Working with Network Data in R: Loading, Visualizing, Heuristically Analyzing, and Statistically Modeling

Momin M. Malik, PhD

ICQCM 2026 Winter Summit

February 26–March 1, 2026, New Orleans, LA

https://www.mominmalik.com/icqcm2026c.pdf

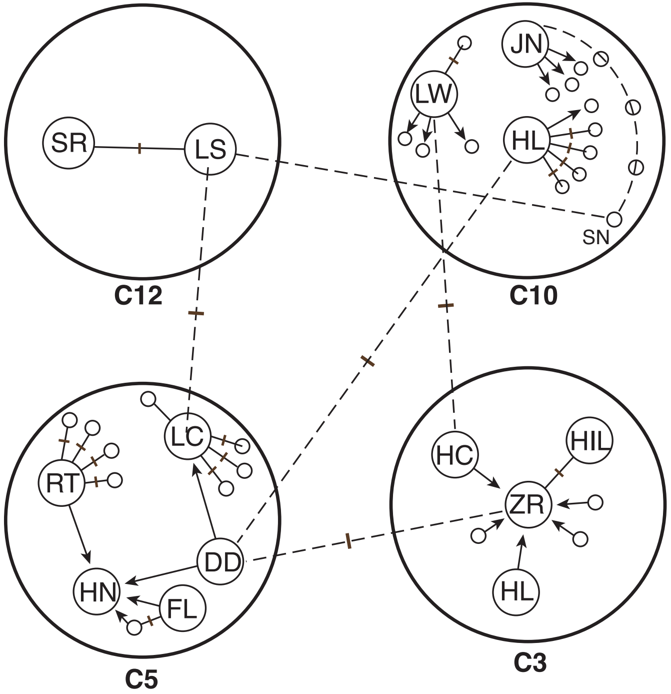
Navajo Weaving: *Replica of a Chip*, by Marilou Schultz (1994)

Outline

- Why networks?
- Representation of networks
- Disclaimers
- Basic network terms
- Heuristic analyses
- Statistical analysis

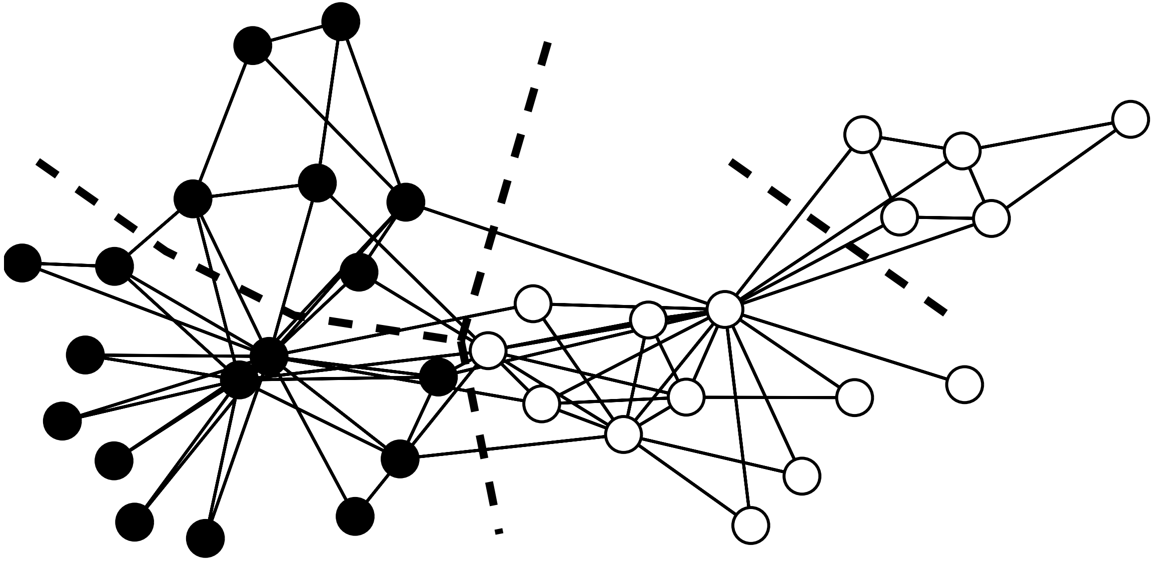
Why networks?

Runaways from “reformatory” for delinquent girls



Moreno (1934); Borgatti et al. (2009)

Competing factions in a karate club



Zachary (1977); Porter et al. (2009)

Why networks?

Represen-
tation of
networks

Disclaimers

Basic
network
terms

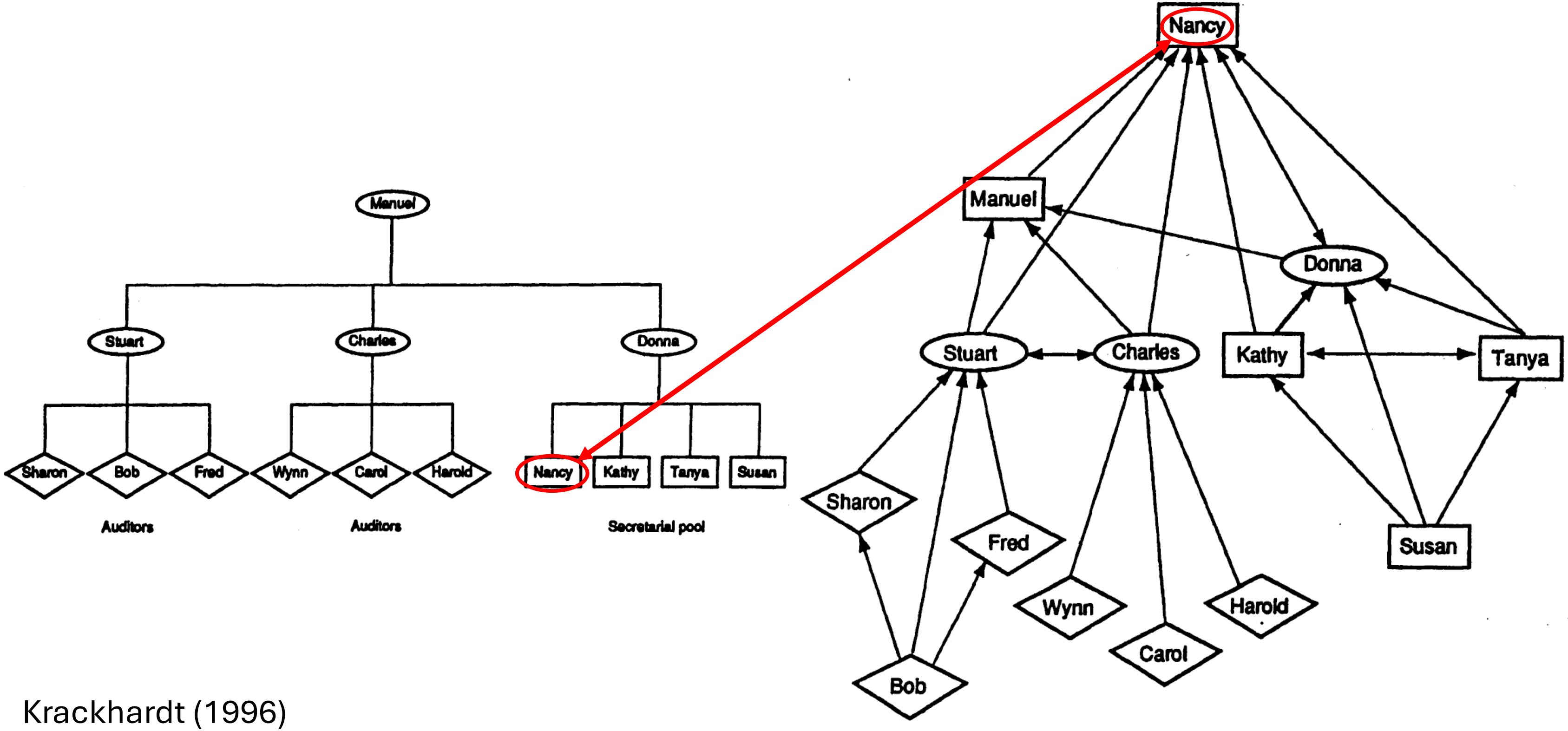
Heuristic
analysis

Statistical
analysis

Conclusions

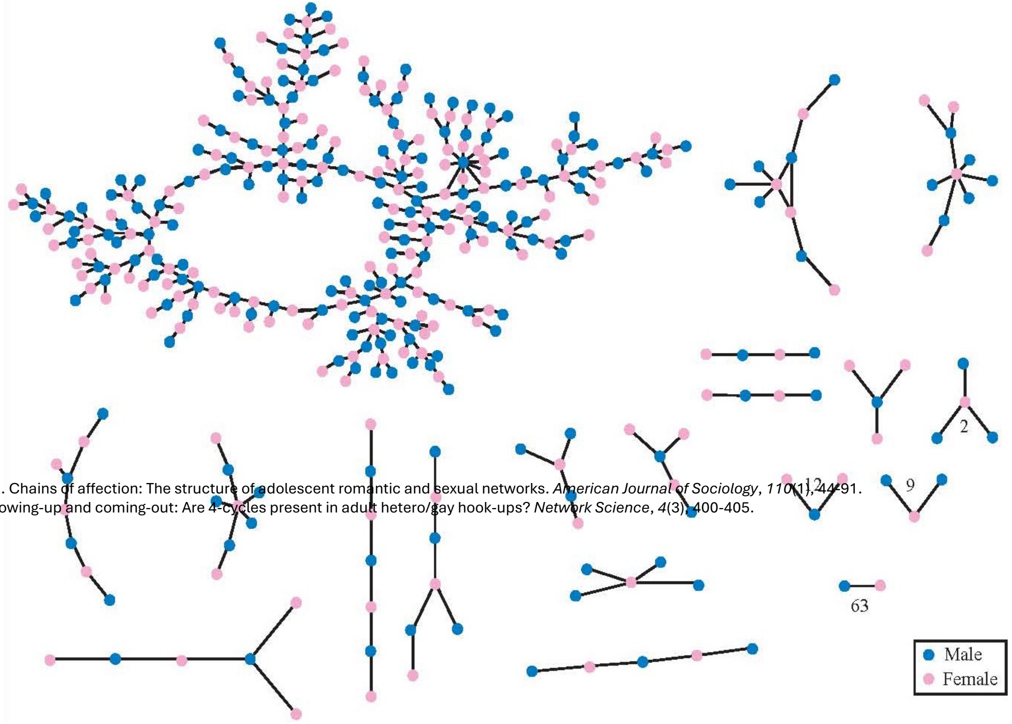
Works cited

Org hierarchy versus informal network



Krackhardt (1996)

