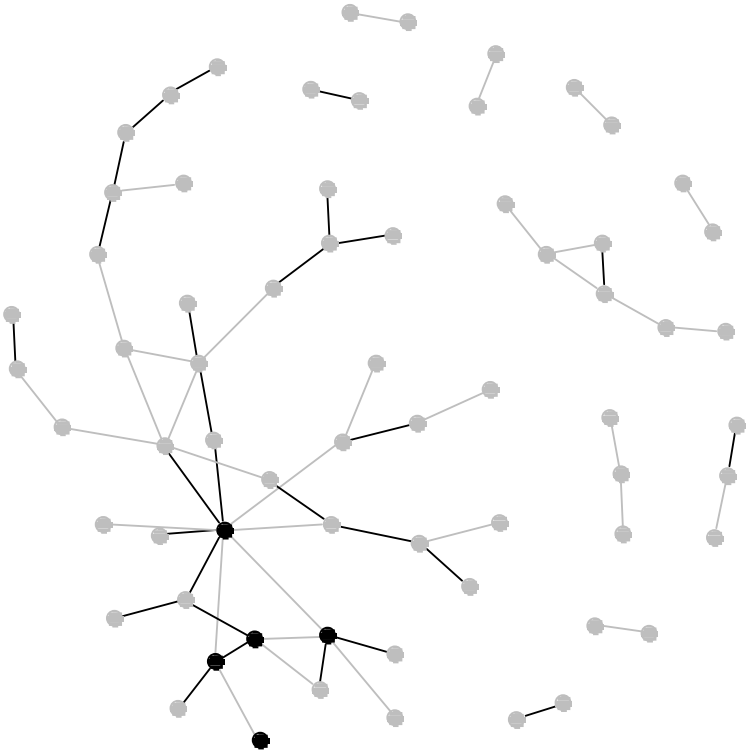
Wasserman and Faust (1994, 4)

- Social network analysis is based on an assumption of the importance of relationships among interacting units.
- Relations, defined by linkages among units are a fundamental component of network theories.
  - Actors and their actions are viewed as interdependent rather than independent.
  - Relational ties (linkages) between actors are channels for transfer or "flow" of resources.
  - Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action.
  - Network models conceptualize structure as lasting patterns of relations among actors.

Graphs, subgraphs, local patterns…

Why do we care about networks as a whole?





Density (Prell 2012, 166-167)

- Density refers to the proportion of ties in a network that are actually present.

- In undirected networks:

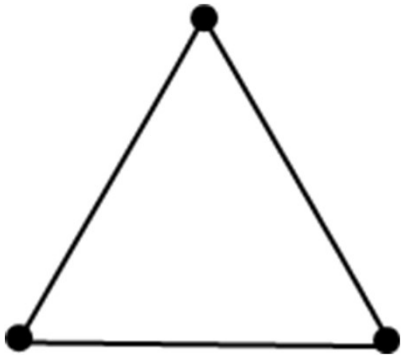
$$d = \frac{L}{n(n-1)/2}$$

Where L refers to the actual number of lines present in the network and n to the number of nodes present in the network.

- In directed networks:

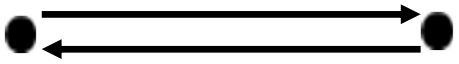
$$d = \frac{L}{n(n-1)}$$

Subgraphs: dyads and triads



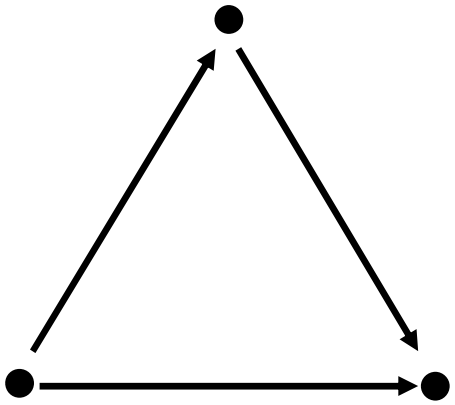
- The friend of my friend is also my friend (Heider)
- Is this pattern common in friendship networks?

