Social Network Analysis Workshop

Glasgow, Scotland - July 2018

Prof James P Curley

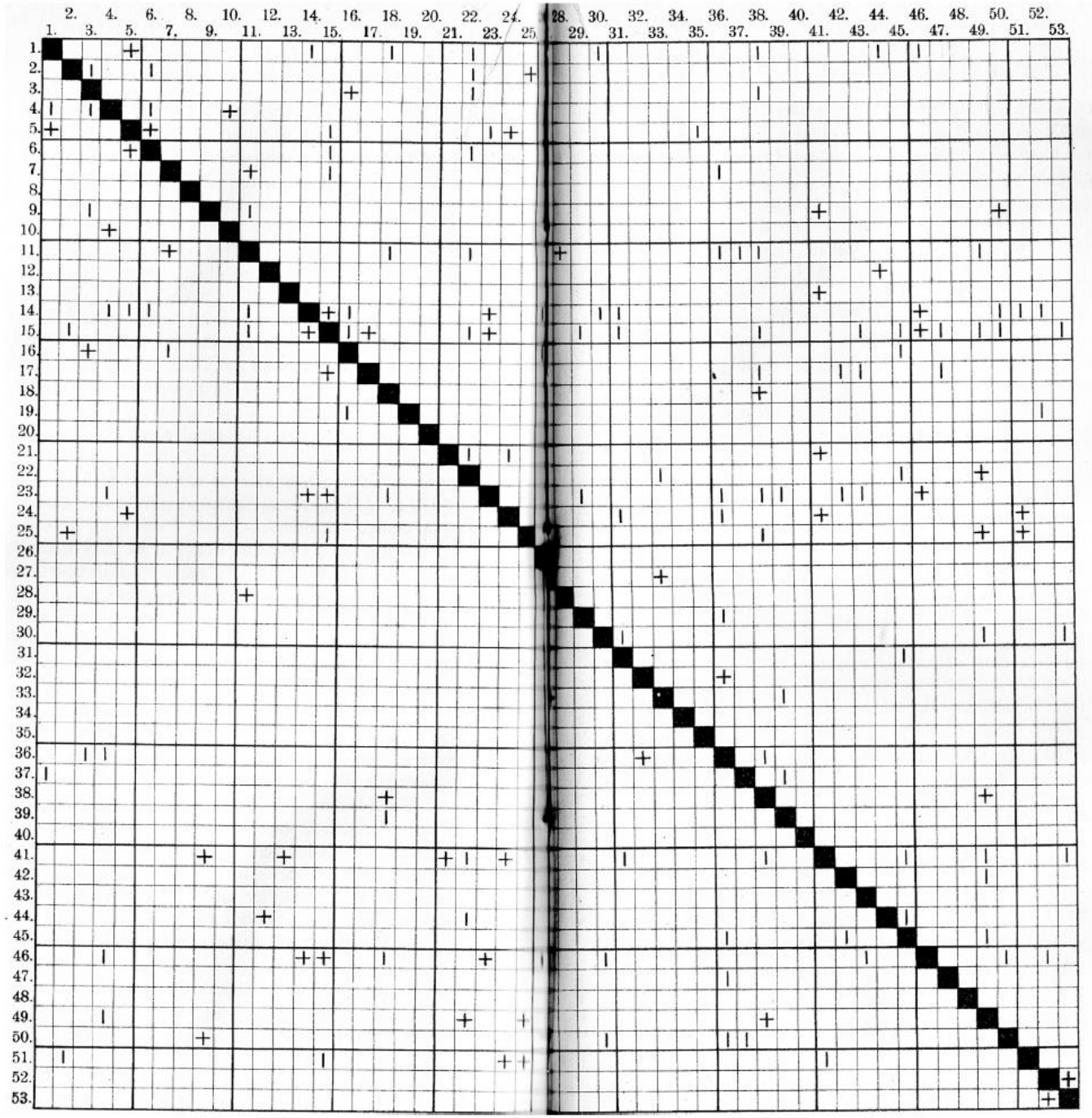
@jalapic

Associate Professor

Psychology Department

University of Texas at Austin

The Study of Social Dynamics



Data from 1880–1881

German teacher Johannes Delitsch published in 1900 in ‘Zeitschrift für Kinderforschung’ (‘Journal of Child Study’).

| = Ingoing friendship relations (“angebotene Freunschaften”) - Child in column befriends child in row.

+ = Reciprocated friendships (“geschlossene Freundschaften”)

Helen Bott 1928

Observation of play activities in a nursery school.

(1) talked to another
(2) interfered with another
(3) watched another
(4) imitated another
(5) cooperated with another

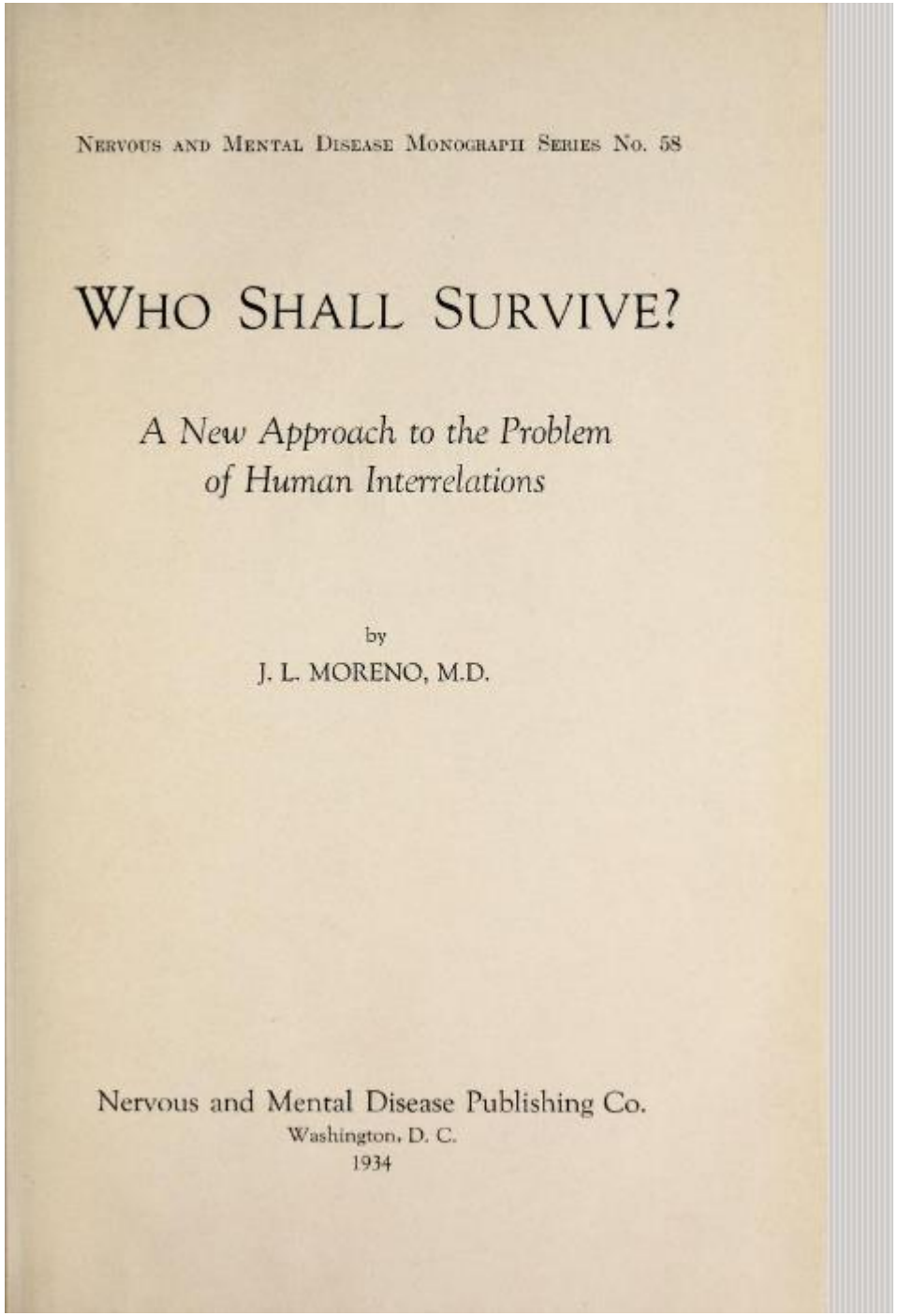
TABLE 2

SOCIAL RESPONSES AMONG ELEVEN CHILDREN UNDER FIVE CATEGORIES

Talking, Interference, Watching, Imitation, Cooperation

Observed behavior of each child towards his fellows

Child	Age (months)	Response	Children played with A	D	E	F*	G	H*	J*	K	M	N*	O*	All Responses	Total
A	26	Ta.		5	2	1			1	15	1		1	64	49
		Int.			5	1			1	5		1	1	12	
		Wa.		1	1	1				1				5	
		Im.			1					3				5	
		Co.								3				3	
D	29	Ta.	4		2	1		2		1	1	1		11	53
		Int.	1		1	1		2		2	5		6	8	
		Wa.	2		3	3	2	1	1	5	5	1		23	
		Im.						2		2	2			9	
		Co.								4				8	
E	31	Ta.	3	4		3	1	1	5	4	4	1	4	24	104
		Int.	6	7	12	12	1	2	1	7	5	7	5	57	
		Wa.		2		1		1	3	1	3	1	2	9	
		Im.		1									2	12	
		Co.											1	2	
E*	36	Ta.		1	3		2		2		3	2	11	23	53
		Int.		1	2		2				3		2	10	
		Wa.			1			1			2	2	6	10	
		Im.					2		1				3	3	
		Co.											1	1	
G	37	Ta.	2		1	1		2		5	3	4	2	16	55
		Int.		2	1	1		3		2	5		4	21	
		Wa.		2		1			1	2	1		3	10	
		Im.									2			3	
		Co.								2	2			5	
H*	38	Ta.	1	2	3		3			5	1	1		3	40
		Int.		2	2	3	3			1	4	4		9	
		Wa.									1			24	
		Im.									1			2	
		Co.												2	
J*	40	Ta.	4		16					1	20	7	5	57	86
		Int.	3	1	6	2				2	3	1	3	18	
		Wa.	1							1	2	2	1	9	
		Im.										1		2	
		Co.												0	
K	45	Ta.	5	2	2	3	4		1		20	2	5	42	73
		Int.	2	1	3	1	2		2		6	1	1	19	
		Wa.												0	
		Im.												2	
		Co.												10	
M	52	Ta.		3	3	1	2			6		3	7	27	69
		Int.		1		2				1			3	7	
		Wa.				1							4	4	
		Im.					1		1				18	21	
		Co.											3	10	
N*	53	Ta.	2	1	3	5	10	1	10	8	6		18	64	104
		Int.	4	2	4	8	2				1		2	22	
		Wa.	5	2		1				1	1		3	12	
		Im.											2	4	
		Co.												2	
O*	54	Ta.	1	1	4	11	2	1	4	3	65	25		115	143
		Int.	1	1	2	1			1	2	7	3		18	
		Wa.				1					1			5	
		Im.									3			1	
		Co.												4	
Responses of Companions		Ta.	21	18	36	26	18	6	18	45	123	41	54	406	855
		Int.	16	18	23	28	12	7	9	21	31	18	18	201	
		Wa.	9	9	7	10	5	2	4	15	19	6	28	111	
		Im.	3	2	4	3	5	2	5	7	8	4	27	70	
		Co.	1	2	1	1	4	2	1	10	10	2	13	47	
Totals			50	49	71	68	42	18	37	98	188	71	140		855



NERVOUS AND MENTAL DISEASE MONOGRAPH SERIES No. 58

WHO SHALL SURVIVE?

A New Approach to the Problem of Human Interrelations

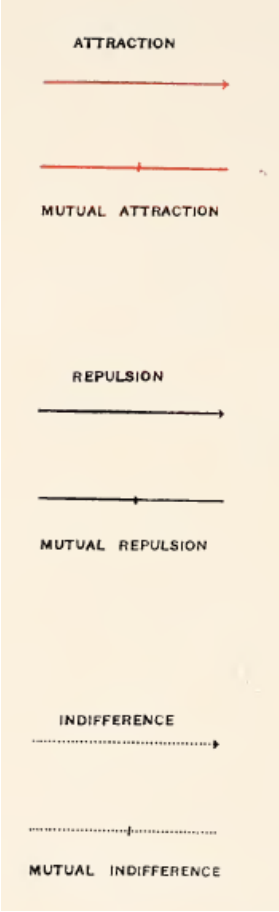
by

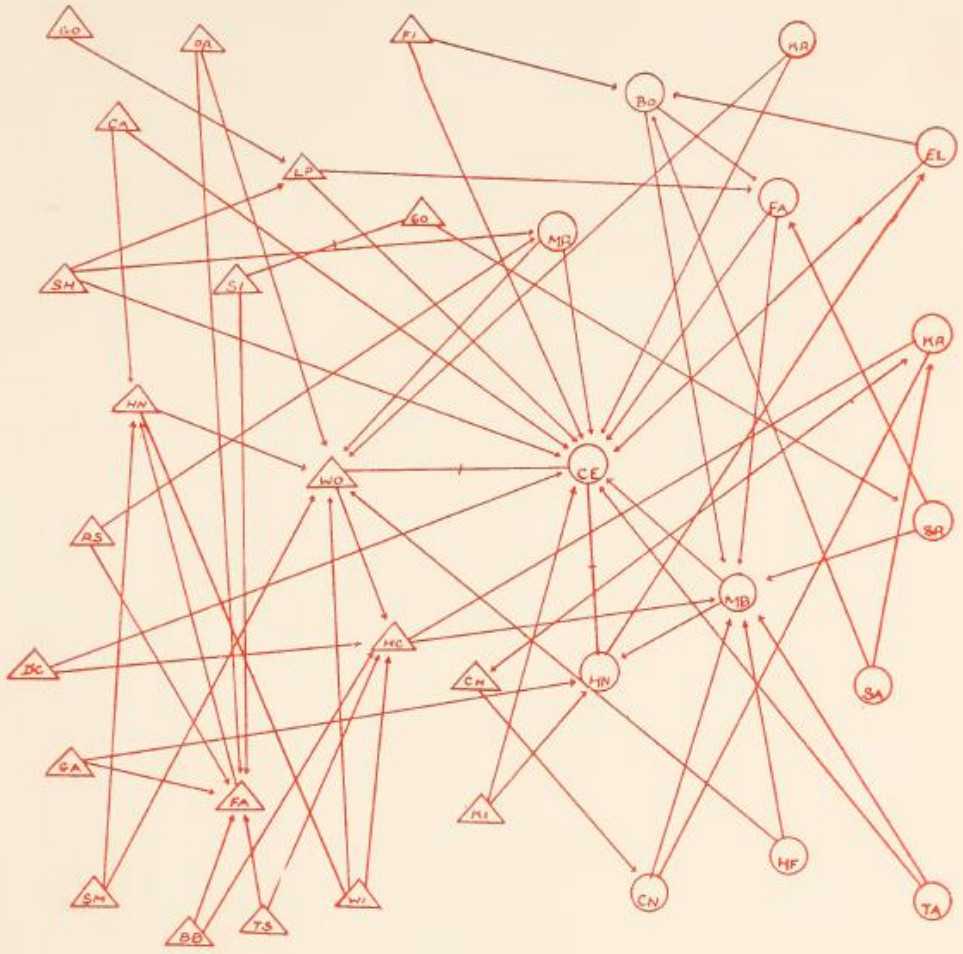
J. L. MORENO, M.D.

Nervous and Mental Disease Publishing Co.
Washington, D. C.
1934

Jacob Levy Moreno

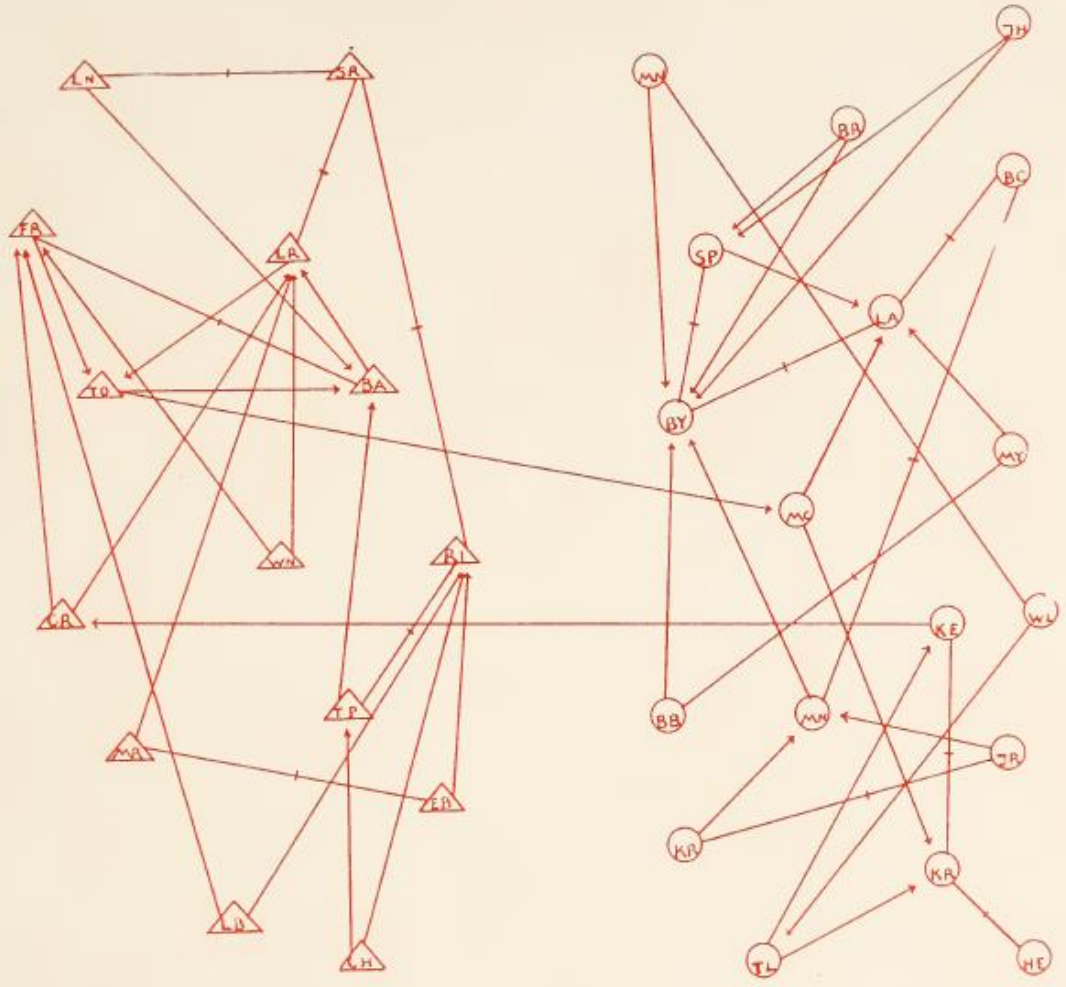
1934





Class Structure, 1st Grade

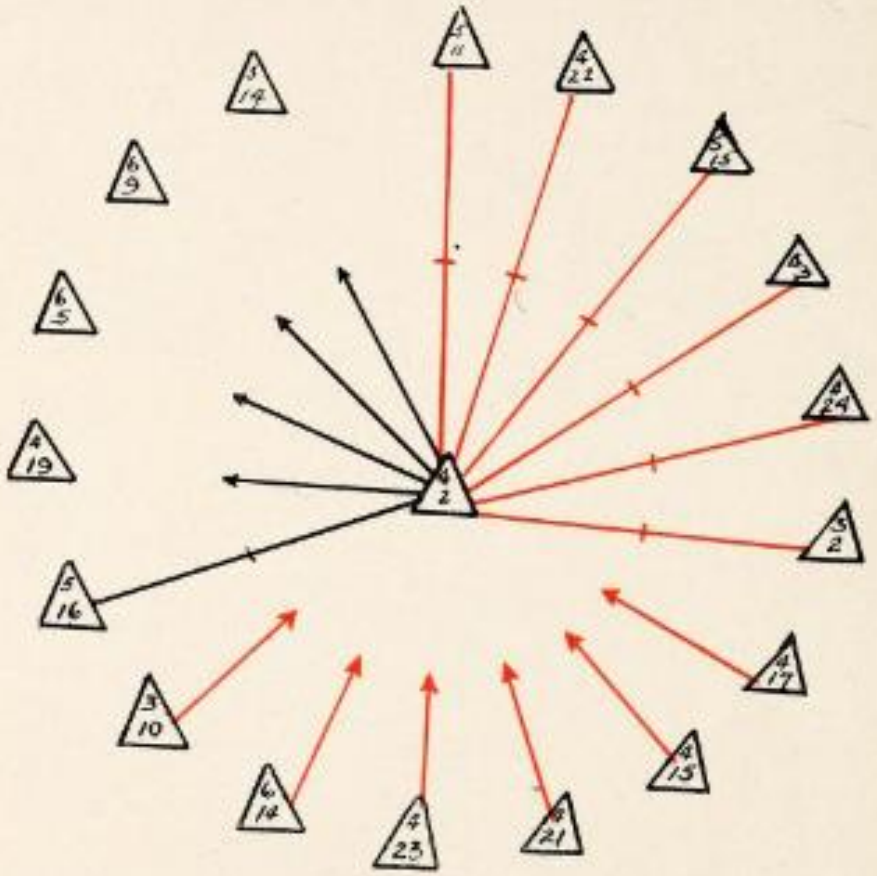
21 boys and 14 girls. *Unchosen*, 18, GO, PR, CA, SH, FI, RS, DC, GA, SM, BB, TS, WI, KI, TA, HF, SA, SR, KR; *Pairs*, 3, EI-GO, WO-CE, CE-HN; *Stars*, 5, CE, WO, HC, FA, MB; *Chains*, 0; *Triangles*, 0; *Inter-sexual Attractions*, 22.



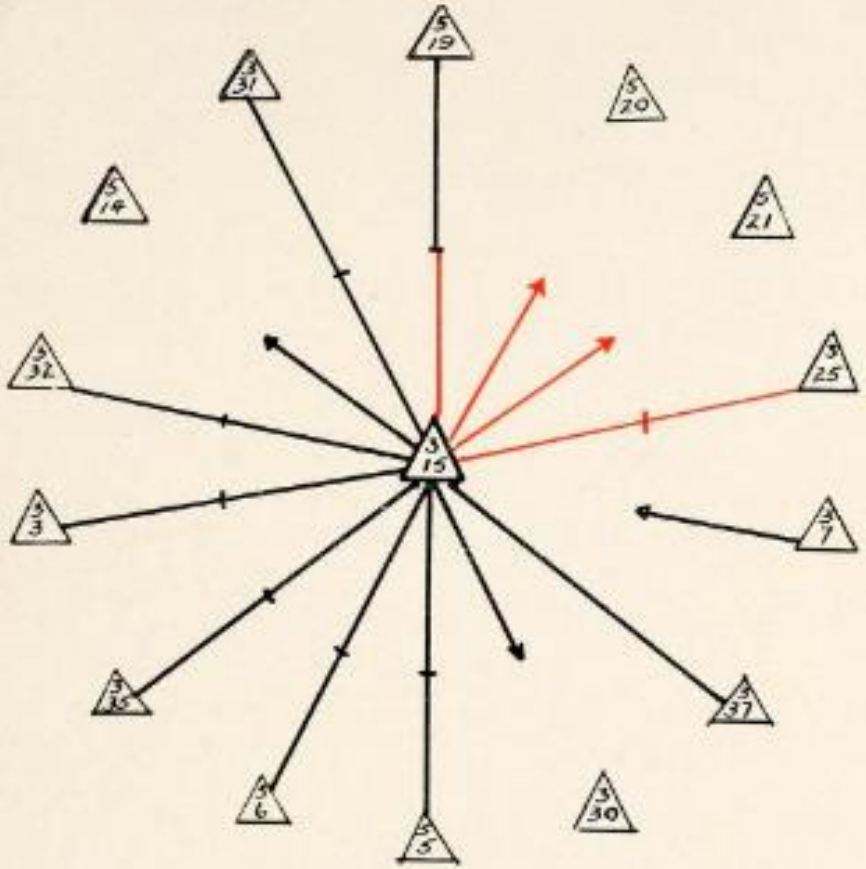
Class Structure, 7th Grade

14 boys and 18 girls. *Unchosen*, 5, WN, CH, LB, JH, BR; *Pairs*, 15, LN-SR, SR-LR, SR-RI, FR-BA, RI-TP, MR-EB, KR-HE, KR-JR, KR-KE, BB-MY, WL-MN, MN-BC, BY-LA, LA-BC, BY-SP; *Stars*, 5, LR, RI, BY, LA, MR; *Chains*, 2, SP-BY-LA-BC-MN-BY, LN-SR-RI-TP-BA; *Triangles*, 0; *Inter-sexual Attractions*, 2.