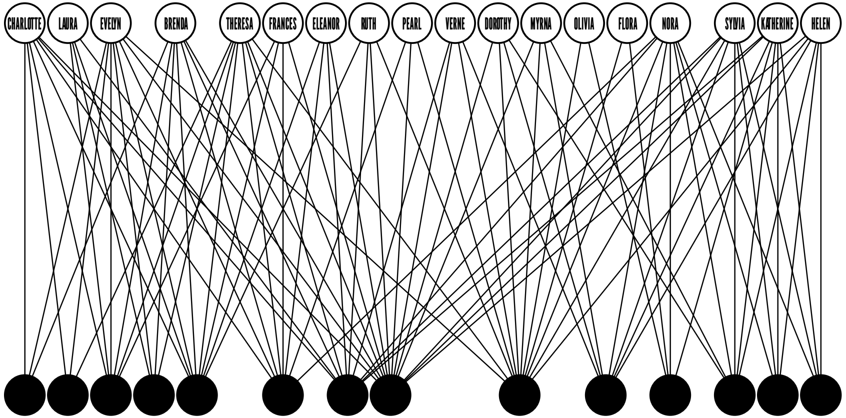
Not dating a ex’s ex in a high school

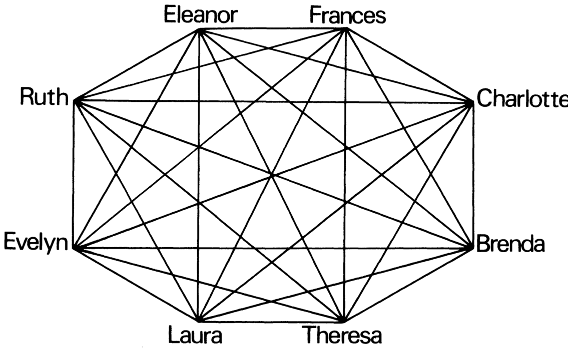


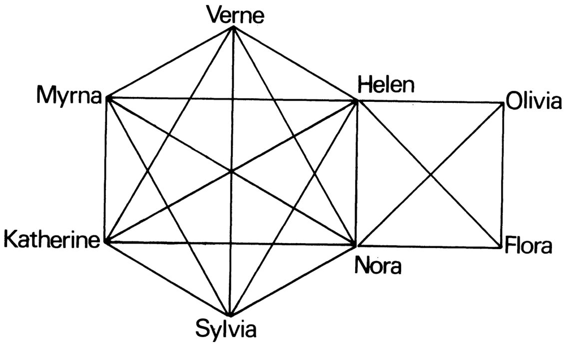
Bearman, P. S., Moody, J., & Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, *110*(1), 44-91.
Marcum, C. S., Lin, J., & Koehly, L. (2016). Growing-up and coming-out: Are 4-cycles present in adult hetero/gay hook-ups? *Network Science*, *4*(3), 400-405.

Bearman et al. (2004); Marcum et al. (2016)

Shared events among southern debutants



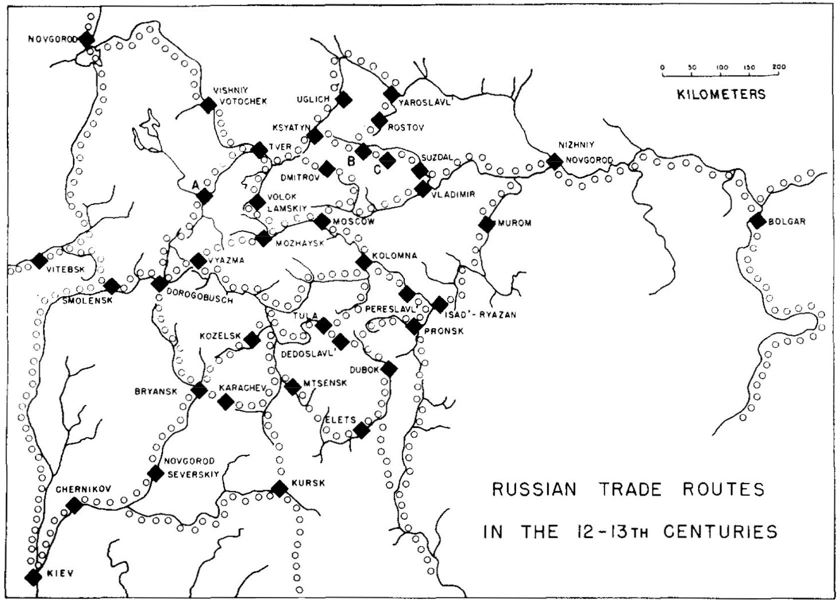


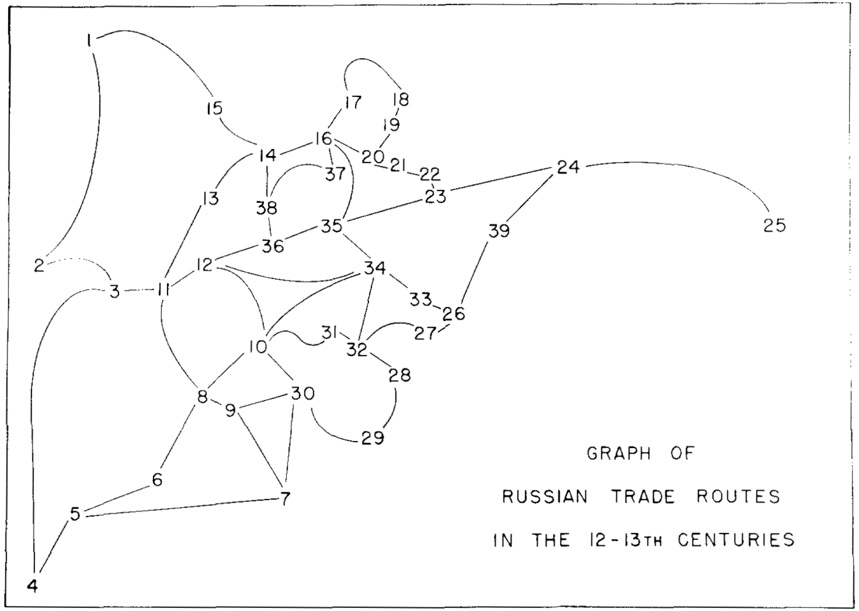


Davis & Gardner (1941); Breiger (1974)

(Same principles was behind recommendations for about a decade)

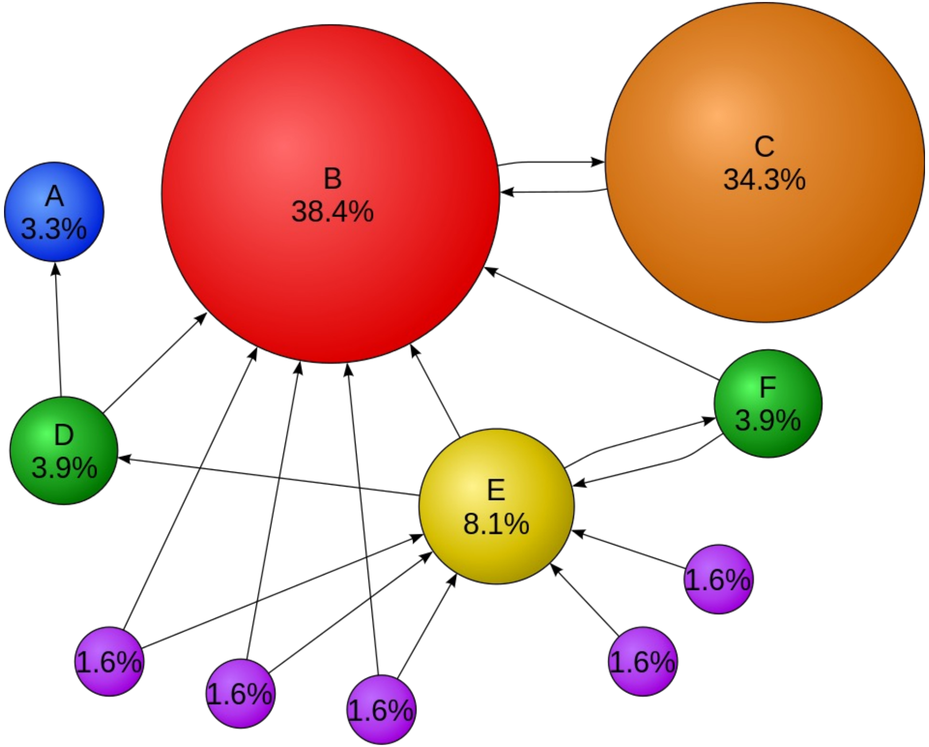
Medieval Moscow at center of river trade (Pitts, 1978/1979)





Pitts (1978/1979)

Similar idea (centrality) behind original Pagerank



Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

Conclusions

Works cited

Notice also problems

- Ella Fitzgerald may have been one of the girls studied among “runaways”! (she was at the same training school at the same time) (Immarigeon, 2014; Cherry, 1995)
- Managers who surveil employee networks may get strategic information
- Surveillance into teens’ sexuality??
- Social network analysis, at least at some points, was used heavily by the US military in counterinsurgency (Knoke, 2013; Mac Ginty, 2010)
- The idea that the most important information is in the *structure*, and we can ignore the content, is limiting

Places used

- Basic research in sociology, anthropology, social psychology, epidemiology
- US military
- Organizational behavior, management science
- “Complexity science”, sociophysics, “computational social science”
- (Are also networks in operations research and optimization, and in computer science research, but those are less relevant)

Theoretical benefits...?

- Networks, in representing relations, have been said to escape reductionism; but they do not (Erikson, 2014)
  - Networks a hallmark of “complex systems”
  - But long tradition of simply displacing the endpoint of reduction from individuals to network processes (seeing people as atoms, and find regular rules determining how they configure)
- Networks lead to lots of exciting math, and are a powerful formal representation, but can also be used qualitatively, and theoretically (e.g., nodes emerge out of the network, rather than vice versa)!
- Statistics on networks: Extremely difficult.

Representations of networks

A “network” can be formed by any sort of relations (but be careful of similarities)

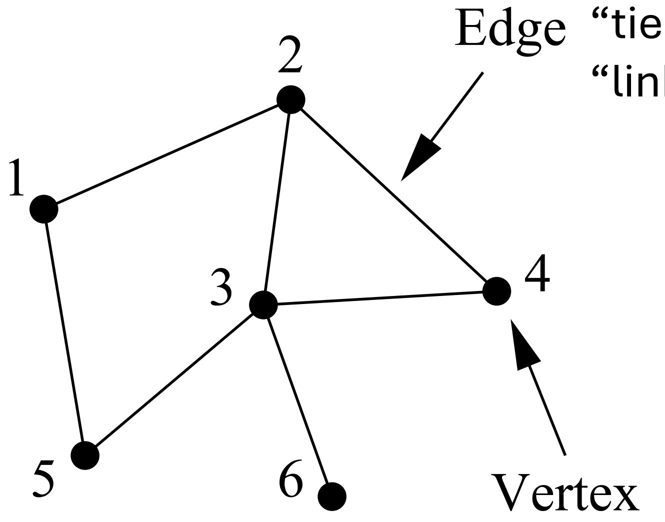
Similarities			Social Relations				Interactions	Flows
Location	**Membership**	**Attribute**	**Kinship**	**Other role**	**Affective**	**Cognitive**	e.g.,	e.g.,
e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	Sex with	Information
Same spatial and temporal space	Same clubs	Same gender	Mother of	Friend of	Likes	Knows	Talked to	Beliefs
	Same events	Same attitude	Sibling of	Boss of	Hates	Knows about	Advice to	Personnel
	etc.	etc.		Student of	etc.	Sees as happy	Helped	Resources
				Competitor of		etc.	Harmed	etc.
							etc.	