Subgraphs: dyad and triad census



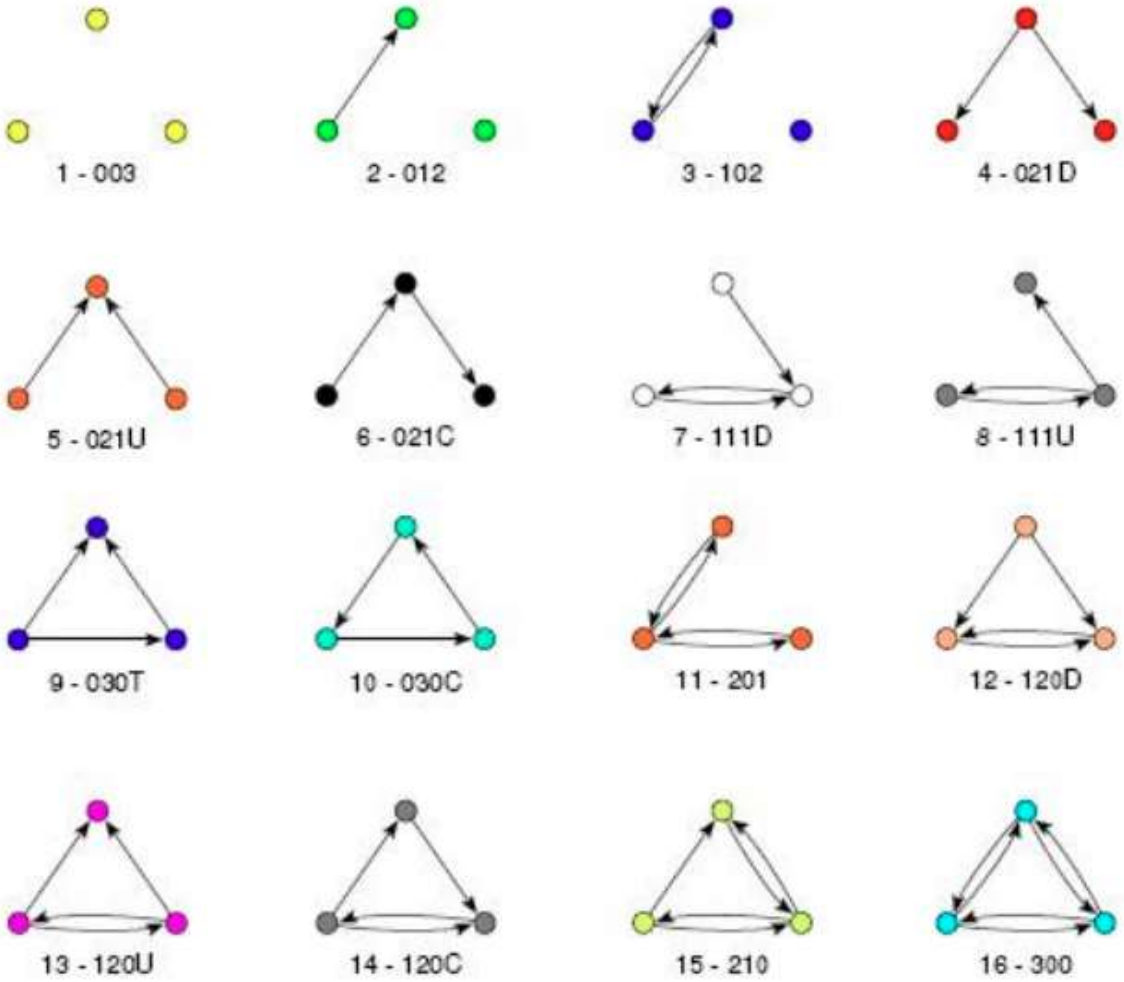
- Is the relationship reciprocal?
- Mutual, asymmetric, or null?

  (e.g., twitter)



- Transitivity
- A to B, B to C, and A to C
- Kinne (2014): “States are more likely to send diplomatic missions to countries where their diplomatic partners also send missions.”