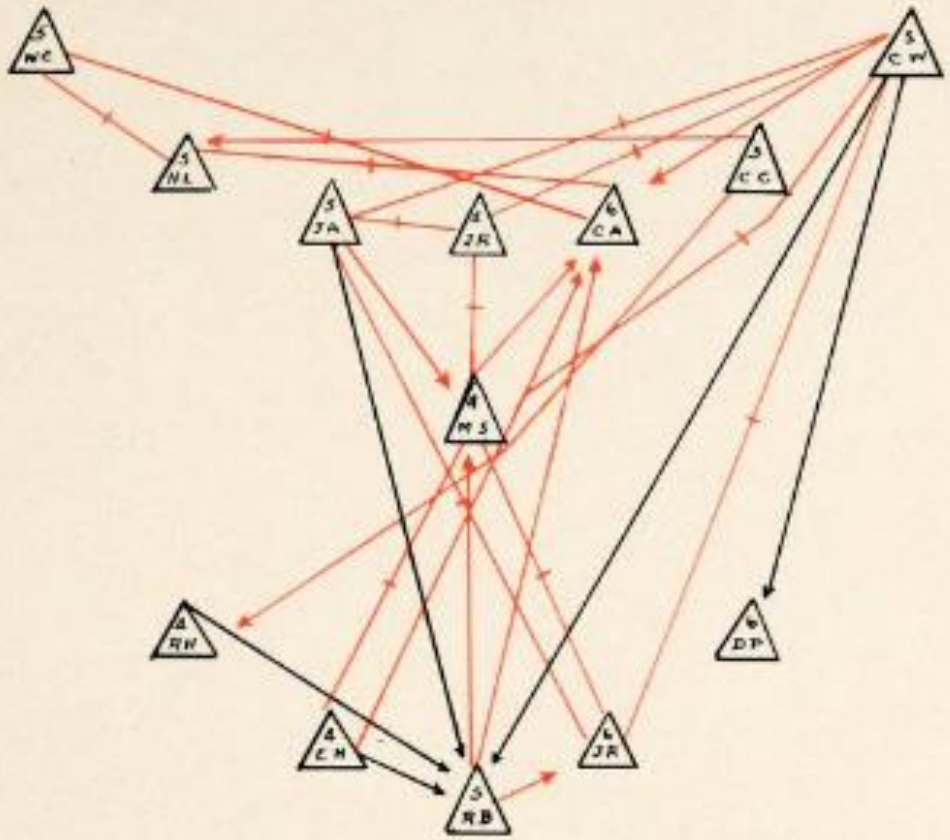
Social atom



The chart shows the position occupied by a typical leader in the school. Note that 4/2 has six reciprocal attractions; six additional attractions toward him by individuals to whom he is indifferent; one mutual dislike and four additional dislikes, each of whom feels neutral toward him. This is a high "batting average." The total number of individuals (17) toward whom relations have been expressed in the course of the survey, distributed over four forms, indicates that 4/2 takes an active attitude toward his fellow students and to life and is not on the "side lines."



The chart depicts an isolated individual. The student 3/15 arouses active feelings—he is the focus of a great deal of antipathy; he likes 3/25, and is accepted by him; he also likes two others, 5/20 and 5/21 who, significantly are new boys and who are indifferent to him, and he also likes 5/19, who, however, dislikes him. This shows that 3/15 has formed only one satisfactory contact in the whole school, while there are seven mutual dislikes, two dislikes for persons who are indifferent to him, and one dislike toward him from a person toward whom he is indifferent.



A Football Team

The chart shows the football team which includes two substitute backs. Note that the quarterback 4/MS is the focus of attraction forces on team, as he should be, but that fullback 5/RB is liked by nobody on the team, and is disliked by four: 4/RH and 4/EH of the back field, and 5/JA and 5/CW of the line. It is easy to see that when 5/RB is running with the ball he is not apt to get the maximum of coöperation in interference and blocking.

(See Map of a School Community, pp. 49–50)

Beiträge zur Sozialpsychologie des Haushuhns.

Von

THORLEIF SCHJELDERUP-EBBE.[1]

Anhang: **Tierpsychologie und Soziologie des Menschen.**

Von

DAVID KATZ.

1. Kapitel.

Die soziale Stellung jedes Individuums.

Thorleif Schjeldrup-Ebbe 1922

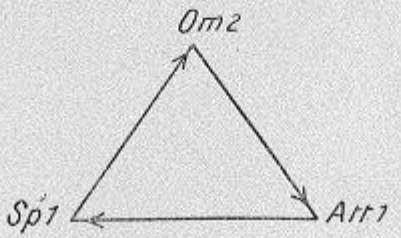
First demonstration of a social hierarchy in animals (chickens)

Demonstrated transitivity as feature of linear social hierarchies

Hackliste Nr. 2. 10 Hennen: Om2, Attl, Spl, Gaml, P2, Lvkl, Sm2, Bll, Gr, Hv2.

Om2	hackt	alle aufser Spl.
Attl	„	alle aufser Om2.
Spl	„	alle aufser Attl.
Gaml	„	P2, Lvkl, Sm2, Bll, Gr, Hv2.
P2	„	Lvkl, Sm2, Bll, Gr, Hv2.
Lvkl	„	Sm2, Bll, Gr, Hv2.
Sm2	„	Bll, Gr, Hv2.
Bll	„	Gr, Hv2.
Gr	„	Hv2.
Hv2	„	keine.

Betrachtet man Hackliste 2, so sieht man, dafs hier keine Henne vorkommt, die Despot über alle anderen Hennen ist. Das ist ein Gegensatz zu dem, was wir in Hackliste 1 fanden. 3 Hennen in Hackliste 2, Om2, Attl und Spl, sind gleich günstig gestellt, indem sie alle 3 jede 9 Hennen hacken. Zu einander verhalten sich die 3 Hennen wie die Hennen Pl, Oml und FrSl in Hackliste 1, insofern Om2 die Attl hackt, welche ihrerseits Spl hackt, während Spl Om2 hackt. Diese 3 Hennen hacken sich gegenseitig im Dreieck, was Fig. 1 zu veranschaulichen sucht, bei der die Richtung der Pfeilspitzen die Richtung bezeichnet, in welcher das Hacken verläuft.



Figur 1

Illustration of Network Intransitivity

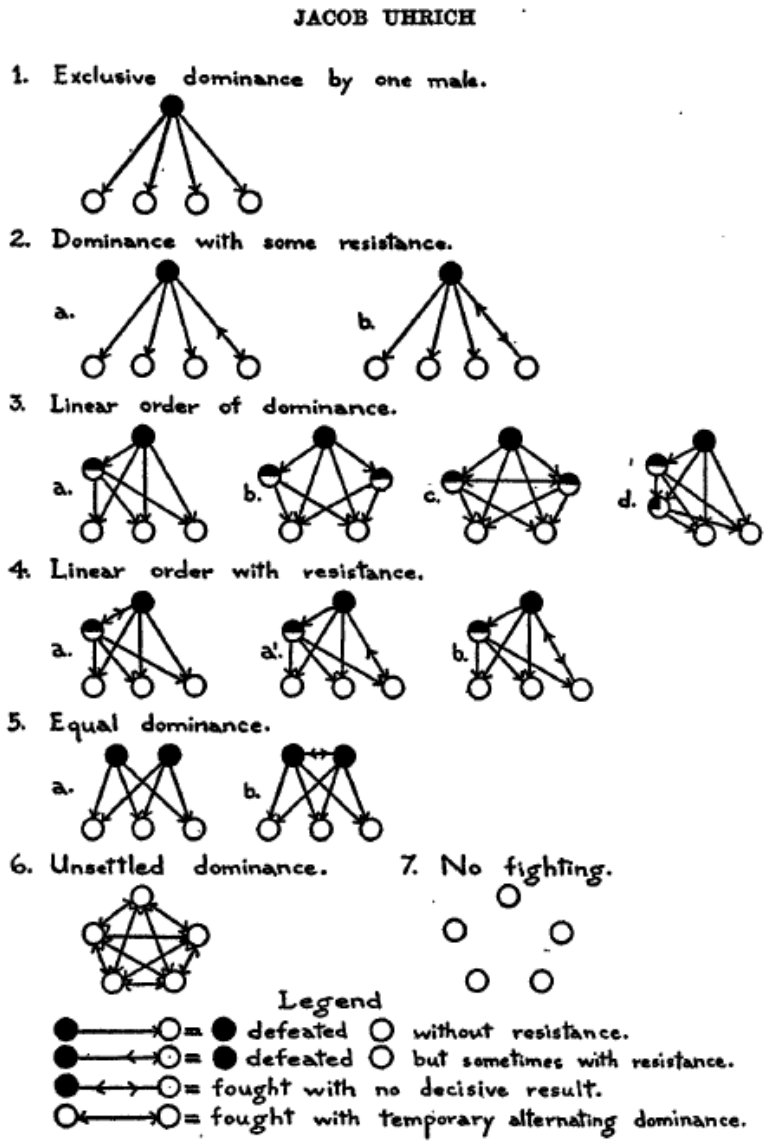


FIG. 1. DIAGRAMS ILLUSTRATING REPRESENTATIVE TYPES OF FIGHTING RELATIONSHIPS AMONG MALE MICE

Jacob Uhrich 1937

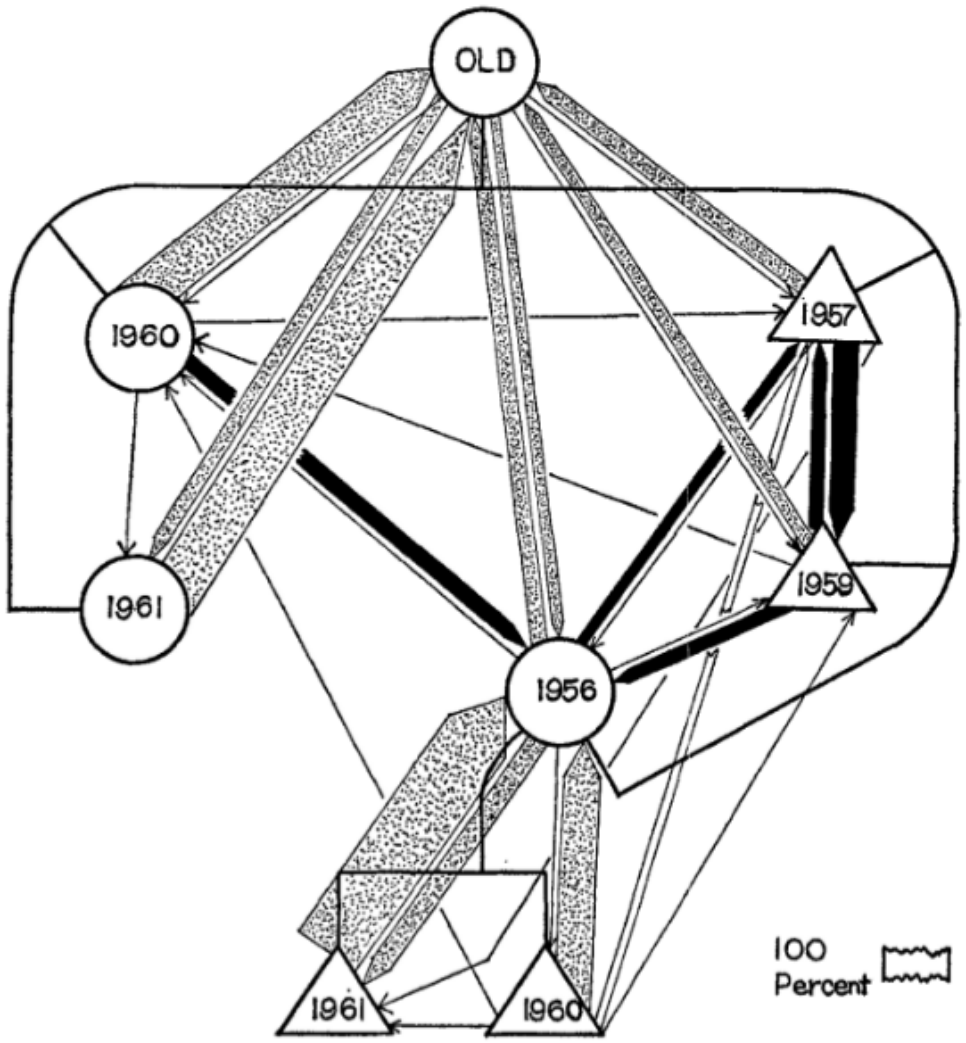
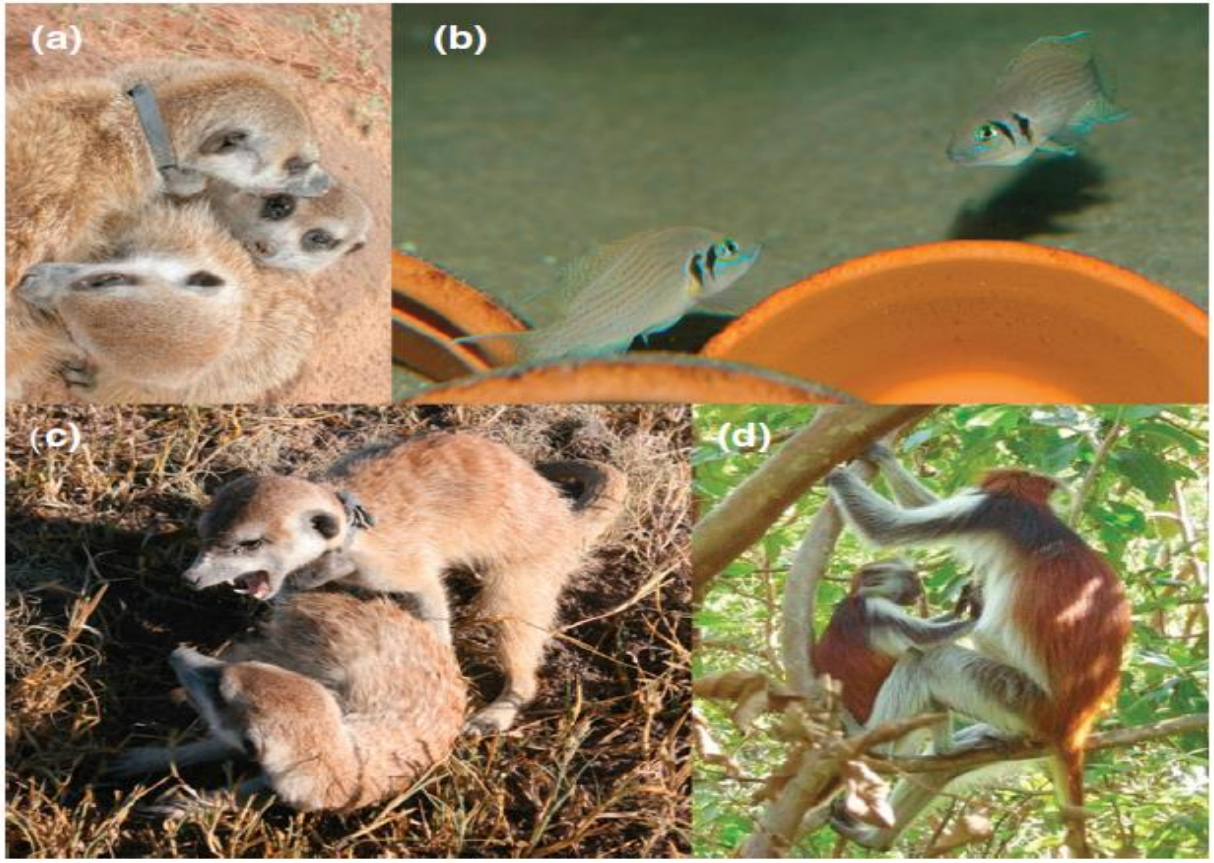


Fig. 3 Grooming relations in genealogy 4 from March 21, 1961 through December 25, 1961. Drawn from the data in table 3. Explanation in text.

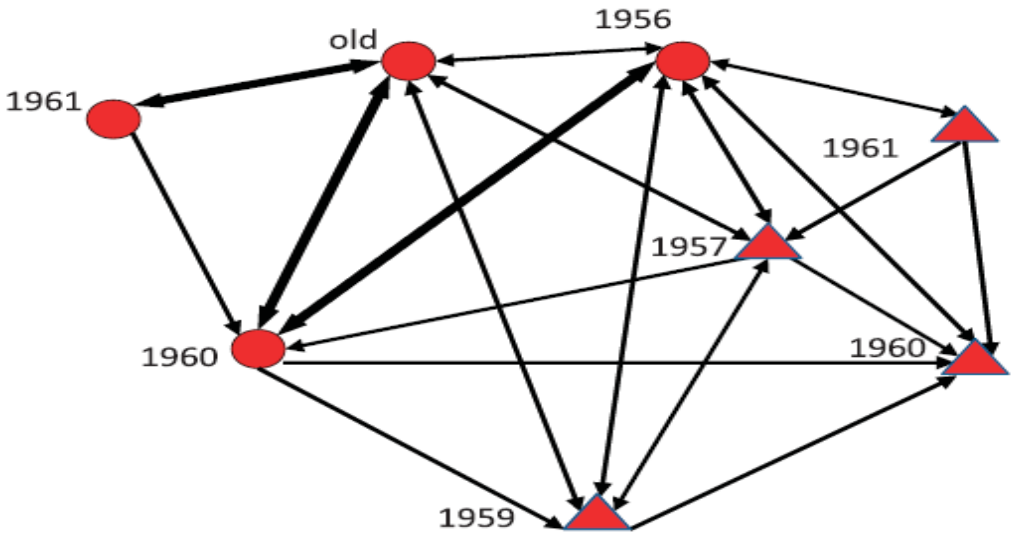
Donald Sade 1965

Explosion of Animal Social/Interaction Network Analyses



TRENDS in Ecology & Evolution

Figure I. A selection of animals showing examples of behavioural interactions that can be directly observed: **(a)** grooming interactions meerkats (*Suricata suricatta*) (credit Joah Madden), **(b)** an aggressive interaction in *Neolamprologus pulcher* (credit Roger Schürch), **(c)** dominance assertion in meerkats (credit Jess Isden), and **(d)** grooming interaction in red colobus monkeys (*Procolobus badius*) (credit Jess Isden).

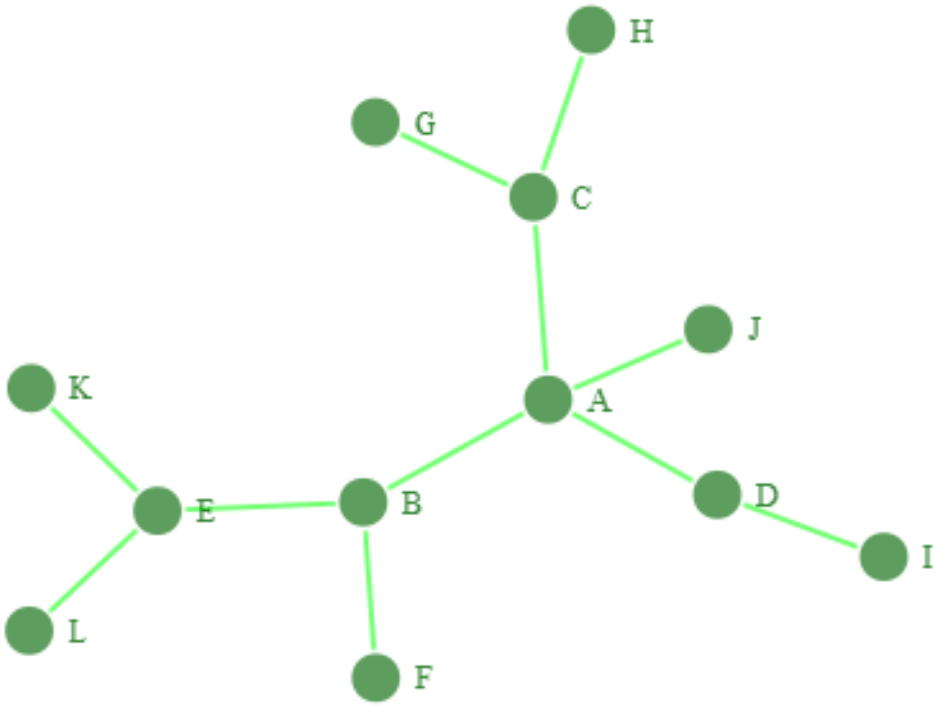


SNA makes it possible to build up larger social units and structures from observations of pair-wise interactions

In social groups, an individual's interaction with another individual is unlikely to be independent of its interactions with others

Croft et al 2011 TREE
Brent et al 2010 Am J Primatol

What is a Social Network ?



Data Structures

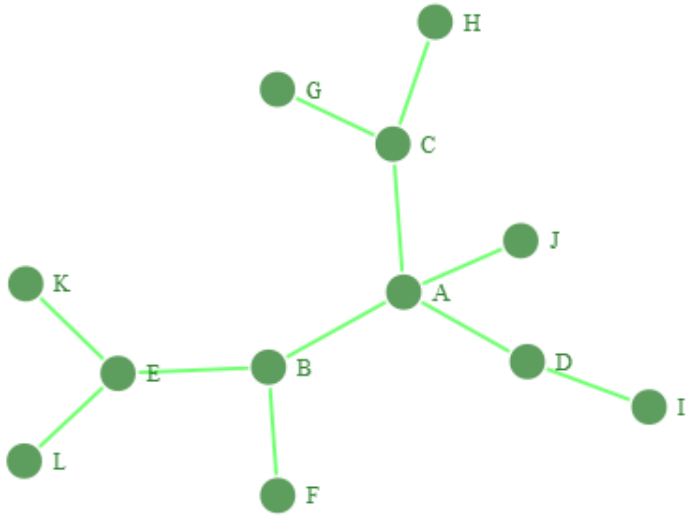
Edgelist

A	B
A	C
A	D
A	J
B	E
B	F
C	G
C	H
D	I
E	K
E	L

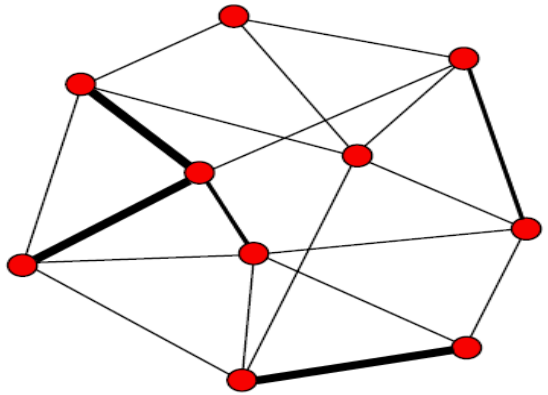
Sociomatrix / Adjacency Matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	.	**1**	**1**	**1**	0	0	0	0	0	**1**	0	0
B	**1**	.	0	0	**1**	**1**	0	0	0	0	0	0
C	**1**	0	.	0	0	0	**1**	**1**	0	0	0	0
D	**1**	0	0	.	0	0	0	0	**1**	0	0	0
E	0	**1**	0	0	.	0	0	0	0	0	**1**	**1**
F	0	**1**	0	0	0	.	0	0	0	0	0	0
G	0	0	**1**	0	0	0	.	0	0	0	0	0
H	0	0	**1**	0	0	0	0	.	0	0	0	0
I	0	0	0	**1**	0	0	0	0	.	0	0	0
J	**1**	0	0	0	0	0	0	0	0	.	0	0
K	0	0	0	0	**1**	0	0	0	0	0	.	0
L	0	0	0	0	**1**	0	0	0	0	0	0	.

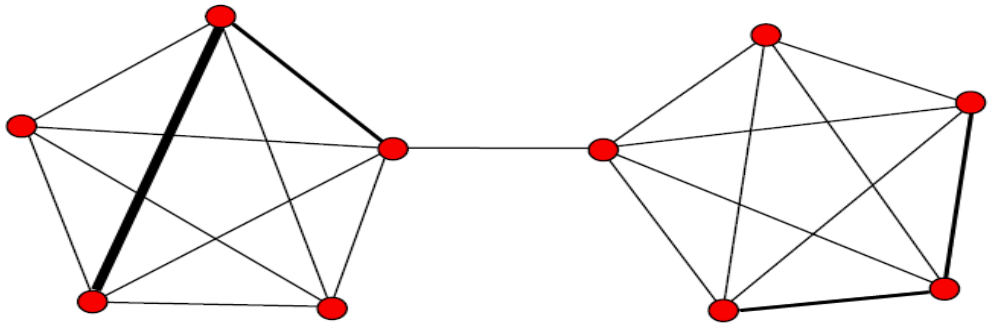
Network / Sociogram



Advantages of SNA



Can measure traditional individual indices of social interaction (e.g. ***degree*** = number of social partners, ***strength*** = frequency of interaction)

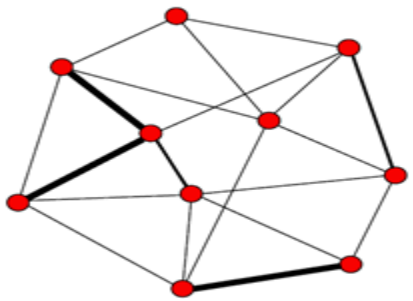


Can ascertain additional measures of an individuals sociality via indirect connections e.g. ***eigenvector centrality*** = number of social partners that one's own social partners additionally possess)

Can also detail social structure of entire network - e.g. ***assortativity*** = tendency of nodes of one class to connect to nodes of the same class which leads to sub-clusters & presence of ***brokers*** (individuals who bridge relationships between clusters)

Across taxa, SNA metrics at the individual, relationship and social structure levels are related to various measures of fitness e.g. social learning, stress, dominance, cooperation

SNA measures



Degree centrality - how many direct ties an individual has with others.