Borgatti et al. (2009)

Getting network data

- Survey instruments: include “name generators”
  - Are now some nice digital tools for this (e.g., Network Canvas)
- Observation
- Administrative data
  - Includes co-authorships or citations (bibliometrics/scientometrics), org charts, attendance
- Digital trace data: Tempting but dangerous!
- Similarity data (including correlations between variables): BAD!!

Key representations: Graph; adjacency matrix; edgelist



(Also called "ties" or "links")

(Also called "nodes")

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Newman (2010)

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

Conclusions

Works cited

Edgelist shows, we go from *n* to *n*²

	Y	X_1	X_2	$\cdots$	X_k
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

	from	to	Y	W_1	W_2	W_3	$\cdots$
e_1	v_1	v_2	y_{12}	$\mathbb{1}(x_{11} = x_{21})$	$x_{12} - x_{22}$	x_{13}	$\cdots$
e_2	v_2	y_3	y_{23}	$\mathbb{1}(x_{11} = x_{31})$	$x_{12} - x_{32}$	x_{13}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
e_{n+1}	v_2	v_1	y_{21}	$\mathbb{1}(x_{21} = x_{11})$	$x_{22} - x_{12}$	x_{23}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$e_{2\binom{n}{2}}$	v_{n-1}	v_n	$y_{(n-1)n}$	$\mathbb{1}(x_{(n-1)1} = x_{n1})$	$x_{(n-1)2} - x_{n2}$	$x_{(n-1)3}$	$\cdots$

Disclaimers

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

Conclusions

Works cited

1: I have a limited perspective

- I mostly stopped working with networks in 2019, so my knowledge may be outdated by a few years
- The social networks world is overwhelmingly white (one of the reasons I increasingly lost interest); there probably is critical race work, but based on my training, I don't know it
  - Jacob Moreno, the first inventor of social network analysis, was purportedly an egomaniac (why SNA died for a few decades after him)
  - Harrison White's 1970s group at Harvard: lots of sexism (Edling, 2009)? Where most of the current field leadership came out of

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

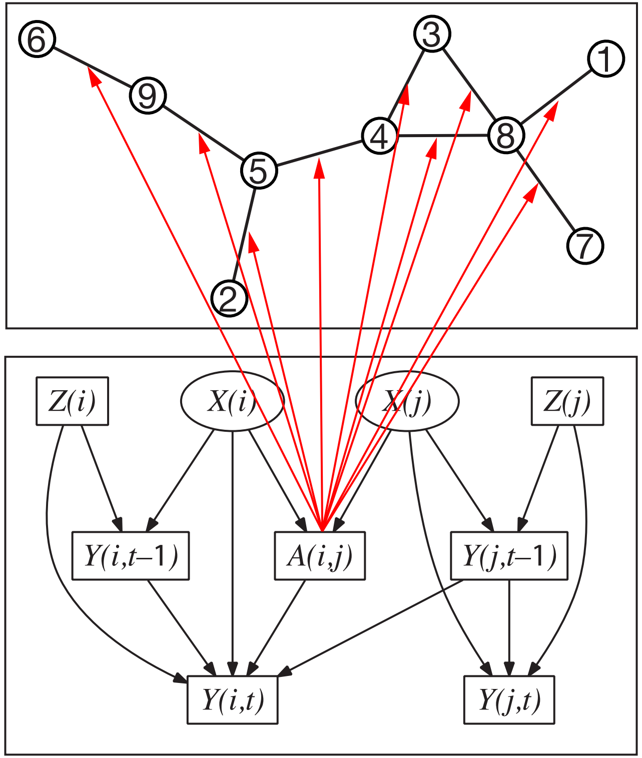
Conclusions

Works cited

2: Graphs are an overloaded representation

Network model (observations)

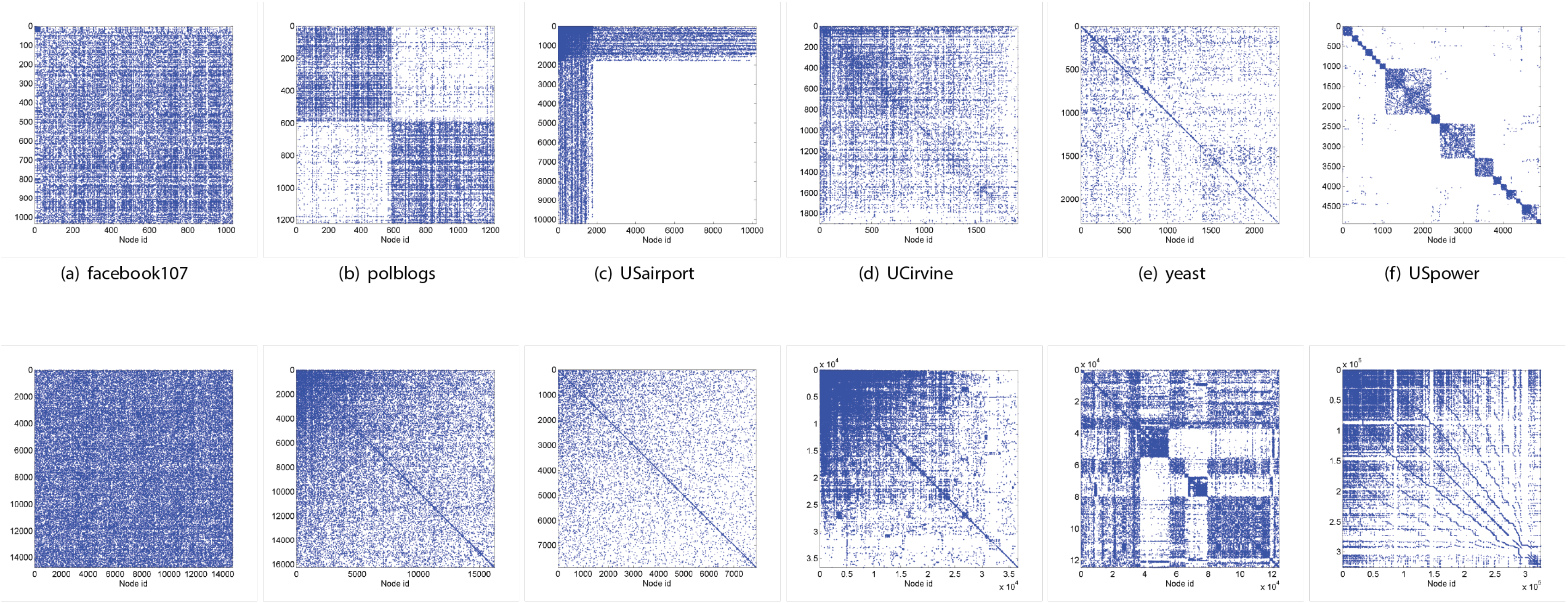
Graphical model (variables), e.g. representing causality



Note: This is also a Markov chain, as it has a tie from *Y*(*t*-1) to *Y*(*t*). Markov chains are themselves graphs, but also have *transition* graphs, which is yet another type of graph.

Bottom: Shalizi & Thomas (2009)

3: Networks are similarities (heatmaps can be used), but similarities are not networks



(a) facebook107 (b) polblogs (c) USairport (d) UCirvine (e) yeast (f) USpower

(g) IMDB (h) cond-mat1 (i) cond-mat2 (j) enron (k) internet (l) www

Caron & Fox (2015)

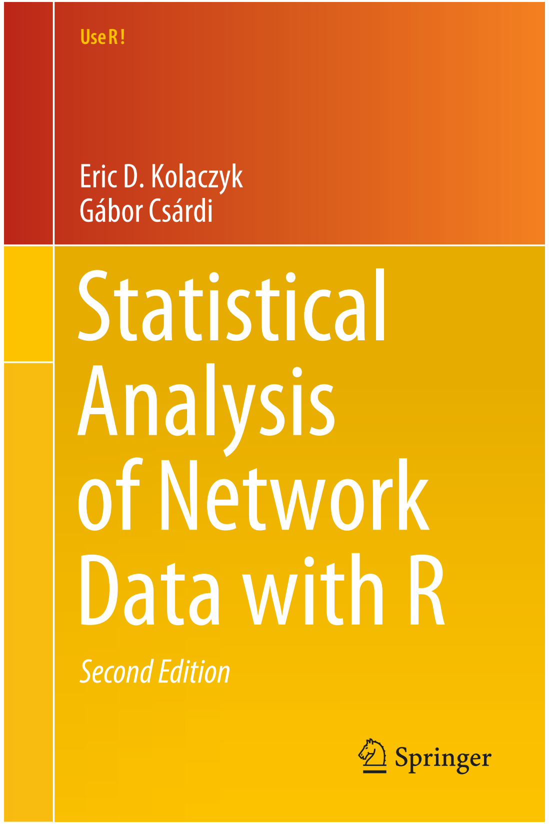
Theoretical benefits...?

- Networks, in representing relations, have been said to escape reductionism; but they do not (Erikson, 2014)
  - Networks a hallmark of "complex systems"
  - But long tradition of simply displacing the endpoint of reduction from individuals to network processes (seeing people as atoms, and find regular rules determining how they configure)
- Networks lead to lots of exciting math, and are a powerful formal representation, but can also be used qualitatively, and theoretically (e.g., nodes emerge out of the network, rather than vice versa)!
- Statistics on networks: Extremely difficult.

