"I always dream of a pen that would be a syringe." — Jacques Derrida

BEHAVIORAL NETWORK SCIENCE

Thomas Hills
University of Warwick

Hohenheim 2023

WARWICK
THE UNIVERSITY OF WARWICK

The Alan Turing Institute

THE ROYAL SOCIETY

Background

- **Computational Social Sciences (Cognitive Science):** learning, memory, language evolution, aging

- **English (BA) and Biology (BS), Biology (PhD):**
  - mathematical biology and neuroscience of cognitive control
  - process and environment (structure)

- **Methods:**
  - **network analysis** (lexical and cognitive structure)
  - **computational modeling** (explanatory process models)
  - **natural language processing** (derive representations from language)

Some of my latest work

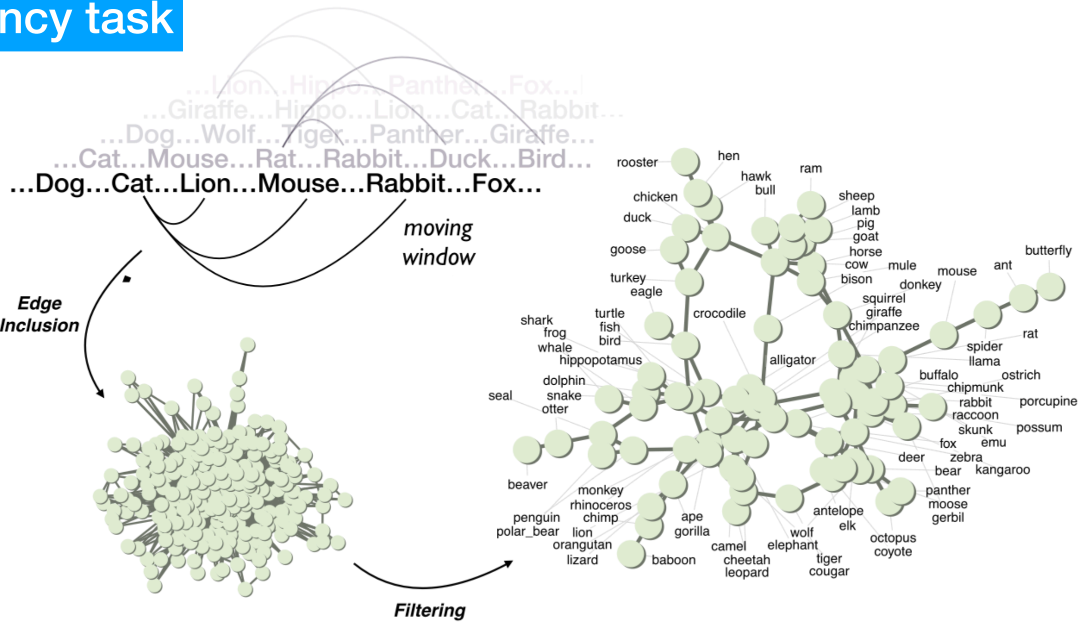
scientific reports

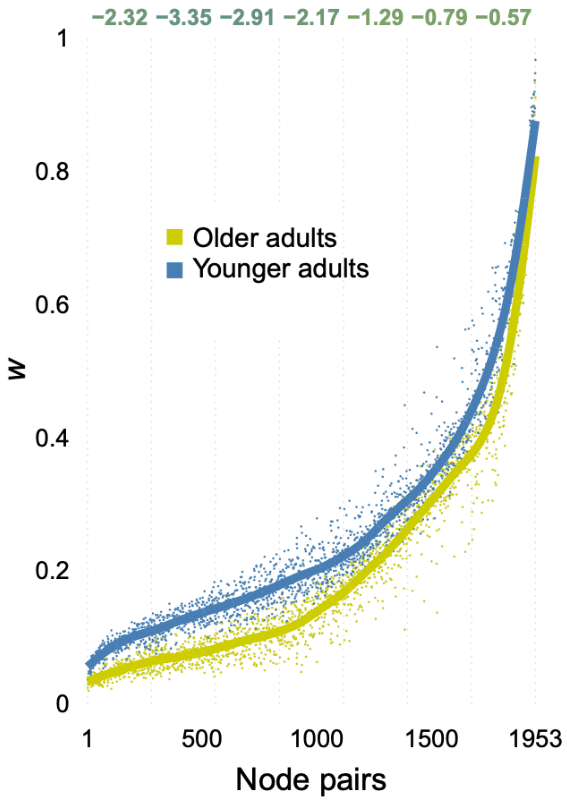
OPEN

Structural differences in the semantic networks of younger and older adults

Dirk U. Wulff[1,2], Thomas T. Hills[3] & Rui Mata[1,2]

Fluency task





Infer that older adult memory is less well connected. Specifically, relationships between words are weaker.

Some of my latest work

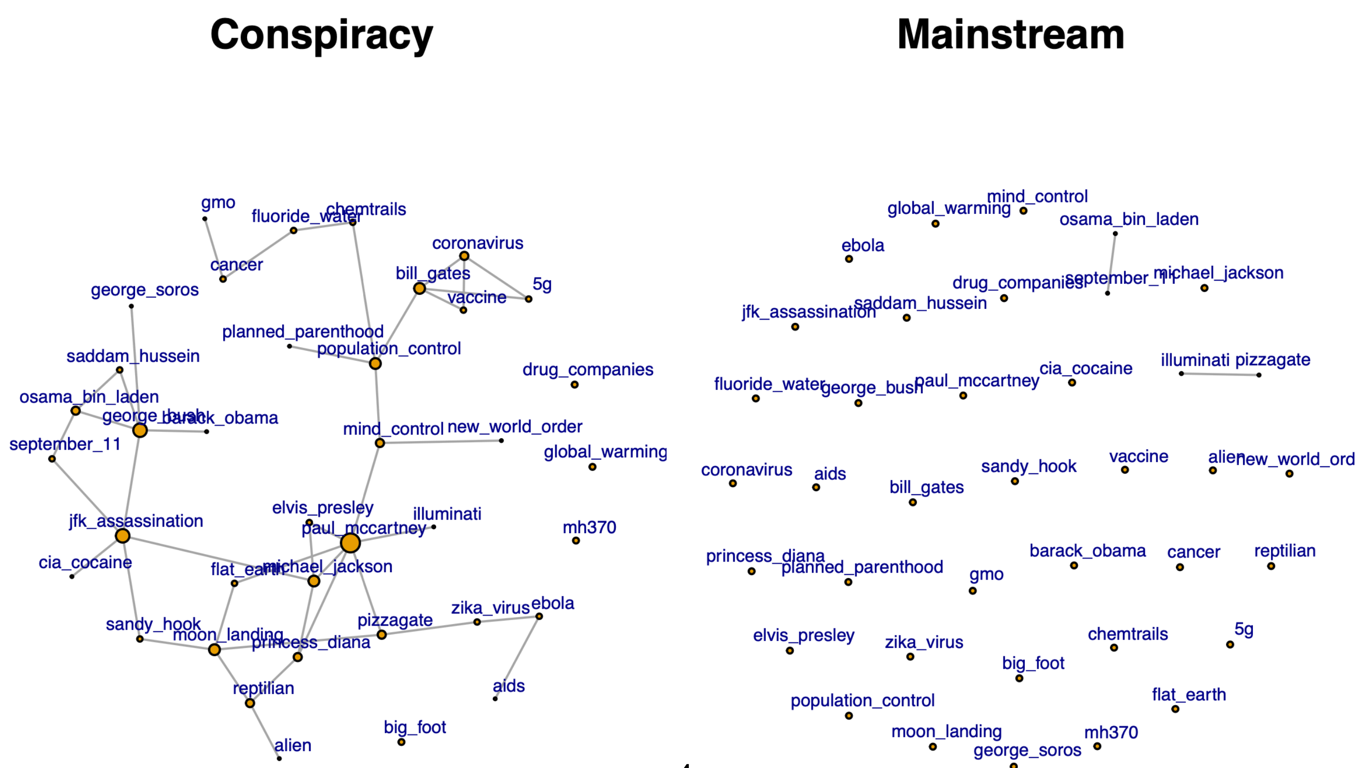
SCIENCE ADVANCES | RESEARCH ARTICLE

NEUROSCIENCE

Interconnectedness and (in)coherence as a signature of conspiracy worldviews

Alessandro Miani[1]*, Thomas Hills[2,3], Adrian Bangerter[1]

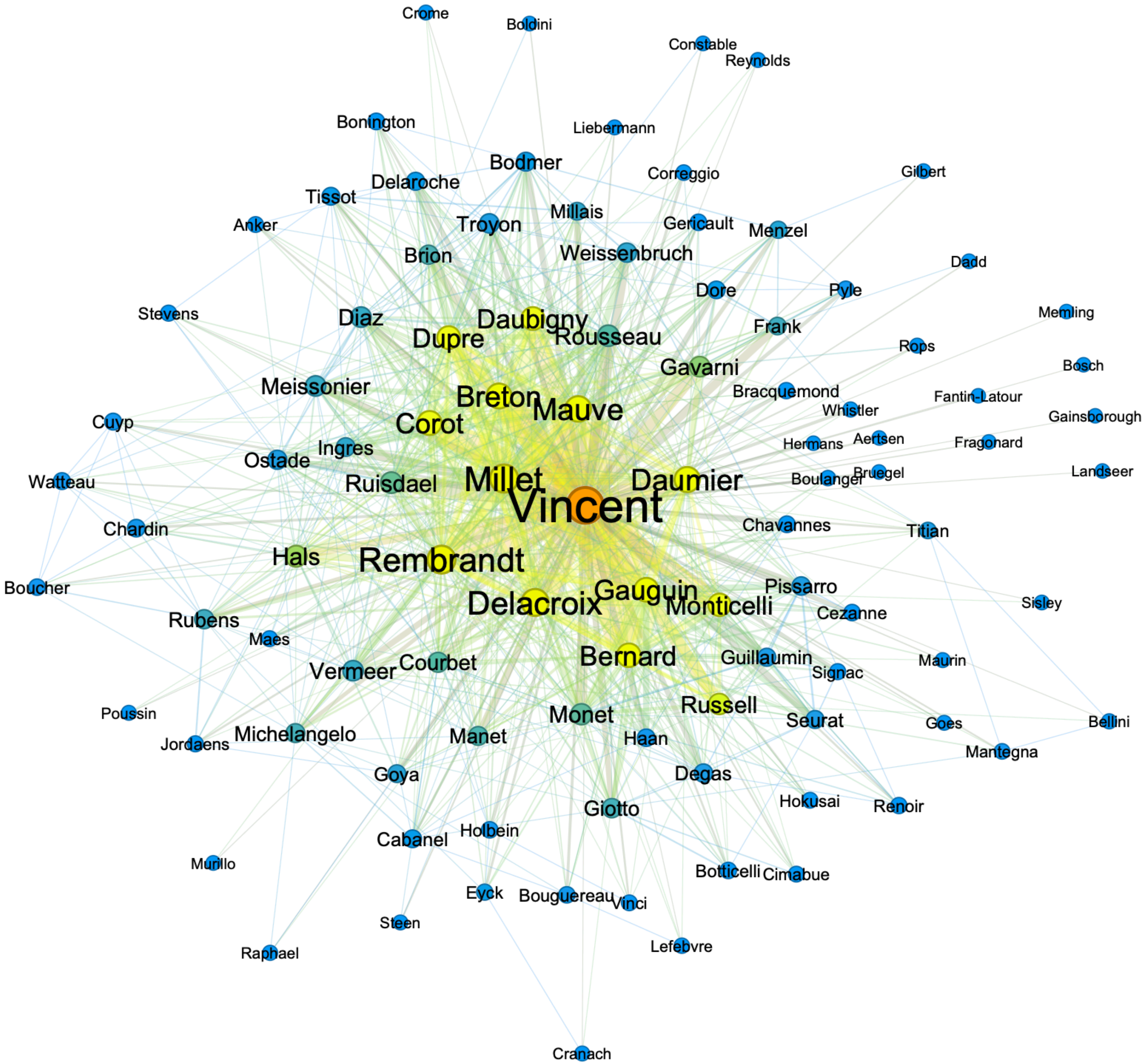
Conspiracy theories may arise out of an overarching conspiracy worldview that identifies common elements of subterfuge across unrelated or even contradictory explanations, leading to networks of self-reinforcing beliefs. We test this conjecture by analyzing a large natural language database of conspiracy and nonconspiracy texts for the same events, thus linking theory-driven psychological research with data-driven computational approaches. We find that, relative to nonconspiracy texts, conspiracy texts are more interconnected, more topically heterogeneous, and more similar to one another, revealing lower cohesion within texts but higher cohesion between texts and providing strong empirical support for an overarching conspiracy worldview. Our results provide inroads for classification algorithms and further exploration into individual differences in belief structures.



Today

Morning Goal: Getting started with network science

Afternoon Goal: Getting our hands dirty with R and your own data (or some of my data on Van Gogh, if you prefer)


Crome
Boldini
Constable
Reynolds
Bonington
Liebermann
Bodmer
Correggio
Gilbert
Delaroche
Tissot
Millais
Anker
Troyon
Gericault
Menzel
Brion
Weissenbruch
Dadd
Dore
Pyle
Stevens
Diaz
Daubigny
Memling
Dupre
Rousseau
Frank
Rops
Gavarni
Bosch
Meissonier
Bracquemond
Fantin-Latour
Breton
Whistler
Mauve
Gainsborough
Cuyp
Corot
Aertsen
Ingres
Hermans
Fragonard
Ostade
Watteau
Ruisdael
Millet
Daumier
Boulanger
Bruegel
Landseer
Vincent
Chardin
Chavannes
Titian
Hals
Rembrandt
Boucher
Pissarro
Gauguin
Delacroix
Monticelli
Sisley
Cezanne
Rubens
Maes
Bernard
Guillaumin
Maurin
Vermeer
Courbet
Signac
Poussin
Russell
Monet
Seurat
Goes
Bellini
Michelangelo
Jordaens
Manet
Haan
Mantegna
Goya
Degas
Hokusai
Renoir
Giotto
Holbein
Cabanel
Murillo
Botticelli
Cimabue
Eyck
Bouguereau
Vinci
Steen
Lefebvre
Raphael
Cranach

- Introductions

- What area of research you’re interested in.

- What kind of data do you have/collect?

- What you hope to learn from this course?

One drive folder

This has all the Rcode, readings, slides, and book chapters I will cover.

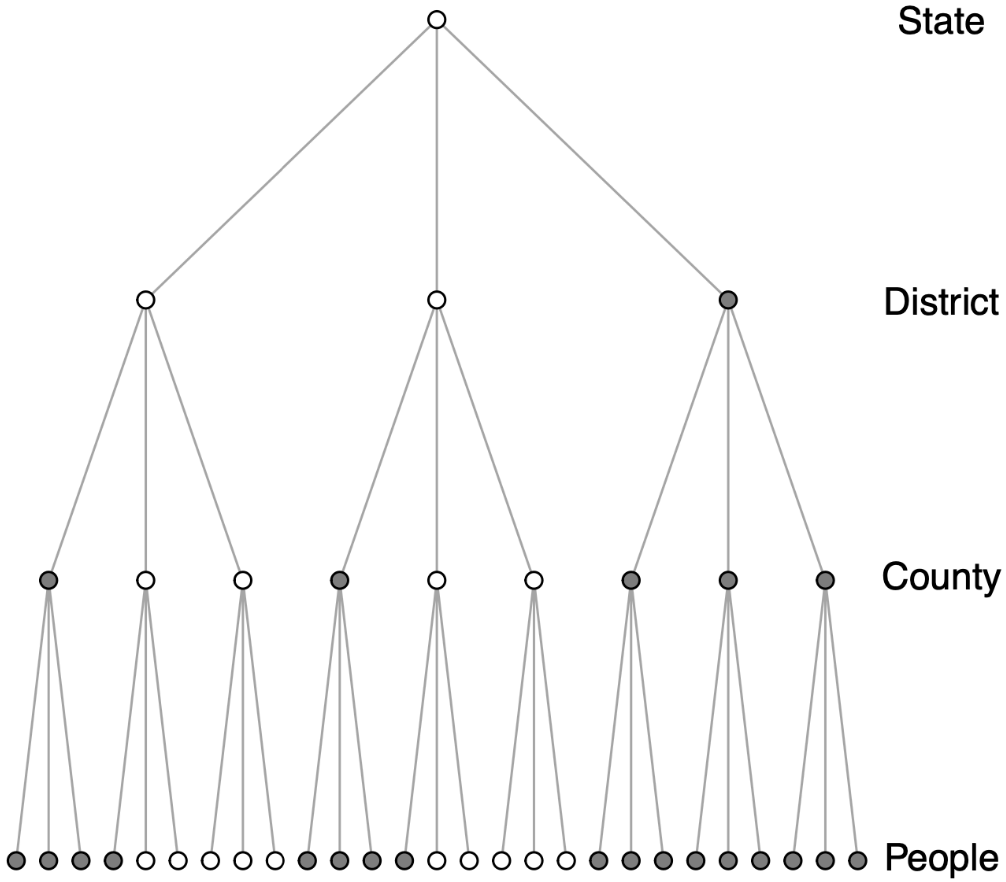
Two problems that are the focus of the this workshop:

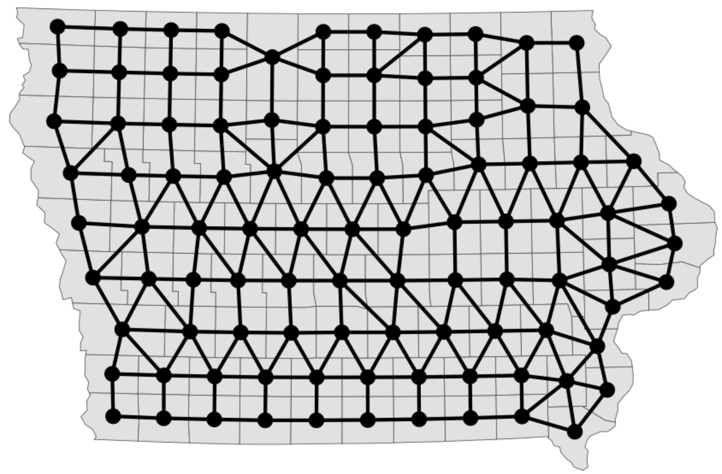
Big problem: How do we think about structure in our data?