Snijders (2012)



Krivitsky and Handcock (2014)

Table 2

MLE parameter estimates for the longitudinal friendship network

Parameter	Formation est. (s.e.)	Dissolution est. (s.e.)
Edges	−3.336 (0.320)***	−1.132 (0.448)*
Homophily (girls)	0.480 (0.269)	0.122 (0.394)
Homophily (boys)	0.973 (0.355)**	1.168 (0.523)*
F→M heterophily	−0.358 (0.330)	−0.577 (0.609)
Primary school	0.650 (0.248)**	0.451 (0.291)
Reciprocity	1.384 (0.280)***	2.682 (0.523)***
Transitive ties	0.886 (0.247)***	1.121 (0.264)***
Cyclical ties	−0.389 (0.133)**	−1.016 (0.231)***

Research design

Network as IV

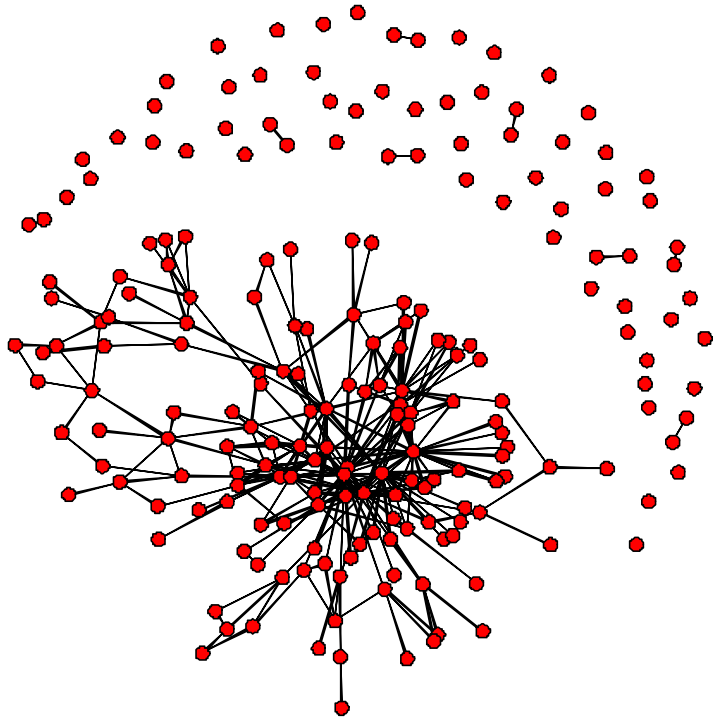
- Friendship -> Behavior
- Centrality -> Promotion

Network as DV

- Homophily -> Friendship
- Personality -> Central actors

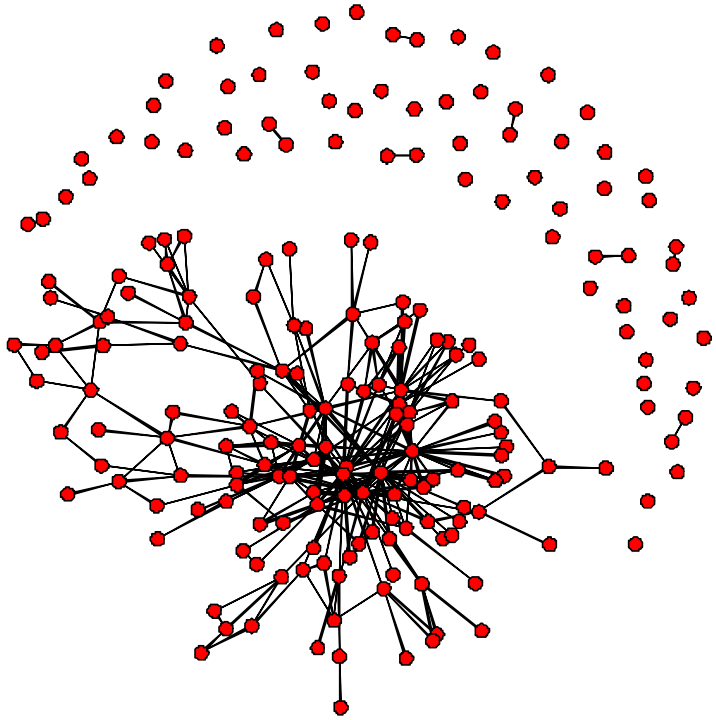
Centrality

- A node's position in a network



Centrality

- Why do we care about central nodes?