Basic network terms

To work through some of these in R, following Kolaczyk & Csárdi (2020)

- Directed and undirected
- Multigraph
- Weighted
- Bipartite; projection
- Ego and alters
- Ego network
- Neighbors, incident, walks
- Connected components (strong and weak)
- Induced subgraph
- Degree distribution
- Node properties vs edge properties
- Graph motifs: *k*-stars, trees, triads
- Triad census
- Community detection
- Density, diameter



Download data

- https://www.stats.ox.ac.uk/~snijders/siena/LazegaLawyers.zip
- Install `igraph` in R
- Optional: download Cytoscape

Other software

- Pajek: Windows-only (I once successfully used Wine for Macs, but it was tricky) GUI software that is really powerful for big networks. Some really nice features (connected components layout), but very user-hostile. Is a book about it
- Gephi: Was abandoned then revived. Lite web version
- Cytoscape: Mostly for biology, but I switched to it when Gephi was abandoned, and still use it

Heuristic analysis

Centralities

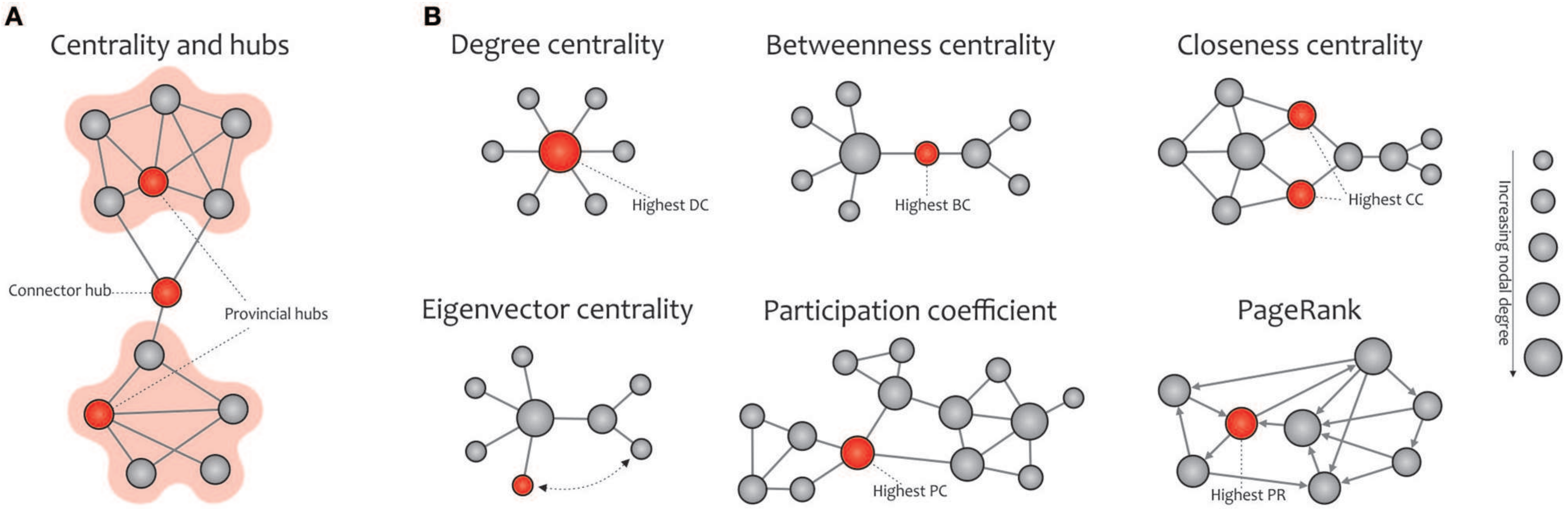
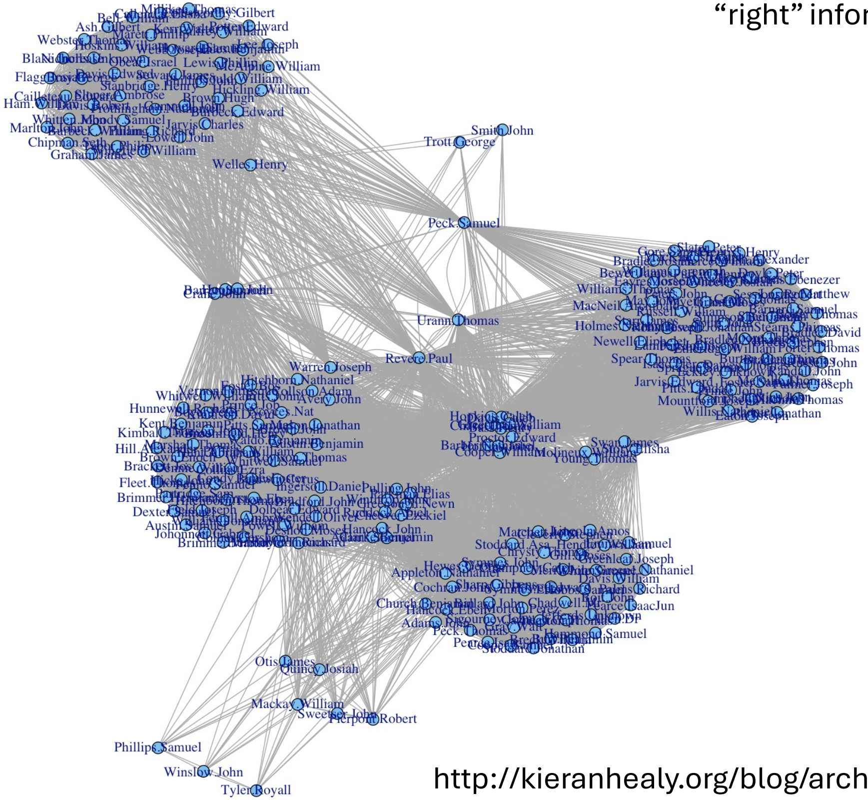
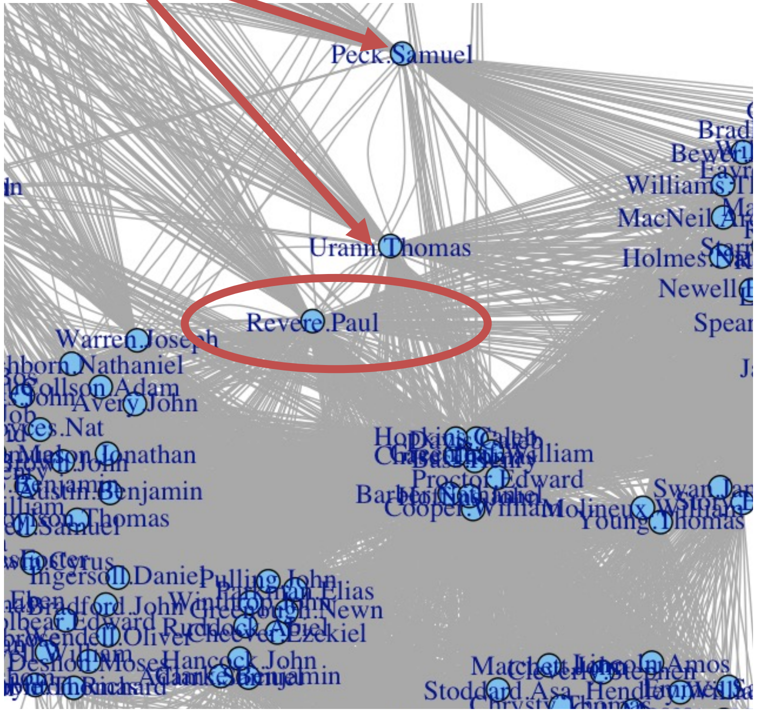


Image: Farahani, F. V., Karwowski, W., & Lighthall, N. R. (2019). Application of graph theory for identifying connectivity patterns in human brain networks: A systematic review. *Frontiers in Neuroscience*, *13*, 585. https://doi.org/10.3389/fnins.2019.00585

Problems: Retrospective, just-so storytelling

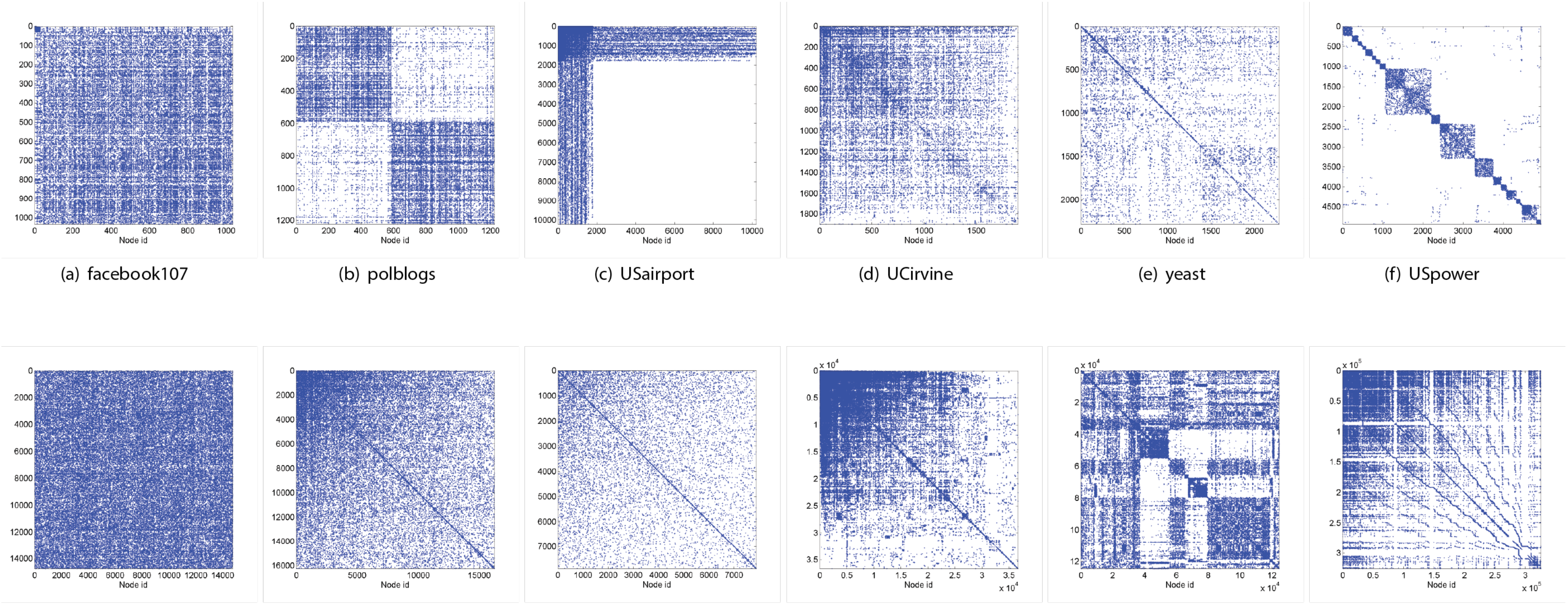


Important but neglected historically? Or is the network diagram not capturing the "right" information?



http://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/

Clustering and “blocks”



(a) facebook107
(b) polblogs
(c) USairport
(d) UCirvine
(e) yeast
(f) USpower
(g) IMDB
(h) cond-mat1
(i) cond-mat2
(j) enron
(k) internet
(l) www

Caron & Fox (2015)