Eigenvector centrality **=** number of social partners that one's own social partners additionally possess

Outdegree - the number of ties that originate from the focal animal (e.g., the number of grooming events initiated by the focal animal)

Indegree - the number of ties that are directed at the focal animal (e.g., the number of grooming events that the focal animal receives).

Strength - the number or rate of interactions (e.g., grooming events) between a pair of animals

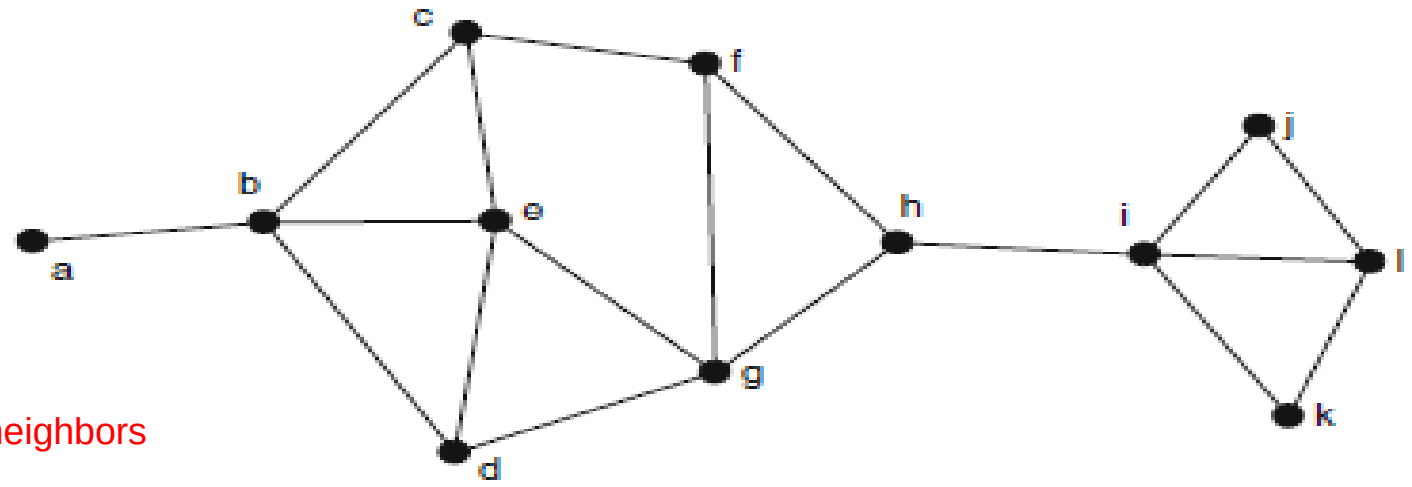
Average path length (L) - the average distance between pairs of individuals within a network . High L scores indicate that some individuals only interact with others very indirectly.

Compactness - how directly connected individuals in the network are (controlling for the size of the network)

Density (D) - how many potential ties between individuals are actually present. Expressed as a percentage, with high scores indicating 'saturated' networks (all possible ties are present) and low scores indicating 'sparse' networks

Clustering coefficient (C) – how 'cliquish' the network is. High C scores indicate that on average, focal individuals are surrounded by other individuals that are well connected to each other

Betweenness - the extent to which an individual is between all other individuals within the network.



Number of immediate neighbors

Degree to which an individual's immediate neighbors are connected

Average number of connections between individual & any other individual

The number of shortest paths between pairs of individuals that pass through an individual

Table 1 Some individual-based measures for the network in Fig. 3

Node/ individual	Node degree	Path length (to/from a node)	Clustering coefficient	Node betweenness
a	1	3.64	0	0
b	4	2.73	0.333	10.5
c	3	2.36	0.333	5.8
d	3	2.36	0.667	4.3
e	4	2.27	0.5	5.3
f	3	2.09	0.333	8.8
g	4	2	0.333	18.2
h	3	2	0.333	28
i	4	2.36	0.333	24.5
j	2	3.18	1	0
k	2	3.18	1	0
l	3	3.09	0.667	0.5
Mean values	3	2.61	0.486	8.83

Krause et al 2009 Behav Ecol Sociobiol

What sort of questions can be asked & answered using SNA?

Individual Level

- What individual characteristics (personality/sex/age/etc.) predict an individuals position within a SN?
- How does the introduction/removal of specific individuals alter the SN?
- How do animals of similar/dissimilar strains/genetics assort in a SN?
- How does an individual's position in the network affect their future reproductive success, well-being, health, *etc.* ?
- Does an individual's position in the network correlate with brain specific phenotypes?

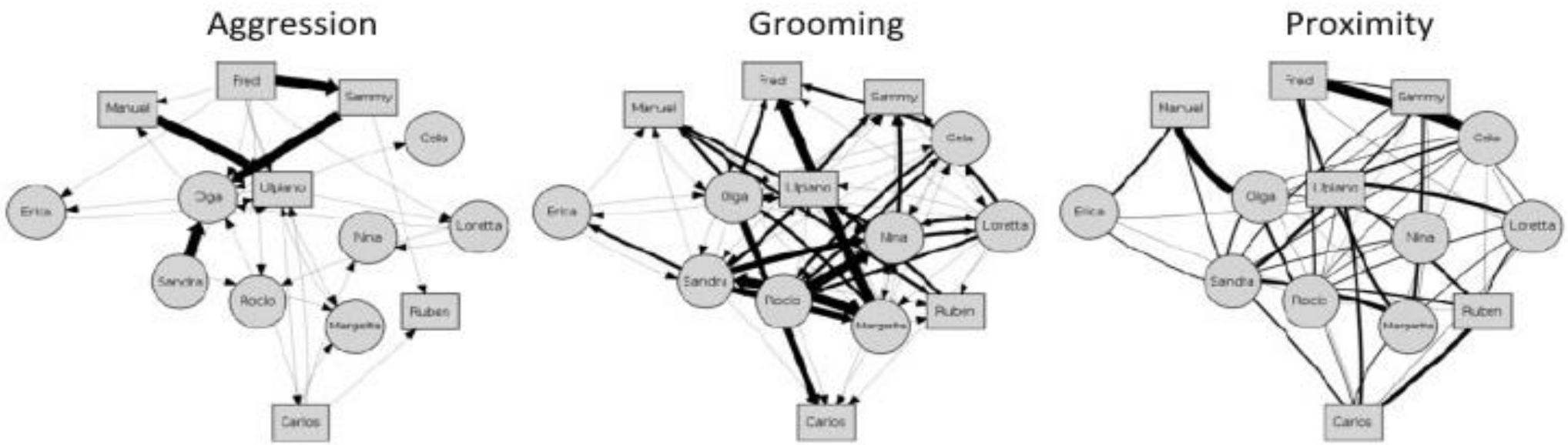
Network Level

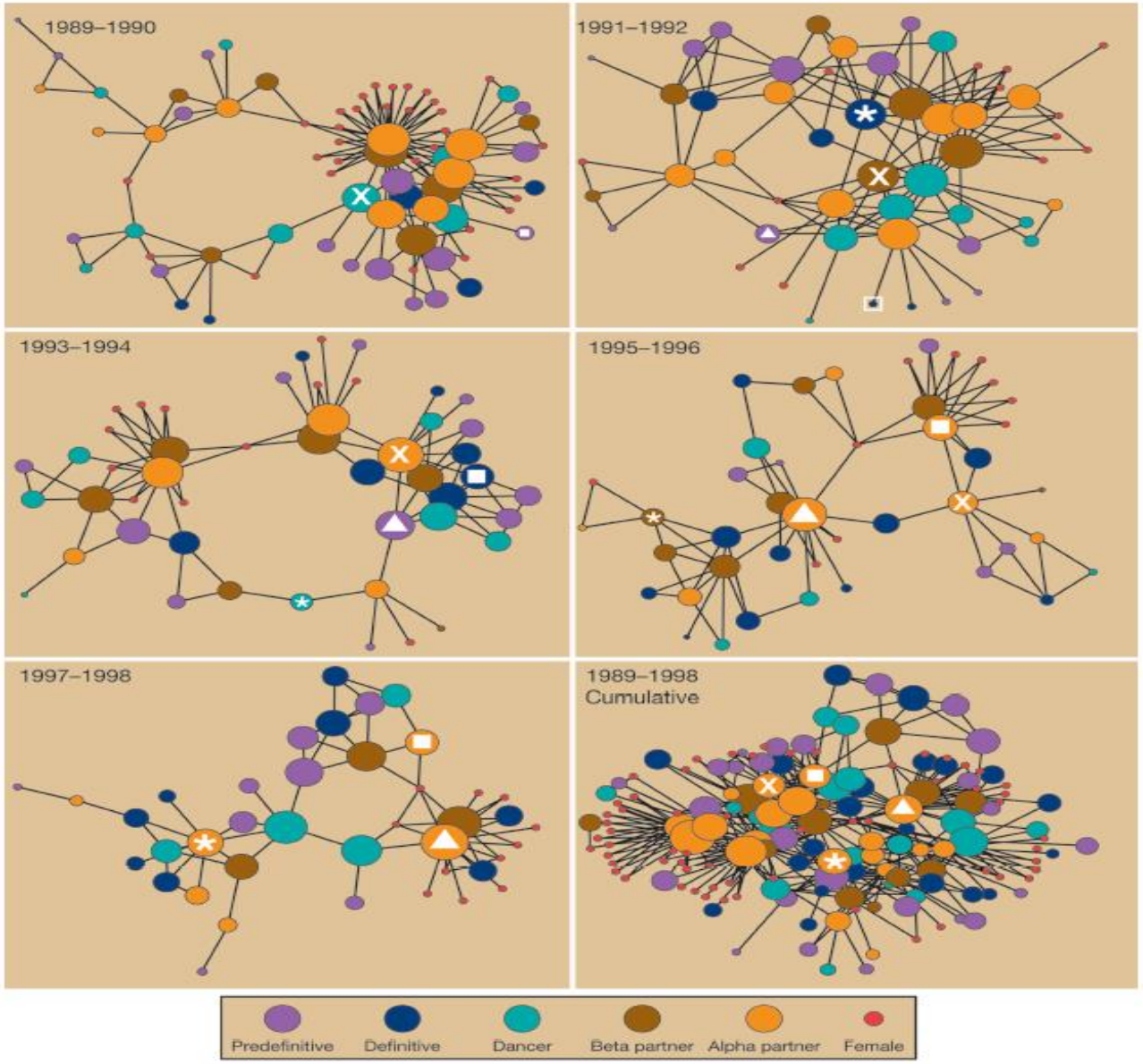
- How does the number & characteristics of individuals alter the composition of the SN?
- How does the external environment (e.g. physical space) alter the composition of the SN?
- How do different network types facilitate the spread of information, disease, cooperation, *etc.*?
- Do young animals establish different social networks to old animals?

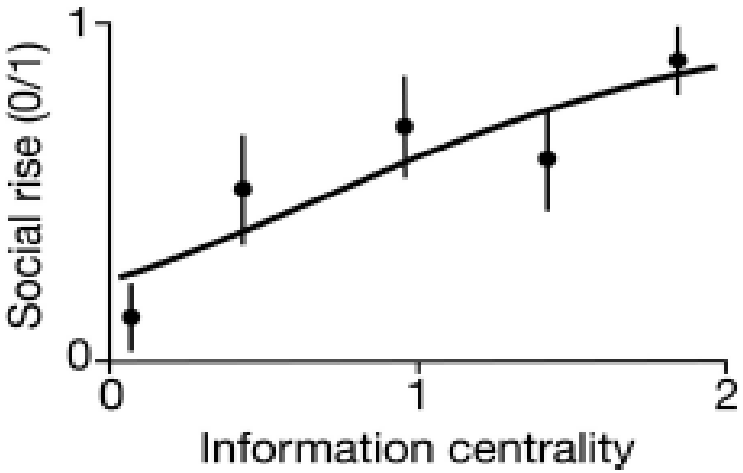
Aggression, Grooming and Group-Level Cooperation in White-Faced Capuchins (*Cebus capucinus*): Insights From Social Networks

MARGARET C. CROFOOT[1–3], DANIEL I. RUBENSTEIN[3*], ARUN S. MAIYA[4], AND TANYA Y. BERGER-WOLF[4]
[1]*Division of Migration and Immuno-ecology, Max Planck Institute for Ornithology, Radolfzell, Germany*
[2]*Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panamá*
[3]*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey*
[4]*Department of Computer Science, University of Illinois at Chicago, Chicago, Illinois*

American Journal of Primatology 73:821–833 (2011)







The centrality of young male long-tailed Mannakins predicts their social rise 4.8 years later – Centrality of adult males did not predict social status

McDonald 2007 PNAS

Using R packages to analyze social networks

- igraph
- statnet
- sna
- network
- intergraph

Packages for global manipulation of network data

- keyplayer
- Rsiena
- ergm
- tsna
- timeordered

Packages modeling specific features of social networks

- asnipe
- assortnet
- anidom
- compete

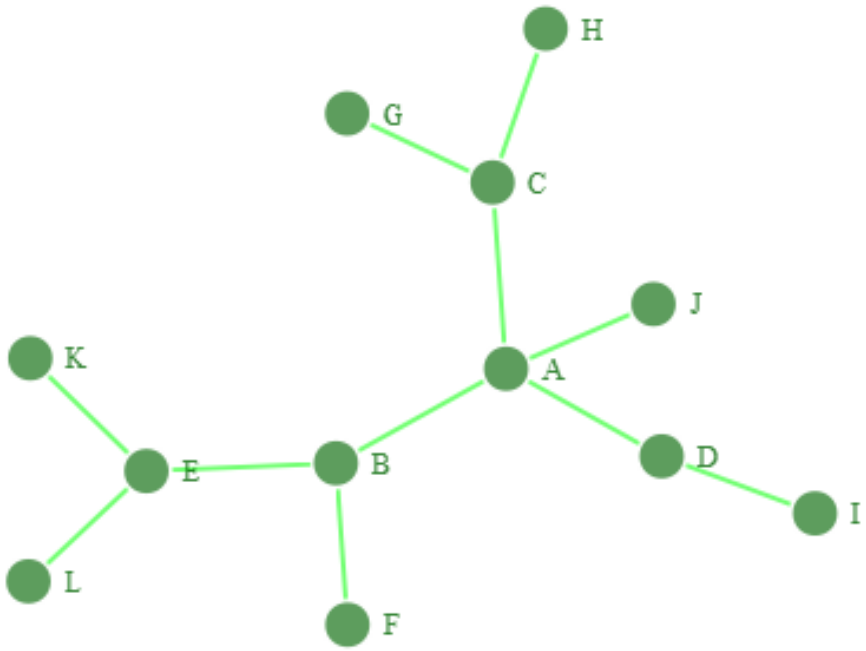
Packages that are more specific for animal behavior

- ….
- many others

(there are many ! – these are just some examples)

Visualizing Networks

Social associations that are observed between individuals over time …



… how to best visualize these sociotemporal patterns ?

What to visualize ?

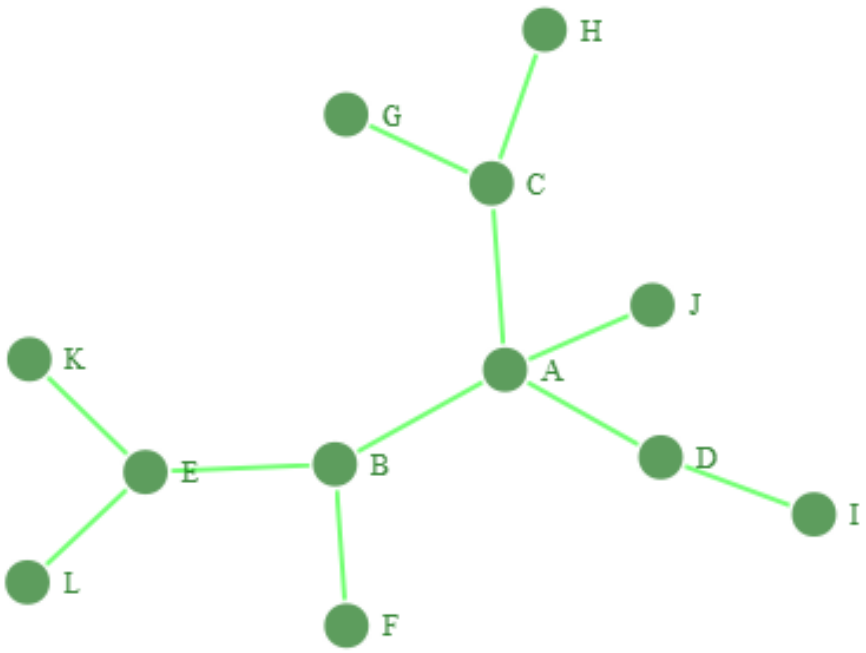
What questions are you interested in?

What do we want to communicate ?

Who is the audience ?

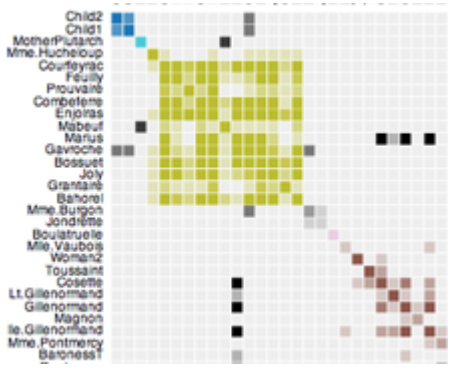
How do you want it to be viewed ?

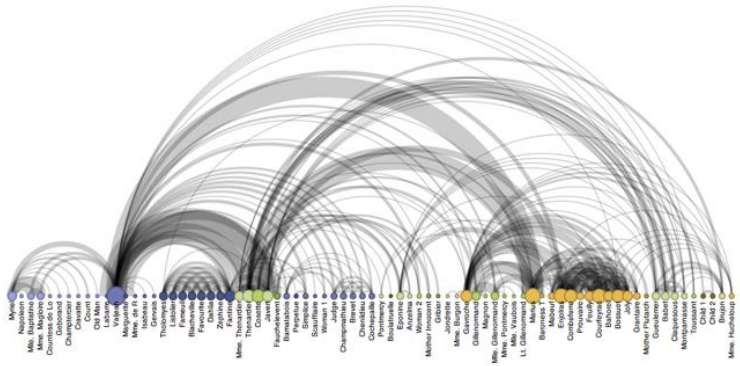
What works best for your particular workflow?

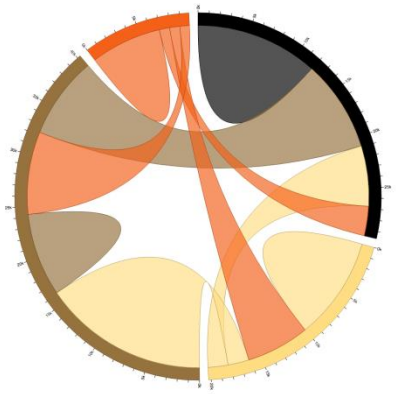


What to visualize ?

- Sociomatrices
- Social Networks
- Social Atom
- Chord diagrams
- Arc diagrams
- Hierarchical trees
- Timeline/Time-ordered networks
- etc.







How to visualize ?