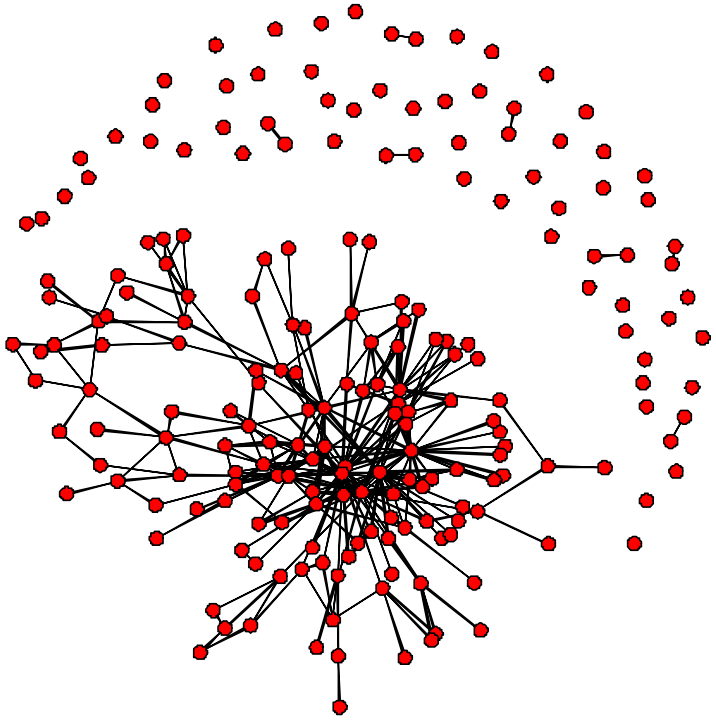


Centrality

Why do we care about central nodes?

- Relational ties (linkages) between actors are channels for transfer or “flow” of resources.
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action.

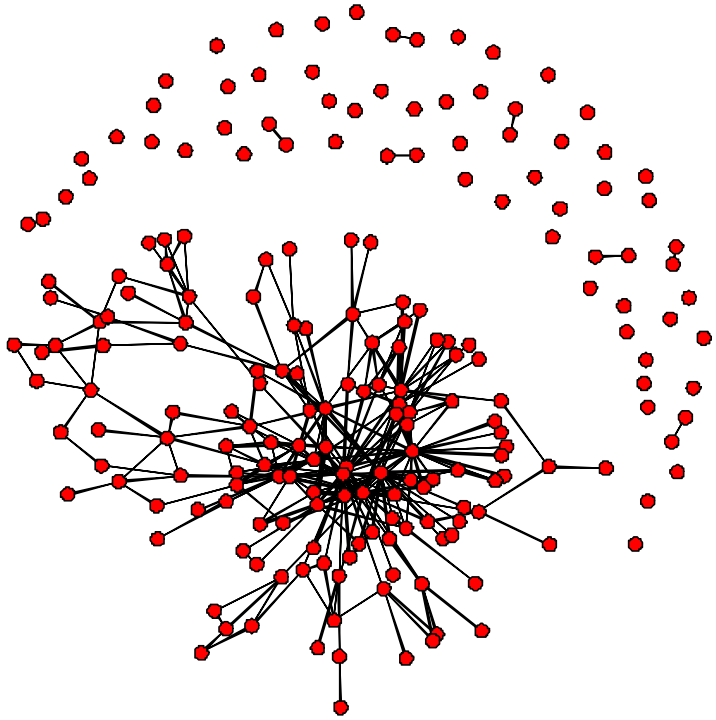
  (Wasserman and Faust 1994)



Centrality

- Why do we care about central nodes?

- In terms of information flows?
  - Get the information earlier than others
  - Control the flow of information



Centrality

- What does it mean for an actor to be central?

Centrality

- What does it mean for an actor to be central?

- Central nodes:
  - Prominent, influential
  - Having great visibility, prestige, power, control, autonomy
  - Leaders, gatekeepers

- "For non-negative relations such as friendship or trust, centrality tends to be viewed as a positive thing for nodes, providing actors with opportunities to influence others and receive flows (i.e., information)."