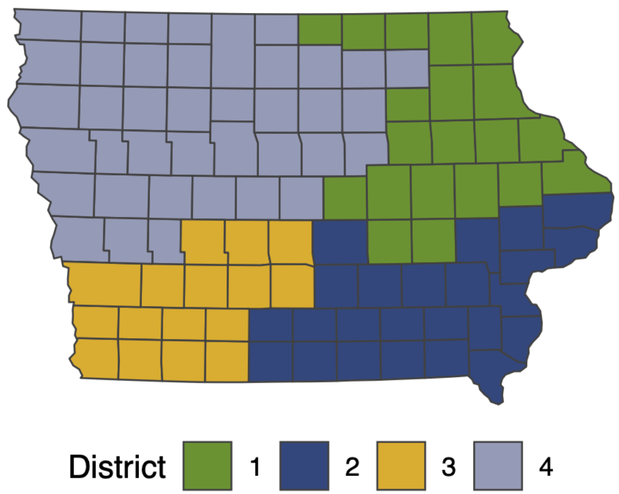
Small question: What are the methods for thinking about quantifying structure in our data?

Three quizzes

How to make the white nodes win?







Friendship paradox

Why do you have fewer friends than your friends do on average?

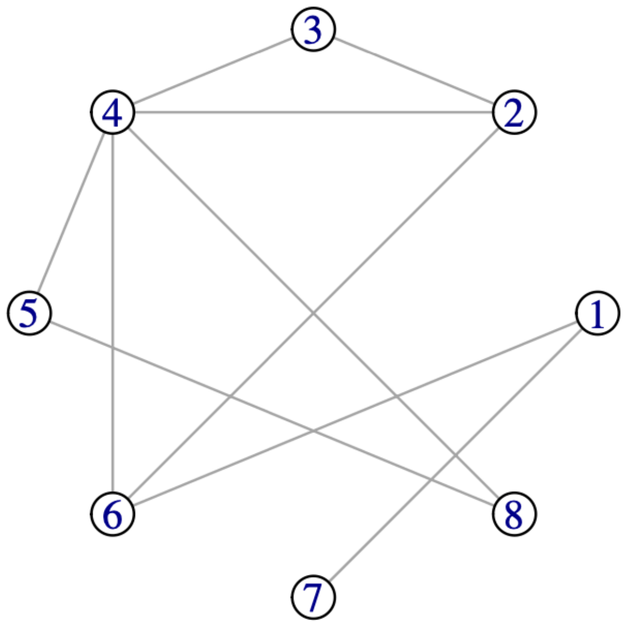


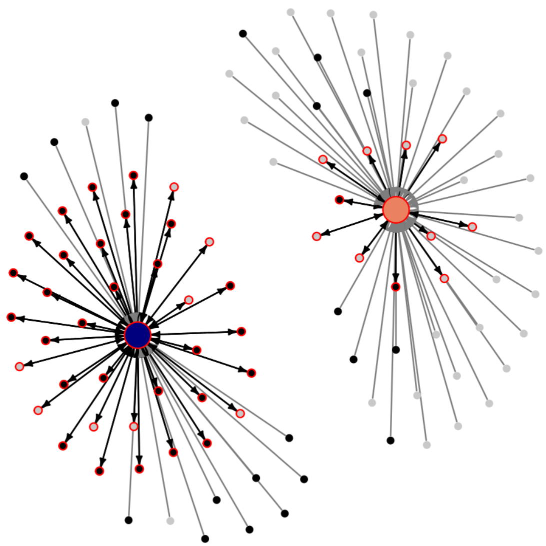
Table 1: The friendship paradox.

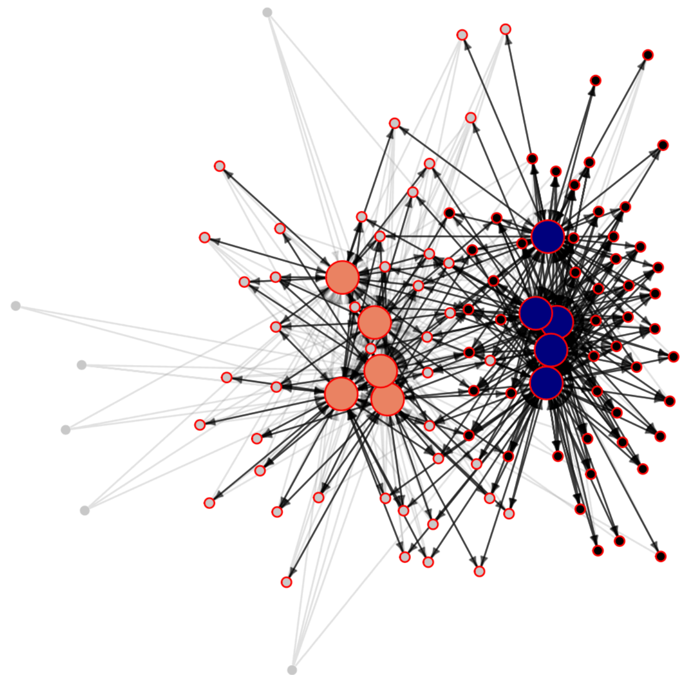
node	myFriends	theirFriends
1	2.0	2.00
2	3.0	3.33
3	2.0	4.00
4	5.0	2.40
5	2.0	3.50
6	3.0	3.33
7	1.0	2.00
8	2.0	3.50
Average	2.5	3.01

Simpson’s paradox

Please explain this: Graduate admissions figures from the University of California, Berkley in 1973: Across all applicants, men were more likely to be admitted than women (men = 44% and women=35%).

Women were more likely to be accepted than men in most of the departments to which they applied. However, they also applied to more competitive departments, which accepted fewer students overall.





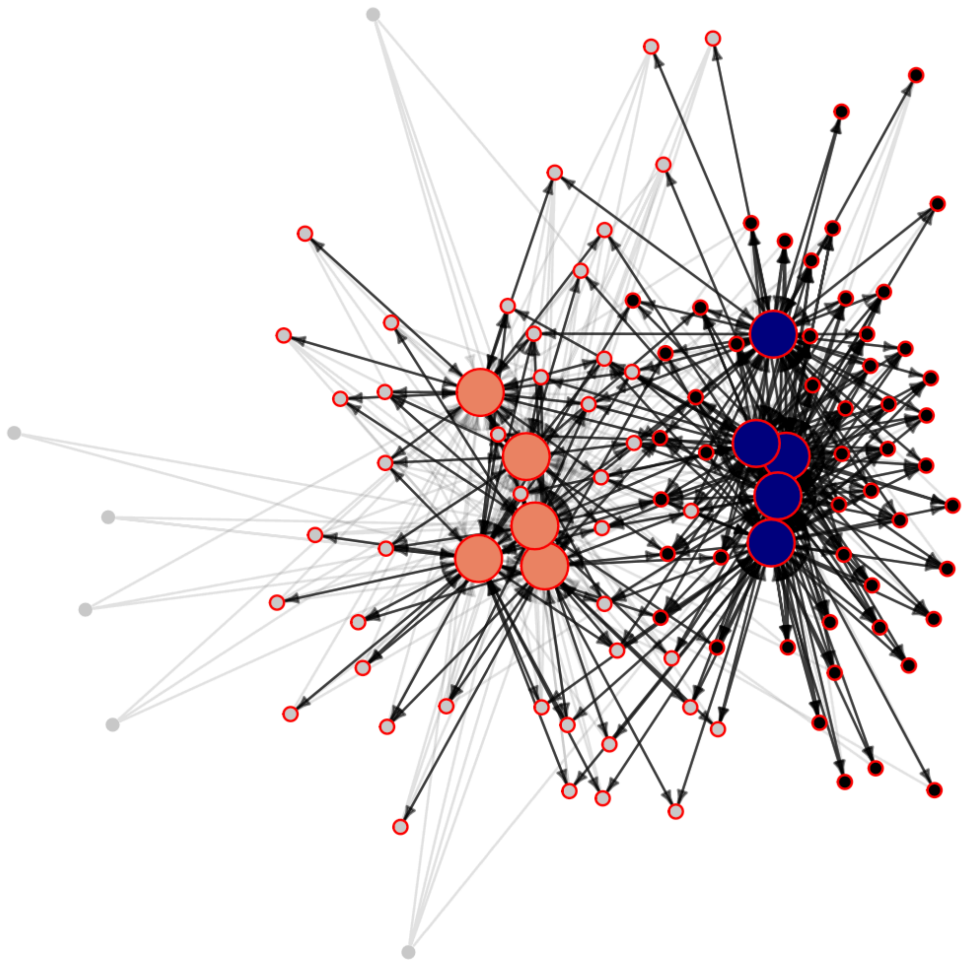


Figure 8: The outcome network for 100 simulated Australian Martu (small nodes) who can each choose multiple strategies (large nodes) from a set of 10 possible strategies, which are high (light) and low (dark) risk. Individuals are divided equally between men (gray) and women (black). Men prefer high risk strategies and women prefer low risk strategies. Each chooses each preferred strategy with probability $p = .8$, choosing the low risk strategy with $p = .2$. Unsuccessful strategies are shown with light gray edges, successful strategies are shown in dark black. Individuals that are successful at least once are circled in red. Low and high risk strategies are assigned a risk level from a beta distribution.

Things I can do with networks

- 1. Measure - data driven
- 2. Model these systems — what motivates this structure?
  - Process model?
  - Agent-based modelling
  - …

Network Basics

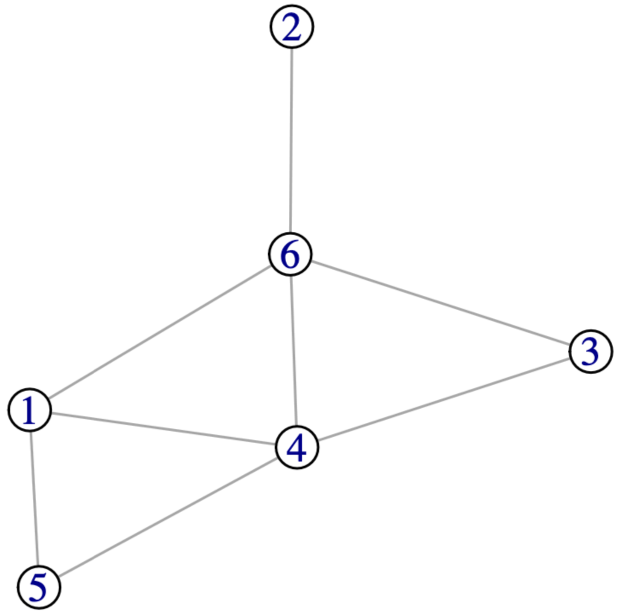
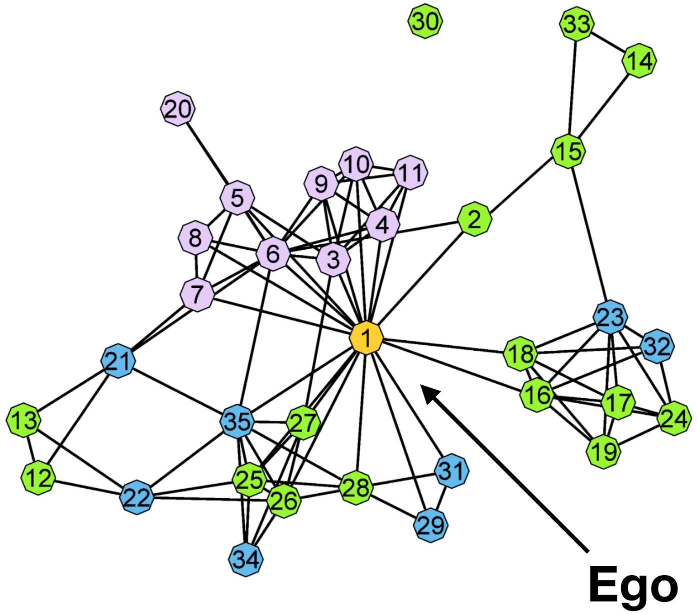
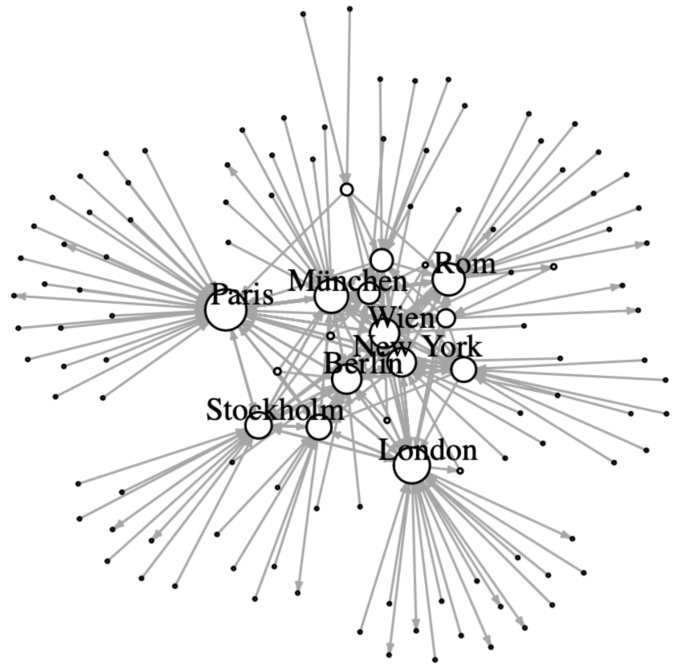


Figure 1: A simple network.

Network Basics



Ego Network



Full Network

Hills & Pachur, 2013

Schich et al., 2014

Nodes and Edges

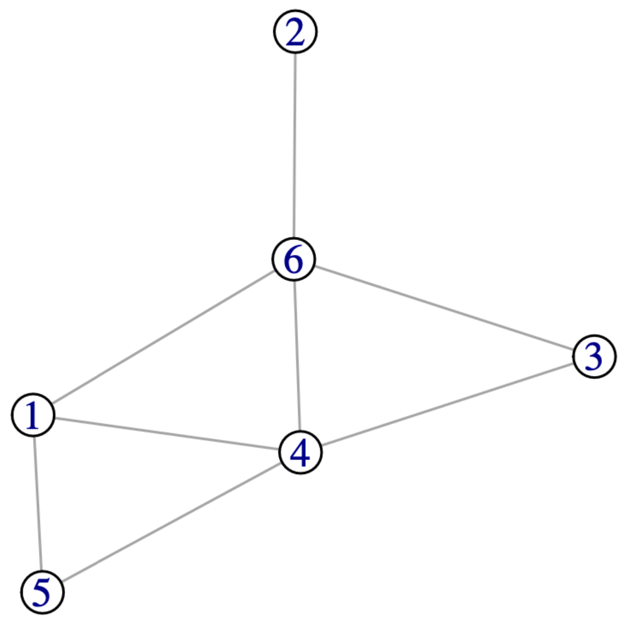
- What are some nodes?
- What are some edges?
- Think about this for your data.

Examples

Network	Nodes	Edges
Social	People	Friendships
Jazz	Musicians	Bands
Jazz	Musicians	Genre
Jazz	Wikipedia pages	hyperlinks
Cognitive Social	People	Remembered relationships
Citation	Articles	Citations
Semantic	Words	Similarity
		Phonology
		Shared features
		Co-occurrence in text
Free Association	Words	Associations
Brain	Neurons	Connectivity
	Modules	Correlated activity
Protein	Proteins	Interactions

Network Basics

- What is a network
- Edge lists
- Adjacency Matrices
- 4 kinds of networks
- Thresholding
- Attributes
- Bipartite Networks
- Multiplex Networks



Representing a simple network

- Edge list
-

V1	V2
1	4
3	4
1	5
4	5
1	6
2	6
3	6
4	6

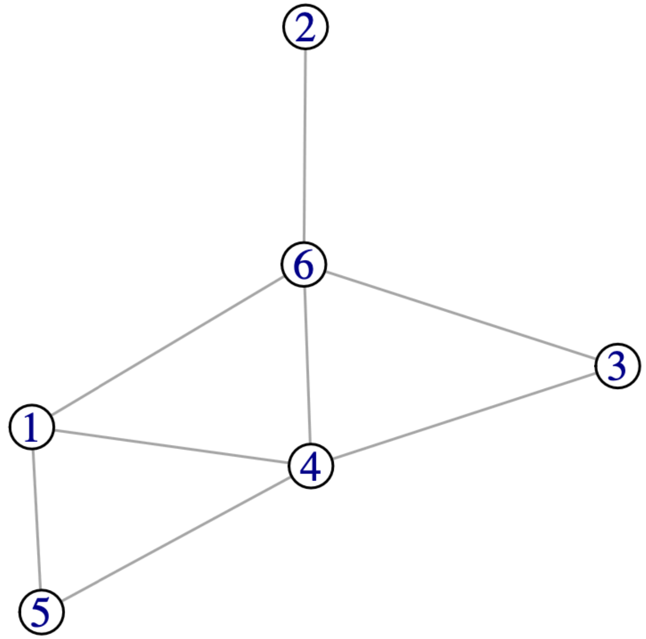


Figure 1: A simple network.