R

Pajek

D3

JointJS

python

UCI

sigmajs

p5*

Gephi

NetDraw

three.js

P

arbor.js

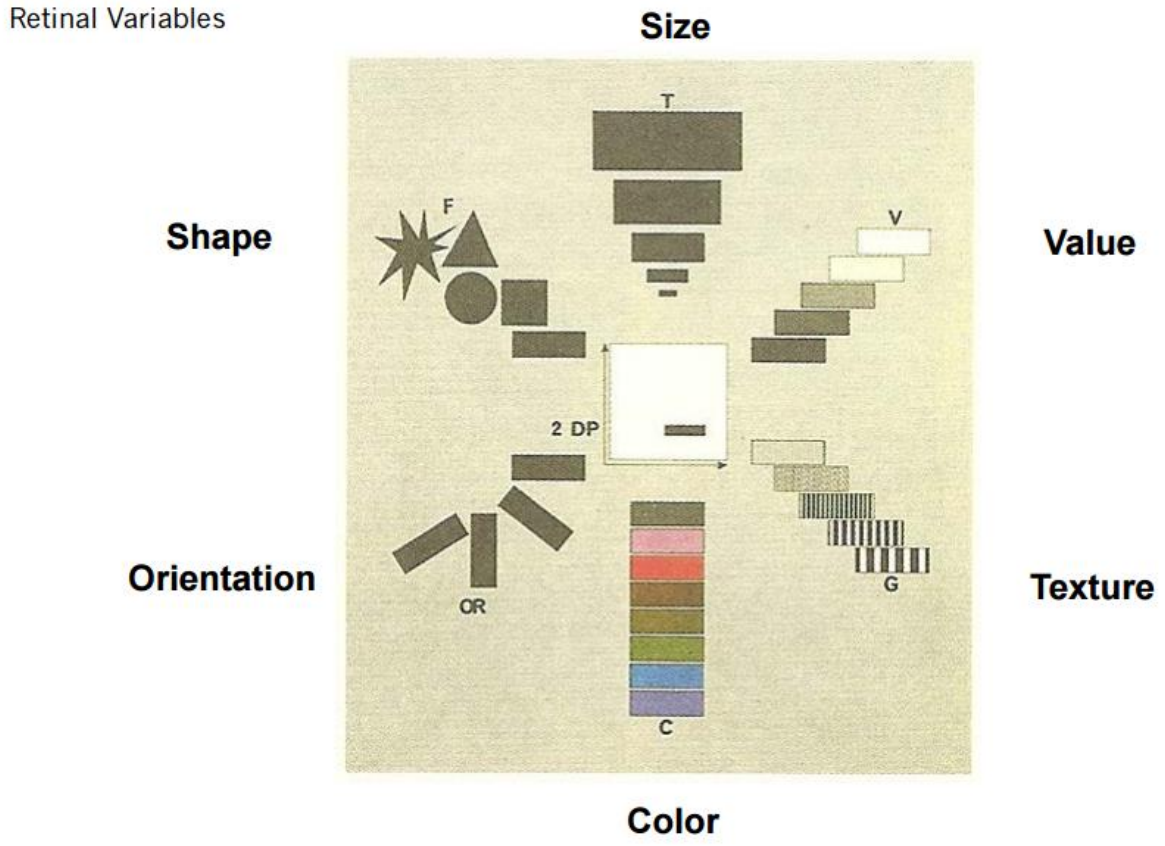
Cytoscape

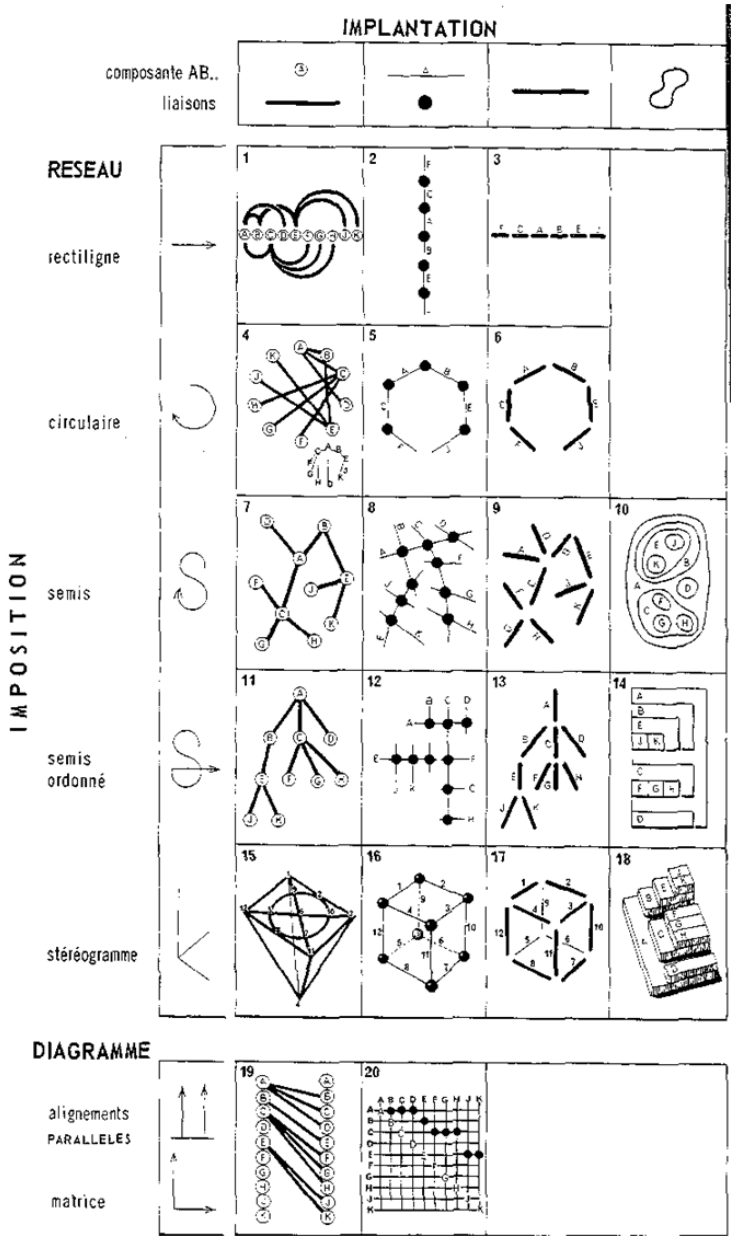
vis.js

Principles of Network Visualization

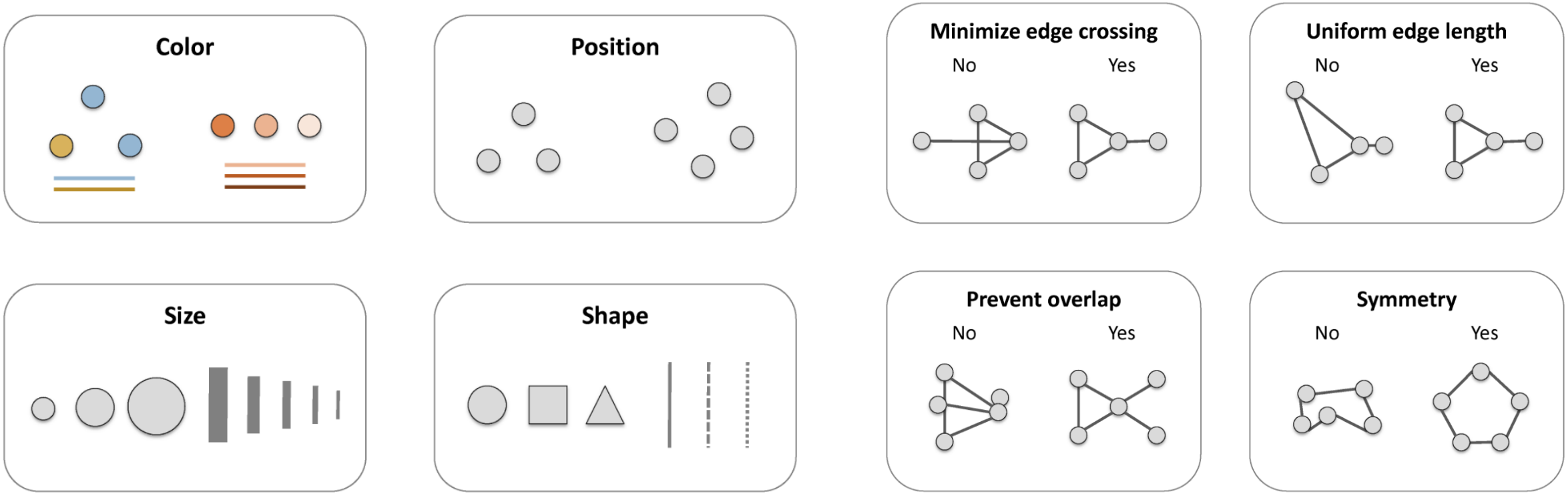
Jacques Bertin 1967

Semiology of Graphics: Diagrams, Networks, Maps



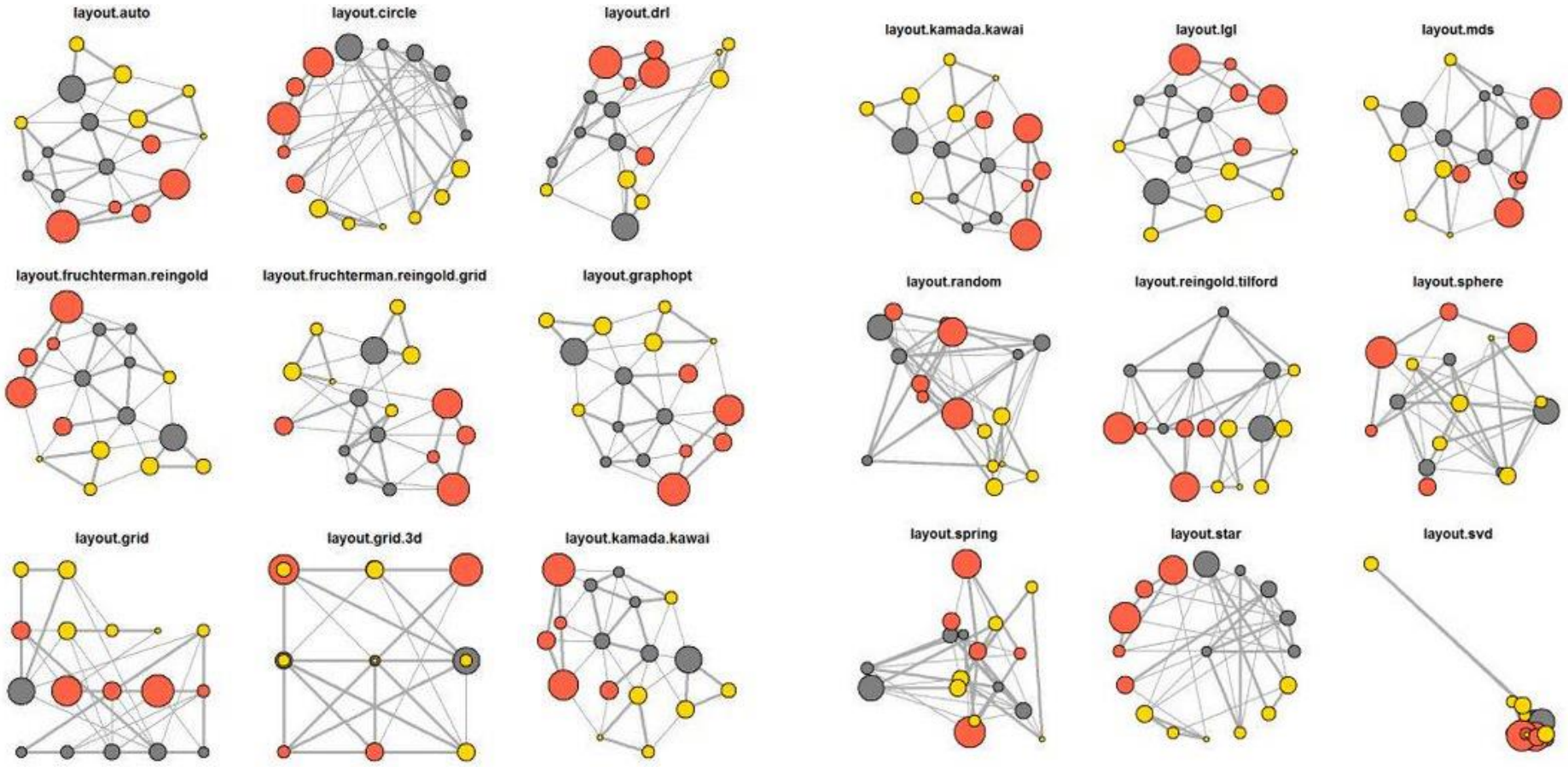


Basic Network Visualization Principles



Katya Ognyanova (Rutgers)

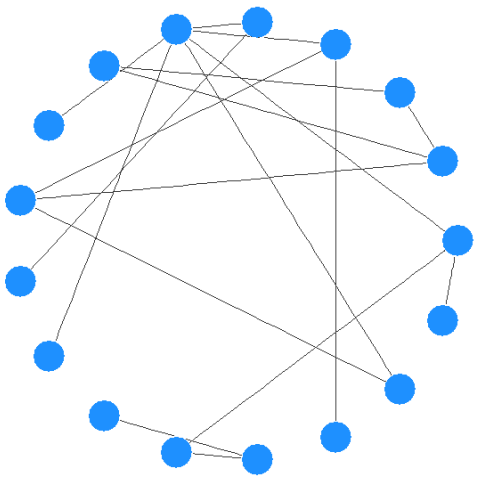
Basic Network Visualization Principles - Layouts



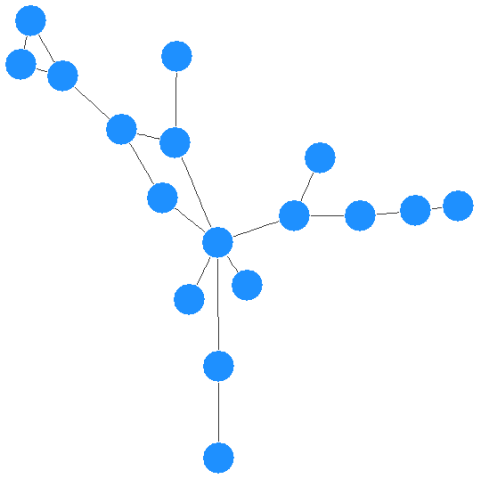
R – igraph package (Gabor Csardi)

Katya Ognyanova (Rutgers)

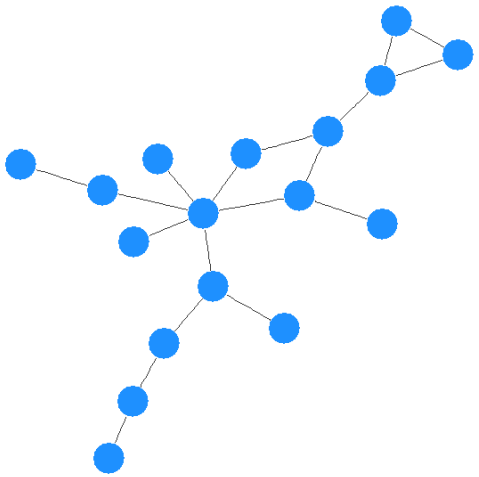
circle



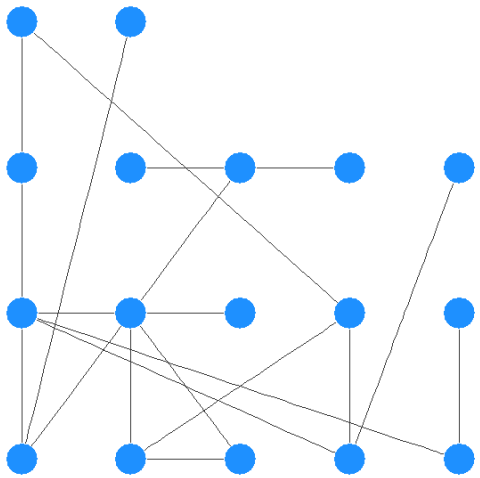
fruchterman-reingold



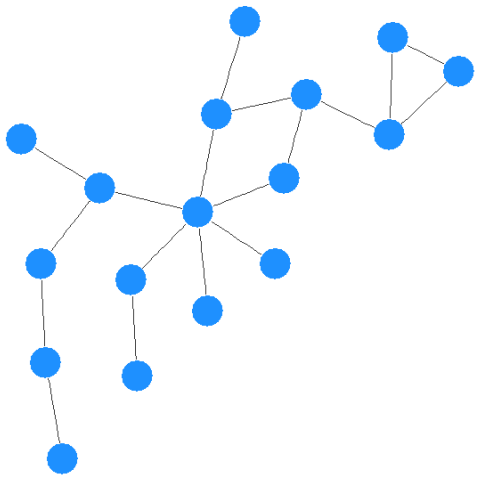
kamada-kawai



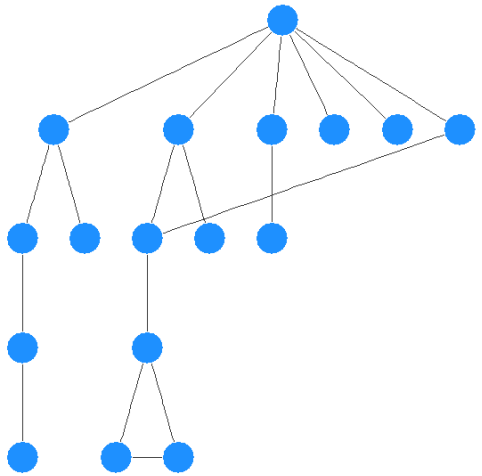
grid



lgl

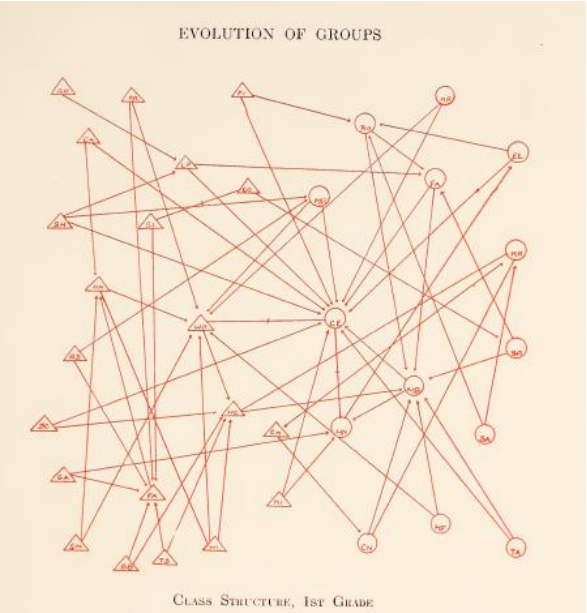


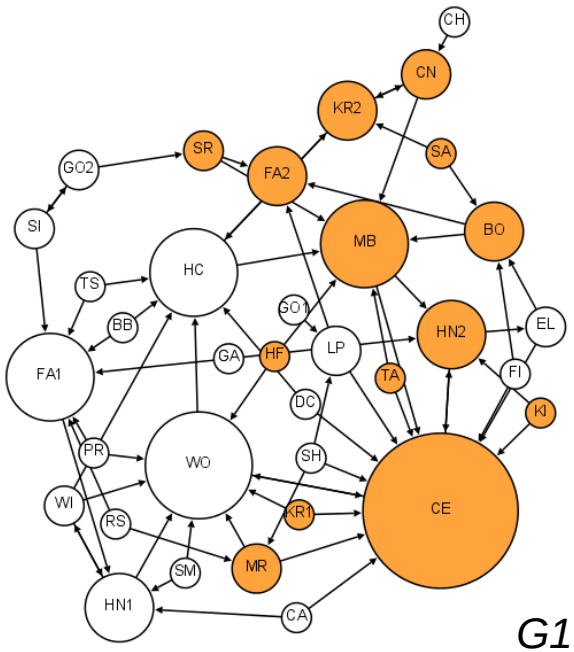
tree



Identifying Network Structures & key nodes

1st Grade





Moreno's original visualizations remade in Gephi by

[Martin Grandjean]

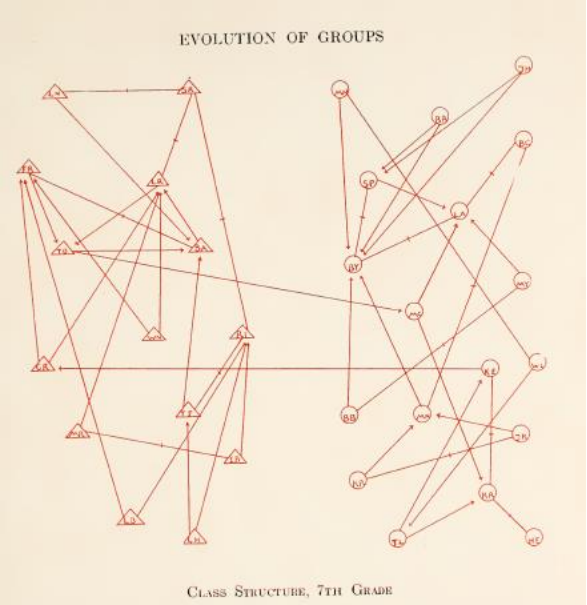
G1: 22 boy-girl edges

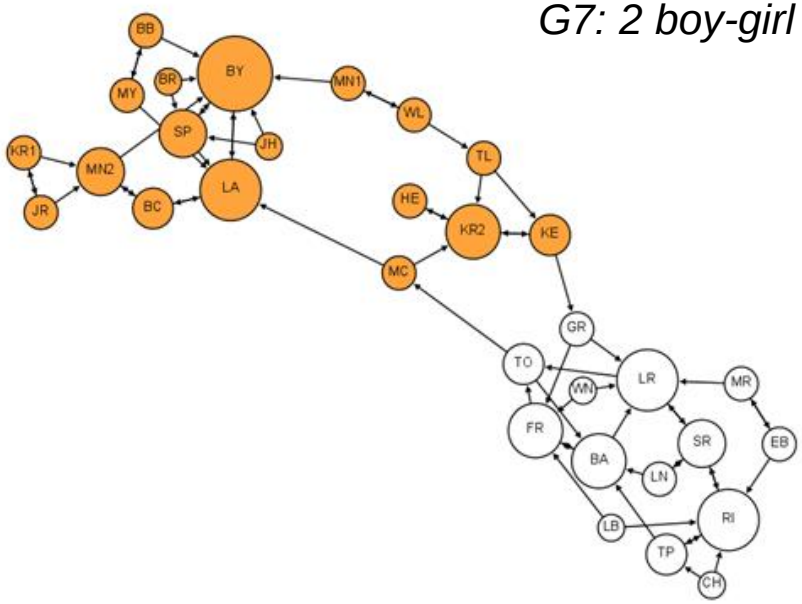
G7: 2 boy-girl edges

Node size – degree

Node color – gender
orange – girls
white – boys

7th Grade



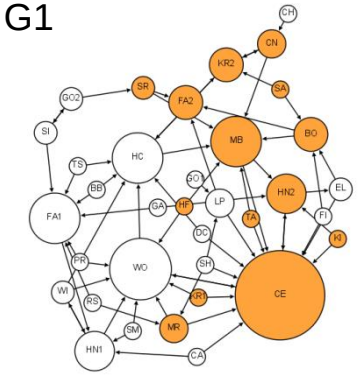


Changes in homophily, reciprocity & community structure across grades

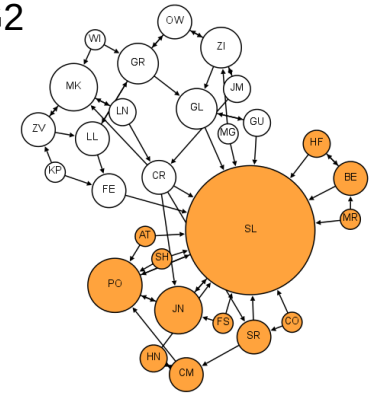
Moreno's original visualizations remade in Gephi by

Martin Grandjean

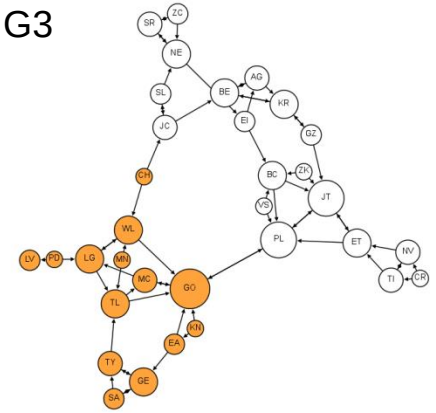
G1



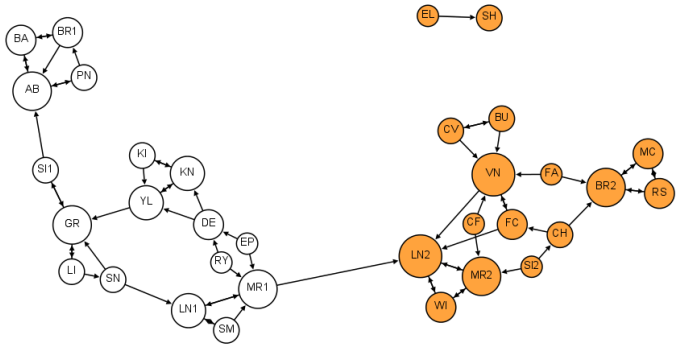
G2



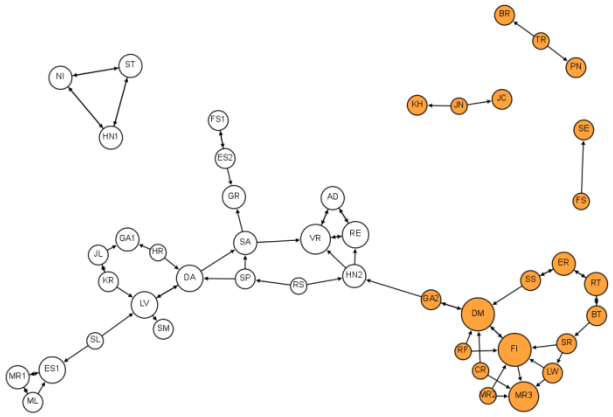
G3



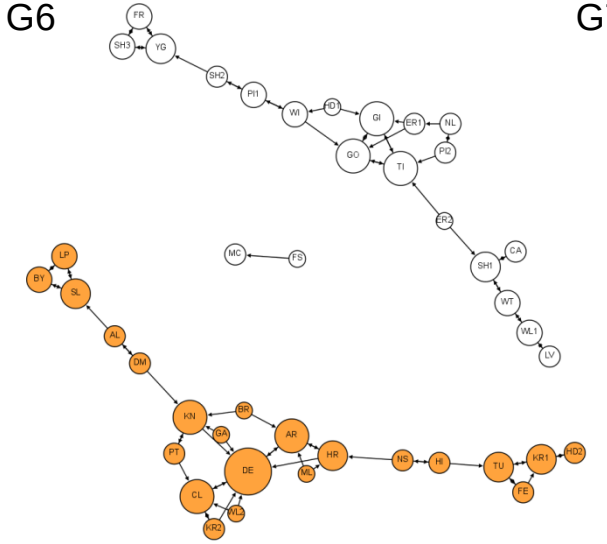
G4



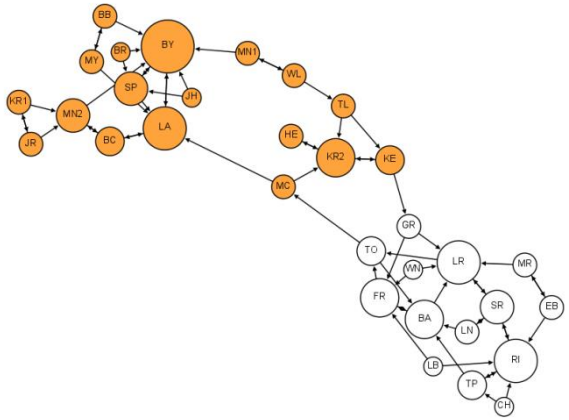
G5



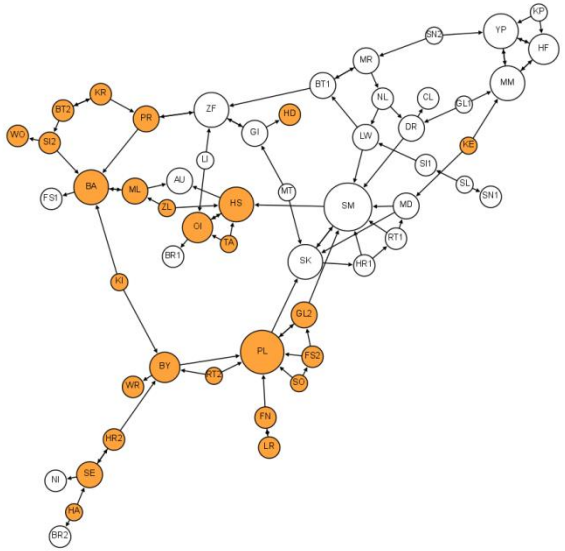
G6



G7



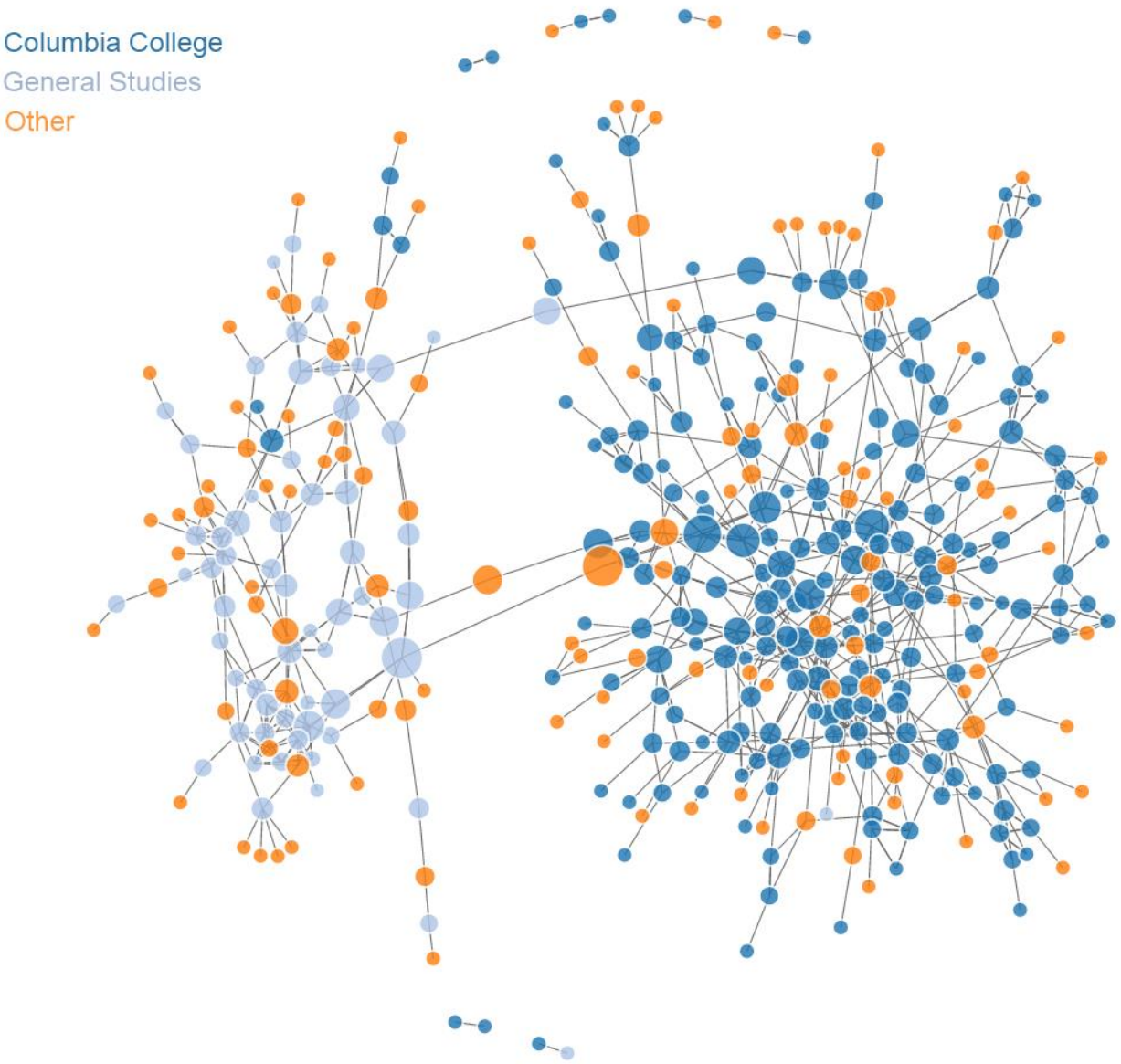
G8



Friendship Network in Columbia Intro Bio Classes

(with Kaytee Turetsky)