Degree centrality

- Total number of ties
- $d_i = \sum_j x_{ij}$

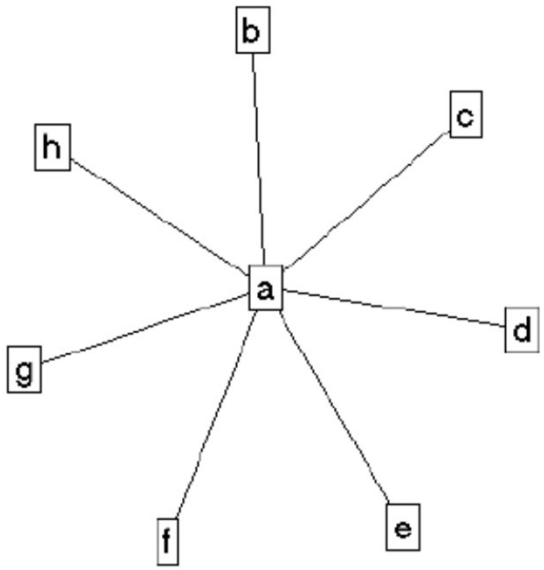
    - In-degree
    - Out-degree

Eigenvector centrality

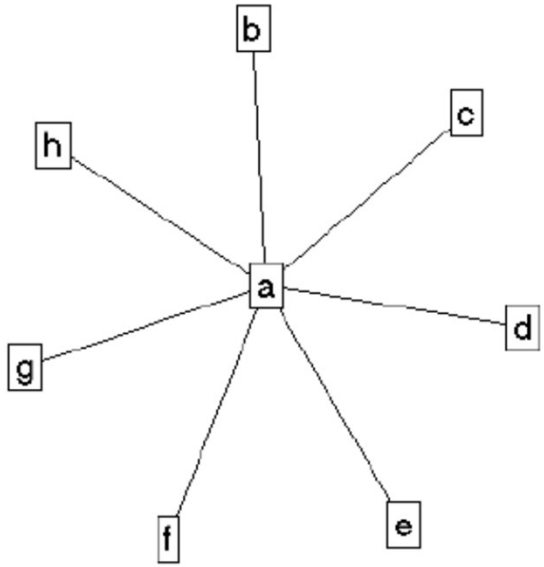
- Proportional to the sum of centralities of the nodes it is connected to
- If one's alters have high degree centrality, then the focal actor has high eigenvector centrality.

- $e_i = \lambda \sum_j x_{ij} e_j$

  - Ego: focal actor
  - Alters: the nodes that eco is connected



Betweenness Centrality



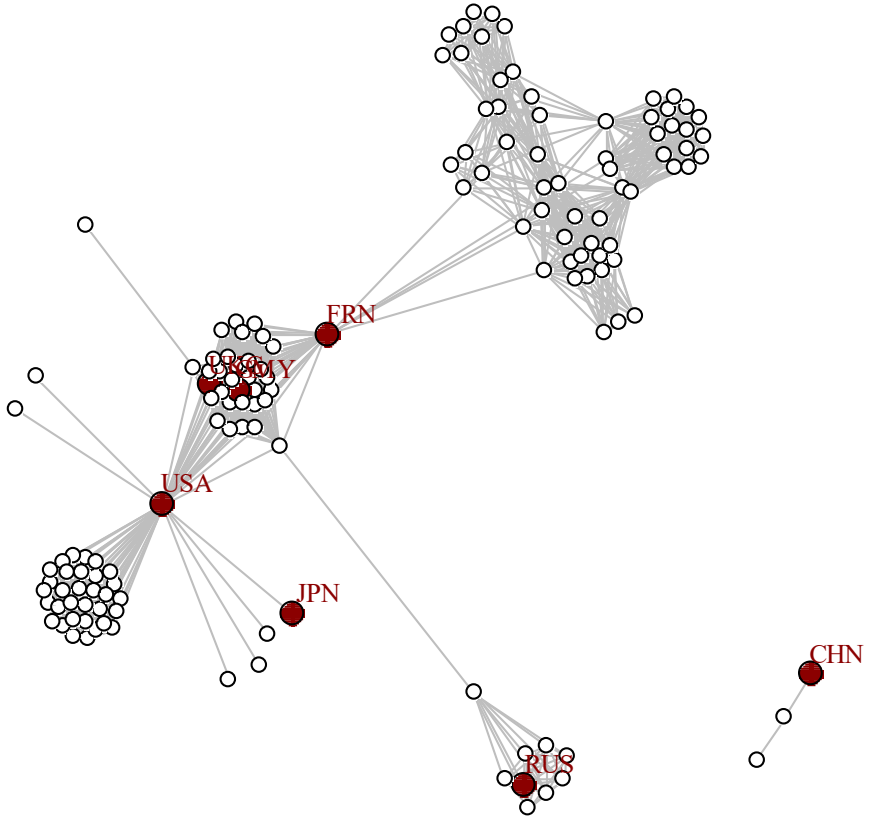
- How often an actor rests between two other actors