Representing a simple network

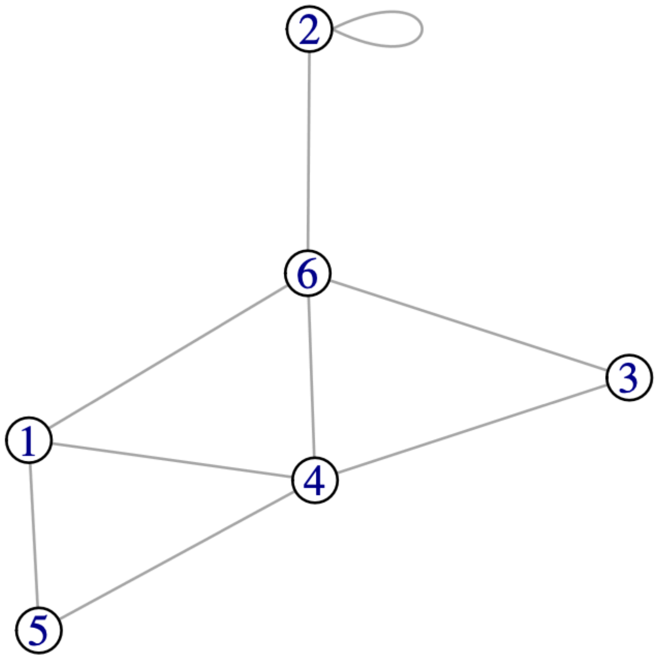
- Edge list

V1	V2
1	4
3	4
1	5
4	5
1	6
2	6
3	6
4	6

- Adjacency Matrix

	1	2	3	4	5	6
1	0	0	0	1	1	1
2	0	0	0	0	0	1
3	0	0	0	1	0	1
4	1	0	1	0	1	1
5	1	0	0	1	0	0
6	1	1	1	1	0	0

Self-loops

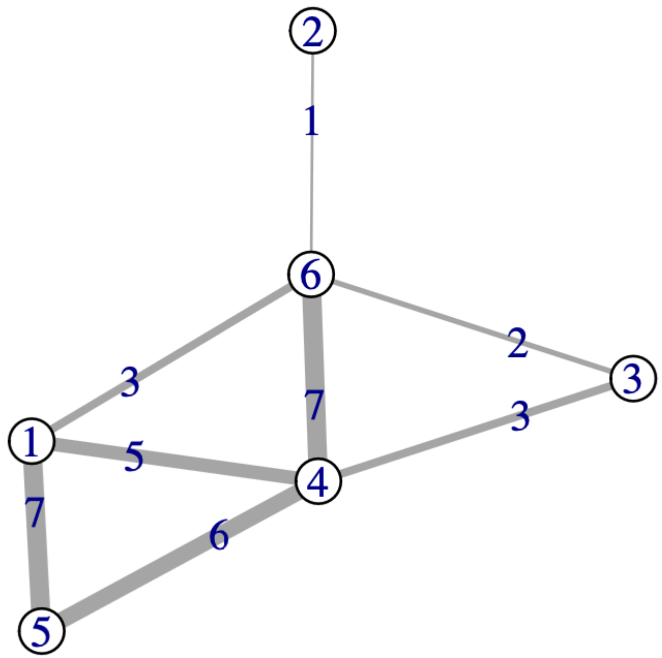


	1	2	3	4	5	6
1	0	0	0	1	1	1
2	0	1	0	0	0	1
3	0	0	0	1	0	1
4	1	0	1	0	1	1
5	1	0	0	1	0	0
6	1	1	1	1	0	0

Diagonal

What might self-loops be good for?

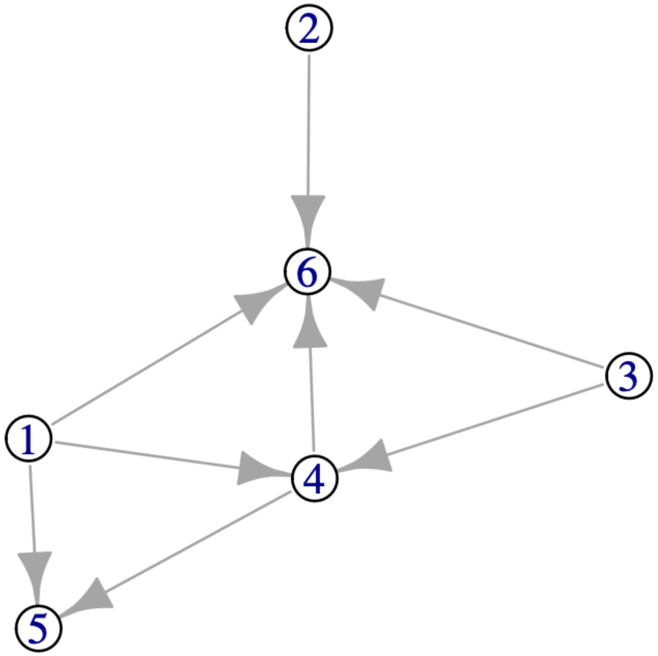
Weighted Networks



	1	2	3	4	5	6
1	0	0	0	5	7	3
2	0	0	0	0	0	1
3	0	0	0	3	0	2
4	5	0	3	0	6	7
5	7	0	0	6	0	0
6	3	1	2	7	0	0

V1	V2	weight
1	4	5
3	4	3
1	5	7
4	5	6
1	6	3
2	6	1
3	6	2
4	6	7

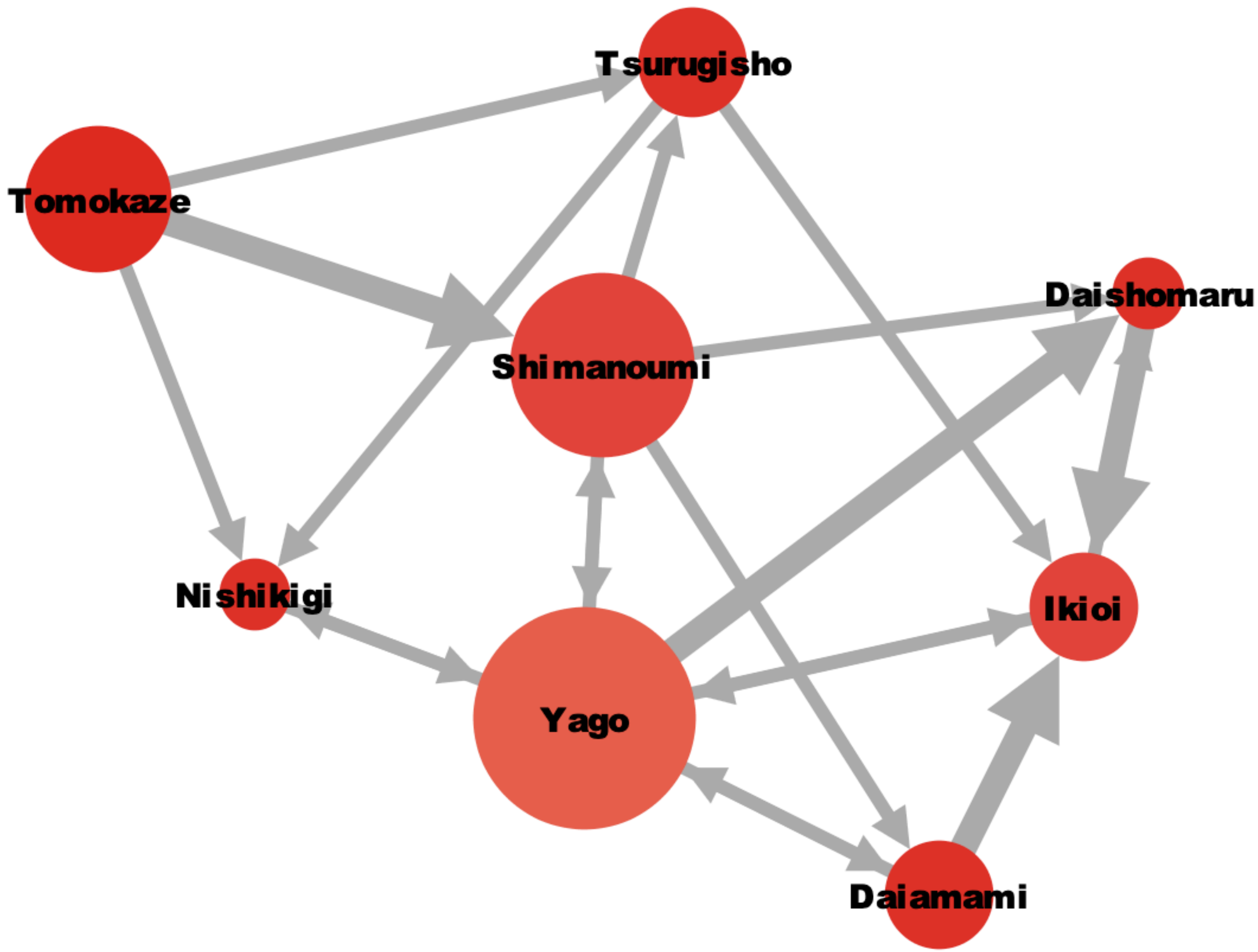
Directed Networks



	1	2	3	4	5	6
1	0	0	0	1	1	1
2	0	0	0	0	0	1
3	0	0	0	1	0	1
4	0	0	0	0	1	1
5	0	0	0	0	0	0
6	0	0	0	0	0	0

V1	V2
1	4
3	4
1	5
4	5
1	6
2	6
3	6
4	6

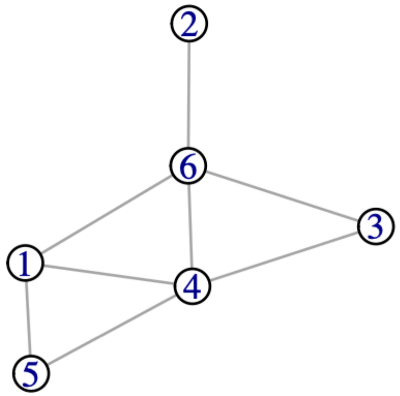
Application of directed network



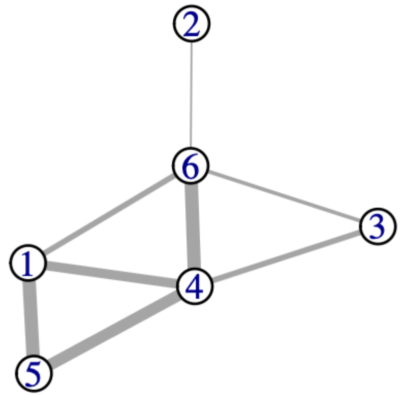
Data from Data.World Sumo Matches 2019

4 kinds of Networks

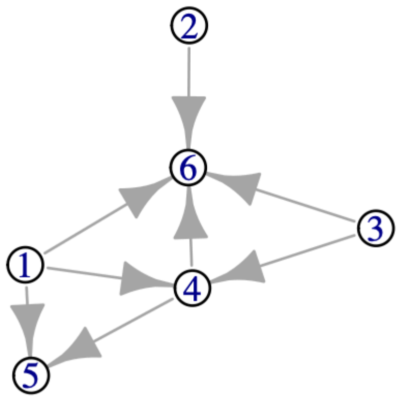
Simple: Unweighted, Undirected



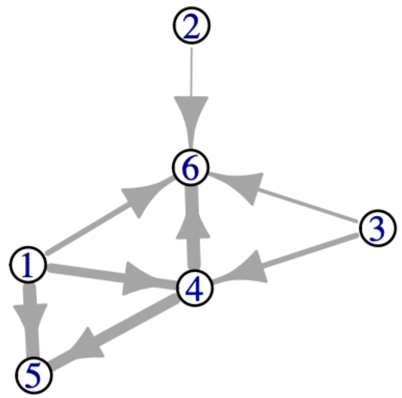
Weighted, Undirected



Directed, Unweighted

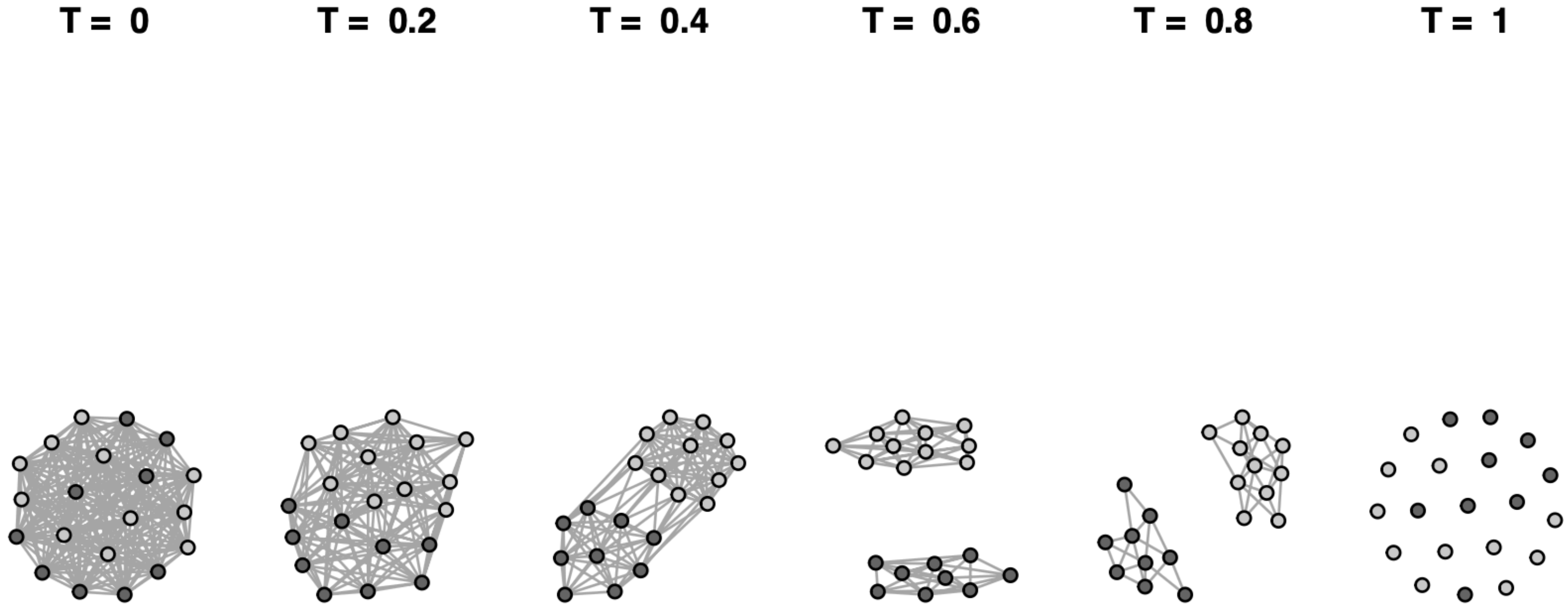


Weighted, Directed



How to get a simple network from weighted data?

- Apply a moving threshold. Keep edges above threshold.



Application of moving threshold to aging networks

- Apply a moving threshold. Keep edges above threshold.

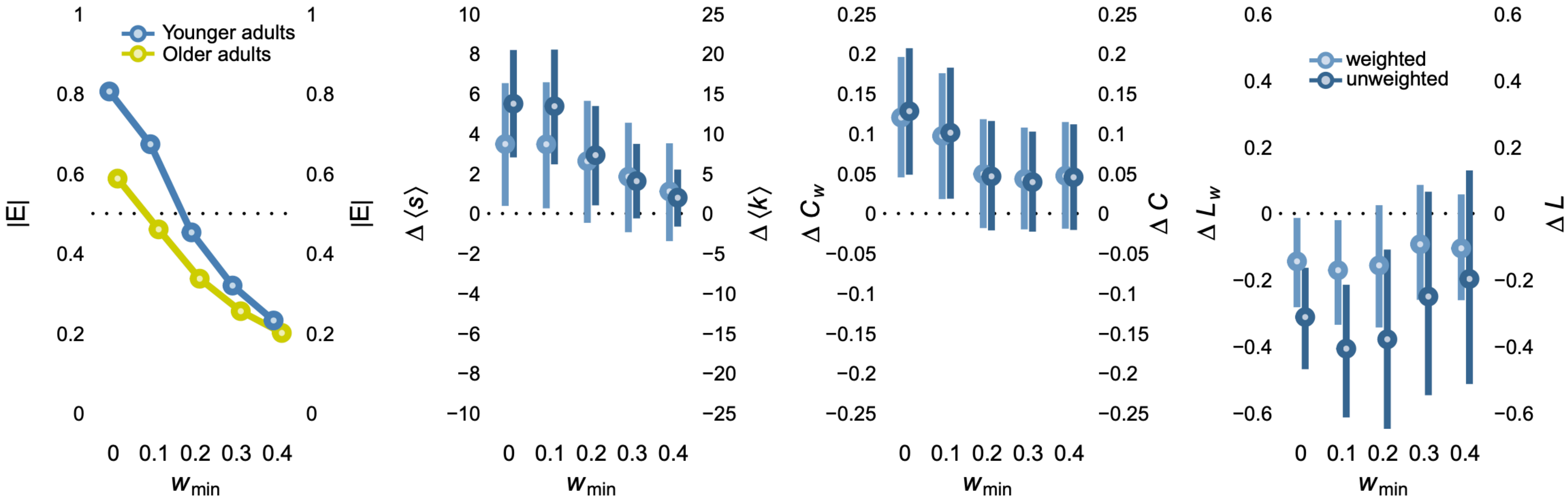


Figure 5. Differences in the macroscopic structure of younger and older adults' similarity rating networks. Blue and yellow circles, in panel 1, correspond to younger and older adults, respectively. In panels 2 to 4, light blue circles and dark blue circles correspond to differences between the younger and older adults' networks derived from weighted and unweighted networks, respectively. Error bars show 95% bootstrapped confidence intervals. Note: $|E|$ - Proportion of edges relative to fully-connected graph; $\Delta\langle s\rangle$, $\Delta\langle k\rangle$ - Differences in average strengths/degrees (unweighted); ΔC_w, ΔC - Difference in average clustering coefficients of weighted/unweighted networks; ΔL_w, ΔL - Difference in average shortest path lengths of weighted/unweighted networks.

How to get a simple network from a directed network?