Made with D3.js & networkD3 R package



Extensions to Network Visualization

Limits to what can be inferred immediately from a static network visualization

Interactivity – users gain more insight through interactivity & can explore data for themselves

Broswer based :

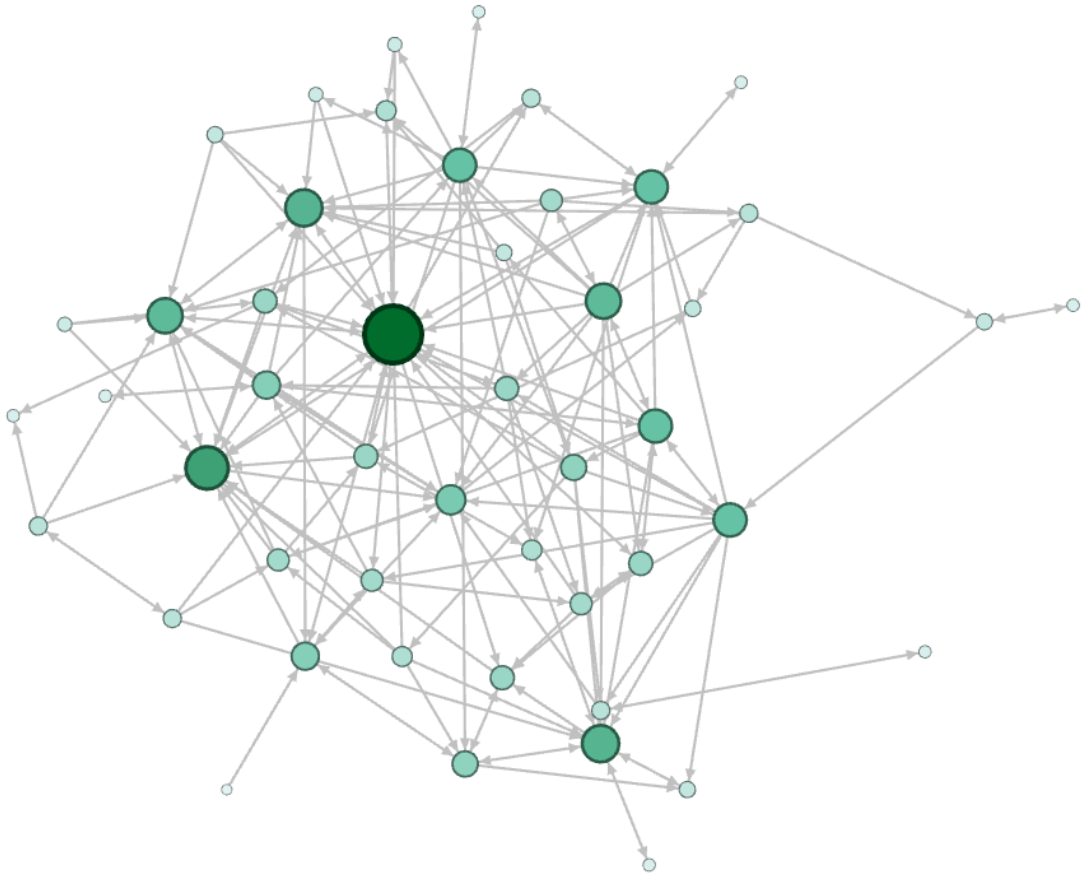
e.g.

- Filtering data
- Mouseover /onClick events
- Linked matrices/networks

Filtering Information - less is more

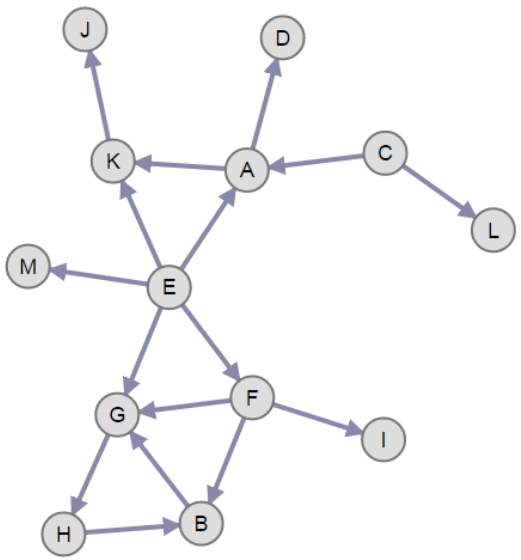
Friendship connections in German school 1880

e.g. identify neighbors in Gephi

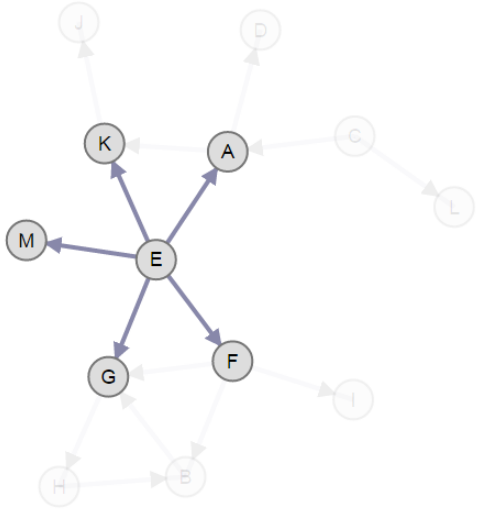


Schmidt has 3 reciprocated friendships, 1 outgoing only & 2 incoming only

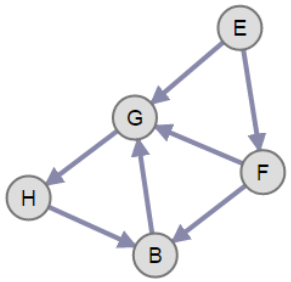
Filtering Information - less is more



e.g. [identify direct neighbors in D3.js]



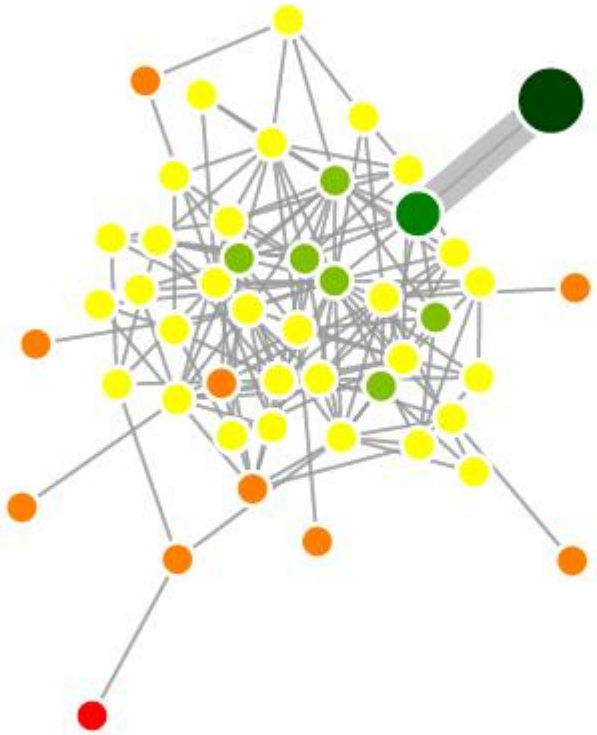
e.g. [identify triads in D3.js]



blockbuilder.org/jalapic

Interactivity

e.g. [find geodesic distances in D3.js]



Who is 1, 2, 3, 4 geodesic distance from Flach ?

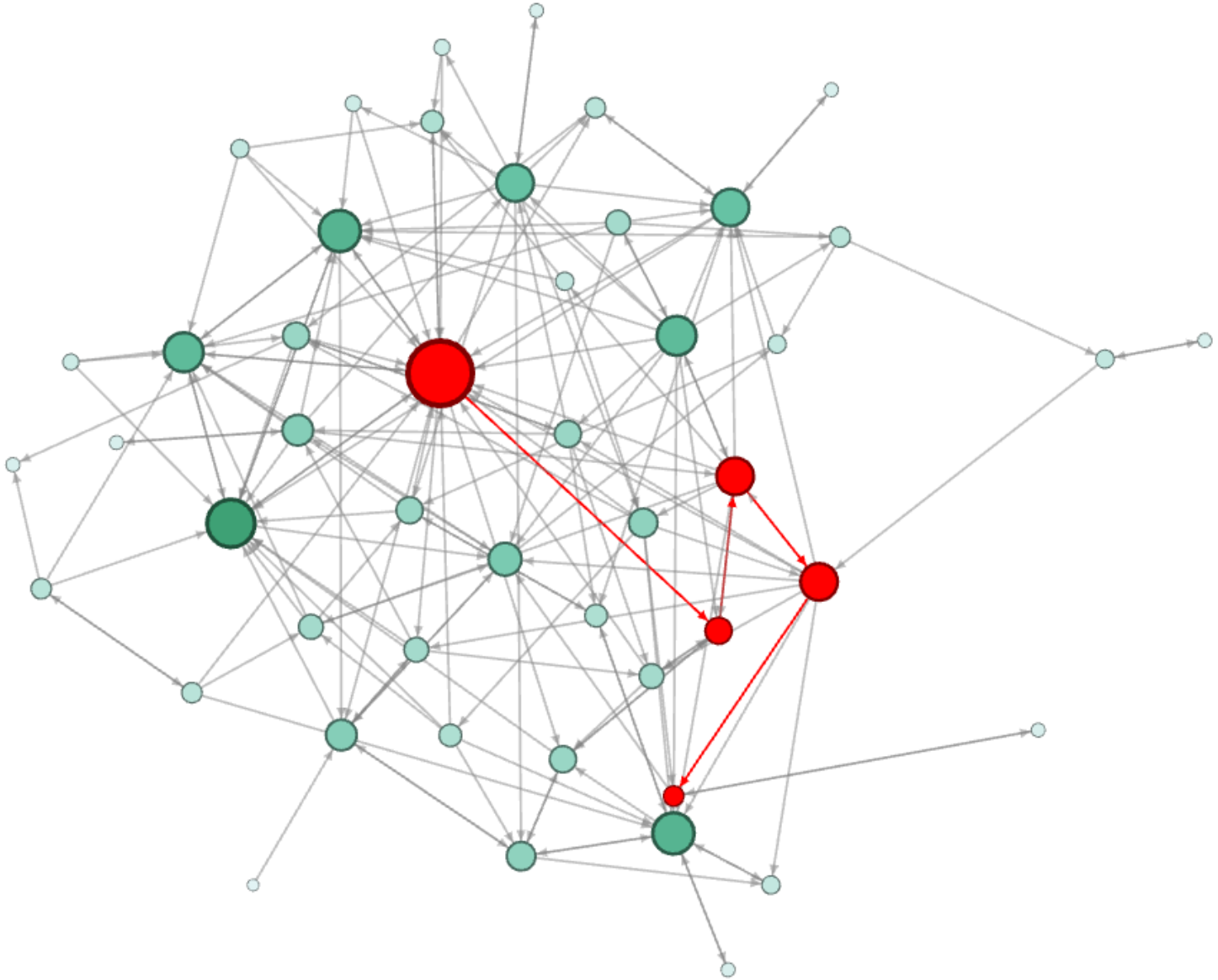
blockbuilder.org/jalapic

Interactivity

e.g. identify shortest paths

Example – Gephi

Friendship connections in German school 1880

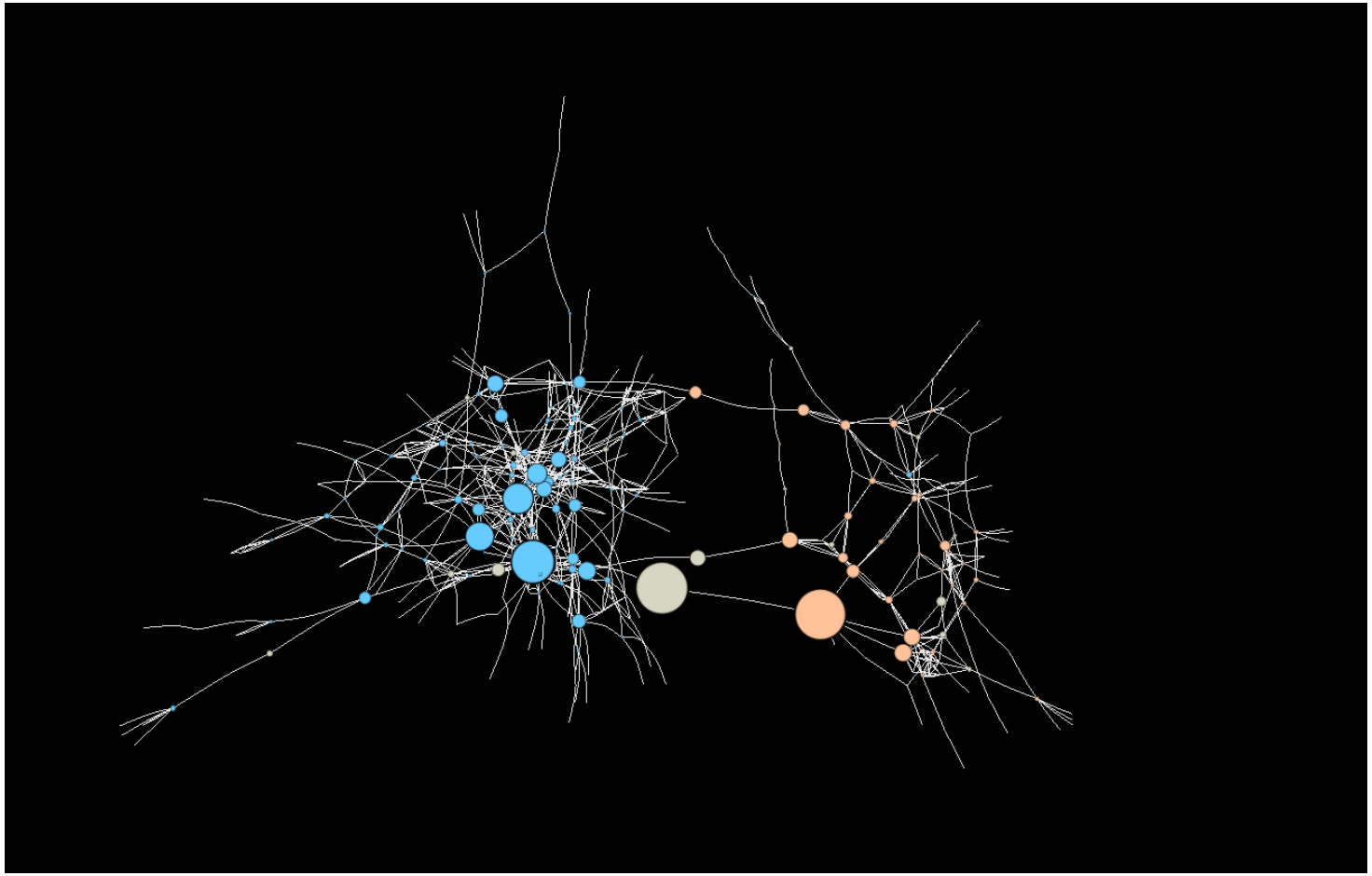


Pfeil → Fritzsche → Fährmann → Stolze → Lasch

Pfeil & Lasch have a geodesic distance of 4

Filtering Information - less is more

e.g. [Only showing nodes of high betweenness]



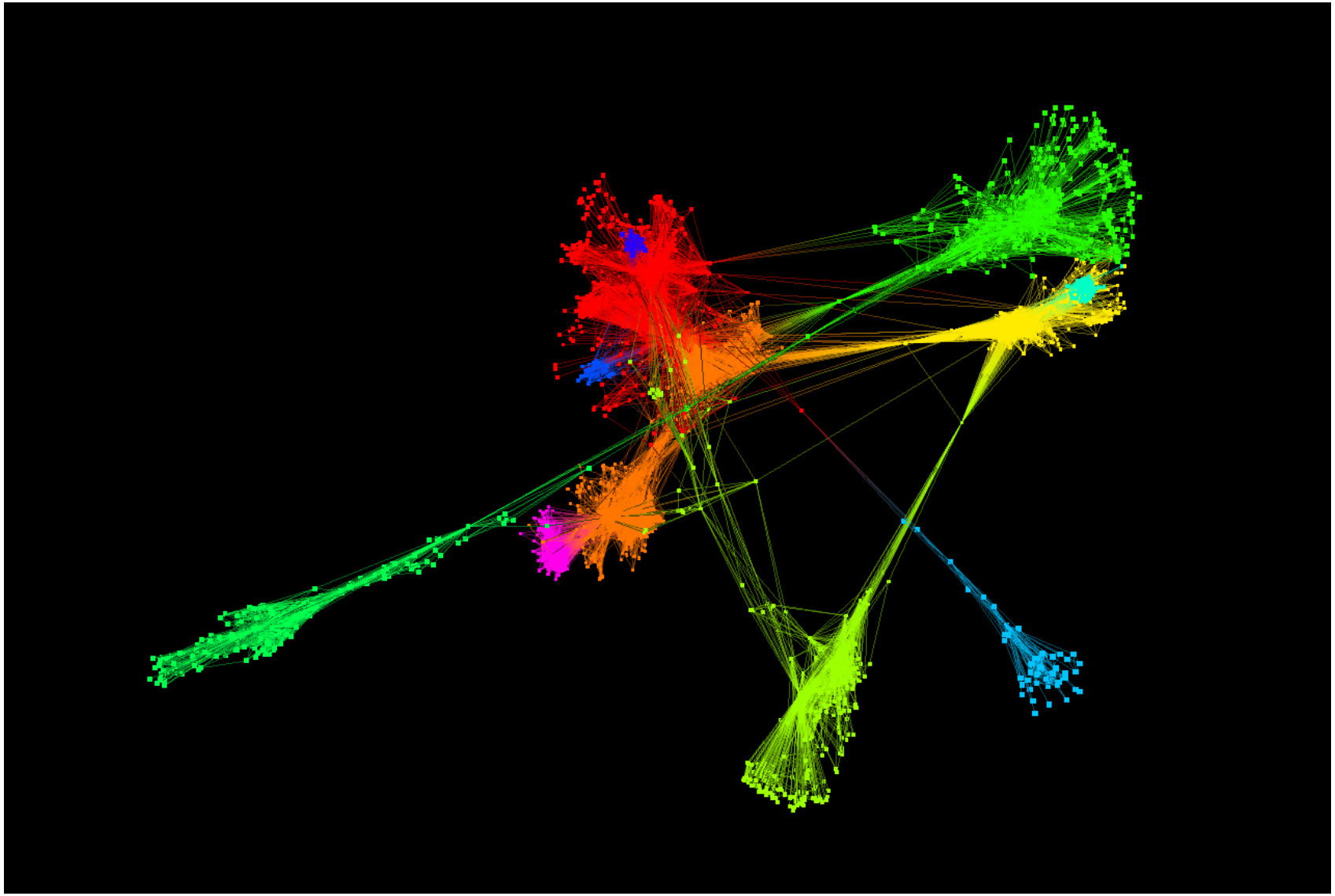
Made using
rthreejs

3D Network Representations

Community structure among friends in Facebook data (Stanford SNAP repository)

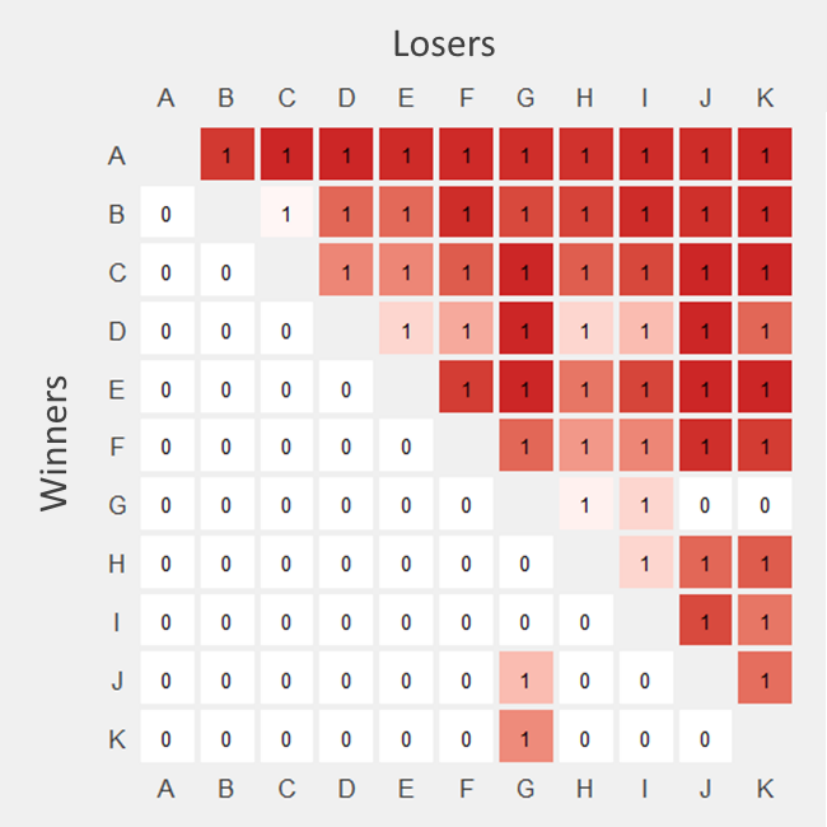
Using rthreejs

Bryan Lewis



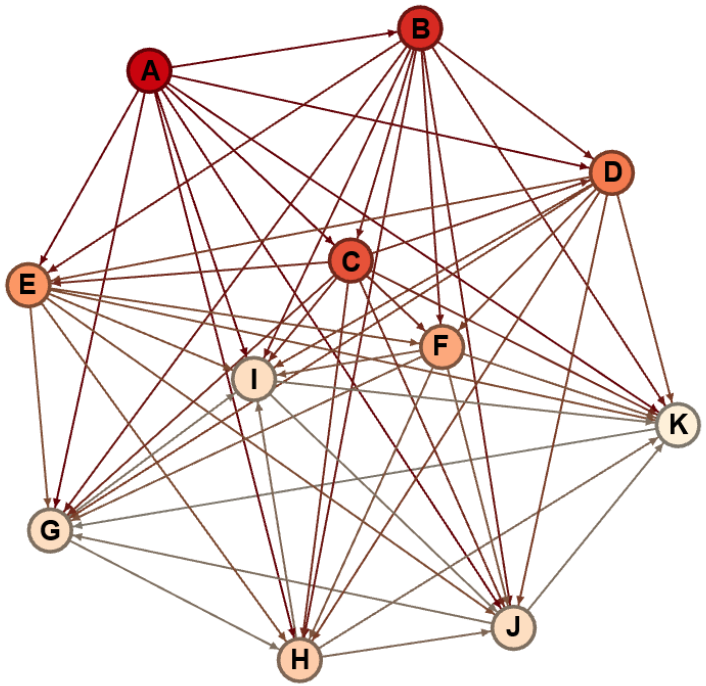
Is a network really the right visualization ?