- $C_B(k) = \sum \frac{\partial_{ikj}}{\partial_{ij}}, i \neq j \neq k,$

Where ∂_{ikj} is the number of geodesics linking actors i and j that pass through node k, and ∂_{ij} is the number of geodesics linking actors k and j

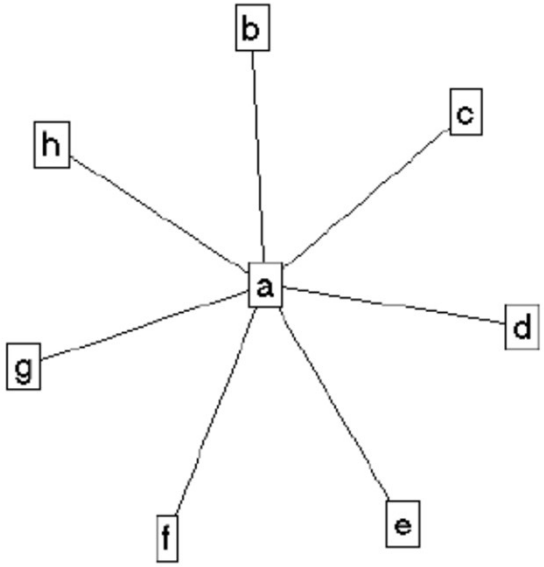
  - Note: geodesic means the shortest path in the network.

Defense pacts in 2008



Closeness Centrality

- Emphasizes an actor's Independence.

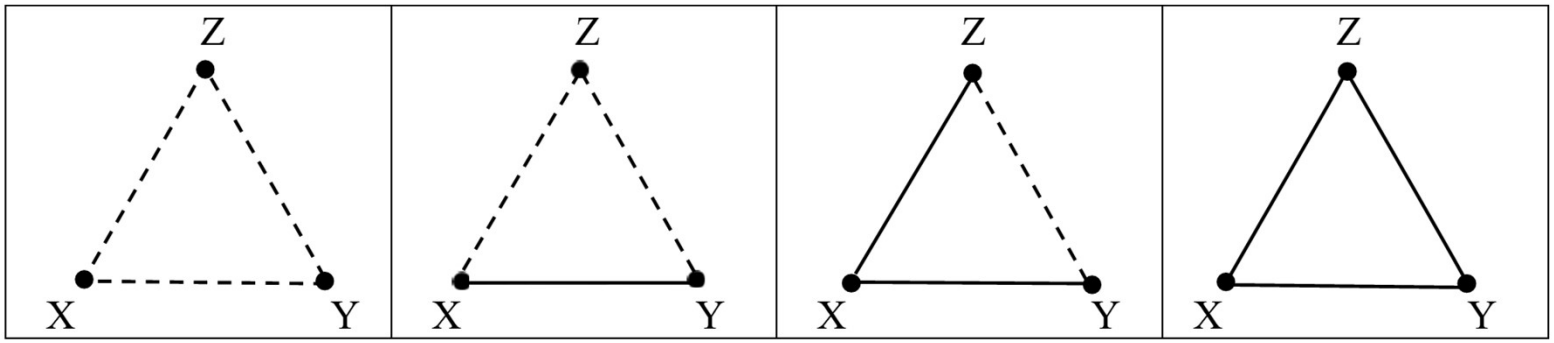


- If an actor is not central, the actor generally needs to rely on others to relay messages through the network. Thus, an actor who is close to many other actors is a very independent actor.