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

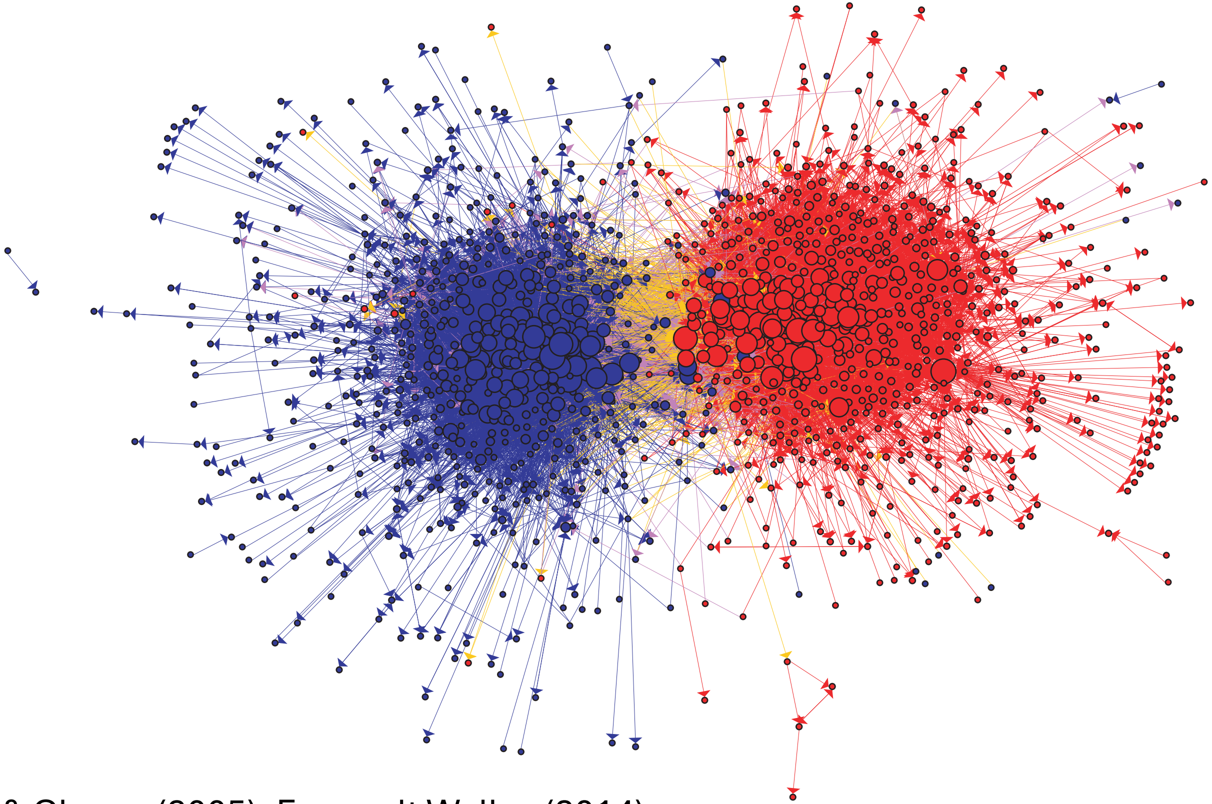
Heuristic analysis

Statistical analysis

Conclusions

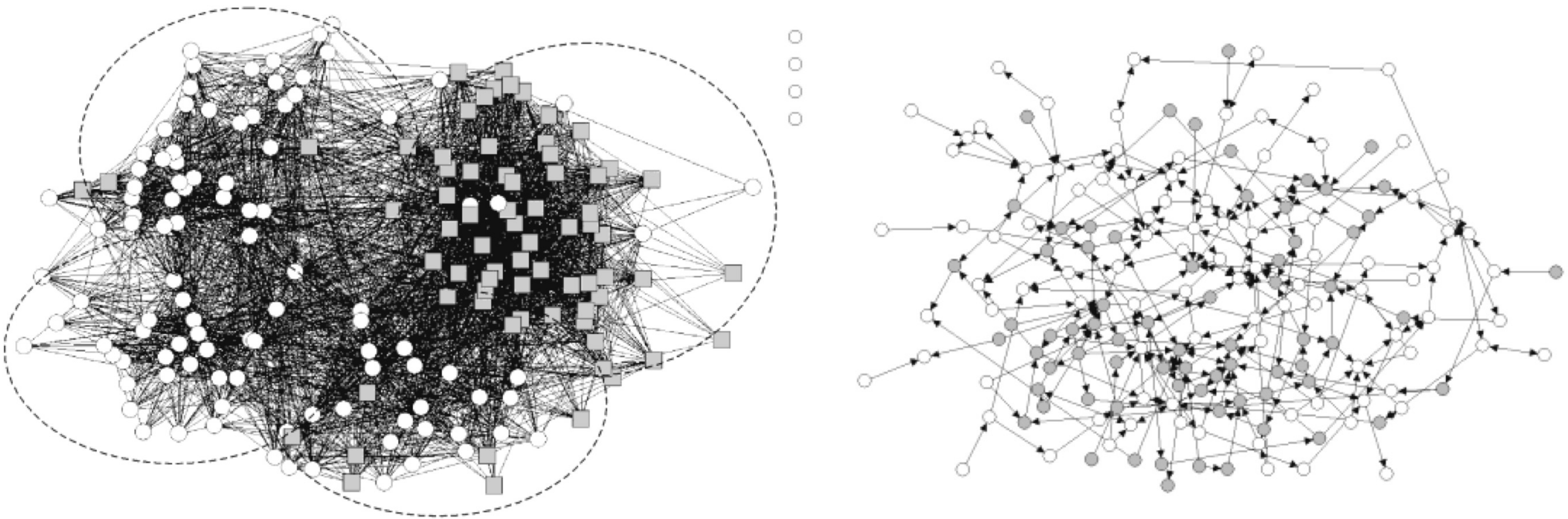
Works cited

Visualization



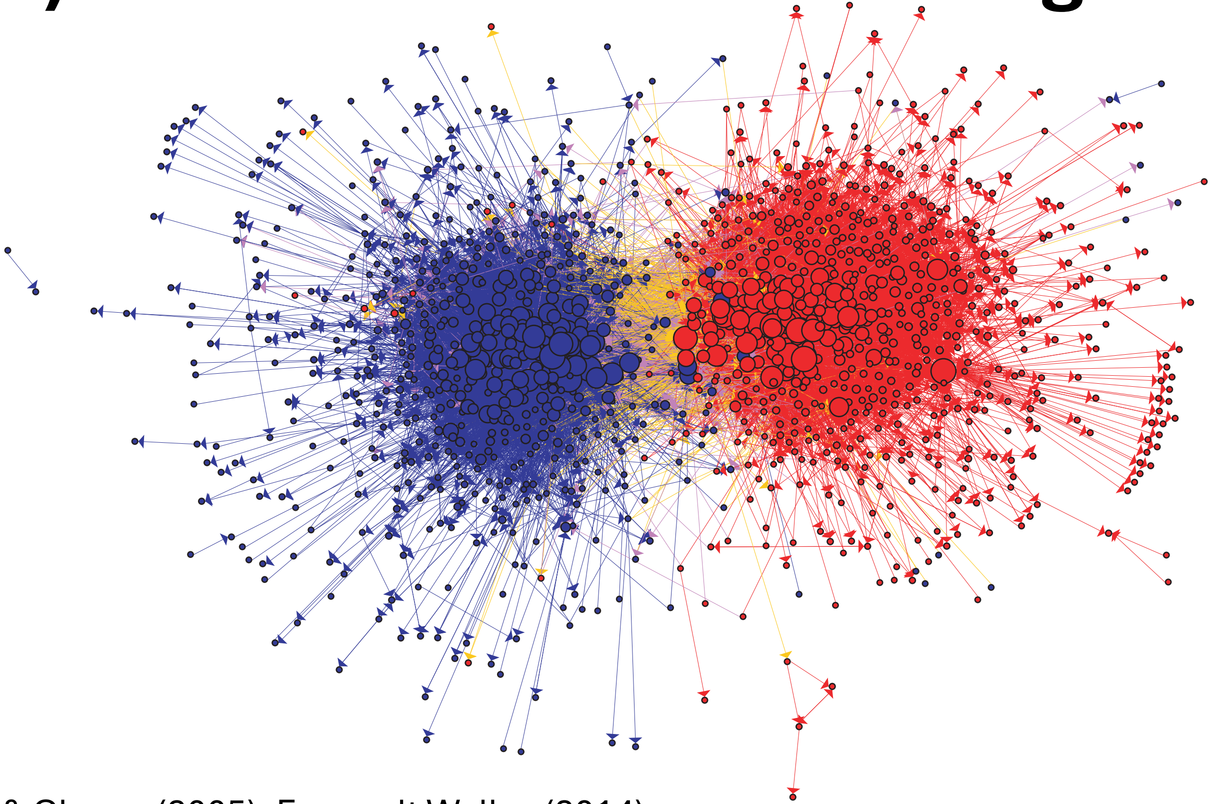
Adamic & Glance (2005); Foucault Welles (2014)

Problems: “Spaghetti”



Durland (2005)

Problems: This image was massaged (17-core) until it showed something



Adamic & Glance (2005); Foucault Welles (2014)