- Keep all edges or only reciprocal edges

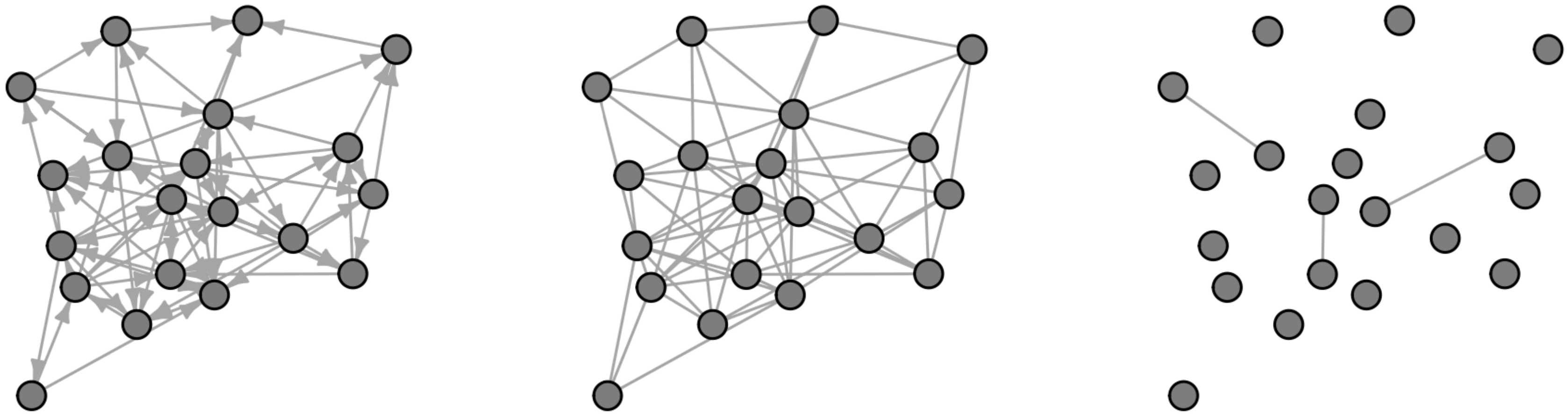
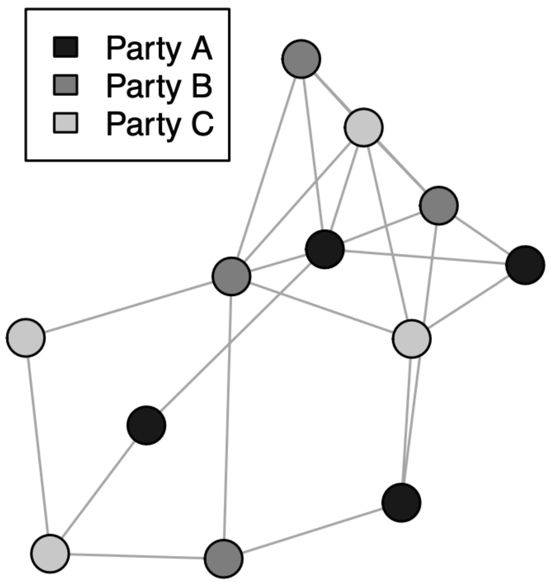


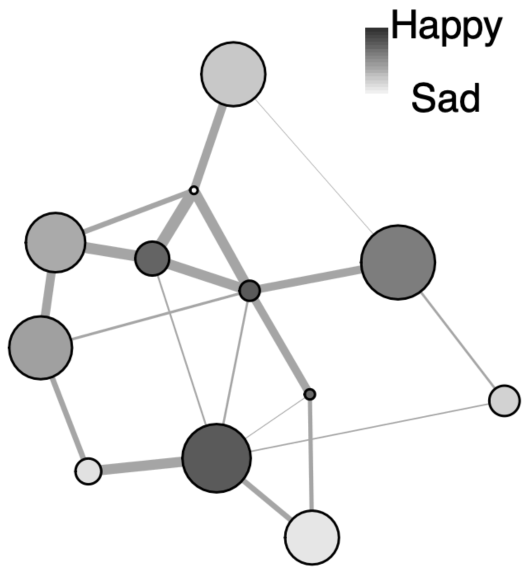
Figure 6: Thresholding directed networks. The network on the left shows the full directed network. The middle network transforms this into a simple, undirected network, making an edge wherever at least one node has a directed edge to the other. The network on the right only keeps reciprocal edges.

Node and edge attributes

Political Party

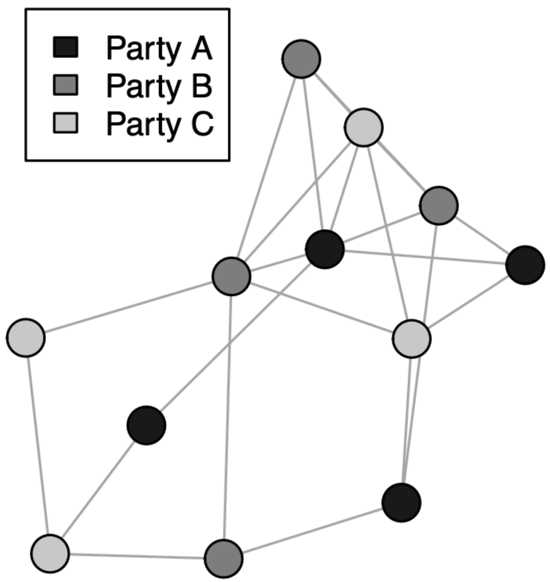


Age

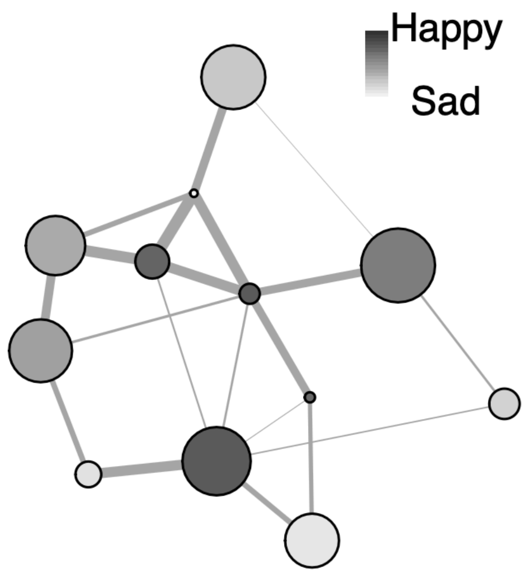


Node and edge attributes

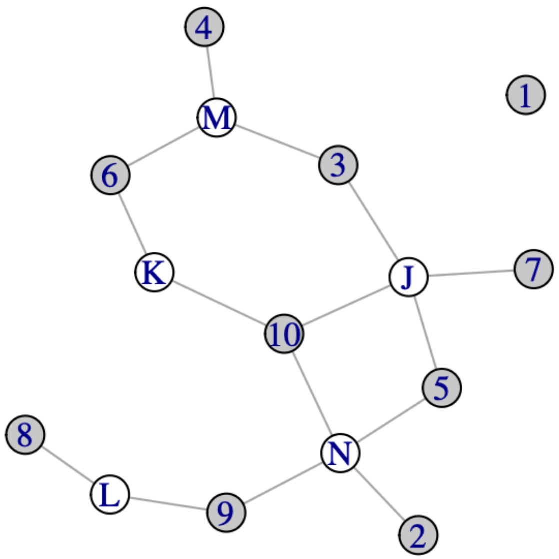
Political Party

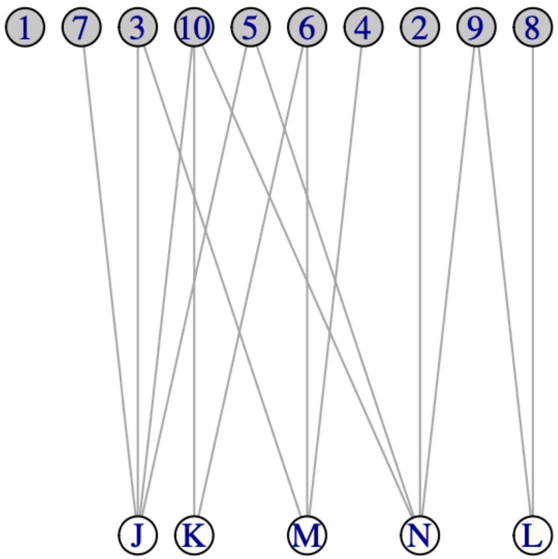


Age



Bipartite Networks





	J	K	L	M	N
1	0	0	0	0	0
2	0	0	0	0	1
3	1	0	0	1	0
4	0	0	0	1	0
5	1	0	0	0	1
6	0	1	0	1	0
7	1	0	0	0	0
8	0	0	1	0	0
9	0	0	1	0	1
10	1	1	0	0	1

Figure 8: A single bipartite network plotted in two different ways.

Bipartite Projections

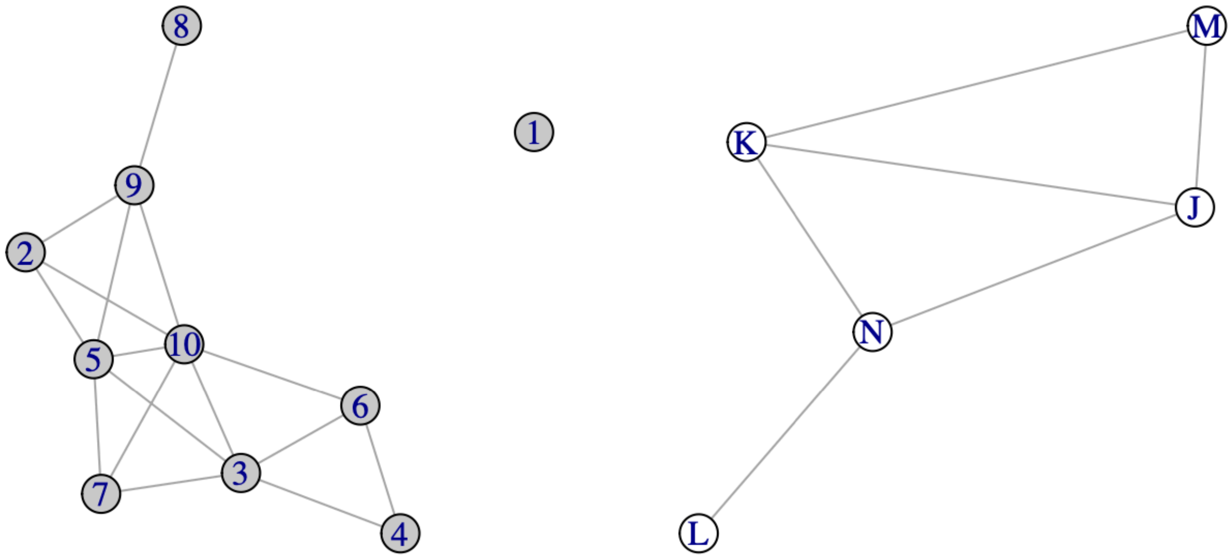
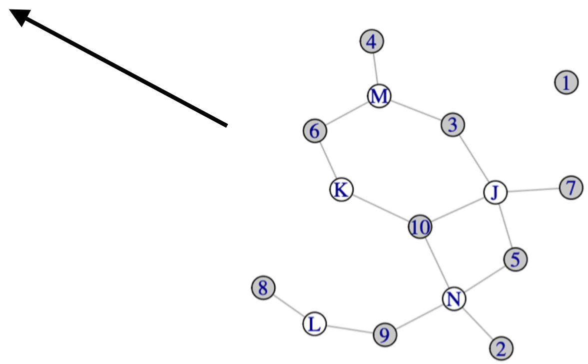


Figure 9: One-mode projections of the two-mode (bipartite) network.

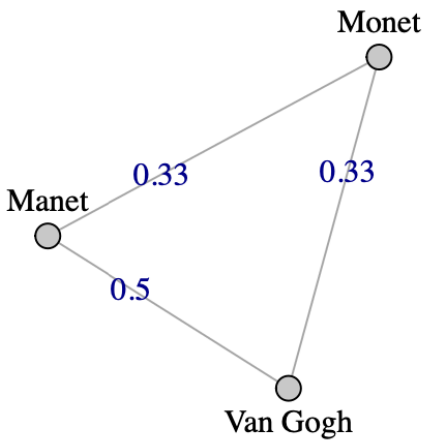


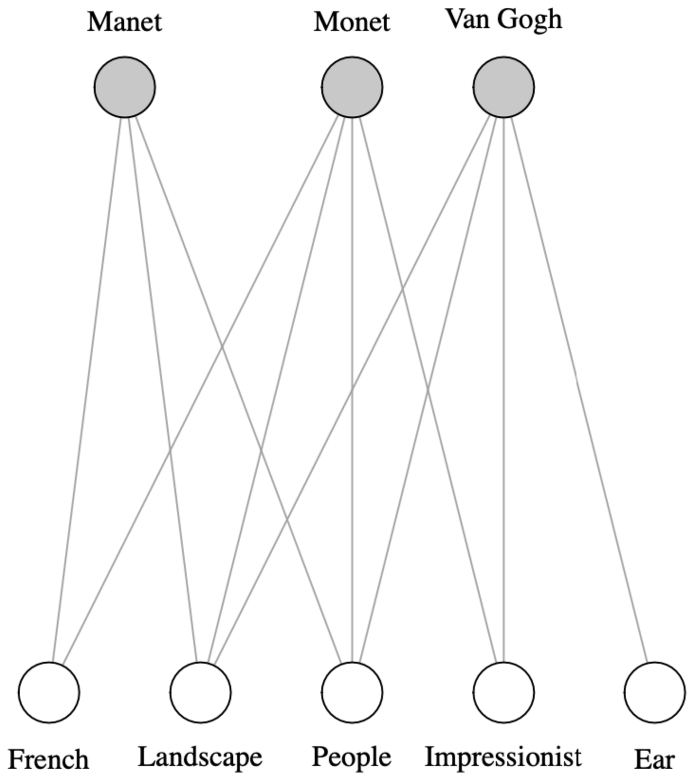
Application of bipartite networks

Figure 1: Paintings from Manet (left), Monet (centre), Van Gogh (right)

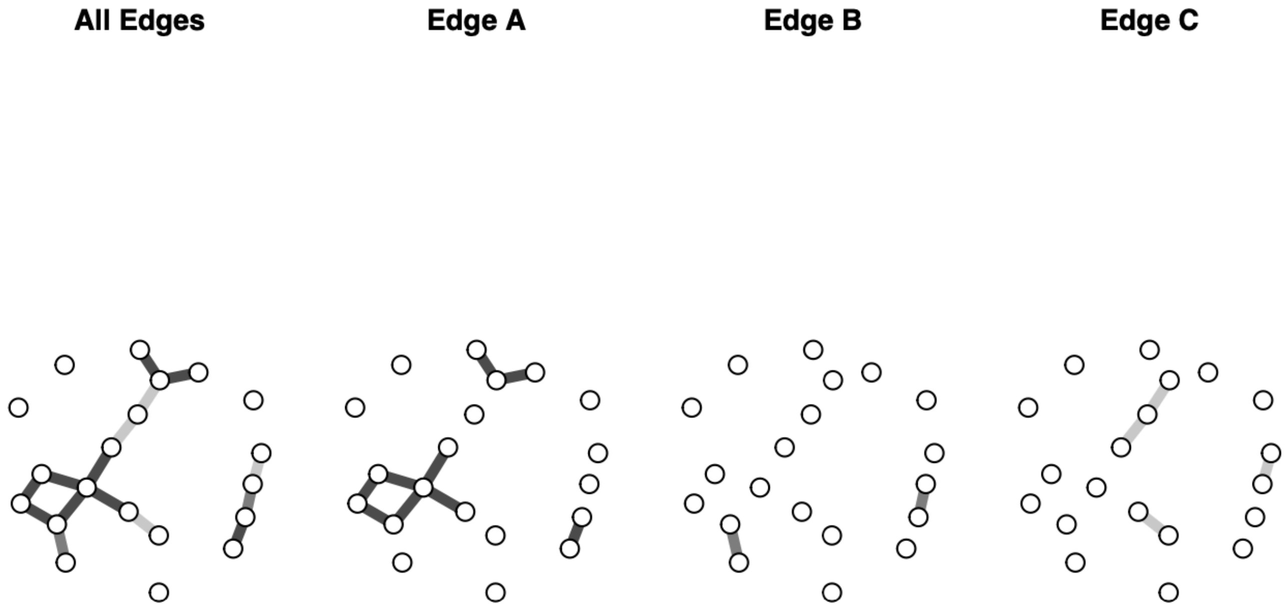
Table 1: A bipartite adjacency matrix with two node types: painters and features.

	French	Landscape	People	Ear	Impressionist
Manet	1	0.1	1.0	0	0
Monet	1	1.0	0.1	0	1
Van Gogh	0	1.0	0.1	1	1





Multiplex Networks



- This afternoon I’ll show you how to make these networks, create, save, and upload data to R.

- I’ll also show you how to turn your own data into a network.

Break

Network Measures

Variable	Definitions
N	Number of nodes
E	Number of edges
L	Average shortest path length
D	Diameter
C_i	Clustering Coefficient
k, k^{in}, k^{out}	Degree, in-degree, and out-degree
b	Betweeness centrality
c	Closeness centrality
x	Eigenvector centrality
r	Assortativity
Q	Modularity

There are so many more (e.g., http://schochastics.net/sna/periodic.html).

G(N,E)

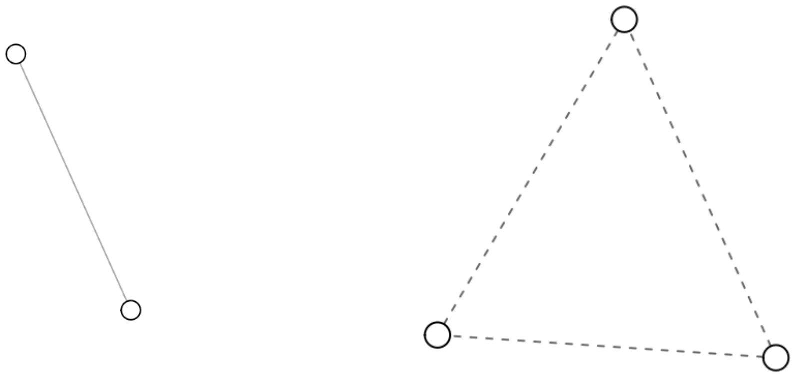
- How many edges in a simple network of N nodes?

	1	2	3	4	5	6
1	0	0	0	1	1	1
2	0	0	0	0	0	1
3	0	0	0	1	0	1
4	1	0	1	0	1	1
5	1	0	0	1	0	0
6	1	1	1	1	0	0

Can you work out a general rule?

Birthday paradox

- What's the probability at least two people in this room share a birthday?
- Why is this a network problem?