- Kinne (2012), economic dependence → Embeddedness

R – Network statistics and centrality

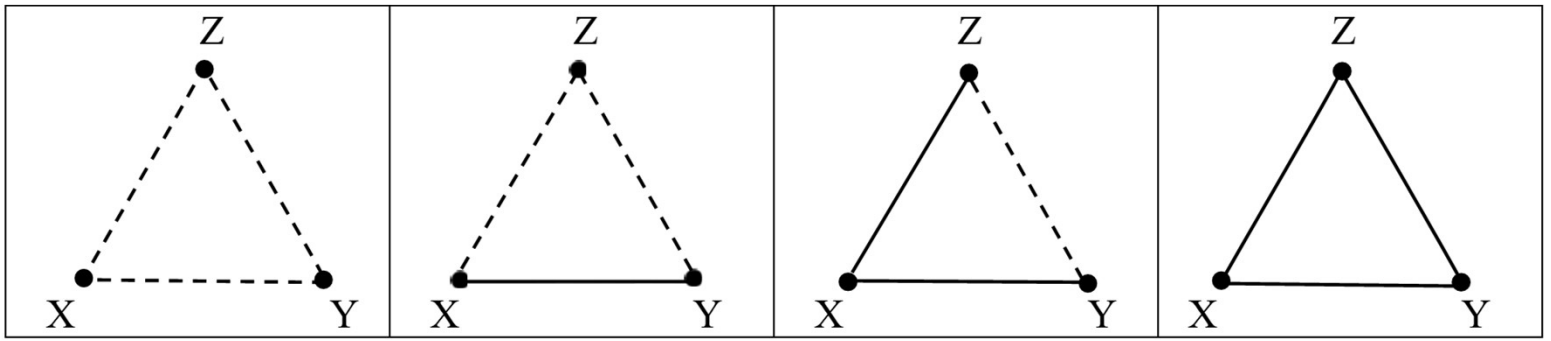
Networks including negative ties



Heider (1946, 1958)

- The friend of my friend is my friend.
- The friend of my enemy is my enemy.
- The enemy of my friend is my enemy.
- The enemy of my enemy is my friend.

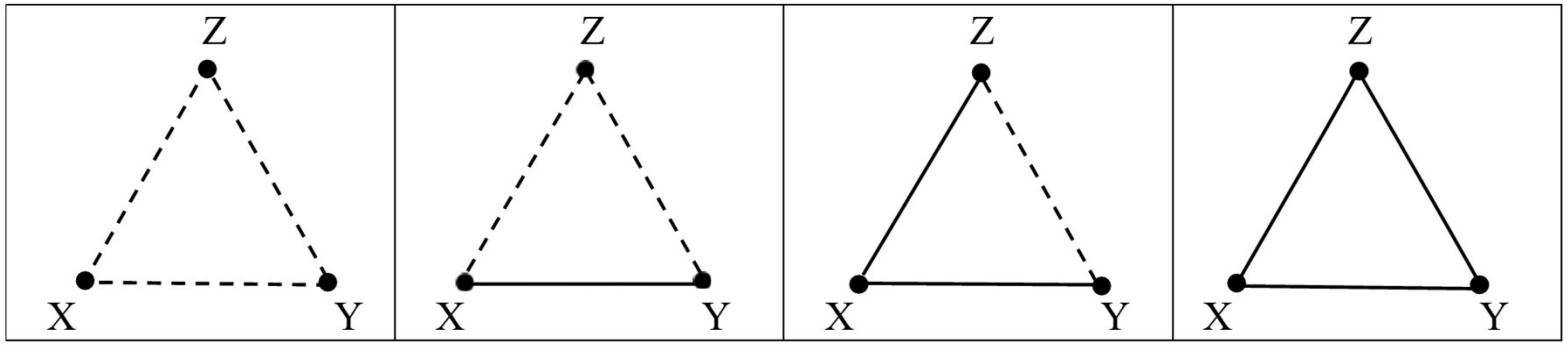
Networks including negative ties



Heider (1946, 1958)

- (2nd) The friend of my enemy is my friend.
- (4th) The enemy of my enemy is my enemy.