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

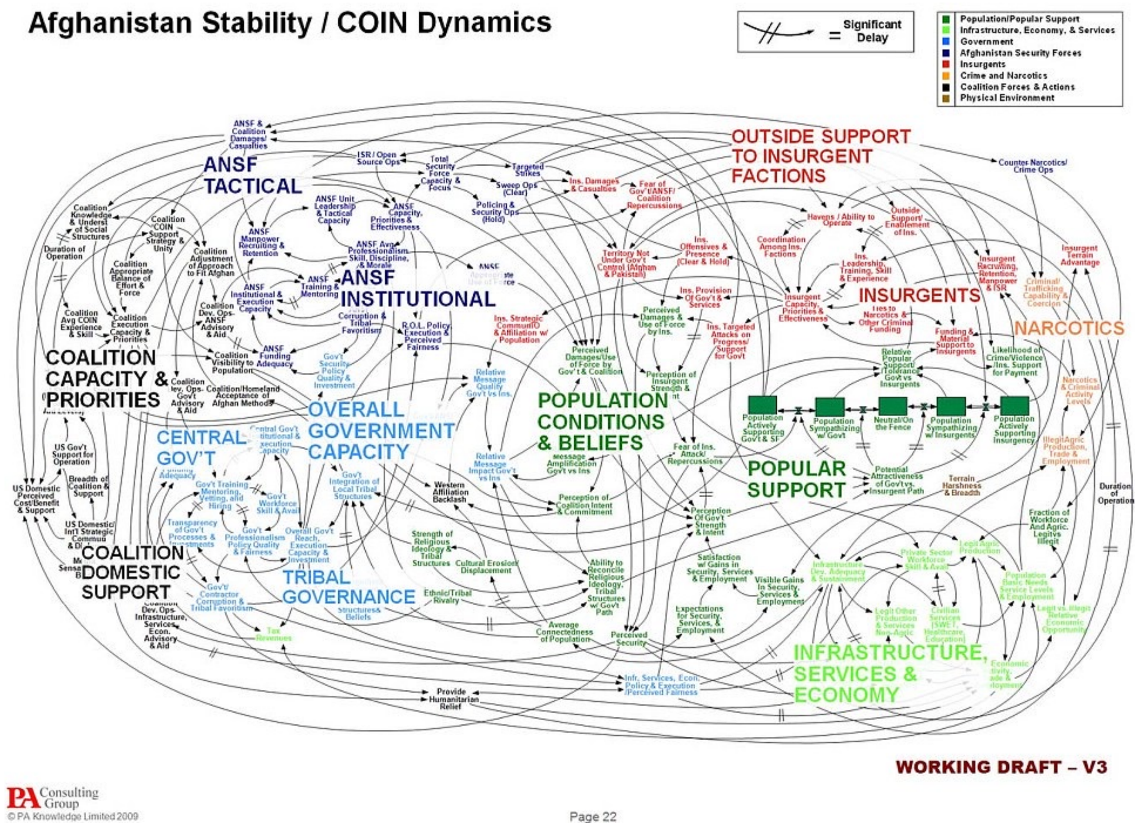
Heuristic analysis

Statistical analysis

Conclusions

Works cited

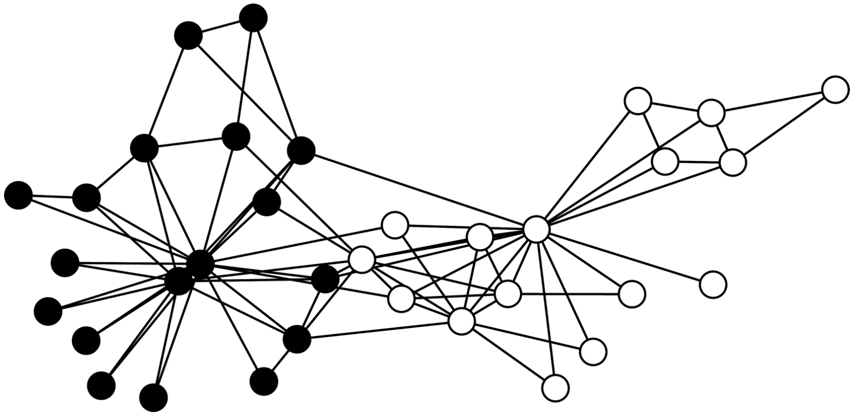
Problems: Bad visualization



Statistical analysis

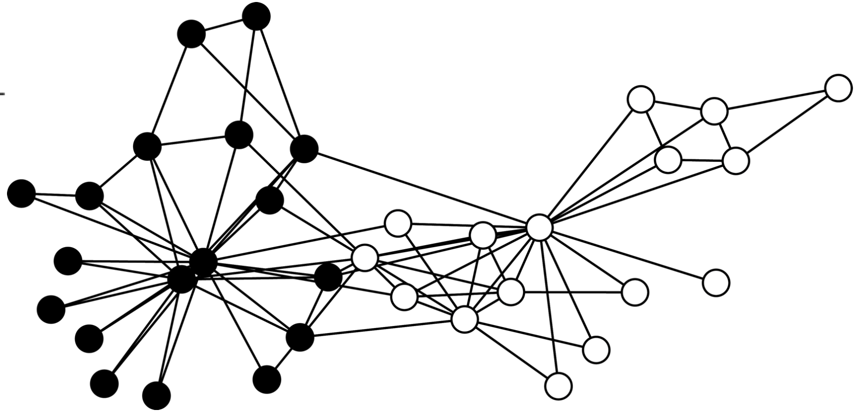
Social scientists: “Obvious” first pass

	Y	X_1	X_2	$\cdots$	X_k
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}



Social scientists: “Obvious” first pass

	Y	X_1	X_2	$\cdots$	X_k	C_d
v_1	y_1	x_{11}	x_{12}	$\cdots$	x_{1k}	d_1
v_2	y_2	x_{21}	x_{22}	$\cdots$	x_{2k}	d_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
v_n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	d_n



The problem for statisticians

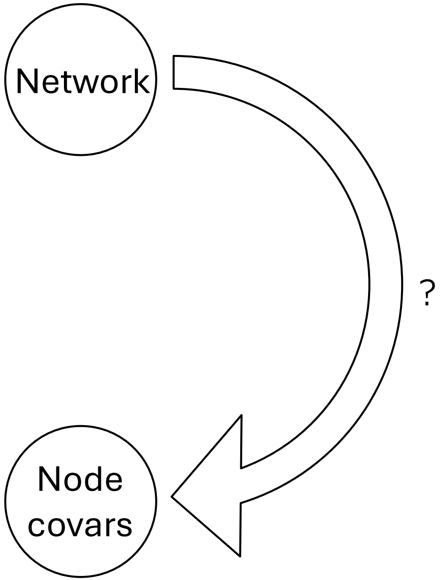
- *Ceteris paribus* (“holding all else constant”) interpretation
- How do we change the (undirected) degree of one node (or some centrality like eigenvector, betweenness, closeness) and hold those of all other nodes constant
- Deeper question: **what are we trying model?**

What are we trying to model?

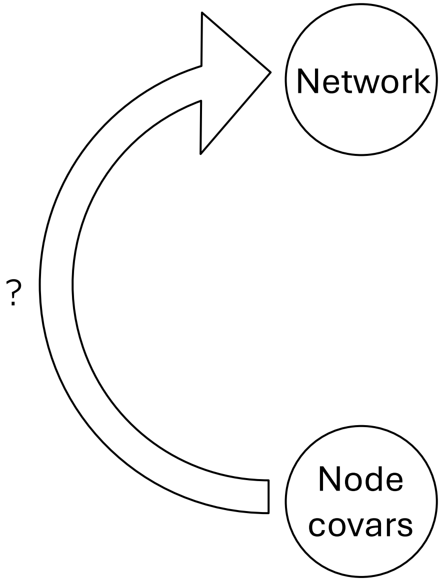
- Centralities are a very crass way of capturing network structure
  - Are a *by-product* of network structure/processes, not what produces them
- Even if we have a directed graph, e.g. an advice network,
  - Modeling in-degree centrality would be getting at who is sought out
  - But not *by whom*
  - Out-degree centrality would be getting at who seeks out advice
  - But not *from whom*
- Model the *process* also to manage dependencies

Network: explanatory or response?

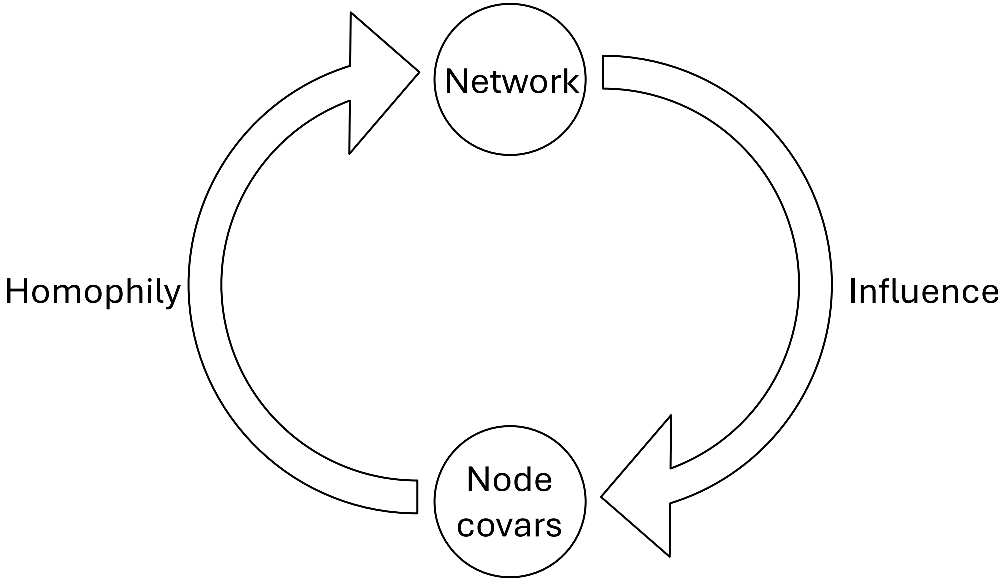
Network as cause? (as explanatory/IV?)



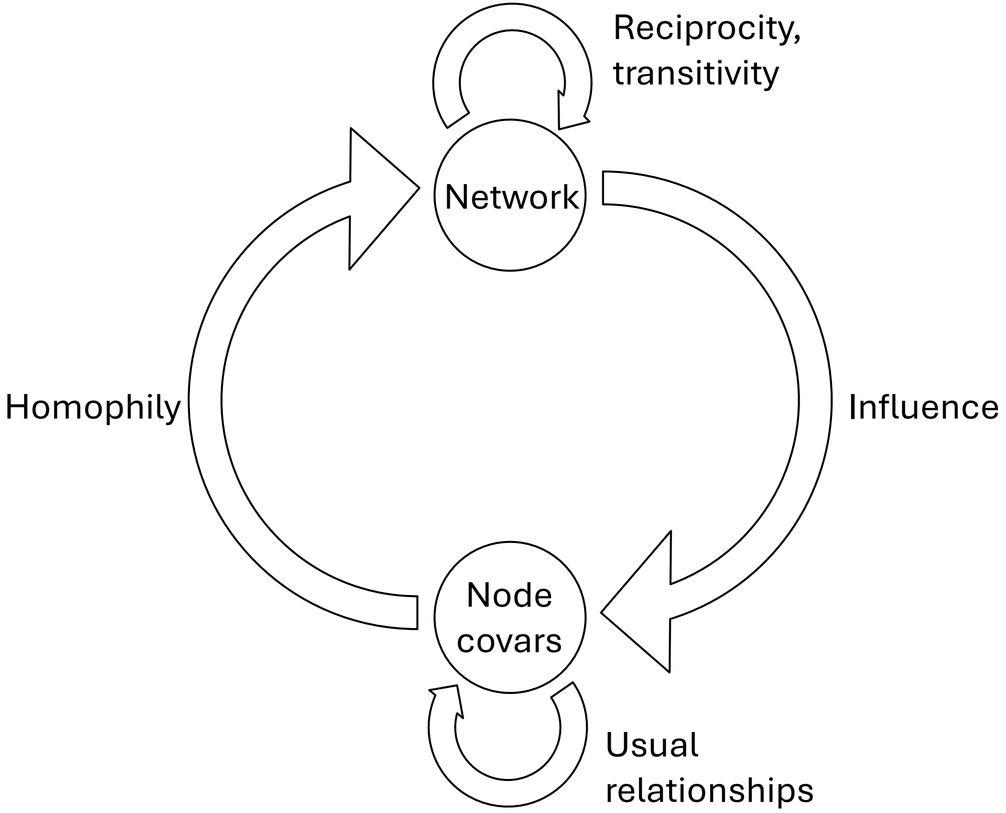
Network as effect? (as response/DV?)



The problem: both happen.



And they aren't the only things.