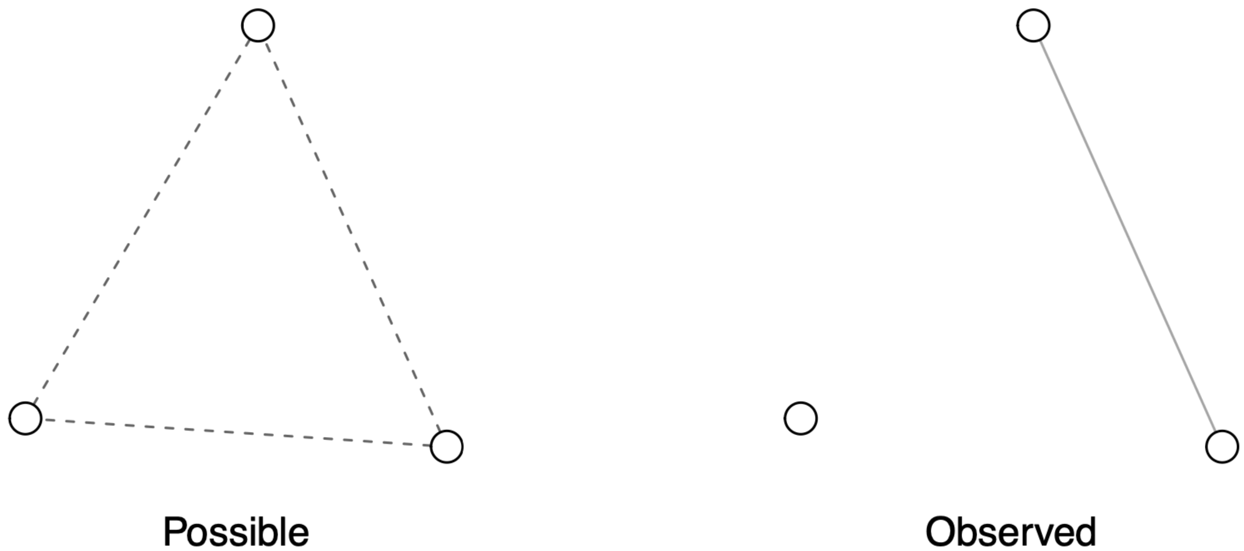
Number of possible edges:
E=N(N-1)/2

Probability of not sharing an edge:
(1-1/365)^E

Density

$$\rho = \frac{2E}{N(N-1)}$$

- The number of observed edges over the number of possible edges.



What does density tell us?

- What kinds of networks are likely to be low density?
- High density?

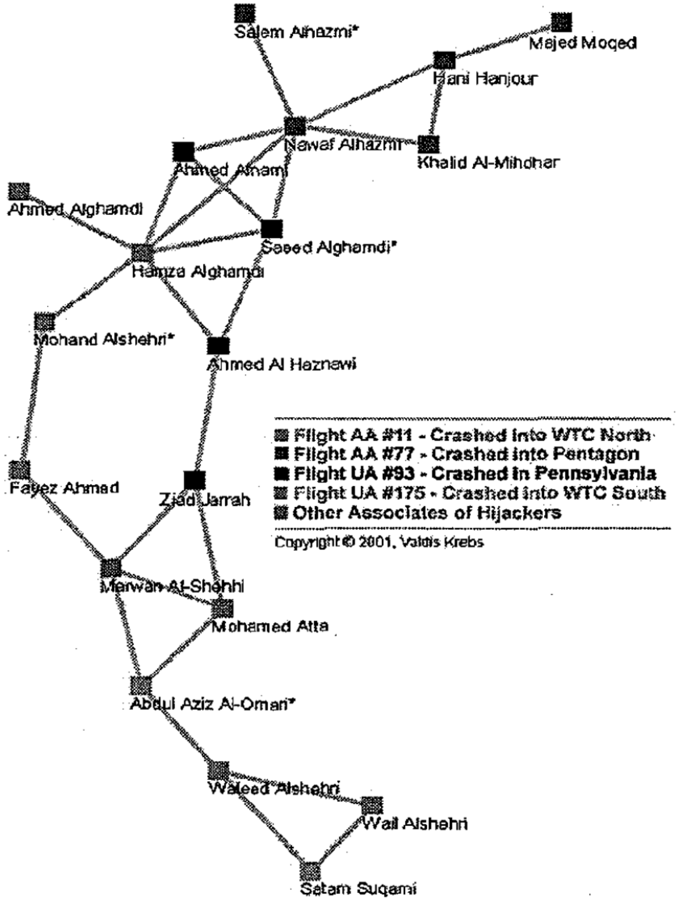
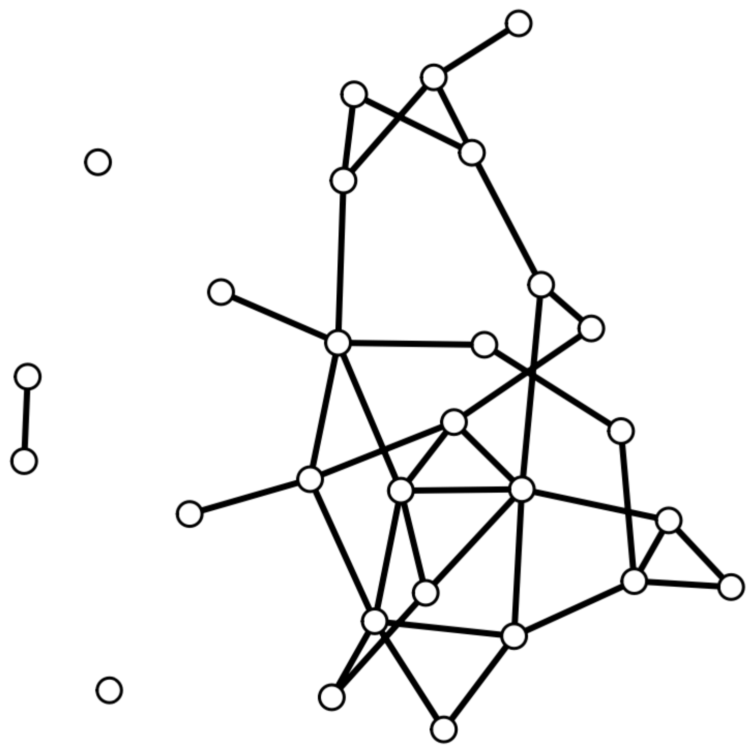


Figure 2 Trusted Prior Contacts

Krebs (2002) Mapping Terrorist Cells

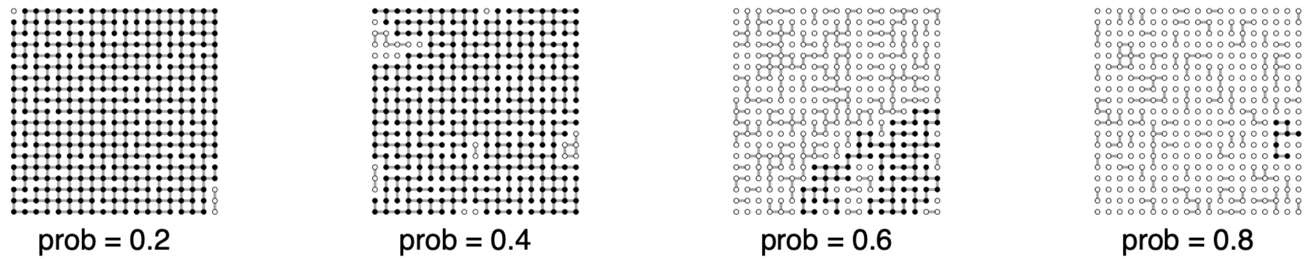
Components

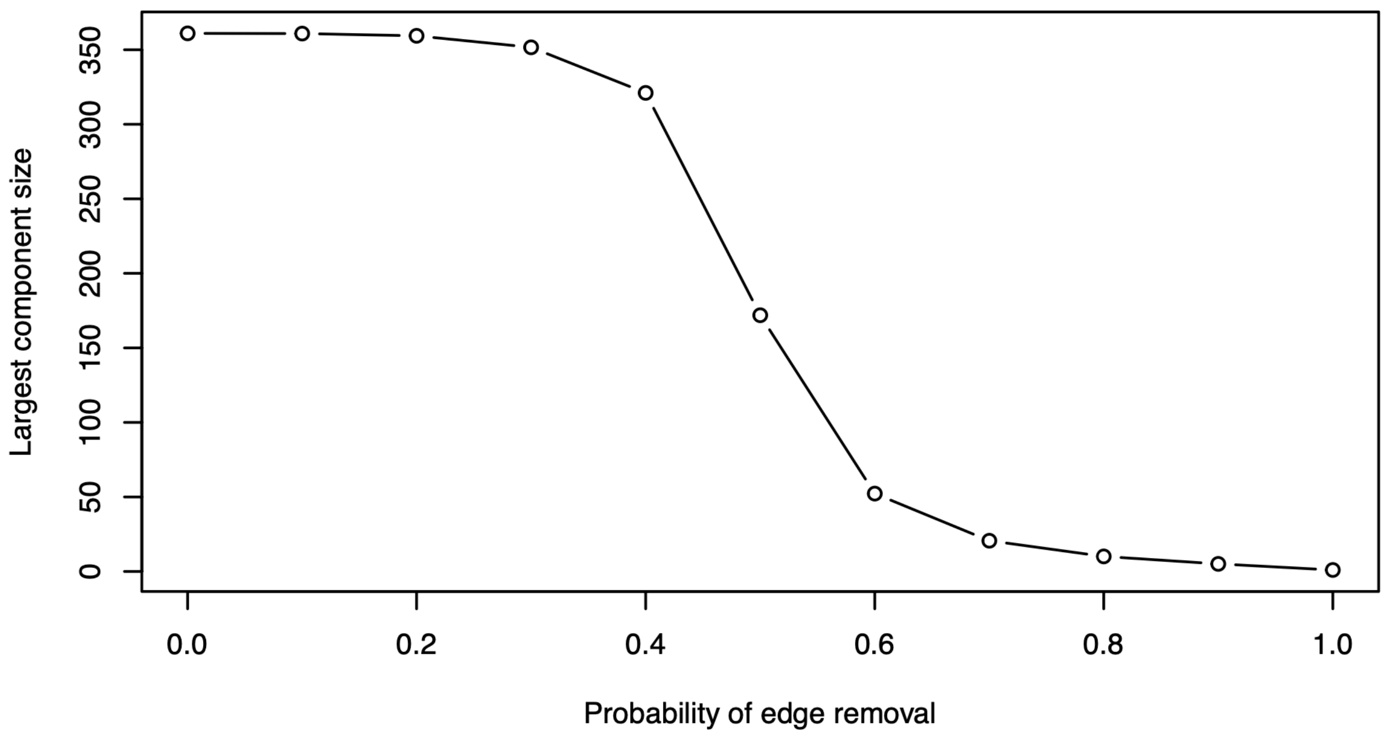
- Component: Collection of nodes that are all ‘reachable’ via a path of edges.



Percolation analysis

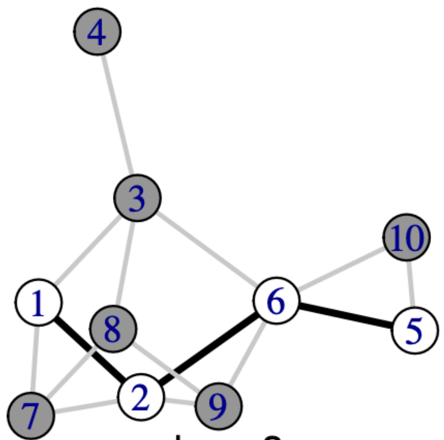
Probability of killing edges in a lattice—how many components?



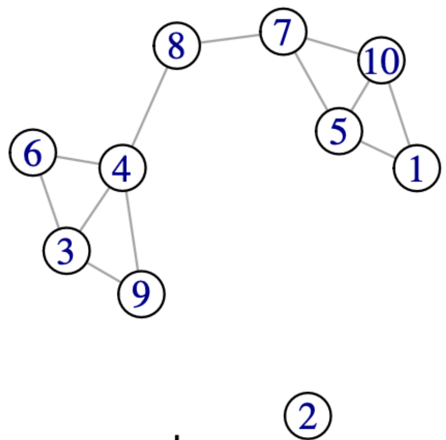


Path Lengths

- Path—a series of contiguous edges.
- Shortest path length (geodesic)
- Diameter—longest shortest path length in a network