Networks including negative ties

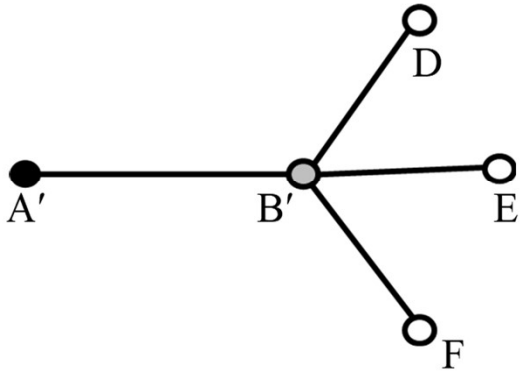


Heider (1946, 1958)

- Cognitive units or entities tend to attain balance, “a situation in which the perceived units and the experienced sentiments co-exist without stress”
  - Is this a common pattern in individuals?
  - How about countries?

Negative centrality

- Focus on A

B: Lower Centrality	B′: Higher Centrality
	

Negative centrality

- Everett and Borgatti (2014)

- “The recipient of a negative tie would be less affected by receiving a negative tie from an actor who gave out a lot of negative ties than if they were the sole recipient of a negative tie from an actor that only gave one negative tie.”

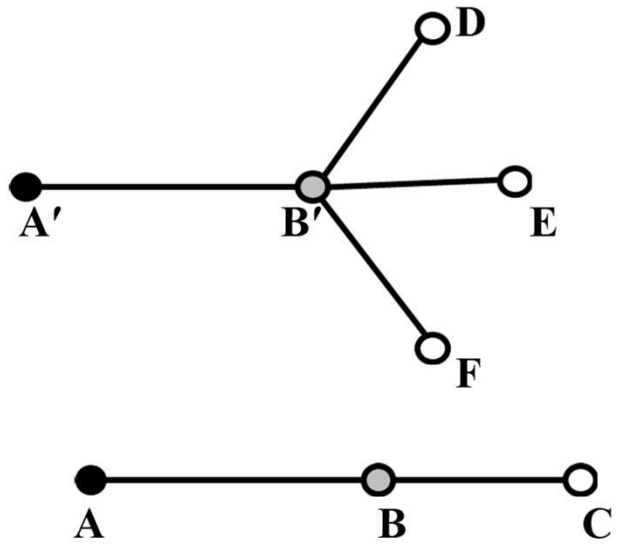
Negative centrality

- Borgatti, Everett, and Johnson (2018: 207-208)

$$PN = \left(I - \frac{1}{2n-2}A\right)^{-1}\mathbf{1}$$

  - A=P-2N, P denotes a positive tie matrix and N denotes a negative tie matrix.
  - If both positive and negative ties are possible, then it ranges from -1 to 2.

Negative centrality



Node	Negative Centrality (unweighted)
B′	0.467
B	0.745
A	0.894
C	0.894
A′	0.933
D	0.933
E	0.933
F	0.933

R – Signnet

References

Borgatti, Stephen P., Martin G. Everett, and Jeffrey C. Johnson. 2018. *Analyzing Social Networks 2nd Edition*.

Heider, Fritz. 1946. "Attitudes and Cognitive Organization." *Journal of Psychology* 21(1): 107-112.

Heider, Fritz. 1958. "Sentiment." In *The Psychology of Interpersonal Relations*. New York: Wiley, pp.174-217.

Kinne, Brandon J. 2012. "Multilateral Trade and Militarized Conflict: Centrality, Openness, and Asymmetry in the Global Trade Network." *Journal of Politics* 74(1): 308-322.

Kinne, Brandon J. 2014. "Dependent Diplomacy: Signaling, Strategy, and Prestige in the Diplomatic Network." *International Studies Quarterly* 58: 247-259.

Krivitsky, Pavel N., and Mark S. Handcock. 2014. "A Separable Model of Dynamic Networks. *J R Stat Soc Series B Stat Methodol* 76(1):29-46.

Prell, Christina. 2012. *Social Network Analysis: History, Theory & Methodology*.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Appplications*.