Technical vocabulary

- “Model misspecification”
  - The wrong functional form, and/or
  - The wrong variables
- Omitted variable bias (OVB)
- Synonymous: “Non-iid data,” “dependent data,” “autocorrelation,” “endogeneity,” “pseudoreplication”

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

Conclusions

Works cited

Dependencies

- From Wikipedia: “Asking two people in the same household whether they watch TV, for example, does not give you statistically independent answers. The sample size, n, for independent observations in this case is one, not two.”
- Networks are an interconnected entity, and **statistics cannot do anything with an n of 1.**
- So we have to make independence assumptions
- See https://www.mominmalik.com/sunbelt2019.pdf

Models of network structure

	Y	X_1	X_2	$\cdots$	X_d
1	y_1	x_{11}	x_{12}	$\cdots$	x_{1d}
2	y_2	x_{21}	x_{22}	$\cdots$	x_{2d}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	y_n	x_{n1}	x_{n2}	$\cdots$	x_{nd}

$\Downarrow$

index	*from*	*to*	Y	W_1	W_2	W_3	$\cdots$
e_1	1	2	y_{12}	$\mathbf{1}(x_{11} = x_{21})$	$x_{12} - x_{22}$	x_{13}	$\cdots$
e_2	2	3	y_{23}	$\mathbf{1}(x_{11} = x_{31})$	$x_{12} - x_{32}$	x_{13}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
e_{n+1}	2	1	y_{21}	$\mathbf{1}(x_{21} = x_{11})$	$x_{22} - x_{12}$	x_{23}	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$e_{2\binom{n}{2}}$	$n-1$	n	$y_{(n-1)n}$	$\mathbf{1}(x_{(n-1)1} = x_{n1})$	$x_{(n-1)2} - x_{n2}$	$x_{(n-1)3}$	$\cdots$

Model the *edges*

- For maybe two years, I didn't realize that you actually transform your data set
- The edges are dependencies between observations
- Problem: the edges are dependent, too!
- Transitivity, reciprocity, Dunbar's number: these are dependencies between dependencies
- Not only are we not measuring important forces, but we assume them away (get OVB!)
- (Useful language: "dyad dependent" vs. "dyad independent")

Exponential[-family] Random Graph Models (ERGMs)

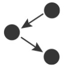
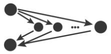
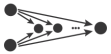
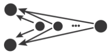
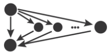
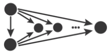
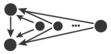
- The crown jewel of 30+ years of research, came out of p2 model
- (Main version treats graphs as the response: graphs as explanatory are called “autologistic actor attribute models” [ALAAMs], aren’t really done)
- Logic: specify a set of sufficient statistics, calculated over whole network
- These can include terms for anything you can think of
- By construction, these are the sufficient statistics for a graph. Question is if there is any weighting of these statistics that can produced the observed graph

Terms in ERGMs

$$S_1(y) = \sum_{1 \le i < j \le n} y_{ij} \qquad \text{number of edges}$$

$$S_k(y) = \sum_{1 \le i \le n} \binom{y_{i+}}{k} \qquad \text{number of } k\text{-stars } (k \ge 2)$$

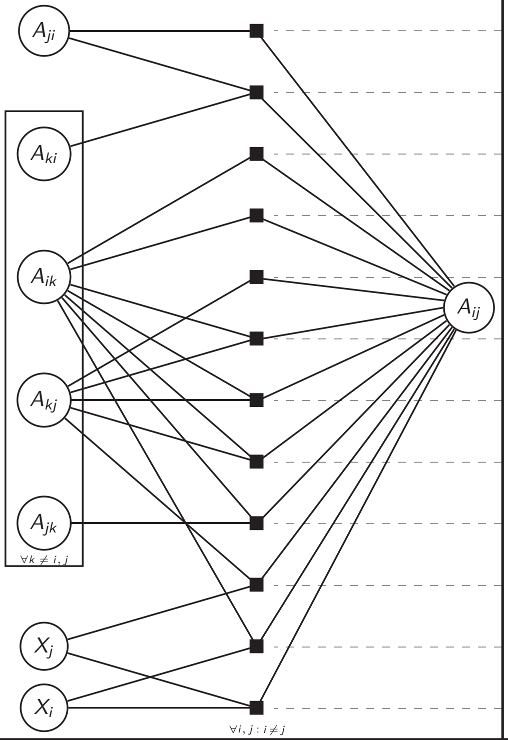
$$T(y) = \sum_{1 \le i < j < h \le n} y_{ij}\, y_{ih}\, y_{jh} \qquad \text{number of triangles}$$

Network statistics	Description	Structural signature
Univariate parameters		
Dyadic parameters		
Reciprocity	Occurrence of mutual ties	
Degree parameters		
Mixed 2-star	Correlation of indegrees and outdegrees	
Alternating-in-star (A-in-S)	Network centralisaiton around indegree	
Isolate	Occurrence of actors with zero indegree and zero outdegree	
Sink	Occurrence of actors with an outdegree of zero and indegree of at least one	
Triangle parameters		
Multiple connectivity (A2P-T)	Multiple paths of indirect connectivity	
Shared out-ties (A2P-U)	Activity based structural equivalence: multiple sets of out-ties to the same third others	
Shared in-ties (A2P-D)	Popularity based structure equivalence: multiple sets of in-ties from the same third others	
Transitive closure (AT-T)	Transitive closure of multiple 2-paths	
Activity closure (AT-U)	Closure of multiple in-2-stars	
Popularity closure (AT-D)	Closure of multiple out-2-stars	

Simpson (2015)

ERGMs as a graphical model

Factor graph



Parameter name	Network Motif	Parameterization	Matrix notation
mutual dyads		$\sum_{i<j} A_{ij}A_{ji}$	$\frac{1}{2}\,\mathrm{tr}\left(\mathbf{AA}^T\right)$
in-two-stars		$\sum_{(i,j,k)} A_{ji}A_{ki}$	$\mathrm{sum}\left(\mathbf{AA}^T\right) - \mathrm{tr}\left(\mathbf{AA}^T\right)$
out-two-stars		$\sum_{(i,j,k)} A_{ij}A_{ik}$	$\mathrm{sum}\left(\mathbf{A}^T\mathbf{A}\right) - \mathrm{tr}\left(\mathbf{A}^T\mathbf{A}\right)$
geom. weighted out-degrees	—	$\sum_i \exp\{-\alpha \sum_k A_{ik}\}$	$\mathrm{sum}\left(\exp\{-\alpha\ \mathrm{rowsum}\left(\mathbf{A}\right)\}\right)$
geom. weighted in-degrees	—	$\sum_j \exp\{-\alpha \sum_k A_{kj}\}$	$\mathrm{sum}\left(\exp\{-\alpha\ \mathrm{colsum}\left(\mathbf{A}\right)\}\right)$
alternating tran-sitive k-triplets		$\lambda \sum_{i,j} A_{ij}\left\{1-\left(1-\frac{1}{\lambda}\right)^{\sum_{k\neq i,j} A_{ik}A_{kj}}\right\}$	$\lambda\ \mathrm{sum}\left(\mathbf{A}\odot\left(1-\left(1-\frac{1}{\lambda}\right)^{\mathbf{AA}-\mathrm{diag}(\mathbf{AA})}\right)\right)$
alternating indep. two-paths		$\lambda \sum_{i,j}\left\{1-\left(1-\frac{1}{\lambda}\right)^{\sum_{k\neq i,j} A_{ik}A_{kj}}\right\}$	$\lambda\ \mathrm{sum}\left(1-\left(1-\frac{1}{\lambda}\right)^{\mathbf{AA}-\mathrm{diag}(\mathbf{AA})}\right)$
two-paths (mixed two-stars)		$\sum_{(i,k,j)} A_{ik}A_{kj}$	$\mathrm{sum}\left(\mathbf{AA}\right) - \mathrm{tr}\left(\mathbf{AA}\right)$
transitive triads		$\sum_{(i,j,k)} A_{ij}A_{jk}A_{ik}$	$\mathrm{tr}\left(\mathbf{AAA}^T\right)$
activity effect		$\sum_i X_i \sum_j A_{ij}$	$\mathrm{sum}\left(\mathbf{X}\odot\mathrm{rowsum}\left(\mathbf{A}\right)\right)$
popularity effect		$\sum_j X_j \sum_i A_{ij}$	$\mathrm{sum}\left(\mathbf{X}\odot\mathrm{colsum}\left(\mathbf{A}\right)\right)$
similarity effect		$\sum_{i,j} A_{ij}\left(1-\frac{\lvert X_i - X_j\rvert}{\max_{k,l}\lvert X_k - X_l\rvert}\right)$	$\mathrm{sum}\left(\mathbf{A}\odot\mathbf{S}\right)$

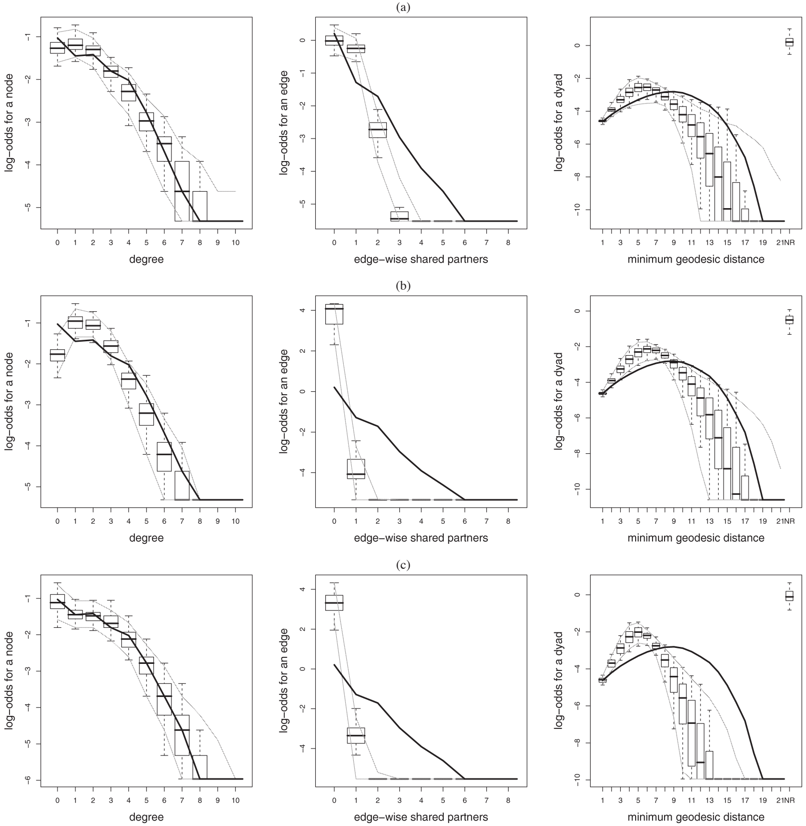
ERGMs: Procedure

- Take the observed graph, do counts of sufficient statistics, and initialize weights of terms (through logistic regression)
- Holding the rest of the graph constant, consider a single edge.
- How would removing this edge (if present) or adding it (if absent) change the count of sufficient statistics? Would a higher/lower count make the graph more likely based on current weights?
- If yes, adjust weights so that the observed graph remains most likely.
- Do this for some time to explore the parameter space (an MCMC procedure)
- At the end: if the terms put in were indeed the “correct” ones, these would be their weights