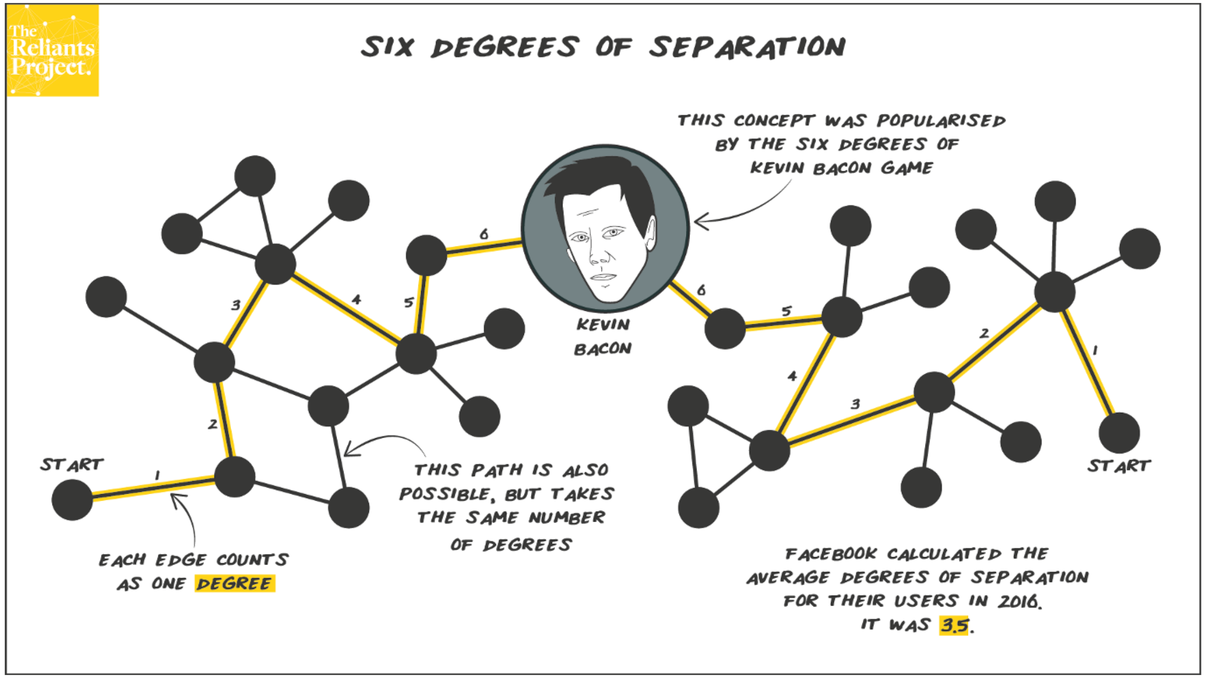


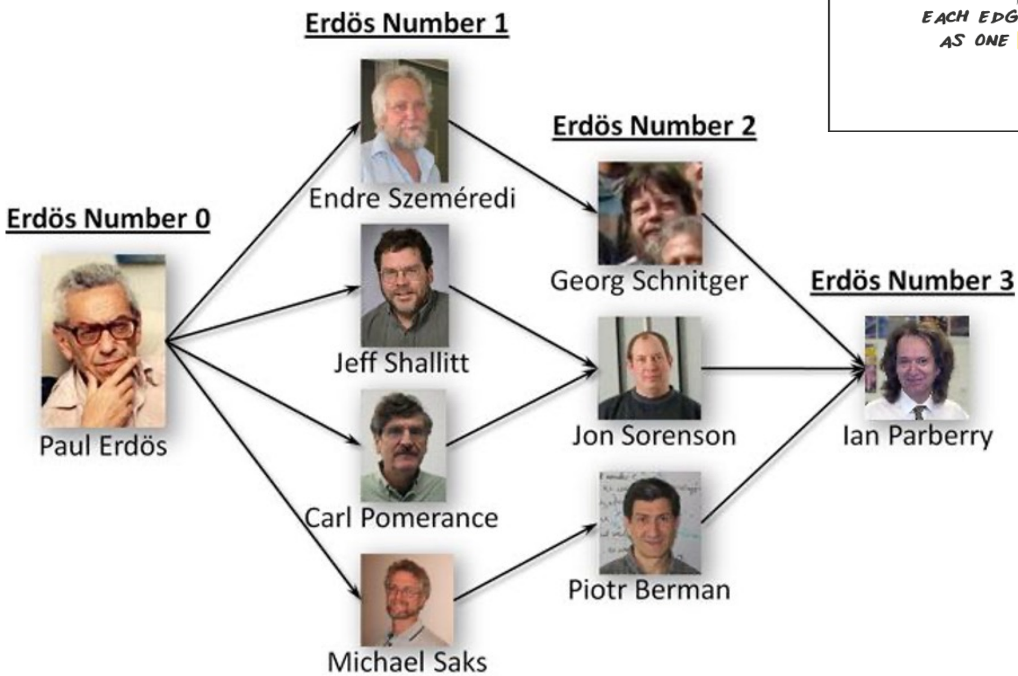
$d_{1,5} = 3$



$d_{4,2} = \infty$

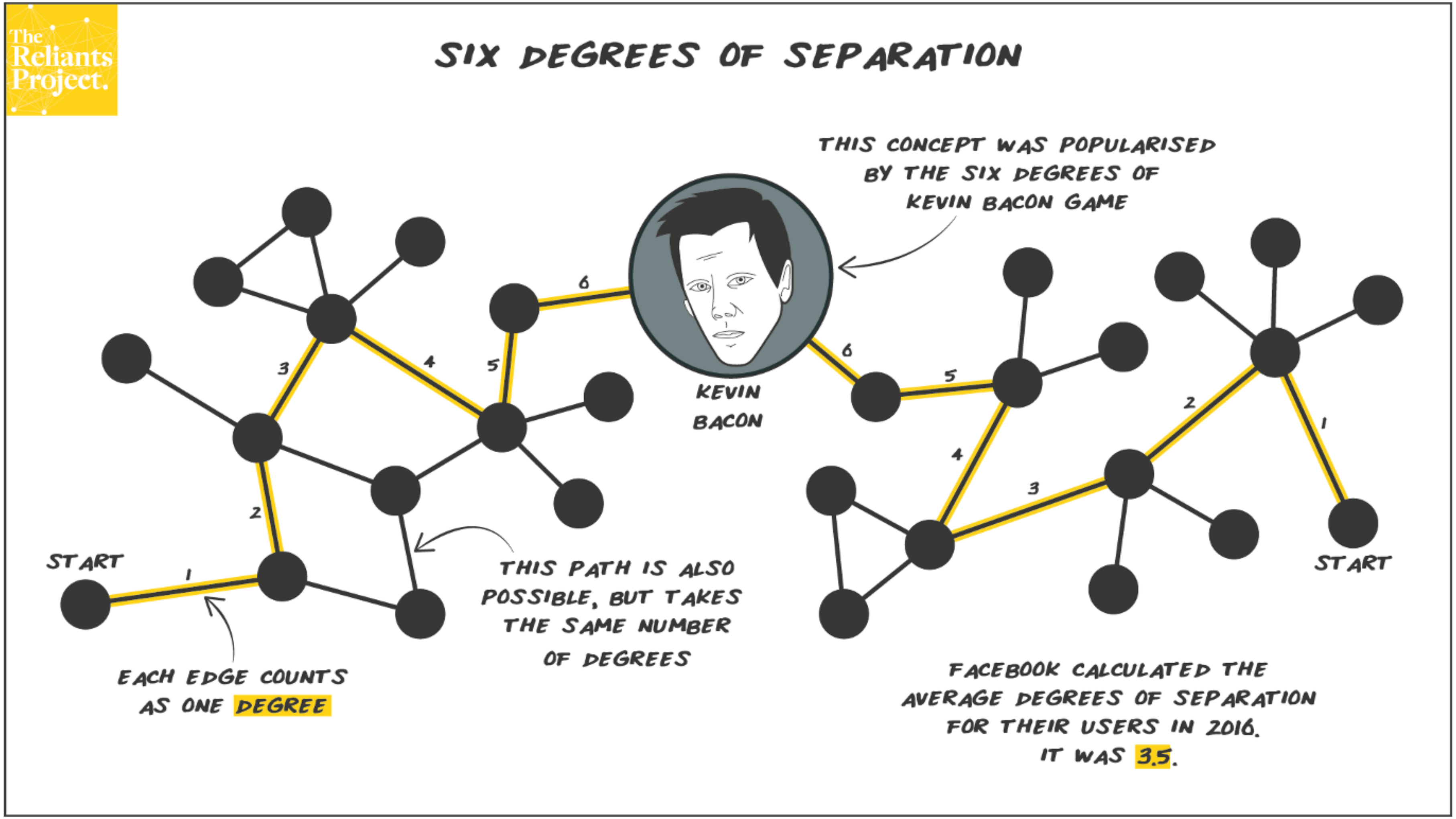
Six Degrees of Separation



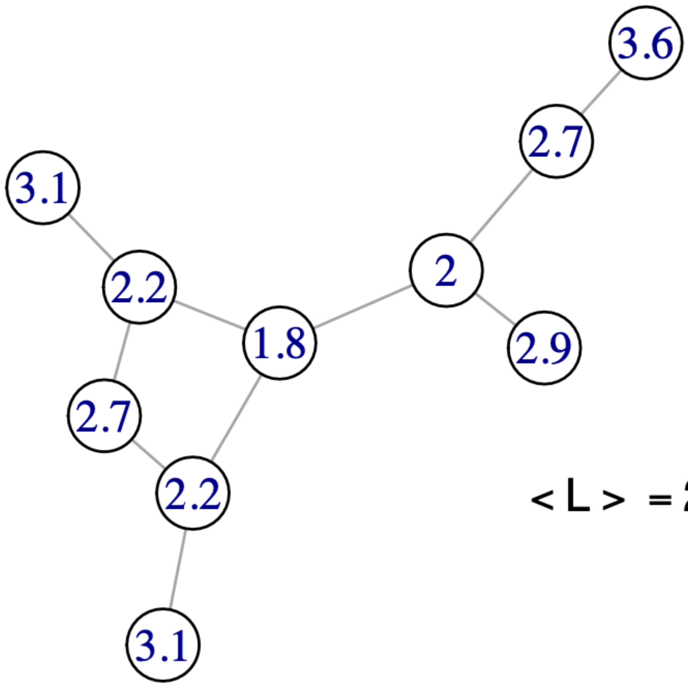


Hills:Adler:Levin:Durrett:Chung:Erdös

Oracle of Bacon

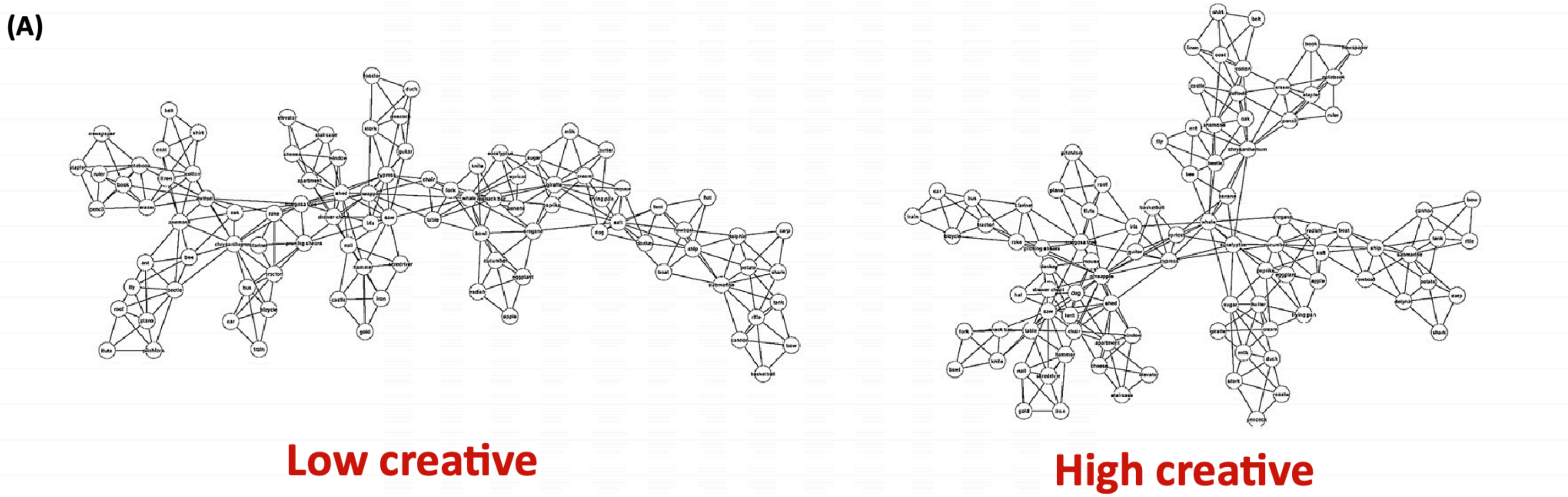


Average shortest path length



< L > = 2.62

Average shortest path lengths are shorter among ideas in more creative people

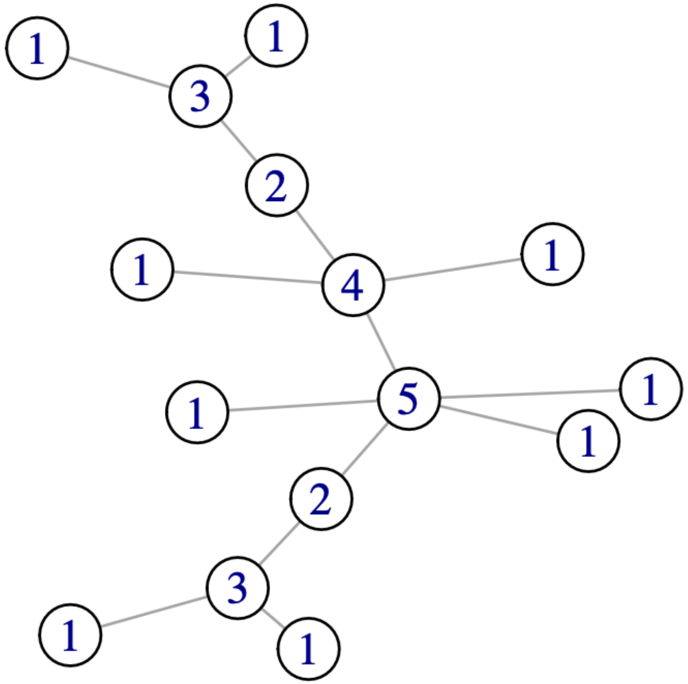


Kenett & Faust (2019)

Centrality

- Centrality is a measure of node importance.
- There are many centrality measures:
- Degree
- Betweenness
- Closeness
- Eigenvector centrality/PageRank
- And many more

Degree centrality



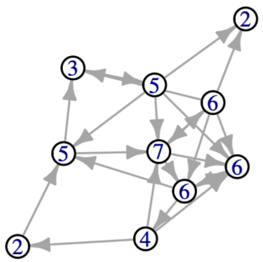
$$k_i = \sum_j A_{i,j} = \sum_j A_{j,i}$$

$$k_i^{in} = \sum_j A_{j,i}$$

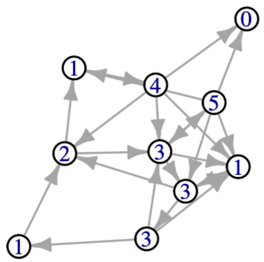
$$k_i^{out} = \sum_j A_{i,j}$$

$$< k >= \frac{1}{N} \sum_{i=1}^{N} k_i$$

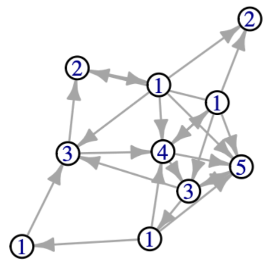
Average degree



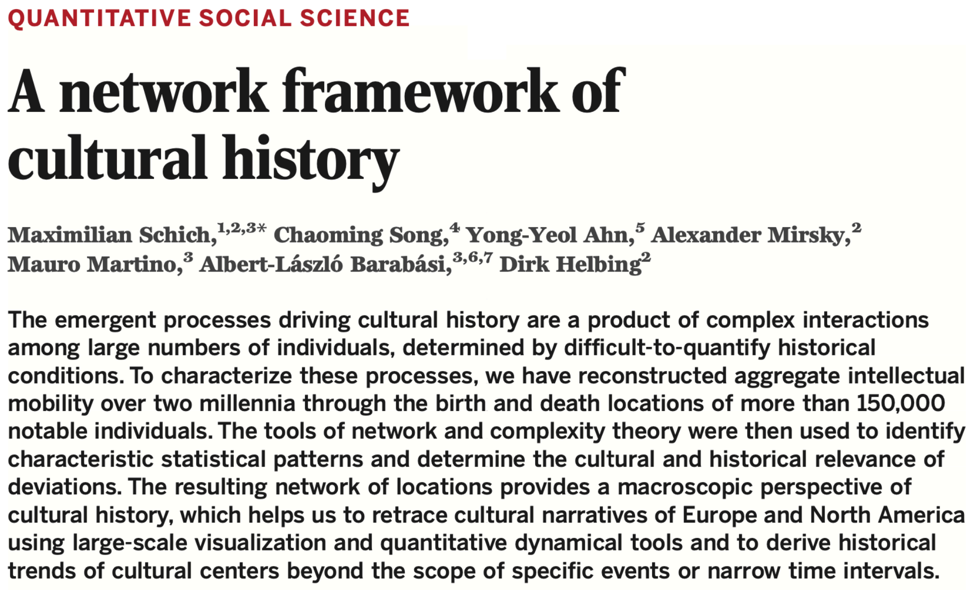
Total



Outdegree



Indegree


QUANTITATIVE SOCIAL SCIENCE

A network framework of cultural history

Maximilian Schich,[1,2,3*] Chaoming Song,[4] Yong-Yeol Ahn,[5] Alexander Mirsky,[2] Mauro Martino,[3] Albert-László Barabási,[3,6,7] Dirk Helbing[2]

The emergent processes driving cultural history are a product of complex interactions among large numbers of individuals, determined by difficult-to-quantify historical conditions. To characterize these processes, we have reconstructed aggregate intellectual mobility over two millennia through the birth and death locations of more than 150,000 notable individuals. The tools of network and complexity theory were then used to identify characteristic statistical patterns and determine the cultural and historical relevance of deviations. The resulting network of locations provides a macroscopic perspective of cultural history, which helps us to retrace cultural narratives of Europe and North America using large-scale visualization and quantitative dynamical tools and to derive historical trends of cultural centers beyond the scope of specific events or narrow time intervals.

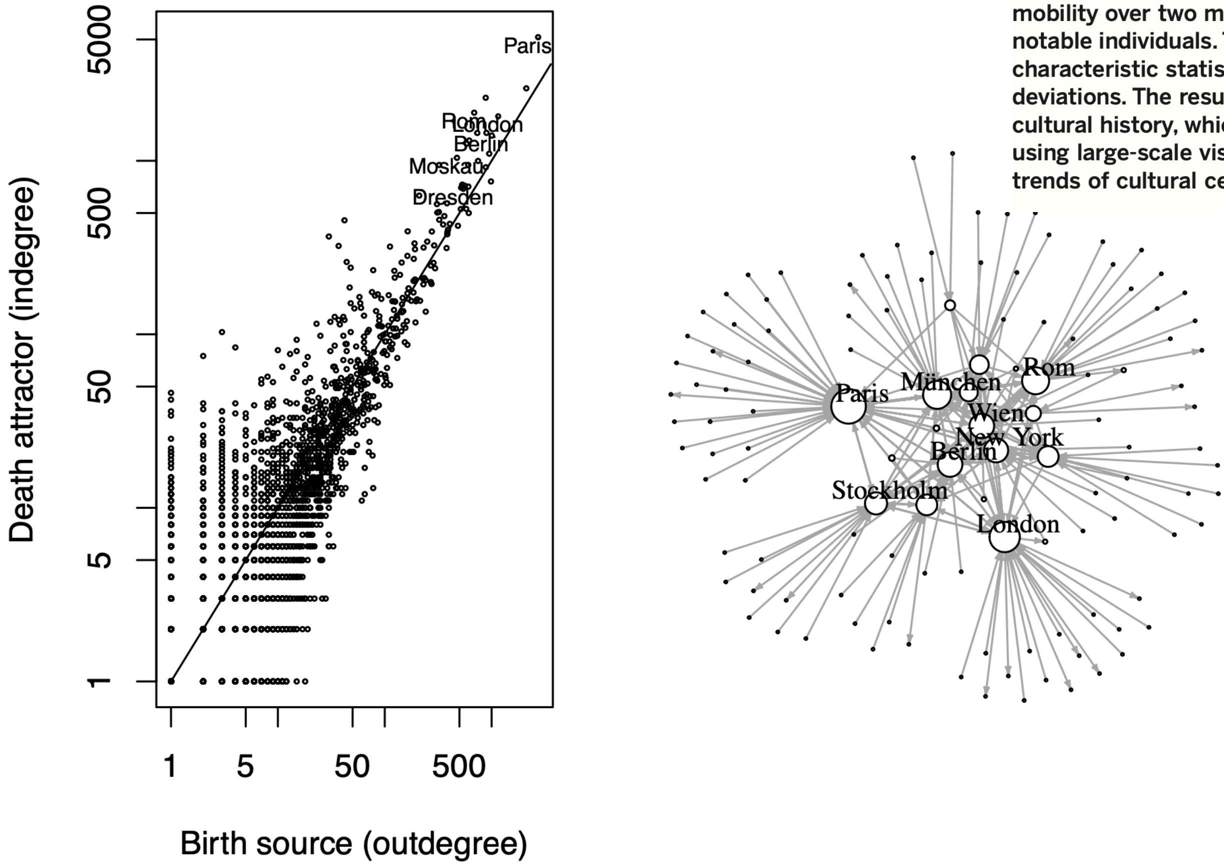


Figure 1: Indegree and Outdegree as birth sources and death attractors for 112,276 notable historical fine artists, c.480bc to 2010ad. There are $n = 46189$ nodes (i.e., places) in the network and 112,276 edges. (Left) Dots represent places. Places above the line are cultural centres where artists tended to be attracted. (Right) A representative subset of the network, showing 13 places with a total degree (in + out) of more than 500 plus a sampling of 117 additional nodes that are connected to them. Data is from the General Artist Lexicon (Beyer, Savoy, and Tegethoff 2016) and the figure is after Schich et al. 2014.