Social Network Analysis of Social Hierarchies



Winners \ Losers	A	B	C	D	E	F	G	H	I	J	K
A		1	1	1	1	1	1	1	1	1	1
B	0		1	1	1	1	1	1	1	1	1
C	0	0		1	1	1	1	1	1	1	1
D	0	0	0		1	1	1	1	1	1	1
E	0	0	0	0		1	1	1	1	1	1
F	0	0	0	0	0		1	1	1	1	1
G	0	0	0	0	0	0		1	1	0	0
H	0	0	0	0	0	0	0		1	1	1
I	0	0	0	0	0	0	0	0		1	1
J	0	0	0	0	0	0	1	0	0		1
K	0	0	0	0	0	0	1	0	0	0	

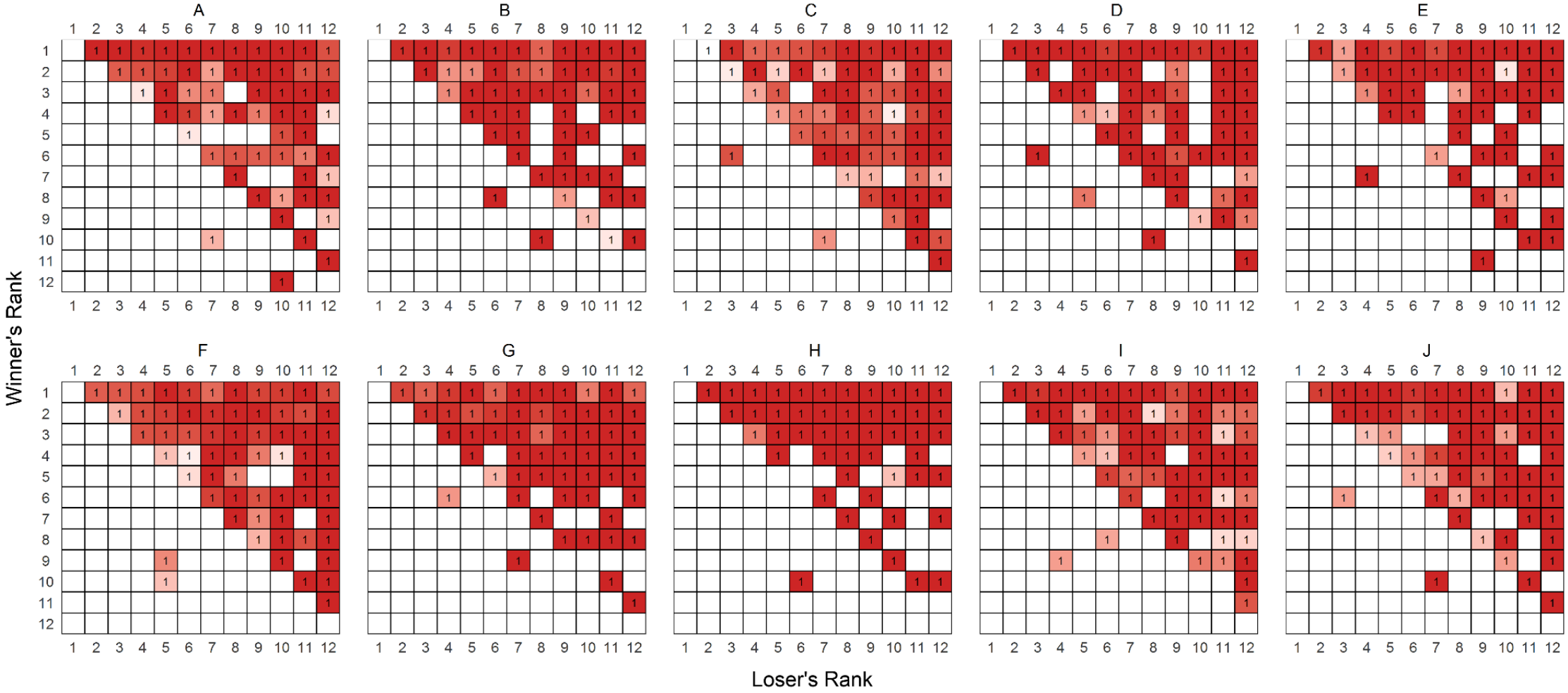
Adjacency Matrix



Network Graph

Comparing many networks by visual inspection may not be that easy or intuitive

Mice have unique ranks and show high directional consistency of behavior

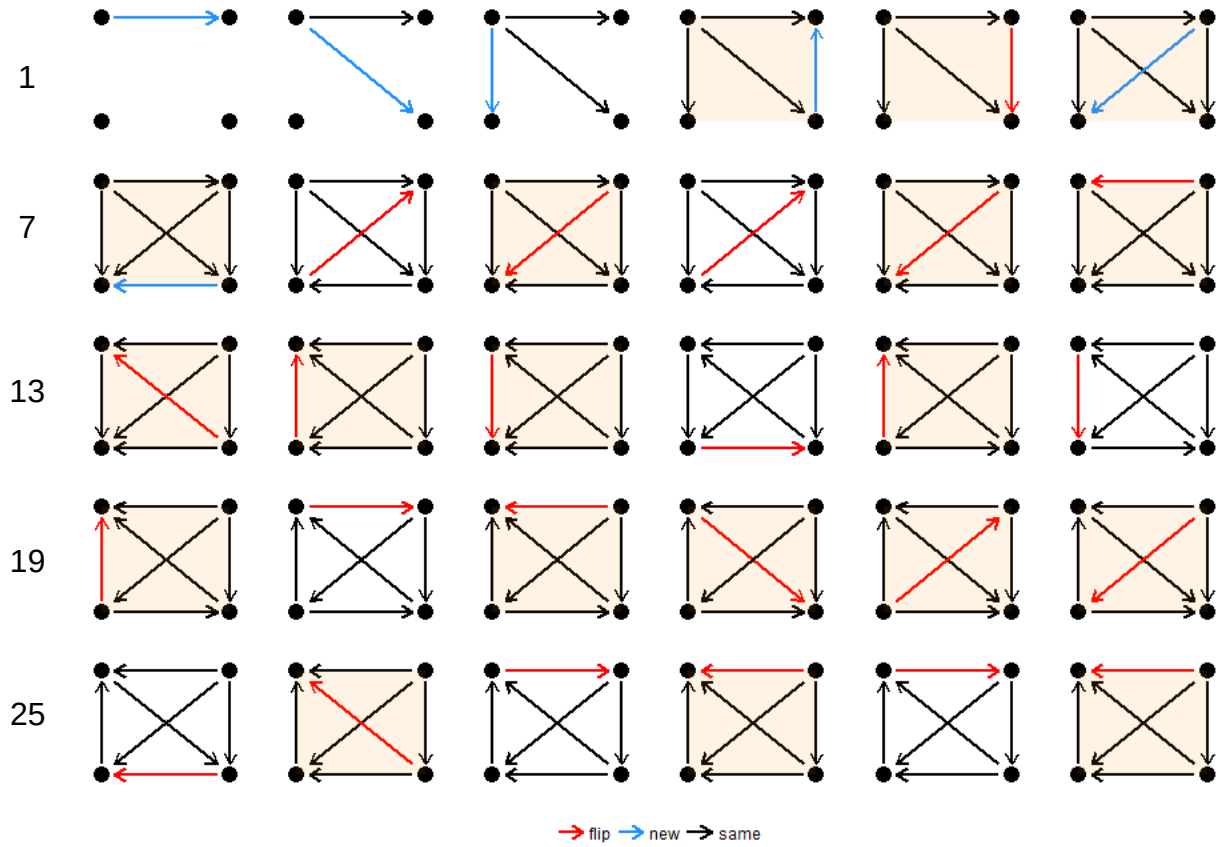


R - ggplot2

Williamson et al. *Anim Behav* 2016.

Small Multiples

Network Formation in African Cichlid Fish



- Every interaction from initial formation through 24 hours
- Proportion of time spent in transitive *vs.* intransitive states
- Likelihood of transitioning between transitivity/intransitivity

Intransitive Transitive

41 unique network types

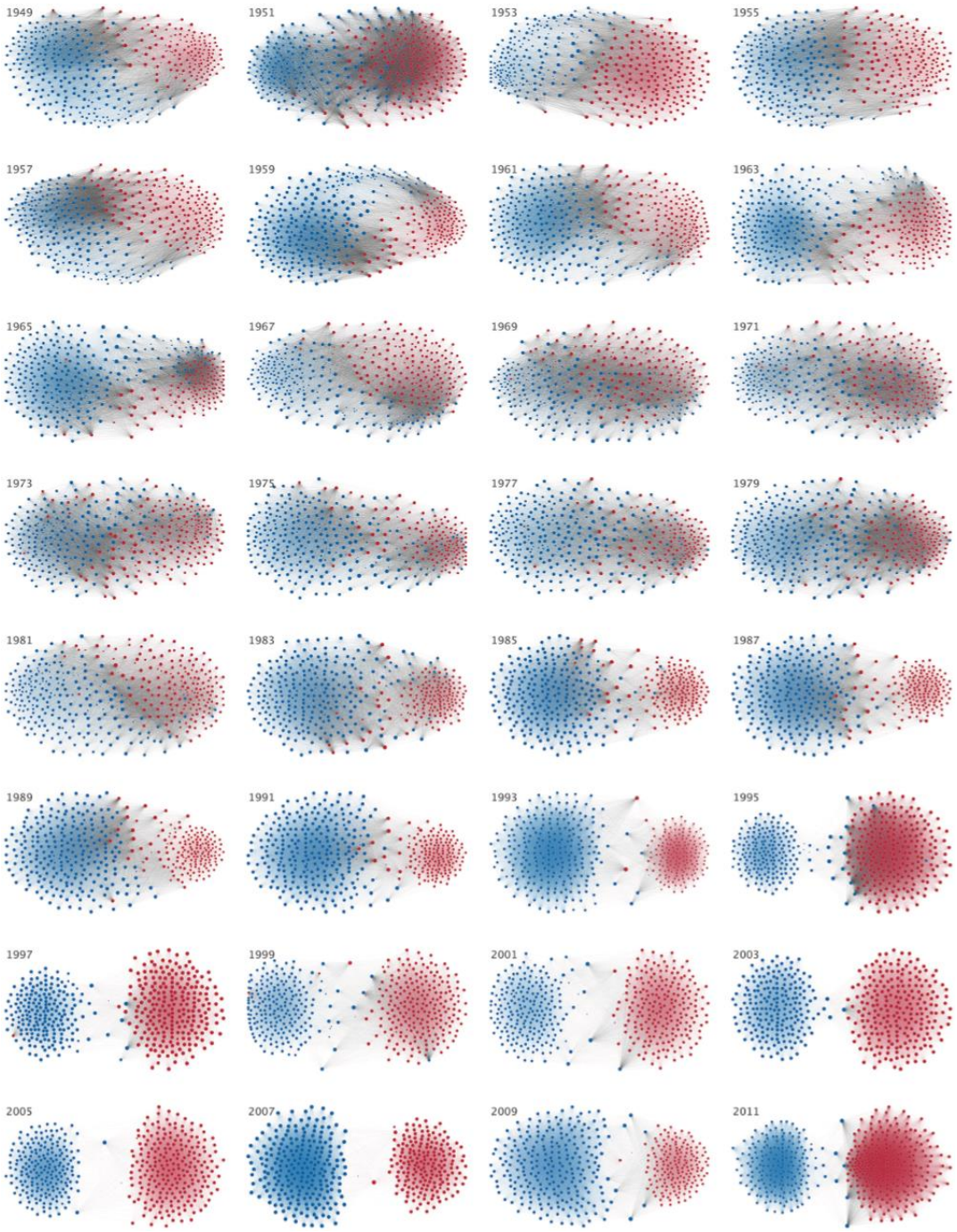
(Collaboration with Prof Ivan Chase – Stony Brook)

R package - hierformR

Small Multiples

The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives

Andris et al. PlosOne 2015



Small Multiples

Progression of flu contagion in the friendship network over time.

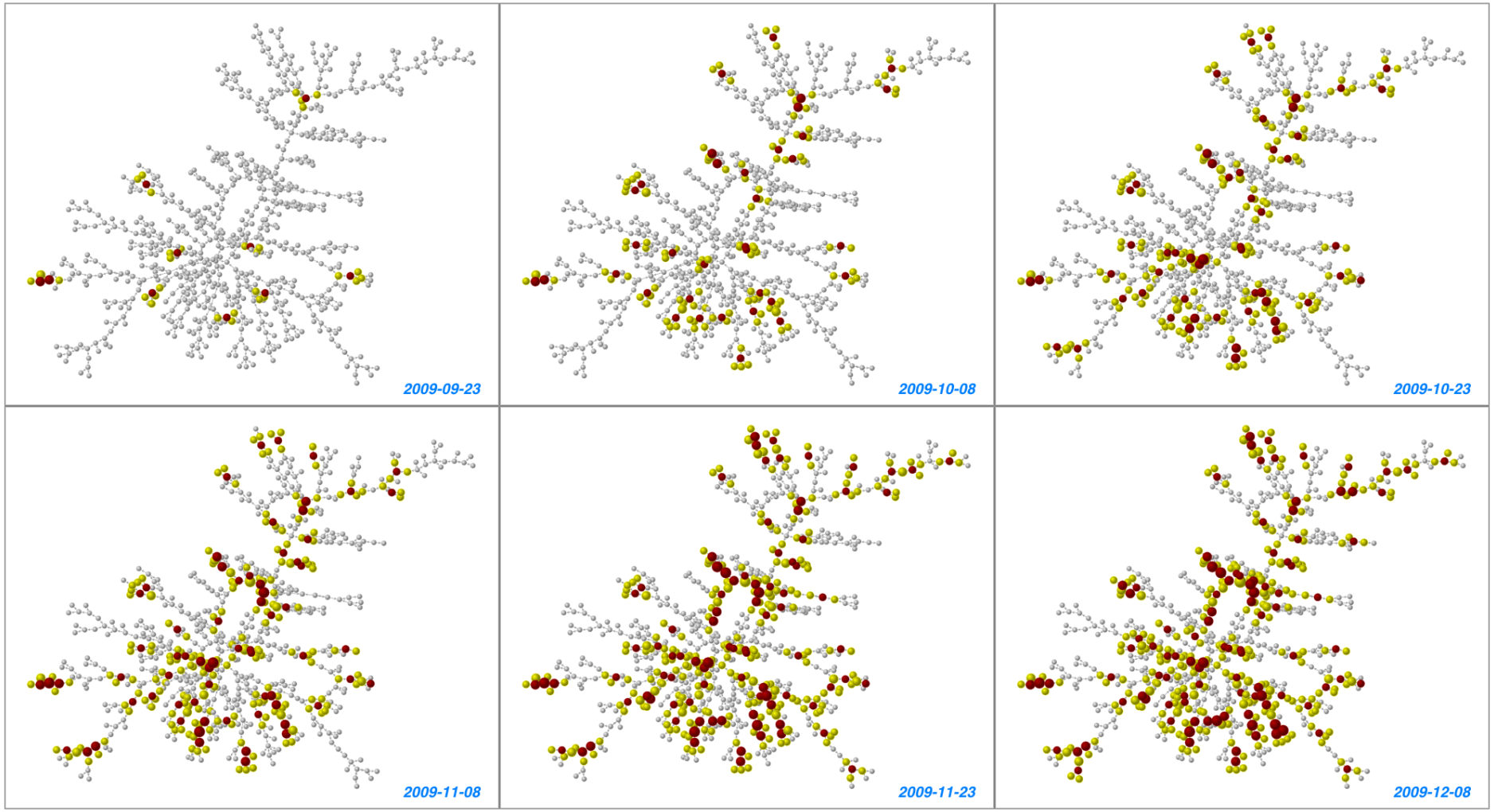
Edges = friendship nomination

White = uninfected individuals/friends

Red = Infected individuals

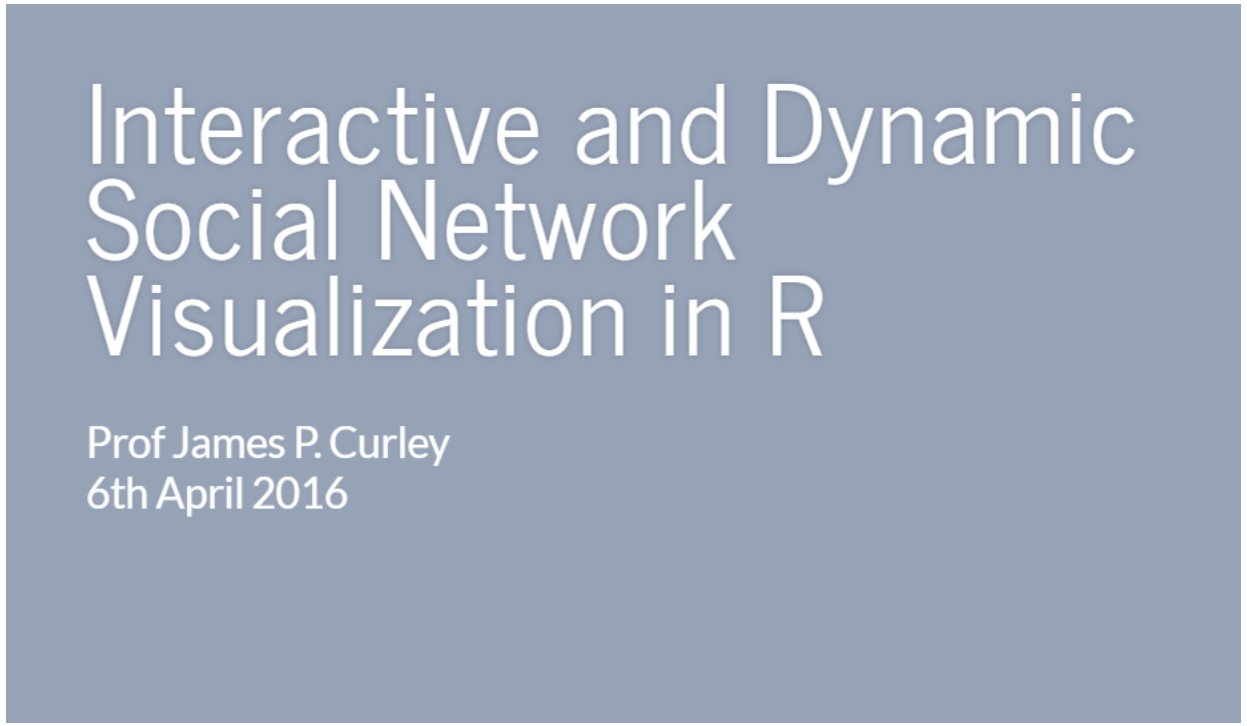
Yellow = friends of infected individuals

Node size = proportional to the number of friends infected.



Christakis & Fowler 2010 PlosOne

Animated Networks using R

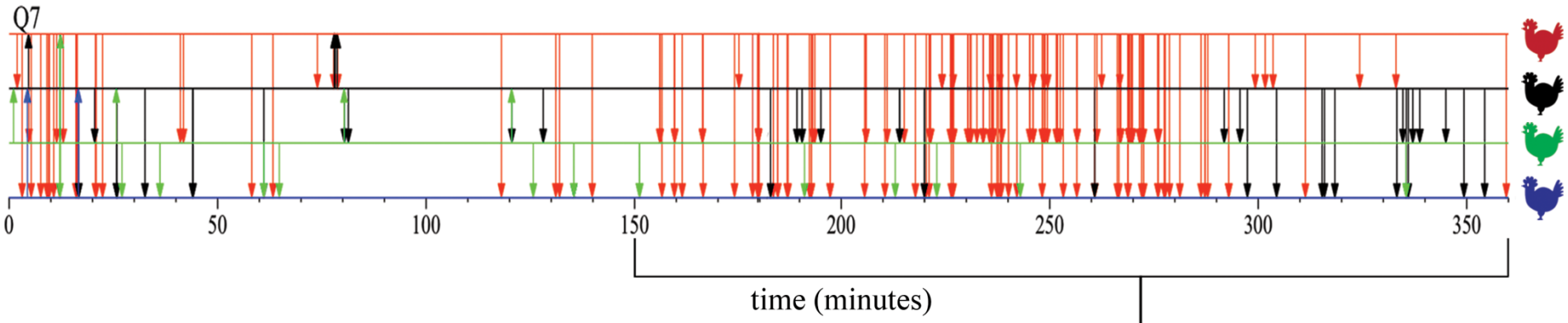


SEE my ONLINE TUTORIAL … http://curleylab.psych.columbia.edu/netviz/netviz5.html

SEE recent animations in rthreejs - http://bwlewis.github.io/rthreejs/animation/animation.html

Timelines

(A) Music Notation Representation of Agressive Interactions in a Group of Hens



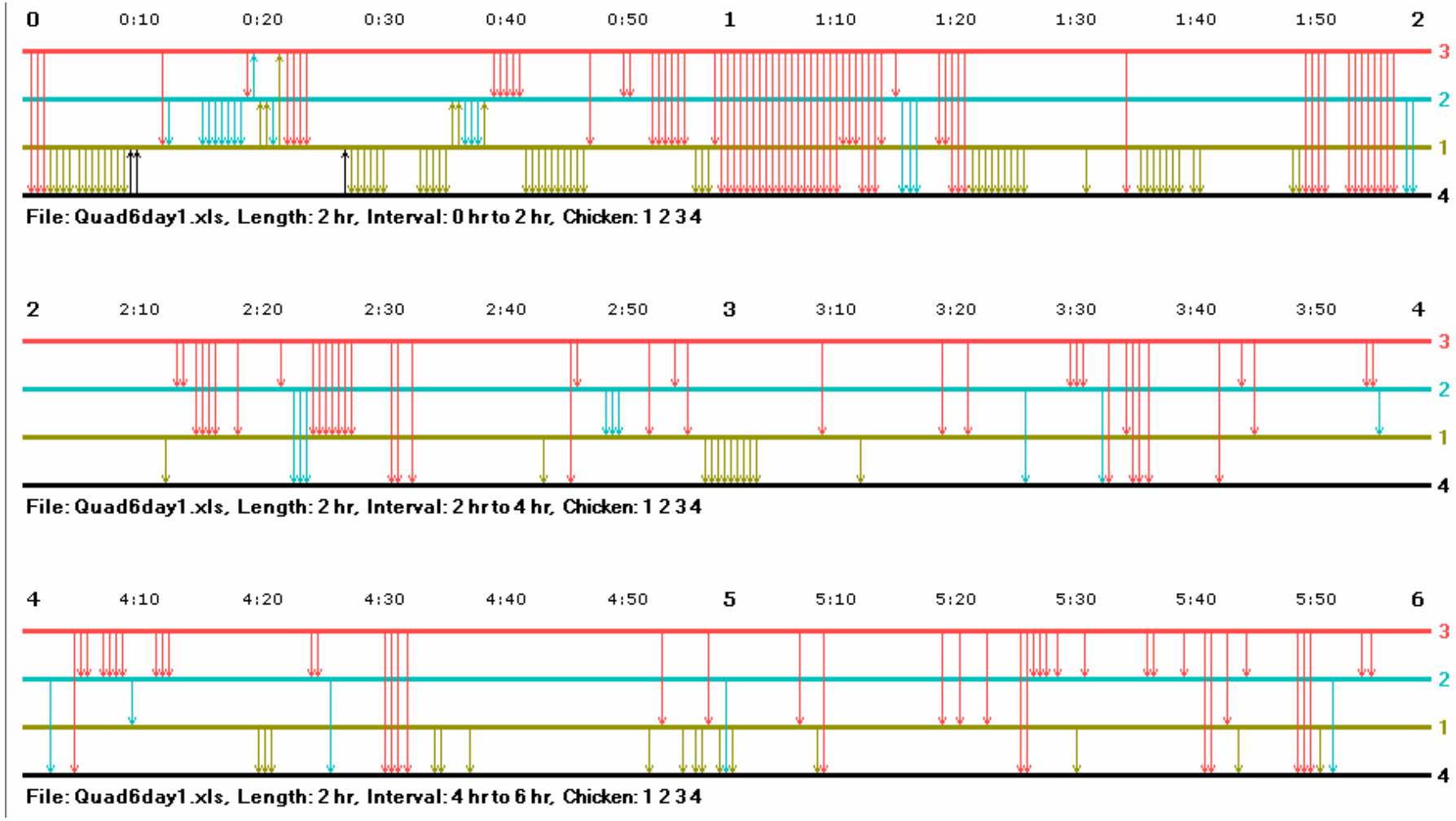
(B) Interaction Matrix

-	17	92	59
0	-	7	8
0	0	-	6
0	0	0	-

(C) Linear Hierarchy

(Ivan Chase)