ERGMs: Goodness-of-fit testing

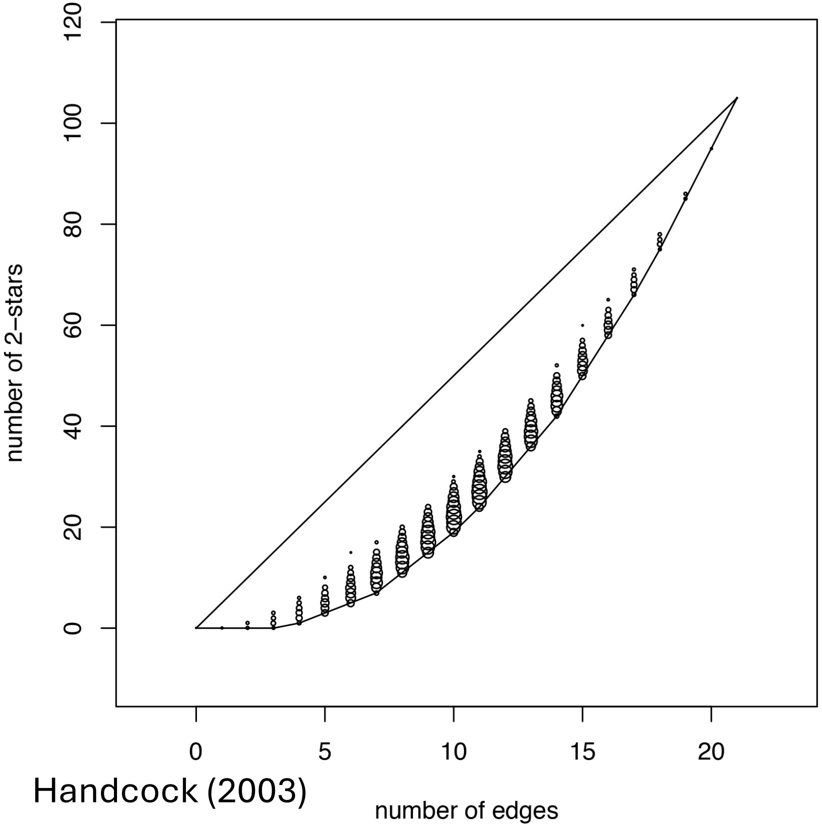
- Excellent goodness-of-fit (GOF) testing framework.
- See if the sufficient statistics that you put into the model can recover the distribution of statistics that were not among your sufficient statistics
- E.g., can density, reciprocity and transitivity alone as sufficient statistics recover the graph's degree distribution?
- Can test with anything (e.g., any subgraph/graph motif density), but should be theoretically important
- Gives a complete framework for finding a parsimonious explanation

Adamic & Glance (2005); Foucault Welles (2014)



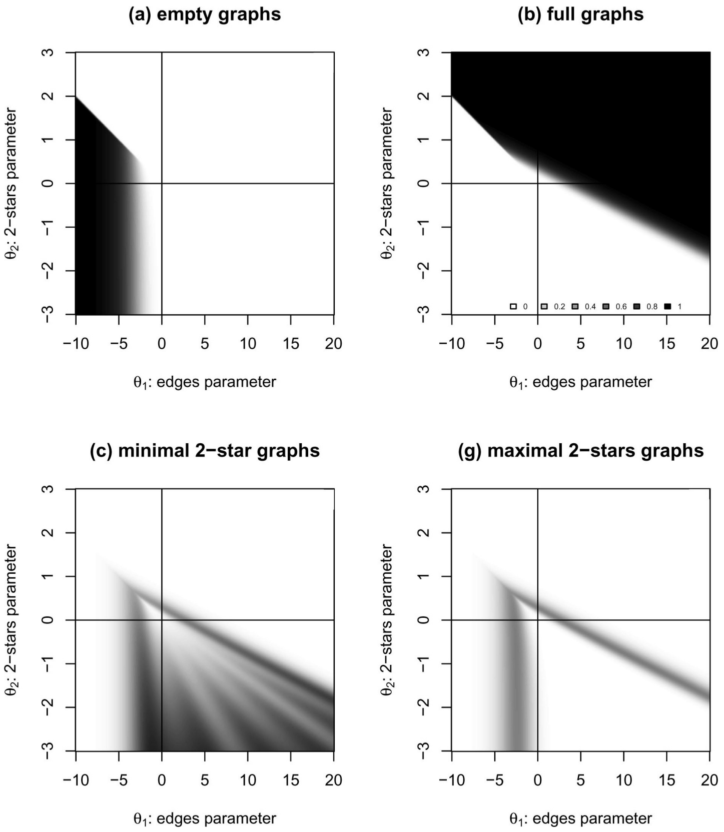
ERGMs: The bad news

- LOTS of problems.
- The space of graphs doesn't play nice with probabilities
- There are only a certain number of graphs of any given size, and only a certain number of graphs with a certain combination of sufficient statistics
- Slight differences in sampling frame can vastly change results (Shalizi & Rinaldo, 2013)



Handcock (2003)

ERGMs: The bad news



- Sometimes, under large portions of the parameter space, the most likely graph is either the complete graph or the empty graph: such specifications are are *degenerate*
- Because the space of graphs is so large, don't know if a model is degenerate or if our MCMC procedure is bad
- Model degeneracy (arguably) has nothing to do with the social phenomena of interest
- Better specifications are (arguably) technical, not sociological, entities: e.g., "geometrically weighted edgewise shared partners"

Models to avoid

- Comparing to an Erdős-Renyi random graph: Don’t do
- QAP: Don’t use
- Configuration model: Don’t use
- Graph neural networks are only random-walk based: Don’t use
- Small world networks, preferential attachment: aren’t actual *models*, just isolated mechanism (and don’t use them as a null model, either)
- Network autocorrelation: On the margin. Treats networks as a nuisance parameter

Models to use

- Latent space models: They "work" (statistically), but I find them useless
- Stochastic Block Models: Use, sometimes
- Relational event models: ERGMs for event data, as legit (or not) as ERGMs
- Stochastic Actor-Oriented Models: Similar to ERGMs but require timesteps, use something like agent-based simulation to fit parameters of a statistical models. Are also version of these for event data. Maybe the most statistically sophisticated and principled model

Conclusions

Larger context

- Networks (network science) had a lot of hype in the 2000s, much of which came from physicists (statistical mechanics) colonizing social science spaces and topics and methods, but that has mostly been eclipsed by ML and AI
- Network analysis is a specialized area, even if it cuts across disciplines; it is possible to have a few tools in your toolkit, but it can be hard to engage with the entirety of the literature without specific study

Take-aways

- Everything is terrible and nothing works... quantitatively
- Qualitative network analysis can use quantitative tools: to tell stories, to visualize networks to talk about them with research participants, to do *interpretive* analysis (heuristics might *coincide*, and so an be used to present, but I think hermeneutics are distinct from heuristics)

Why networks?

Represen-tation of networks

Disclaimers

Basic network terms

Heuristic analysis

Statistical analysis

Conclusions

Works cited

Your next steps

- Work through Kolaczyk & Csárdi, *Statistical Analysis of Network Data with R* (2nd edition). After using `igraph` to manipulate networks, they go over loading older R packages and running ERGMs and other models; all sections are great as a tutorial
- Learn to use Gephi and/or Cytoscape for visual interactive analysis
- Find places to get network data and incorporate into your work

Works Cited

Adamic, L. A., & Glance, N. (2005). The political blogosphere and the 2004 U.S. election: Divided they blog. In *Proceedings of the 3rd international workshop on link discovery* (LinkKDD '05) (pp. 36–43). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/1134271.1134277

Bearman, P. S., Moody, J., & Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, *110*(1), 44-91. https://doi.org/10.1086/386272

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. Science, 323, 892-895. https://doi.org/10.1126/science.1165821

Breiger, R. (1974). The duality of persons and groups. *Social Forces*, *53*, 181-190. https://doi.org/10.2307/2576011

Caron, F., & Fox, E. B. (2015). Sparse graphs using exchangeable random measures. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *79*(5), 1295–1366. https://doi.org/10.1111/rssb.12233

Cherry, F. (1995). One man's social psychology is another woman's social history. In: *The stubborn particulars of social psychology: Essays on the research process* (pp. 68–83). London: Routledge.

Davis, A., & Gardner, B. B. (1941). *Deep south: A social anthropological study of caste and class*. Chicago: University of Chicago Press.

Durland, M. M. (2005). A formative evaluation of the integration of two departments. *New Directions for Evaluation*, *107*, 81-94. https://doi.org/10.1002/ev.163

Edling, C. (2009). We always know more than we can say: Mathematical sociologists on mathematical sociology. *Frontiers of Sociology*, *11*, 345–367. https://doi.org/10.1163/ej.9789004165694.i-450.110

Erikson, E. (2013). Formalist and relationalist theory in social network analysis. *Sociological Theory*, *31*(3), 219-242. https://doi.org/10.1177/0735275113501998

Foucault Welles, B., & Meirelles, I. (2014). Visualizing Computational Social Science: The Multiple Lives of a Complex Image. *Science Communication*, *37*(1), 34-58. https://doi.org/10.1177/1075547014556540

Handcock, M. S. (2003). *Degeneracy in statistical models of social networks*. Working Paper No. 39, Center for Statistics and the Social Sciences, University of Washington.

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258. https://doi.org/10.1198/016214507000000446

Knoke, D. (2013). "It takes a network": The rise and fall of social network analysis in U.S. Army counterinsurgency doctrine. *Connections*, *33*, 1-10.

Krackhardt, D. (1996). Social networks and the liability of newness for managers. In *Trends in Organizational Behavior*, Volume 3 (pp. 159-173).

Krebs, V. E. (2002). Uncloaking terrorist networks. *first monday*, *7*(4). https://firstmonday.org/ojs/index.php/fm/article/view/941/863

Immarigeon, R. (2014). The 'ungovernable' Ella Fitzgerald. *Prison Public Memory Project*, October 29, 2014. https://www.prisonpublicmemory.org/blog/2014/the-ungovernable-ella-fitzgerald

Mac Ginty, R. (2010). Social network analysis and counterinsurgency: A counterproductive strategy? *Critical Studies on Terrorism*, *3*(2), 209–226. https://doi.org/10.1080/17539153.2010.491319

Marcum, C. S., Lin, J., & Koehly, L. (2016). Growing-up and coming-out: Are 4-cycles present in adult hetero/gay hook-ups? *Network Science*, *4*(3), 400-405. https://doi.org/10.1017/nws.2016.3

Moreno, J. L. (1934). *Who shall survive? A new approach to the problem of human interrelations*. Nervous and Mental Disease Publishing Co.

Newman, M. E. J. (2009). *Networks: An introduction*. Oxford: Oxford University Press.

Pitts, F. R. (1978/9). The medieval river trade network of Russia revisited. *Social Networks*, *1*, 285–292. https://doi.org/10.1016/0378-8733(78)90025-4

Porter, M. A., Onnela, J.-P., & Mucha, P. J. (2009). Communities in networks. *Notices of the AMS*, *56*(9), 1082–1166. https://people.maths.ox.ac.uk/~porterm/papers/comnotices.pdf

Shalizi, C. R., & Thomas, A. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, *40*(2), 211–239. https://doi.org/10.1177/0049124111404820

Shalizi, C. R., & Rinaldo, A. (2013). Consistency under sampling of exponential random graph models. *The Annals of Statistics*, *41*(2), 508–535. http://dx.doi.org/10.1214/12-AOS1044

Simpson, C. R. (2015). Multiplexity and strategic alliances: The relational embeddedness of coalitions in social movement organisational fields. Social Networks, 42, 42–59. https://doi.org/10.1016/j.socnet.2015.02.007

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, *33*(4), 452-473. https://doi.org/10.1086/jar.33.4.3629752