Weighted degree (strength)

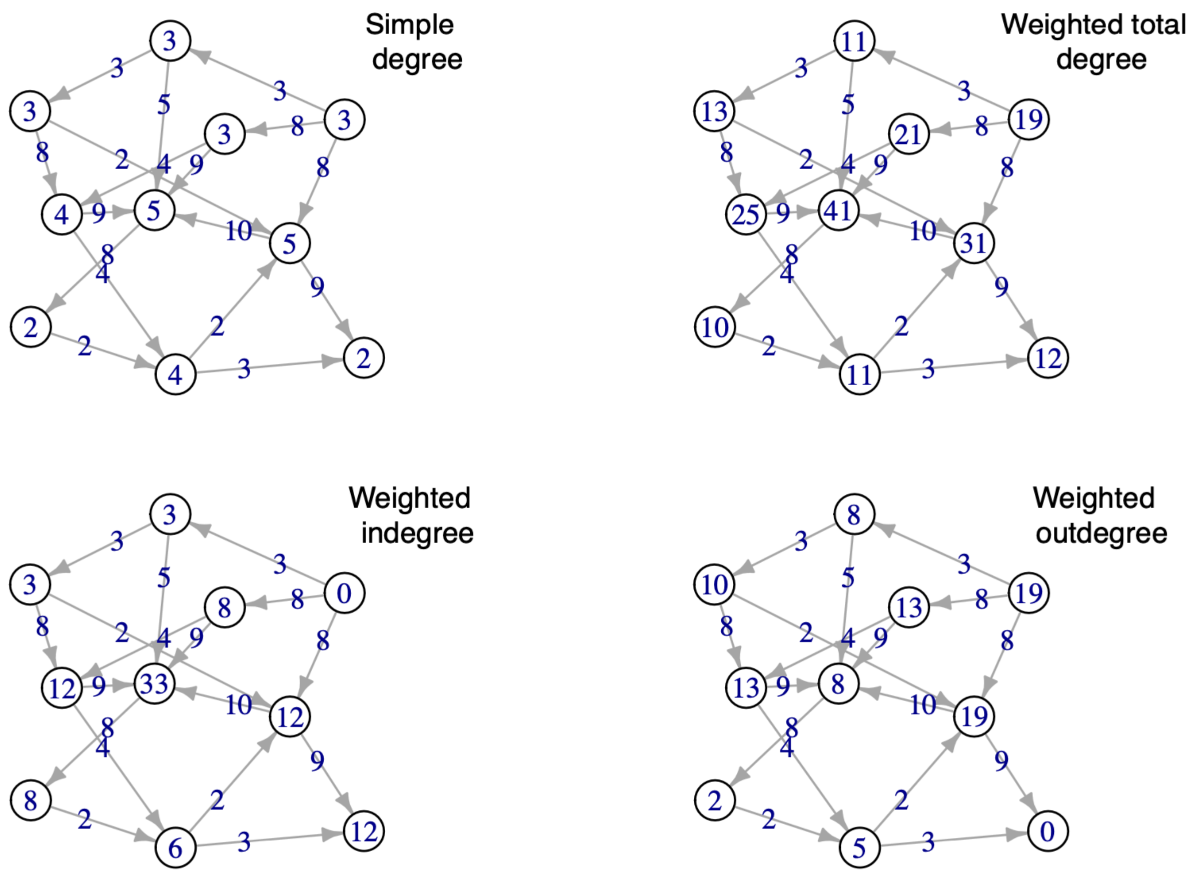
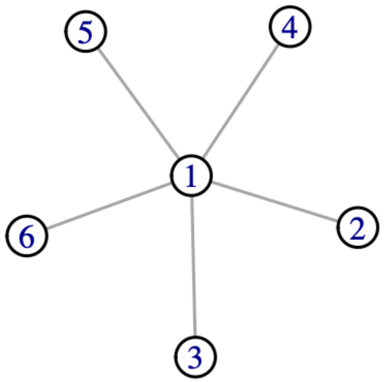


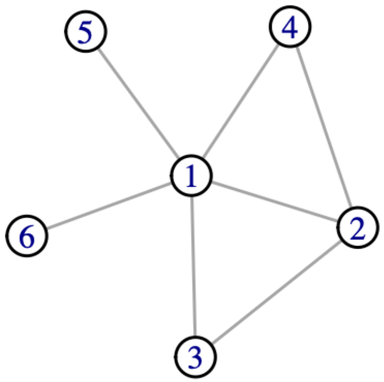
Figure 9: Various ways of computing the degree in directed weighted networks. Nodes are labeled with their relevant degree.

Clustering Coefficient

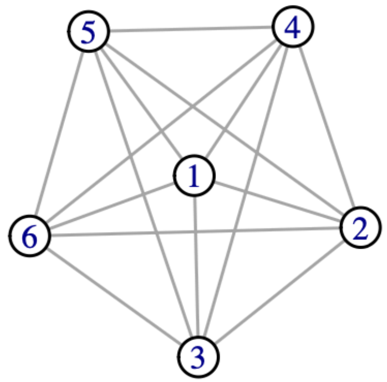
The clustering coefficient has two forms. The first is a node-level or local clustering coefficient. This measures the proportion of a node's neighbors that are connected by an edge.



C=0



C=.2



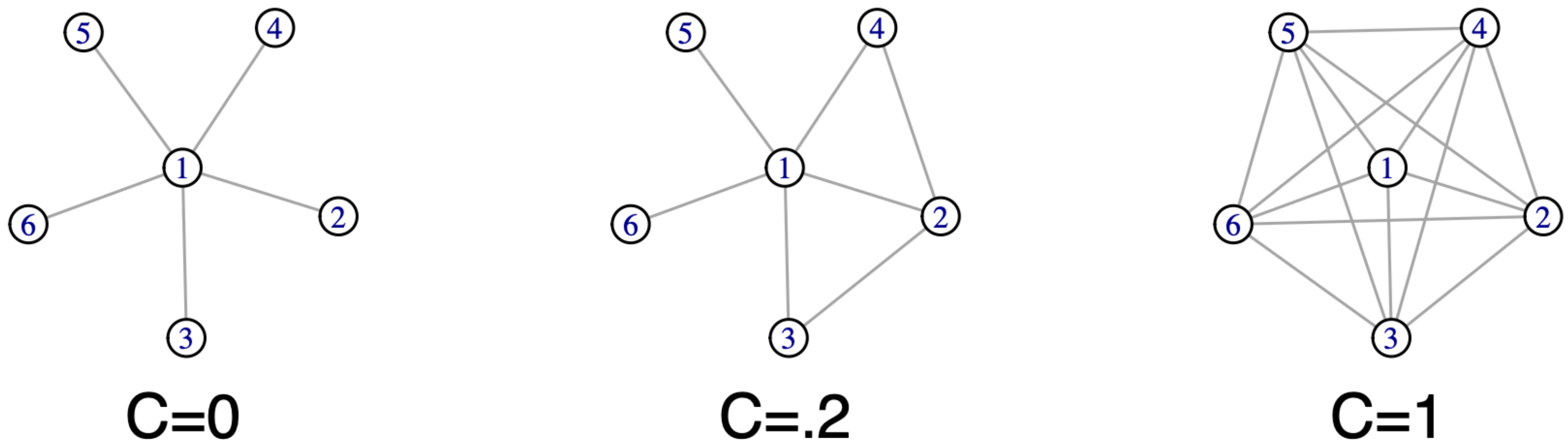
C=1

$$C_i = \frac{2e}{k_i(k_i - 1)}$$

Clustering Coefficient

(Node level)

The clustering coefficient has two forms. The first is a node-level or local clustering coefficient. This measures the proportion of a node's neighbors that are connected by an edge.



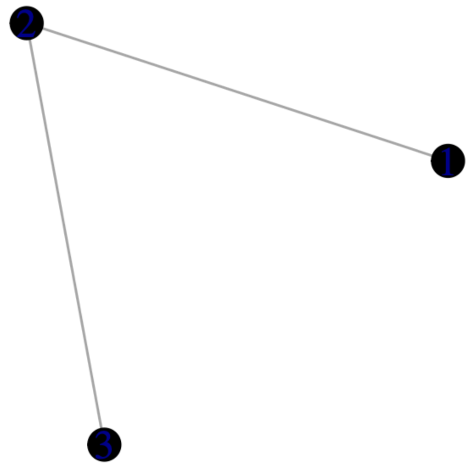
$$C_i = \frac{2e}{k_i(k_i - 1)}$$

One can compute the average C over all nodes

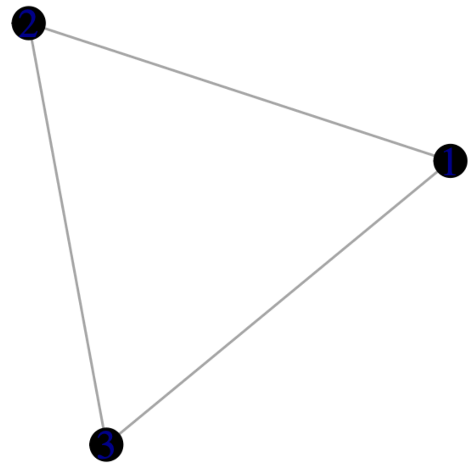
Clustering Coefficient

(Transitivity: graph level)

Transitivity measures the proportion of triplets in the network that are transitive (i.e. a triangle).



Intransitive triplet



Transitive triplet

$$T = \frac{3\Delta}{\Lambda}$$

Transitivity is a graph level metric

Clustering coefficient and transitivity can diverge (node vs. graph level view)

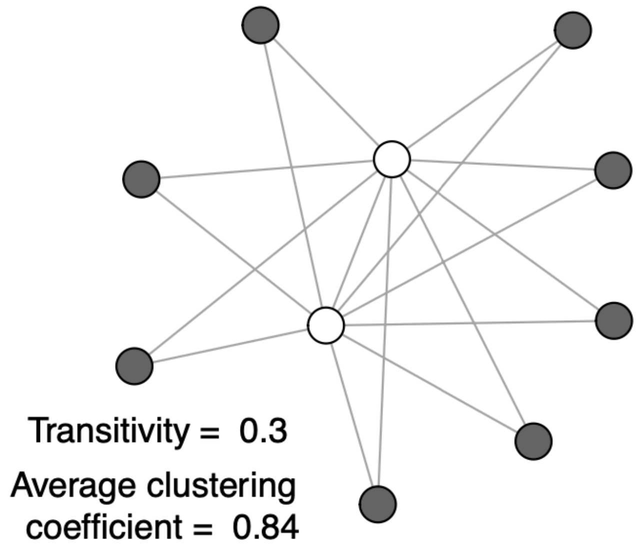
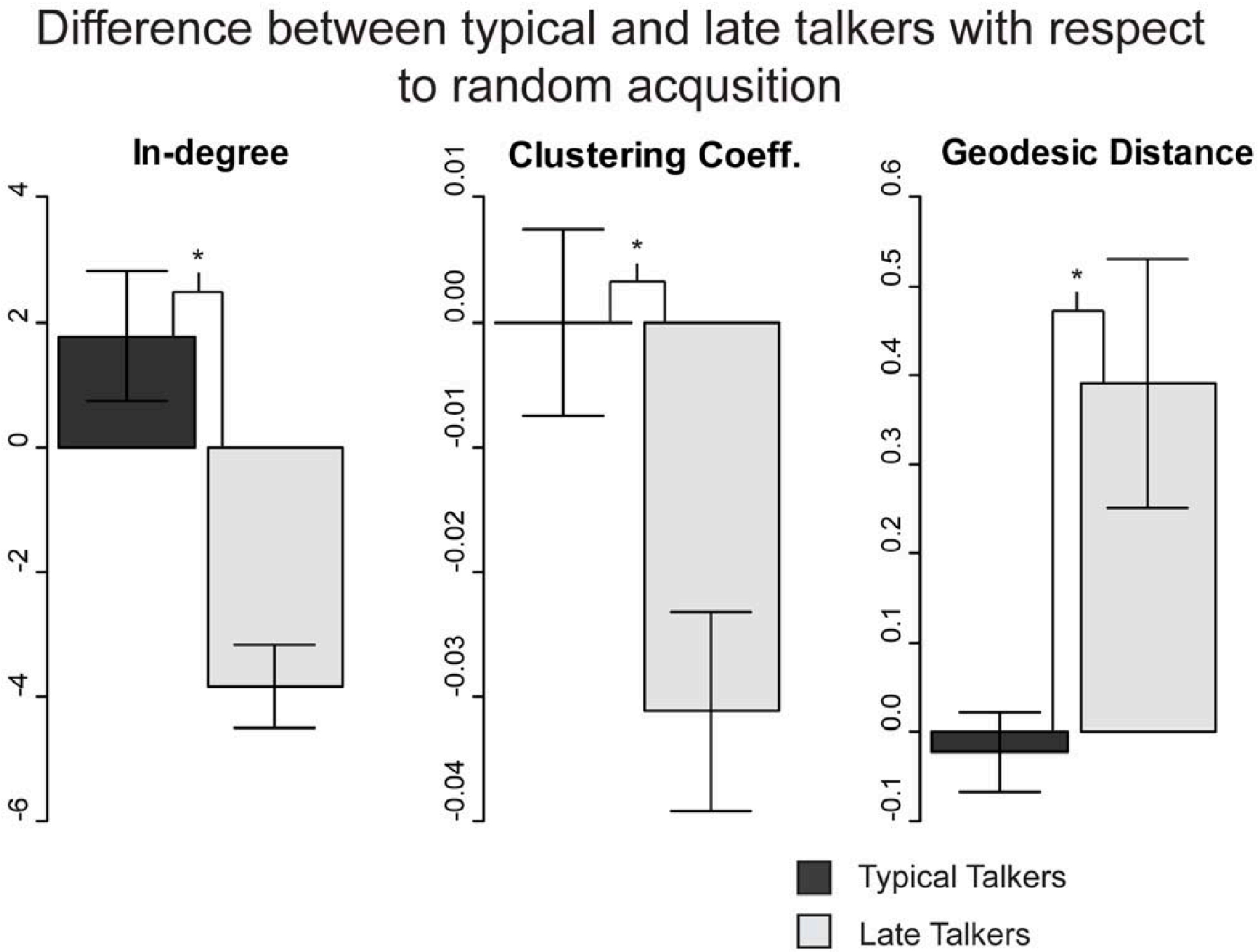


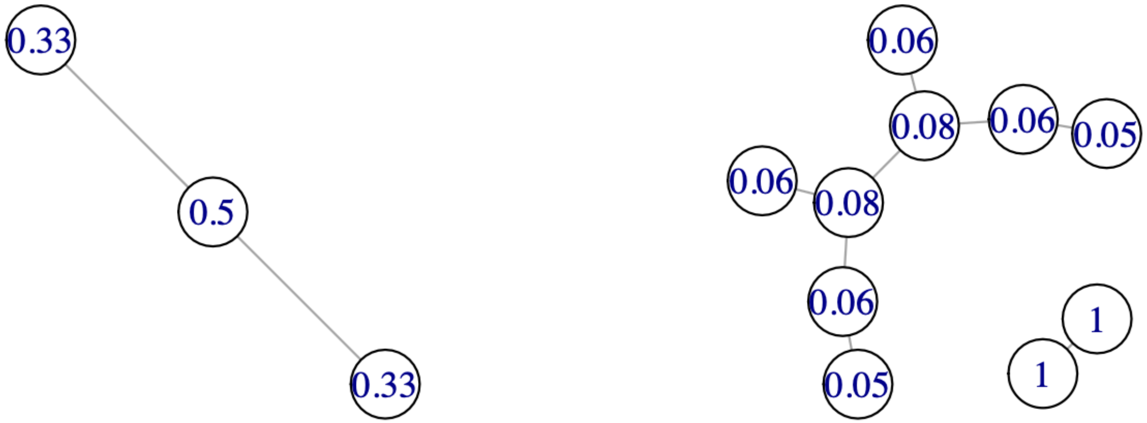
Figure 12: The wheel network demonstrates the difference between transitivity and average clustering coefficient. As the outer nodes increase, the average clustering coefficient approaches 1 and the transitivity approaches 0.

Clustering coefficient and the words children learn



Late talkers have lower degree and lower clustering coefficient and have average shortest path length (ASPL = geodesic distance)

Closeness centrality

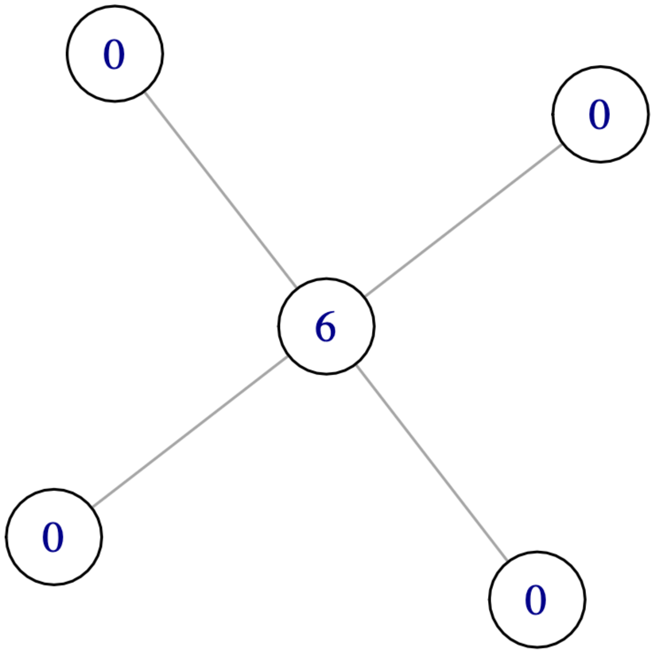


Each node is labeled with its closeness centrality

$$c_i = \frac{1}{\sum_{j=i}^{N} d_{ij}}$$

Betweenness centrality

The betweenness centrality for a node i is the number of shortest paths between all other pairs of nodes that pass through node i.

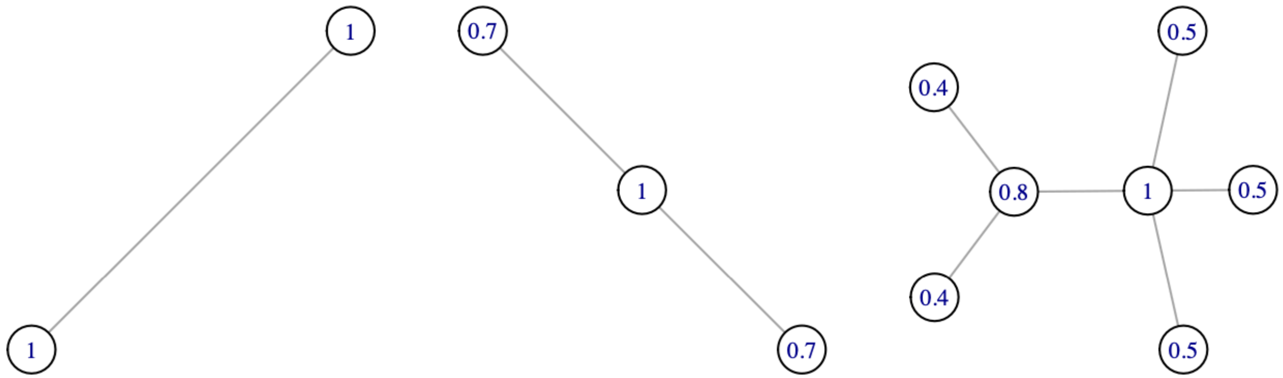


$$b_i = \sum_{i \neq j \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

6 possible paths, all pass through the central node.