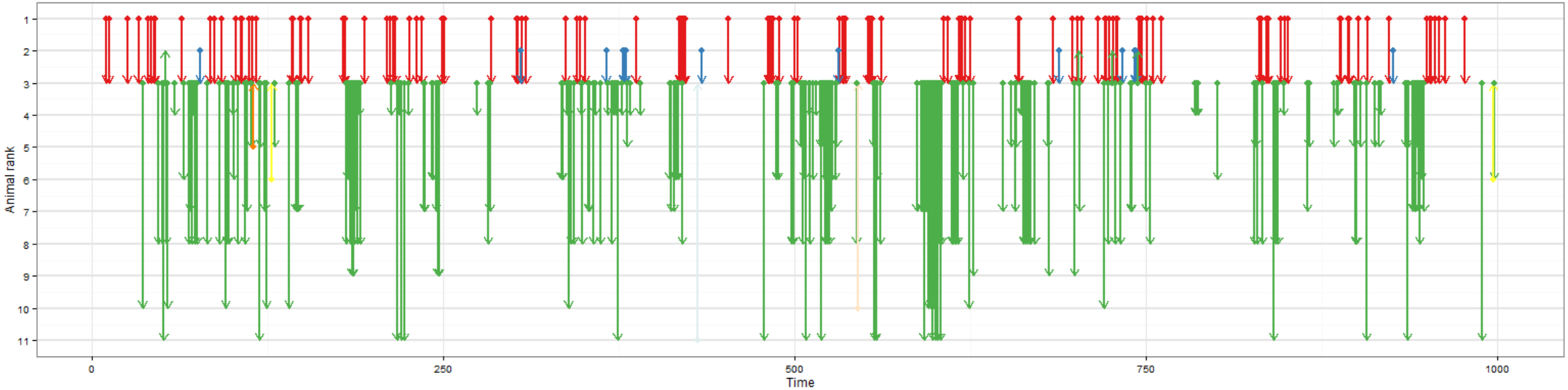
Timelines



Curley 2016 – “musicnotationR” R package

Timelines

Shiny App: https://jalapic.shinyapps.io/music/



Filter all aggressive interactions (incoming & outing) for animal Rank #3 - high directional consistency of behavior

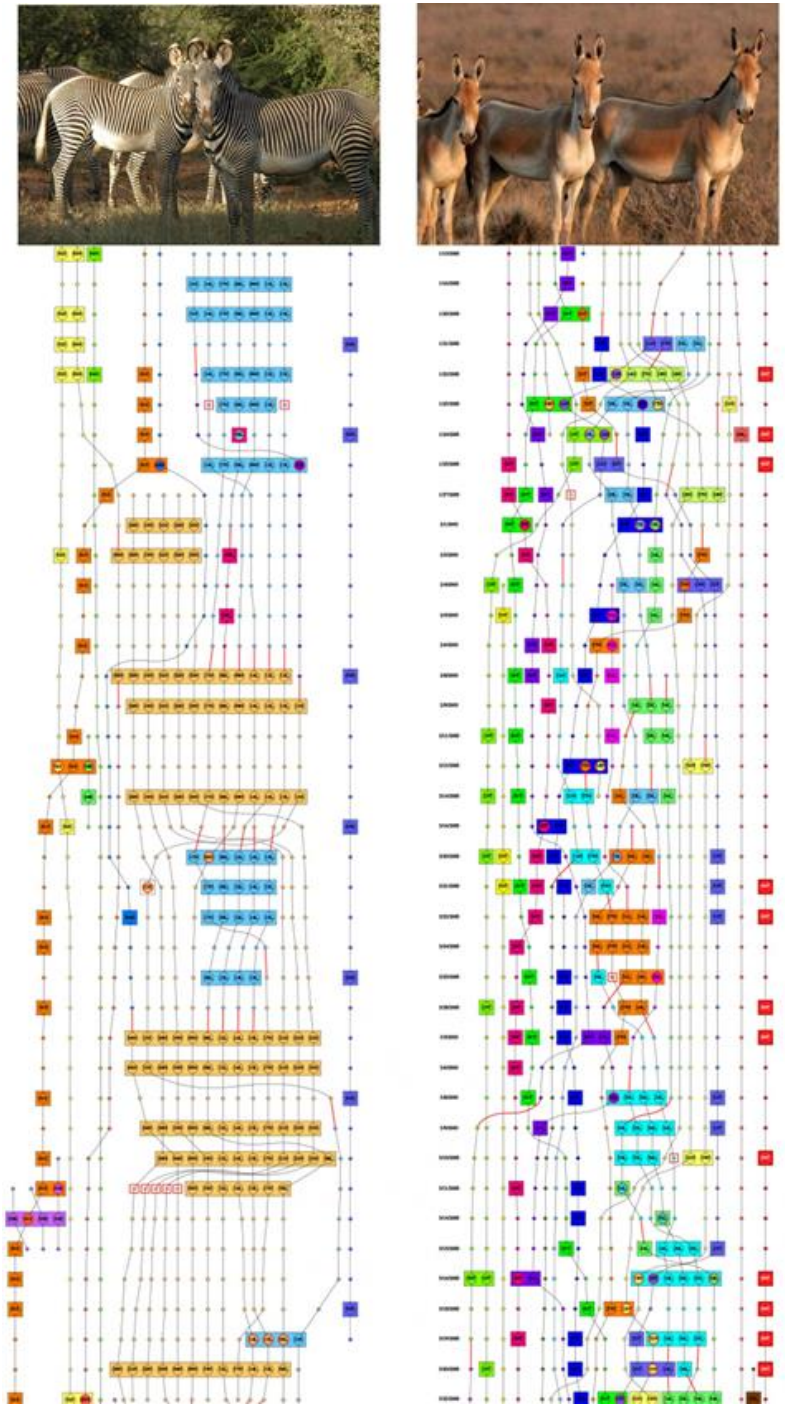
Colors represent unique individuals

Curley 2016 – “musicnotationR” R package

Timelines

Dan Rubenstein- Princeton
PlosOne 2015

Community membership over time in zebras vs onagers



Conclusions

- Lots of tools exist to visualize networks
- It’s easy to make really attractive network visualizations
- It’s also easy to make really bad network visualizations

- **What question do you want to explore or communicate to an audience?**
- **What’s the most effect way of achieving that?**
- **Learn the programming language / software that will facilitate this !**

Links to examples

Katya Ognyanova tutorials - http://kateto.net/

Martin Grandjean work - http://www.martingrandjean.ch/social-network-analysis-visualization-morenos-sociograms-revisited/

My friendship visualization - http://curleylab.psych.columbia.edu/friendship.html

My friendship visualization 2 - http://curleylab.psych.columbia.edu/friends.html

Identifying neighbors in D3js - http://blockbuilder.org/jalapic/0020bebb012760de0d54

Identifying triads in D3js - http://blockbuilder.org/jalapic/aeb4b1085f800d31b4be

Identifying geodesic distance in D3js - http://blockbuilder.org/jalapic/0eba1241acdf67c62f5b87357864d662

rthreejs example by Bryan Lewis - http://jalapic.github.io/facebook

Linked Matrices in D3js - http://curleylab.psych.columbia.edu/socdom1.html

D3js animated network example - http://blockbuilder.org/jalapic/cf1454d62c7efa48d96ee4d803856dfa

A network consists of:

Nodes = individuals, groups, classes of individuals etc.
Node attributes = phenotypic traits of the individual e.g. gender, age, group membership

Edges = relationship between two nodes,
Weighted = frequency of interaction between nodes, or strength of association/relationship

Types of Network

Binary Matrix vs Valued Matrix vs Thresholded

Directed vs Undirected

Static vs Dynamic (Temporal)

Edgelists can be:

- pairwise events
- undirected or directed
- with time of event
- with on/off times
- weighted

Adjacency matrices (sociomatrices) can be:

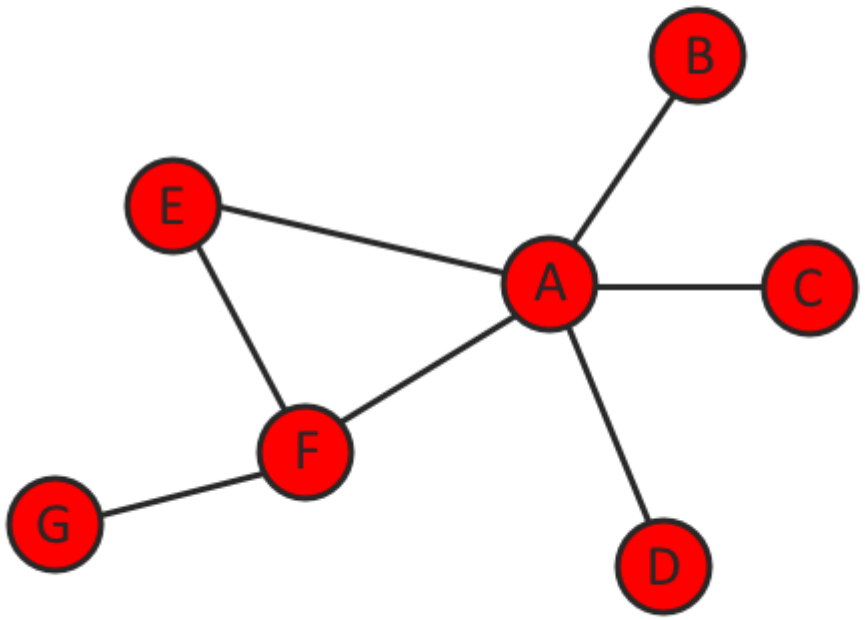
- frequency of events
- comembership
- association indices
- measure of relatedness
- measure of distance
- nominations (ranks)
- rates of interaction

Both edges and nodes can have attributes:

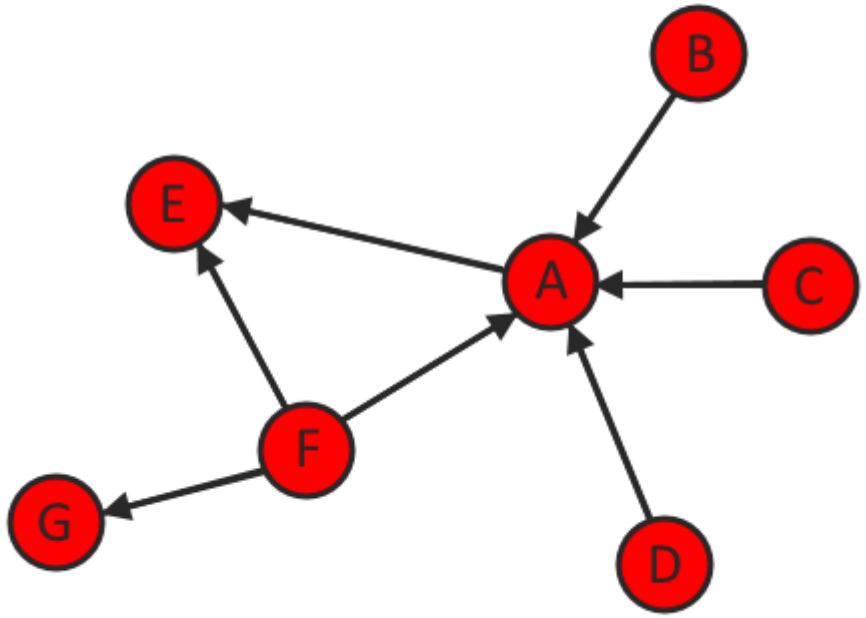
e.g. EDGES - frequency, weight, type, positive/negative

e.g. NODES – age, sex, group affiliation, personality etc.