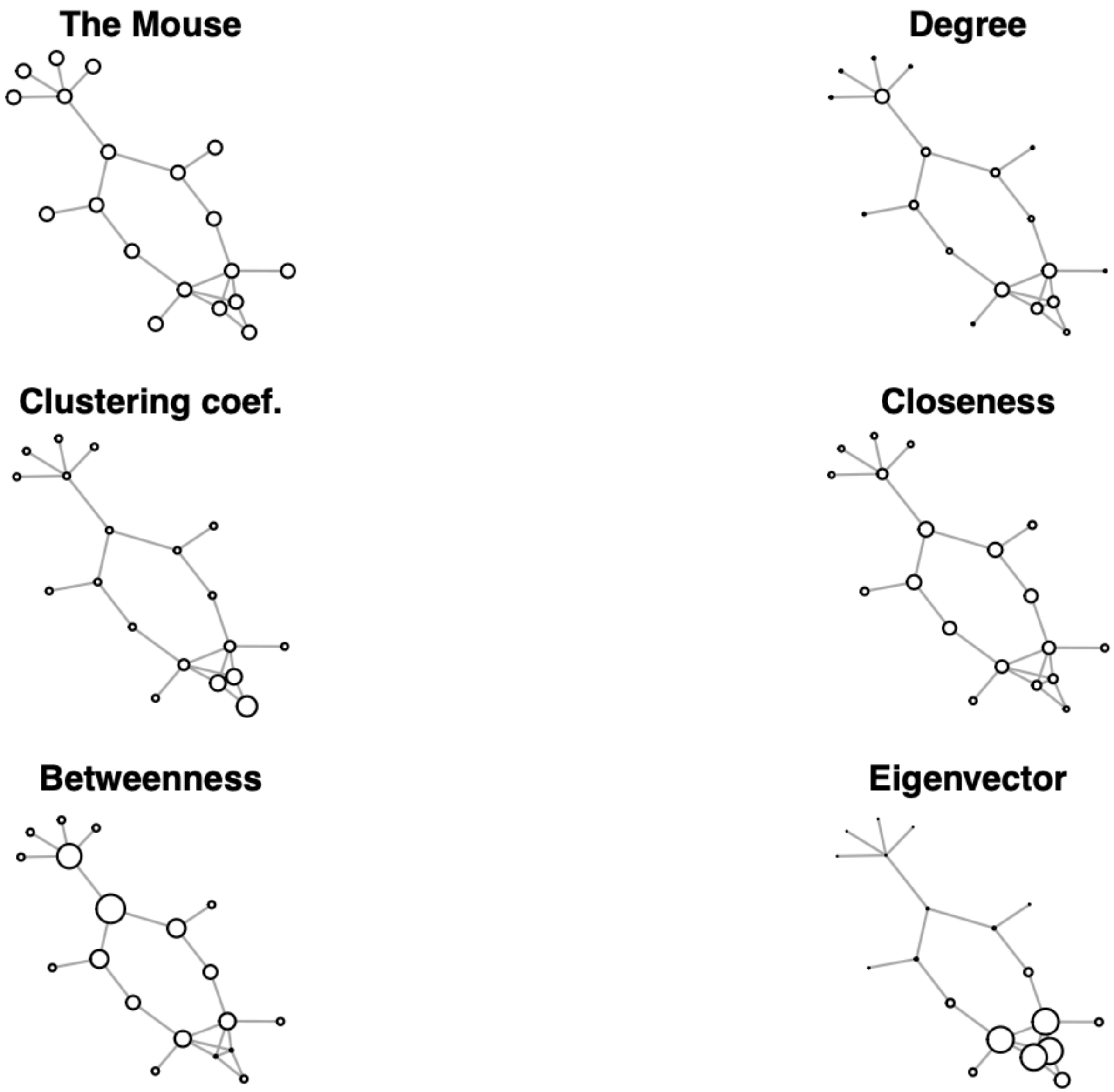
Eigenvector centrality

Eigenvector centrality is analogous to prestige. To be prestigious, one must receive prestige from other nodes. The more prestigious the nodes one receives prestige from, the more prestige one receives. The definition is recursive: It requires that we know how prestigious each node is before we can compute the prestige of any node.



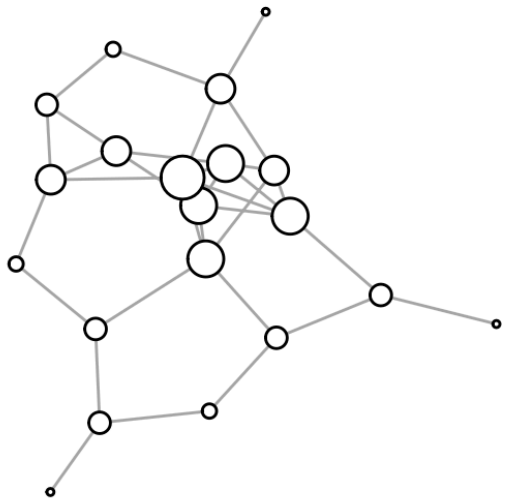
This measure is the basis of PageRank and Katz centrality—both look at how nodes recursively give and receive 'value' to their neighbours.

Measures of centrality

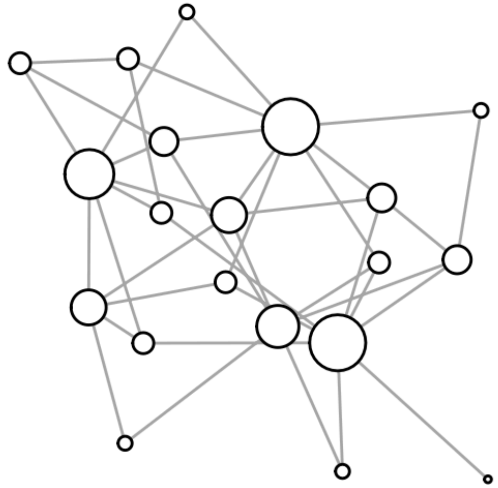


Assortatitivity

Assortativity evaluates the degree to which nodes with similar properties connect with each other. In social networks, this is known as homophily: “birds of a feather flock together.”



r=.45

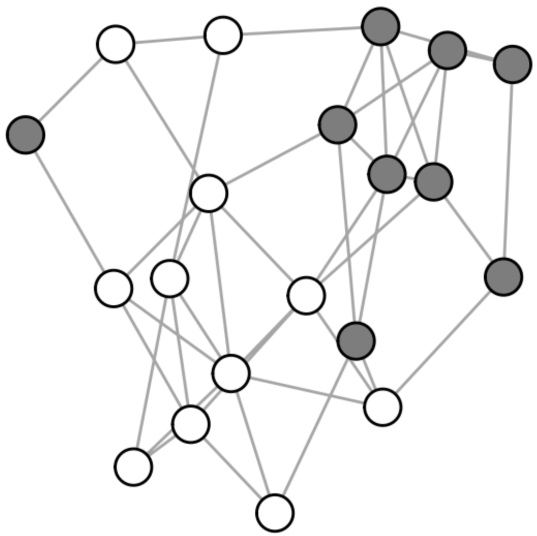


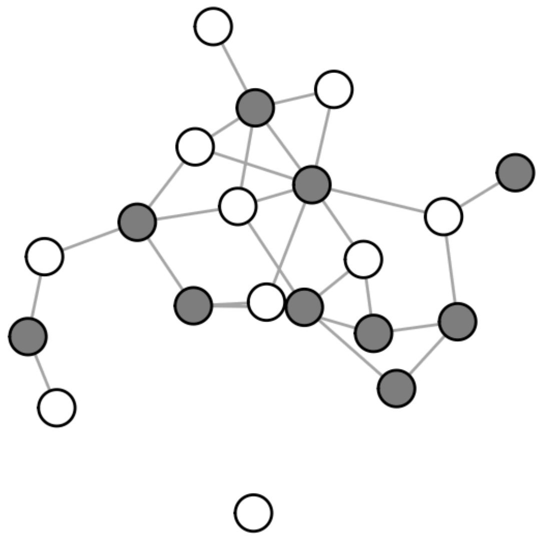
r=-.62

To evaluate assortativity we compute an assortativity coefficient: the Pearson correlation between pairs of connected nodes in the network with respect to the value in question. To do this, generate an edge list from the network, replace the node labels with the value for each node, and take the correlation of the two columns of values.

Assortatitivity

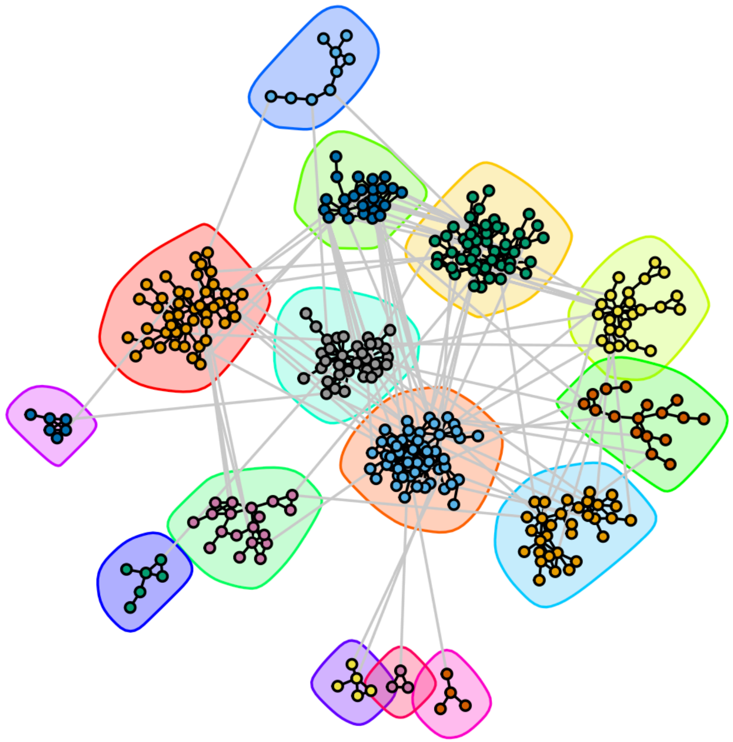
Assortativity by colour (a node attribute)





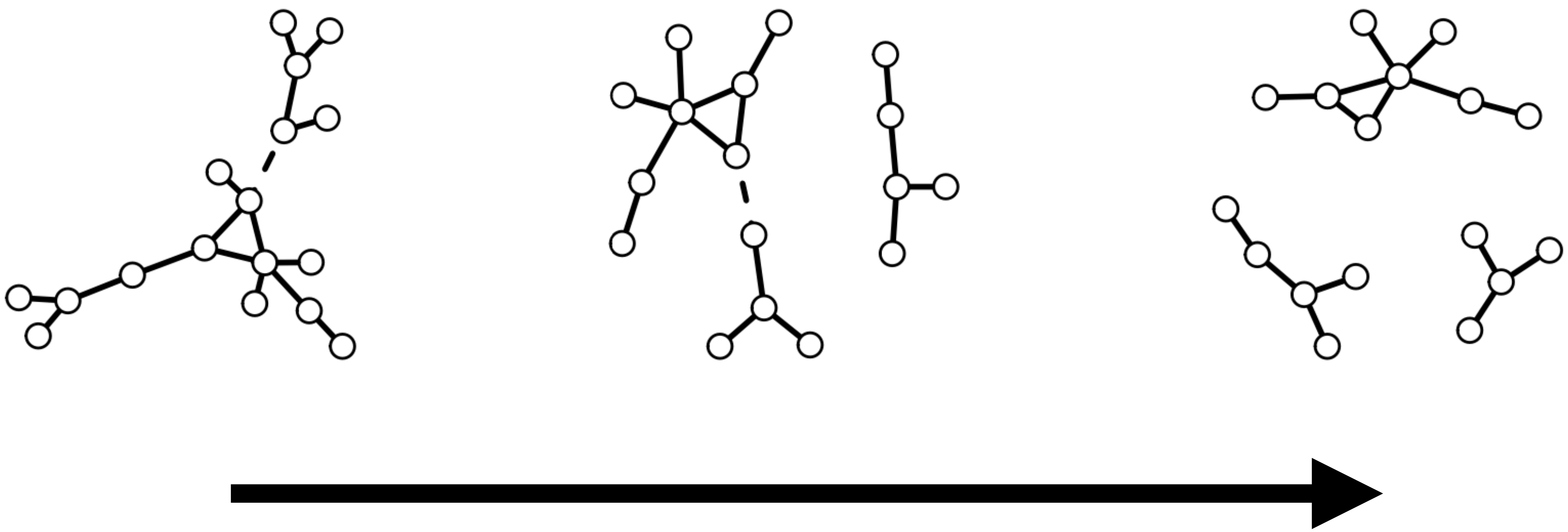
Community Detection

Communities can be detected by identifying clusters of nodes that are more well connected to one another than they are to members of other communities. A division of the network into a set of communities is called a *partition*.

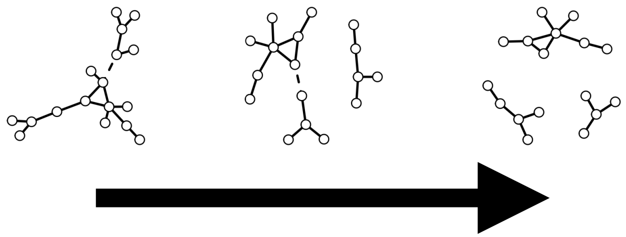


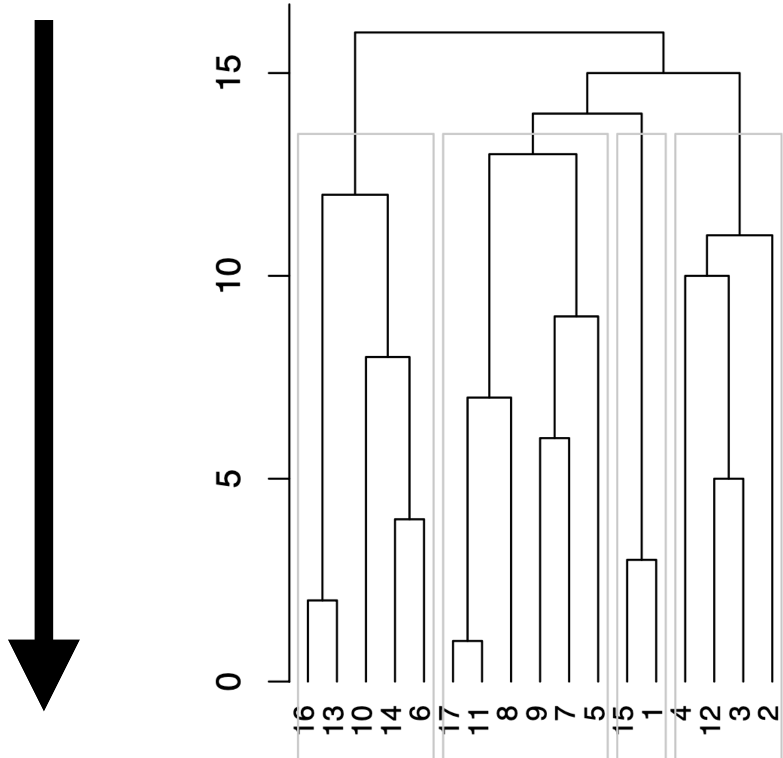
Girvan-Newman Method

The Girvan Newman Method (Girvan and Newman 2002) (or edge betweenness method) is based on the observation that edges connecting separate communities have high edge betweenness: shortest paths between members of different communities will pass through edges with high edge betweenness.



Girvan-Newman Method





How to pick the best partition?

Modularity

Modularity, Q, is a measure of the difference between the observed links within communities and the expected links within the same communities if all edges were distributed at random.

$$Q = \frac{1}{2m}\sum_{i.j}[A_{ij} - \frac{k_i * k_j}{2m}]\delta(c_i, c_j)$$

Observed **Expected**

High modularity means more observed than expected.
Choose partition with highest modularity.

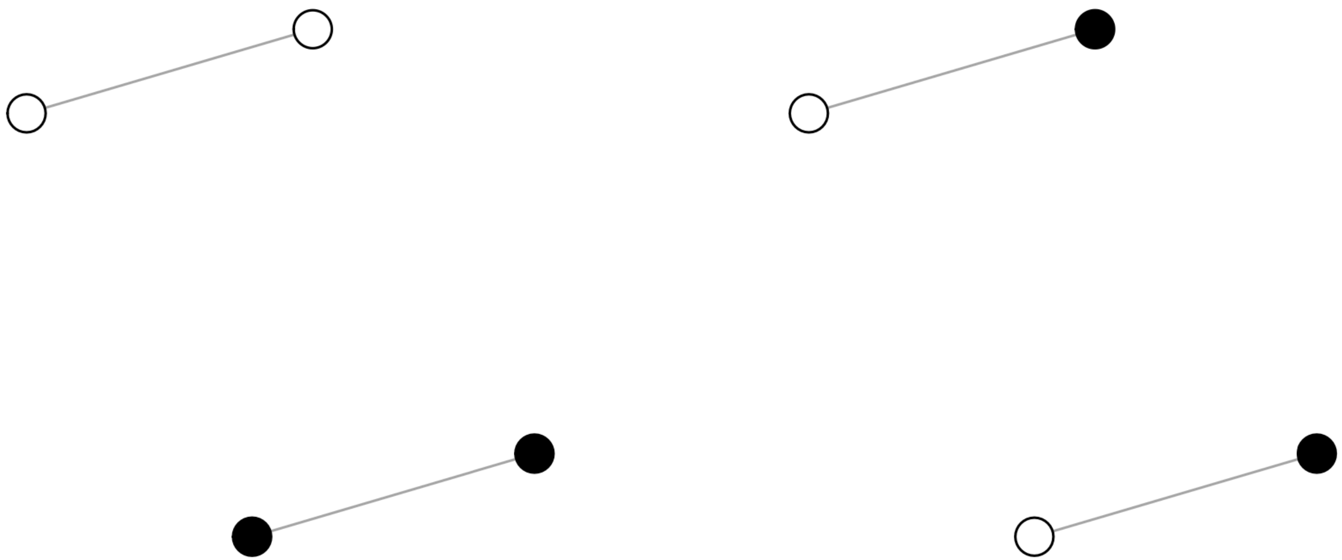


Figure 21: Modularity: Two possible network architectures for 4 nodes with 2 links, each node with degree 1, and 2 communities. The network on the left has a Q=0.5. The network on the right has Q=-0.5.

Community detection

(Many methods)

Girvan Newman | **Louvain** | **Walktrap** | **Clique Percolation**