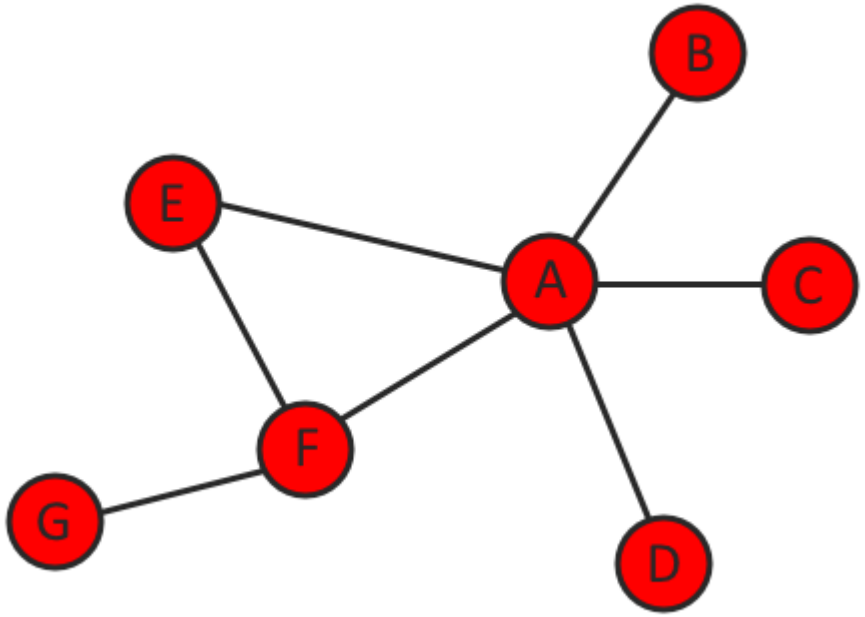
Undirected



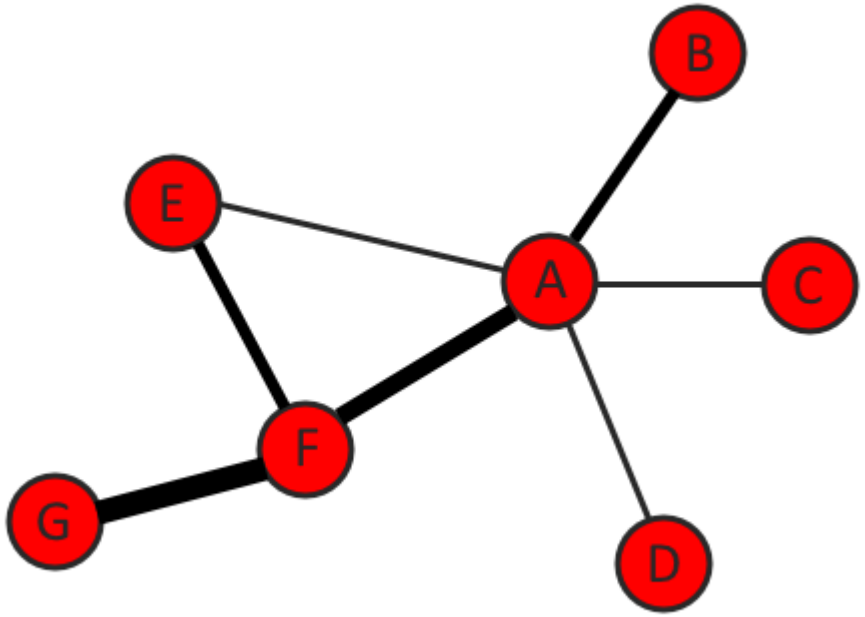
Directed



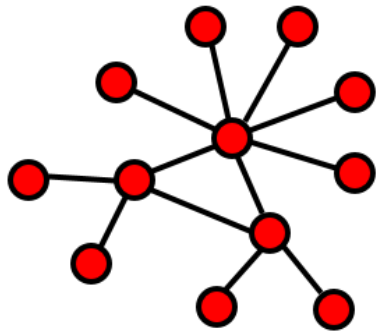
Unweighted



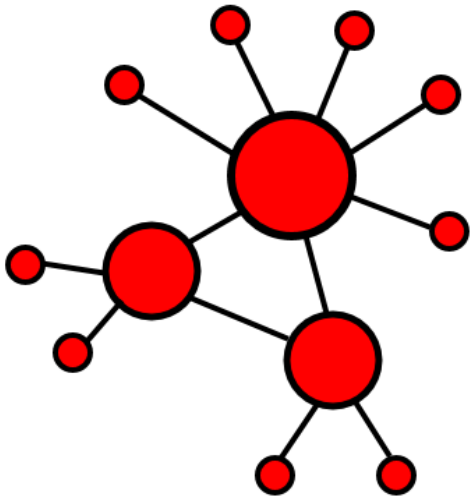
Weighted



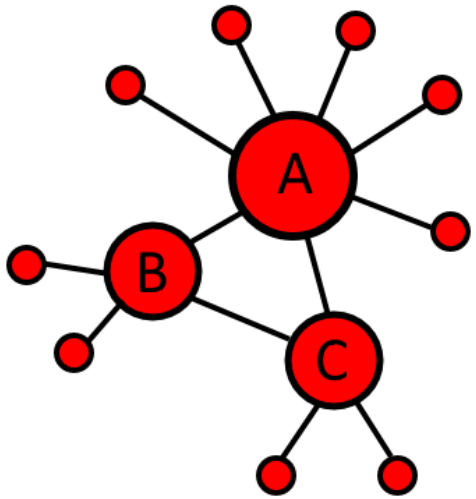
default



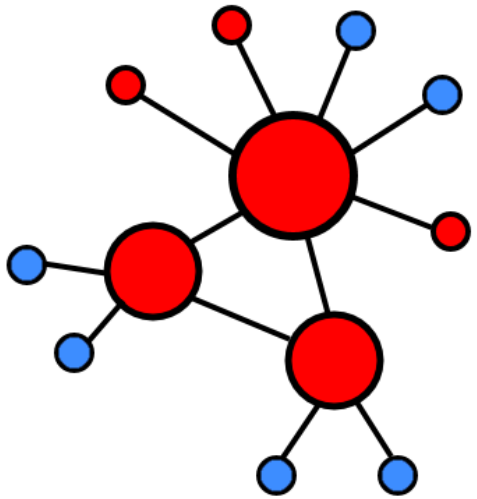
size



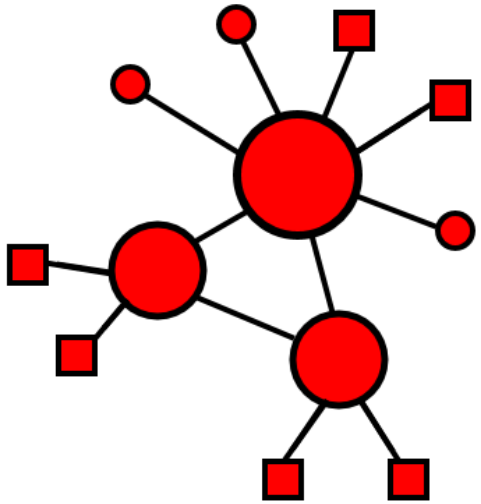
labels



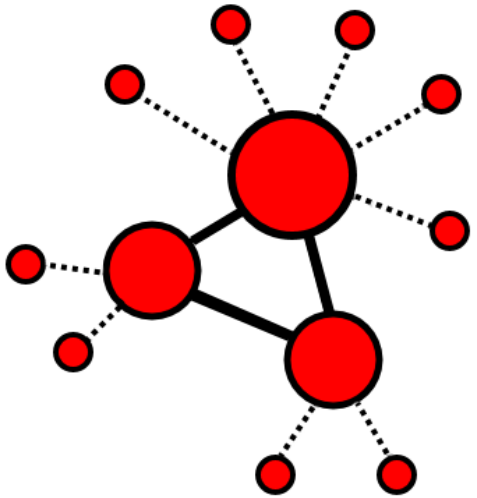
color



shape



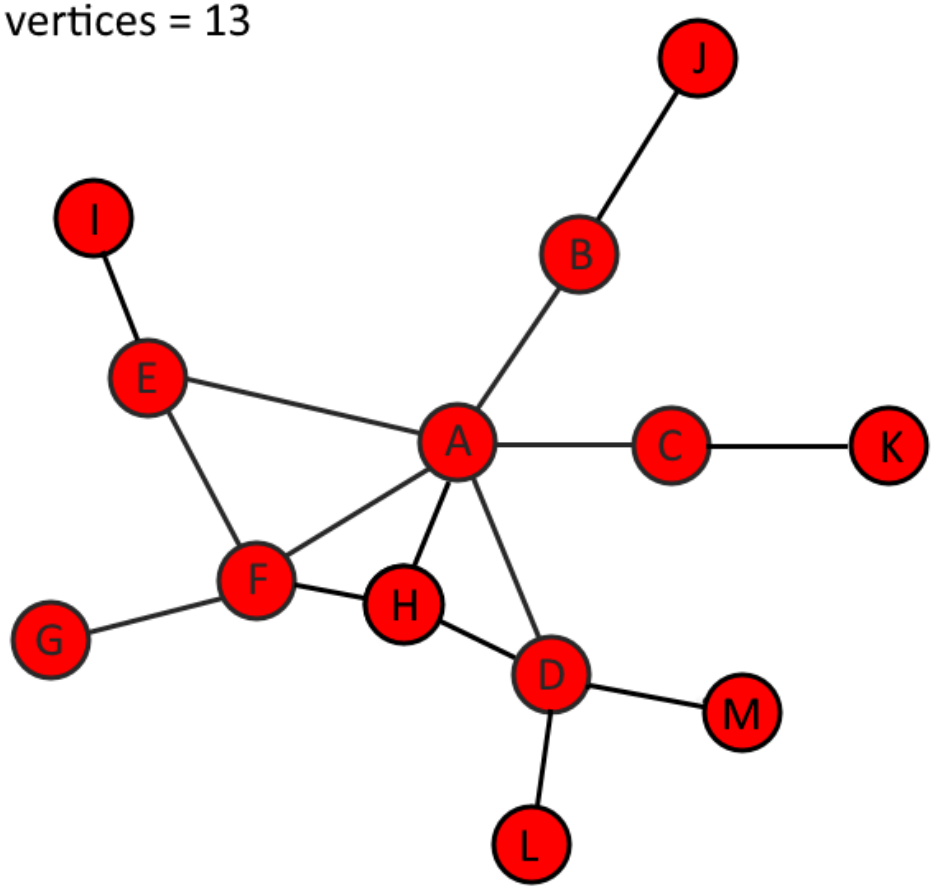
edges



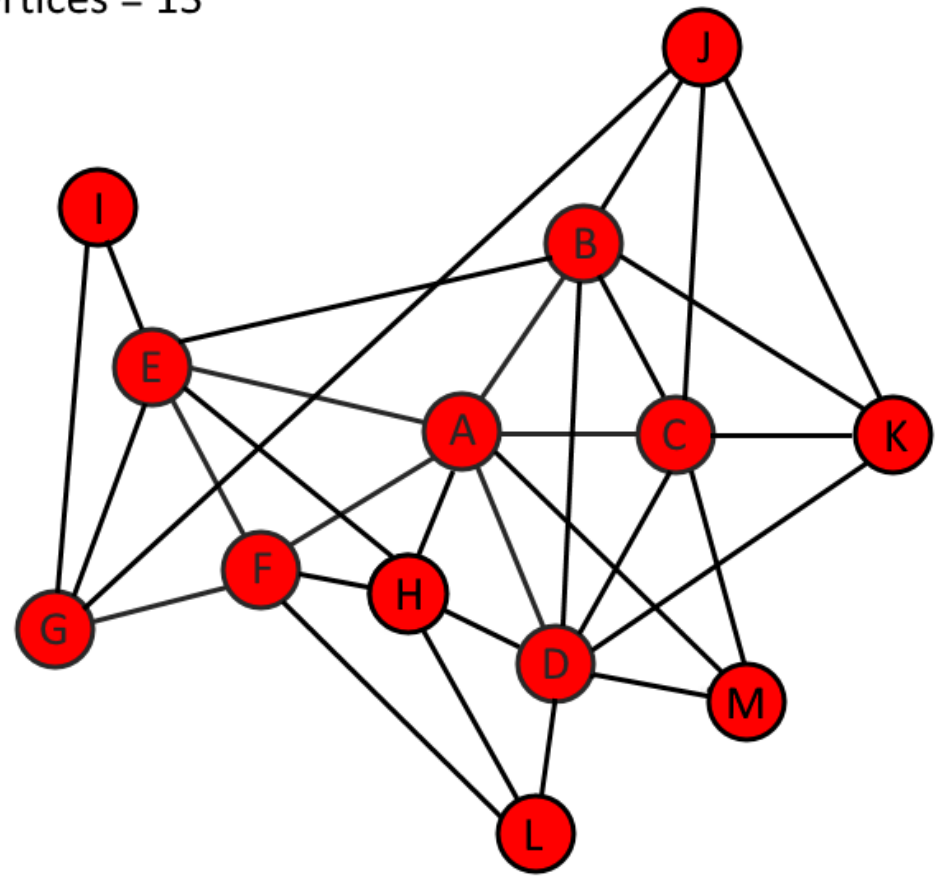
Measures of Graph Cohesion

Graph Density

density = 0.19
edges = 15
vertices = 13

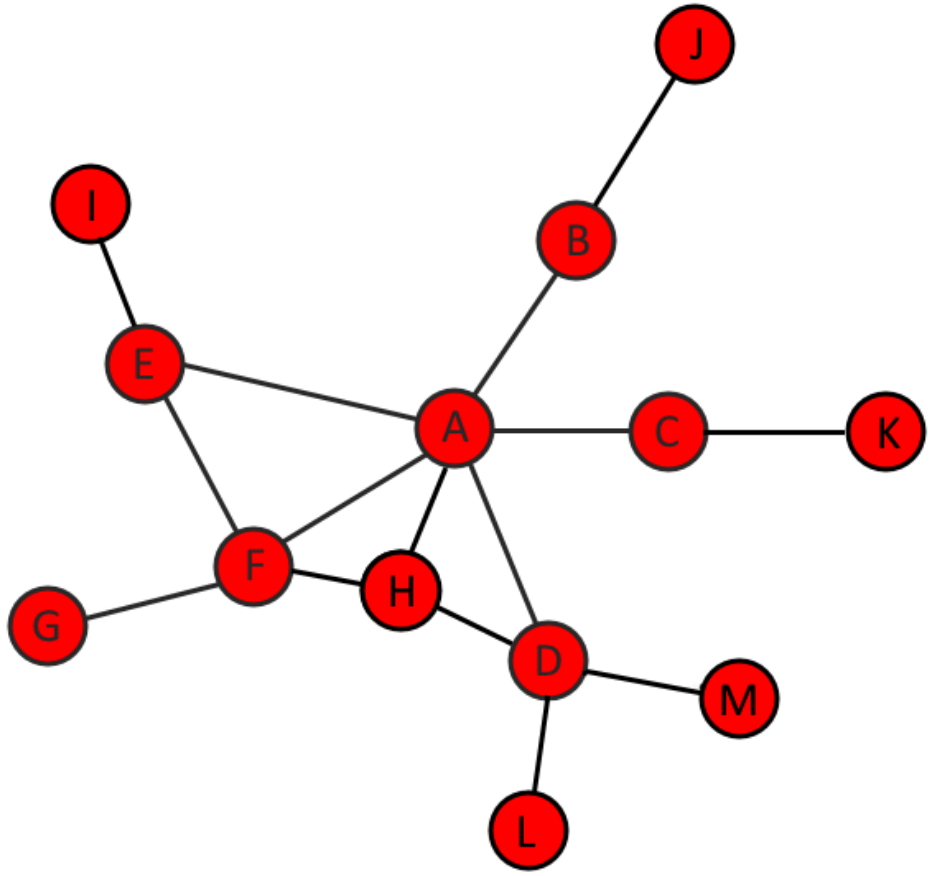


density = 0.38
edges = 30
vertices = 13

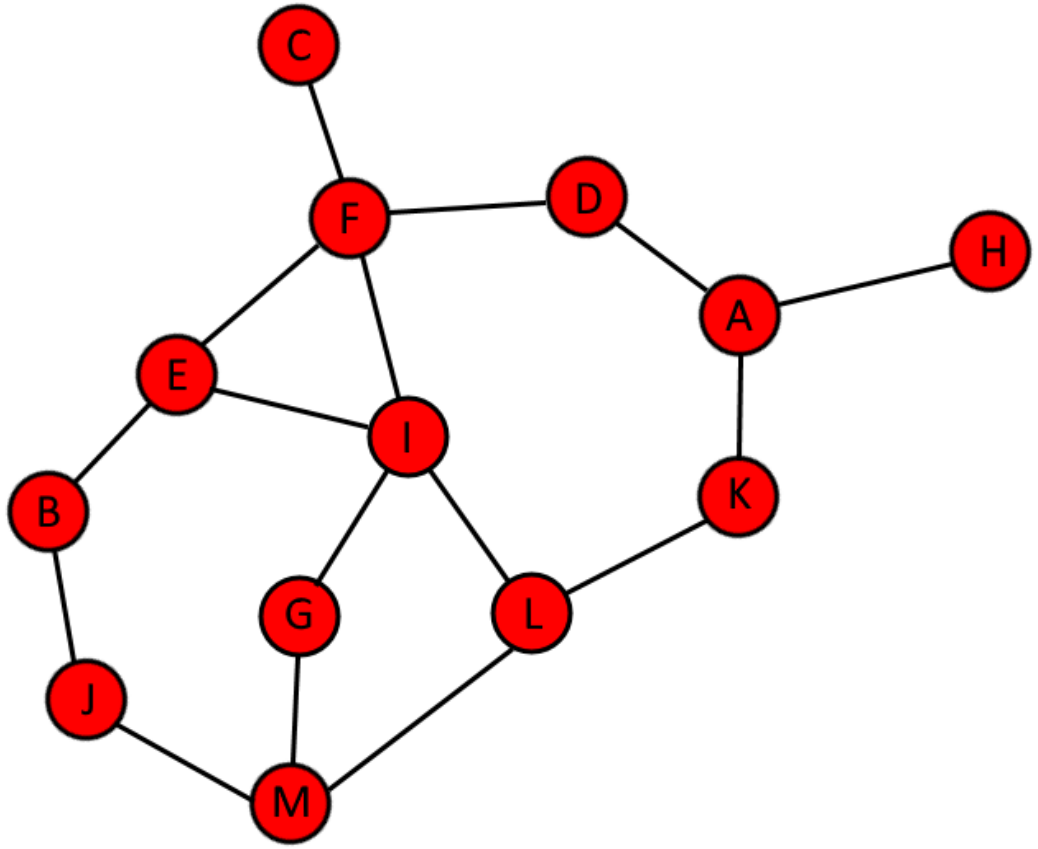


Density = the proportion of possible edges in the network.
In valued network equates to tie strength

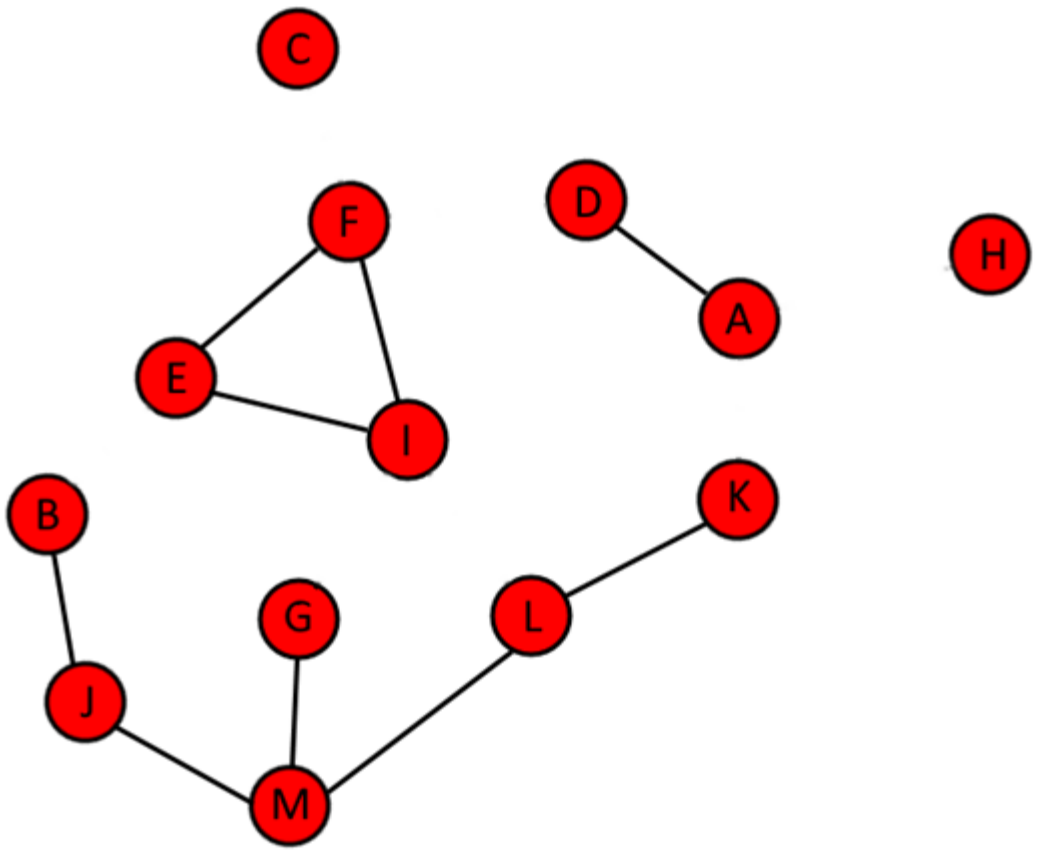
density = 0.19
vertices = 13



density = 0.21
vertices = 13



Components



CR = 0.33

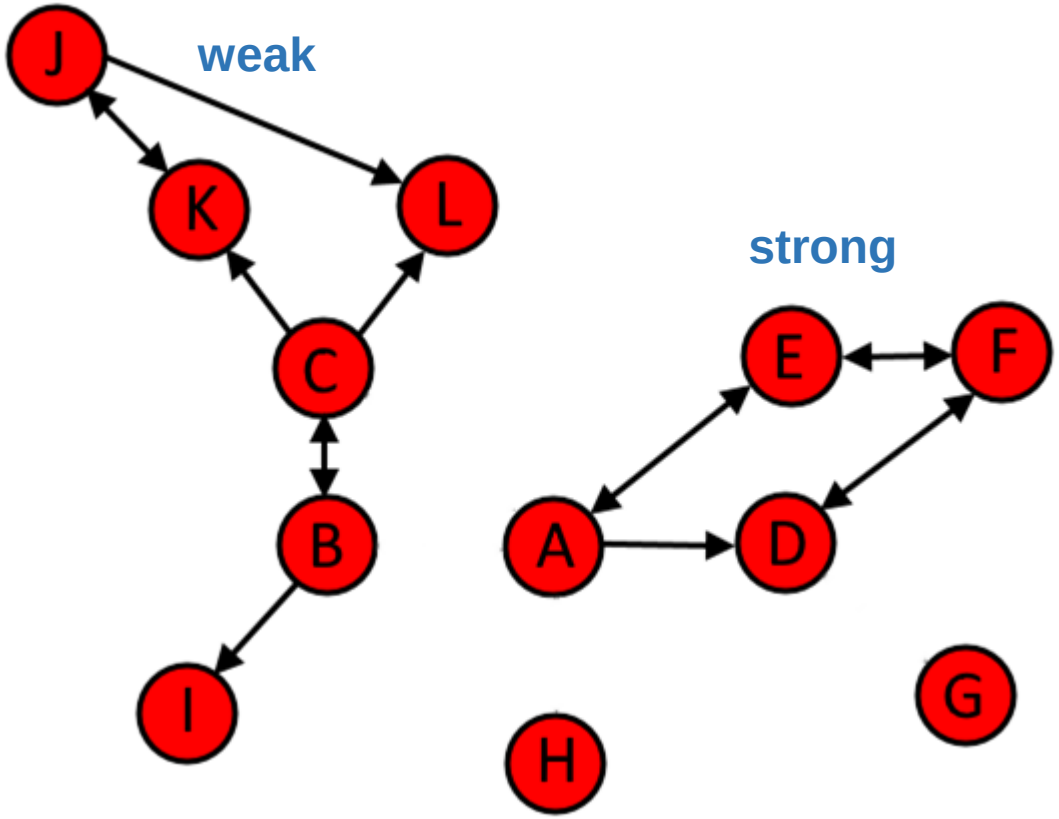
Giant component – component with largest number of nodes (typically used when component is large fraction of graph)

Isolates can be components of size 1

Component Ratio = CR = c - 1 / n -1 (c is number of components, n is number of nodes in the graph)

Maximum is 1 when every node is an isolate. Minimum of 0 when whole network is one component. An inverse measure of cohesion.

Components

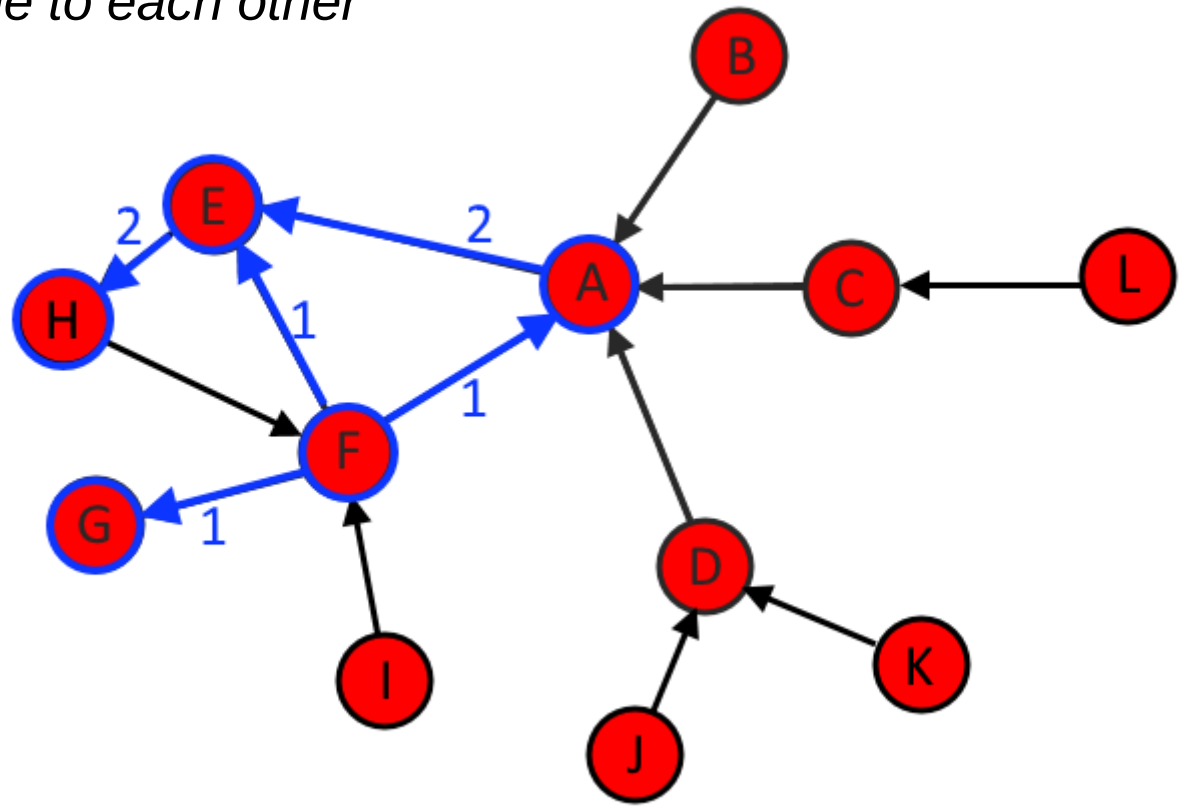


Strong component - in directed graphs, all nodes can reach each other via directed paths

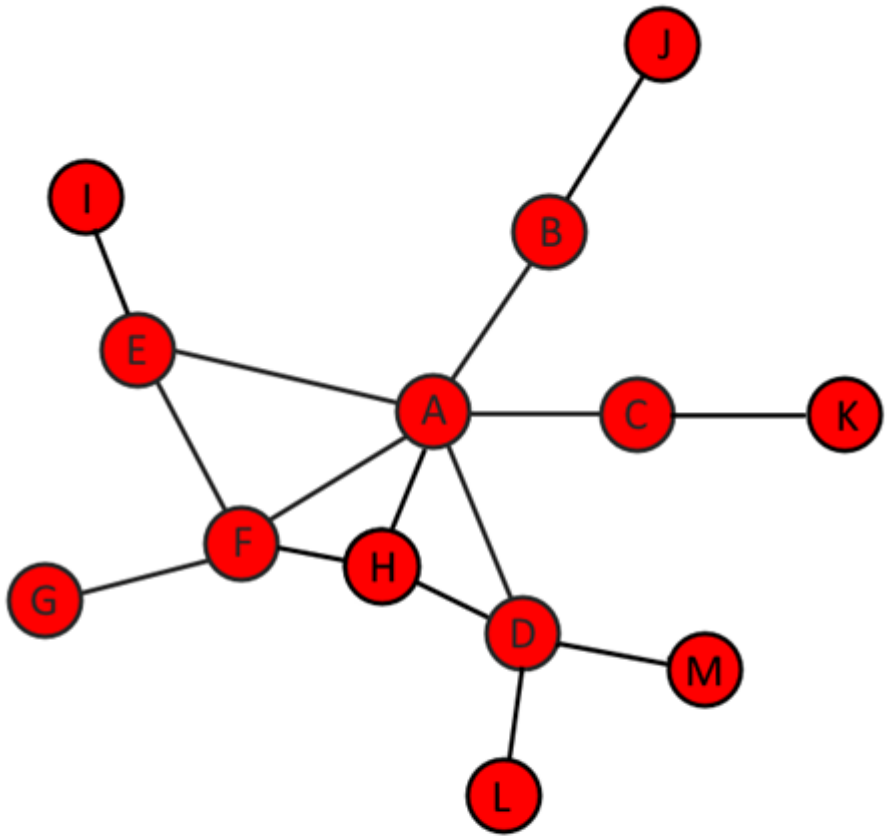
Weak component – in directed graphs, all nodes cannot reach each other via directed paths

Geodesic Distance

Distance from each node to each other



Paths



Diameter = 4
Average path length = 2.47

A path is a walk in which each other actor and each other relation in the graph may be used at most one time

A trail between two actors is any walk that includes a given relation no more than once (the same other actors, however, can be part of a trail multiple times.

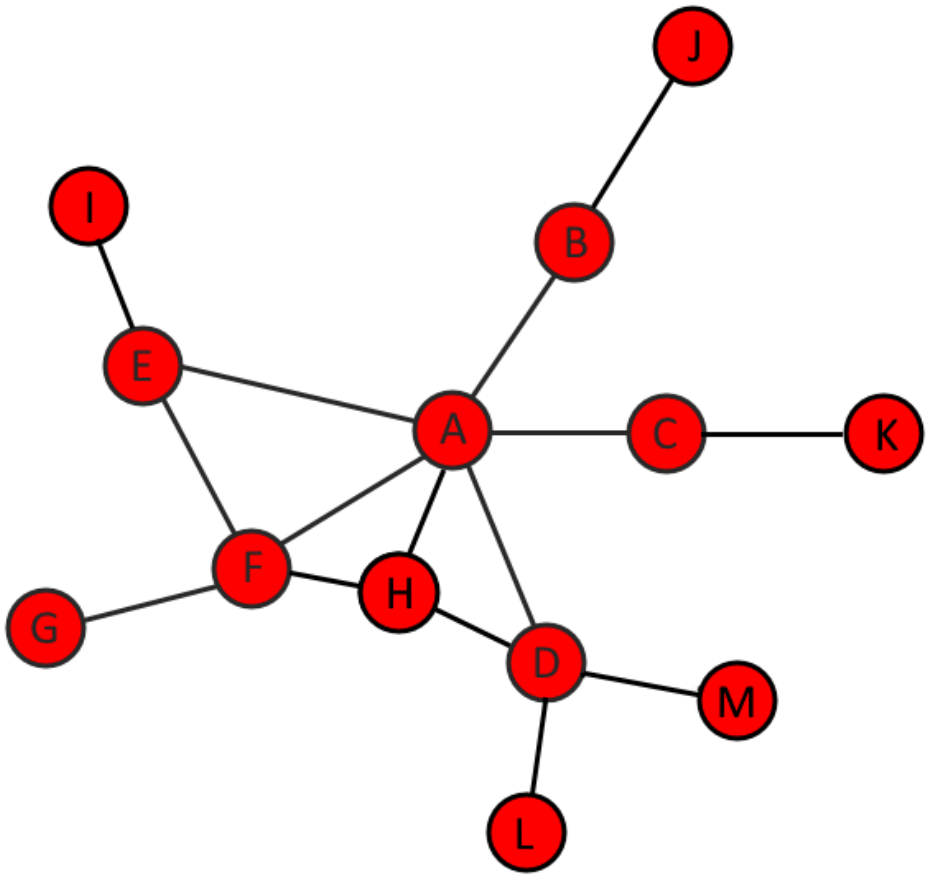
Undirected network - either direction
Directed network - can only go via direction of edge

<u>Shortest path</u> – the shortest number of steps between any pair of node

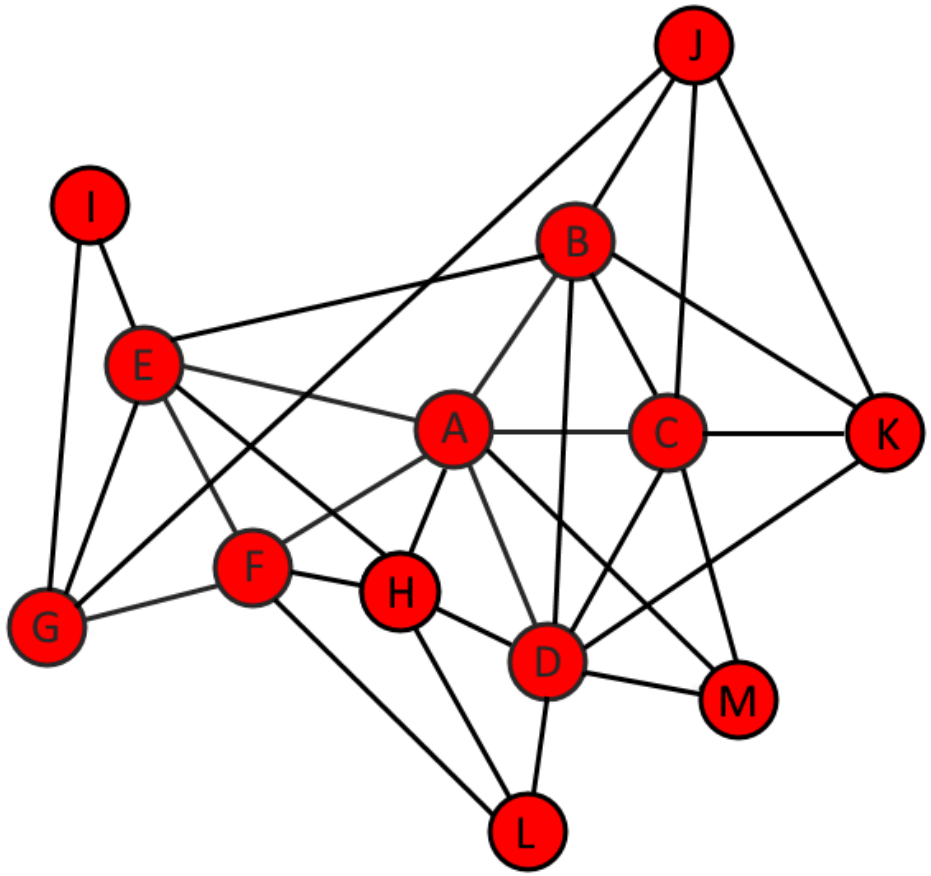
<u>Diameter</u> – the longest shortest path in the graph

<u>Average path length</u> – the average shortest path between all pairs of nodes

average path length = 2.47

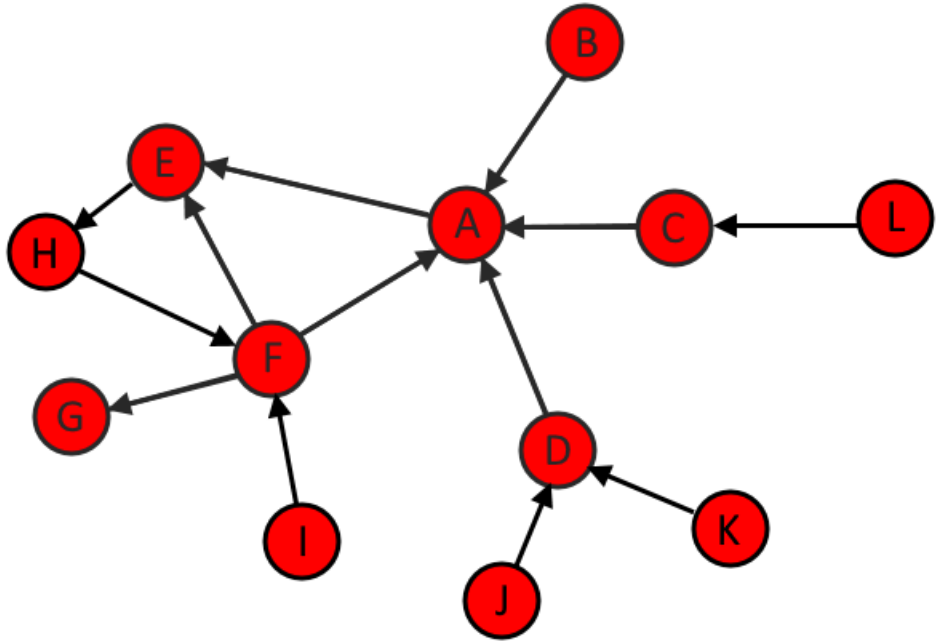


average path length = 1.81



Directed Connectedness

What proportion of pairs of nodes can reach each other?:

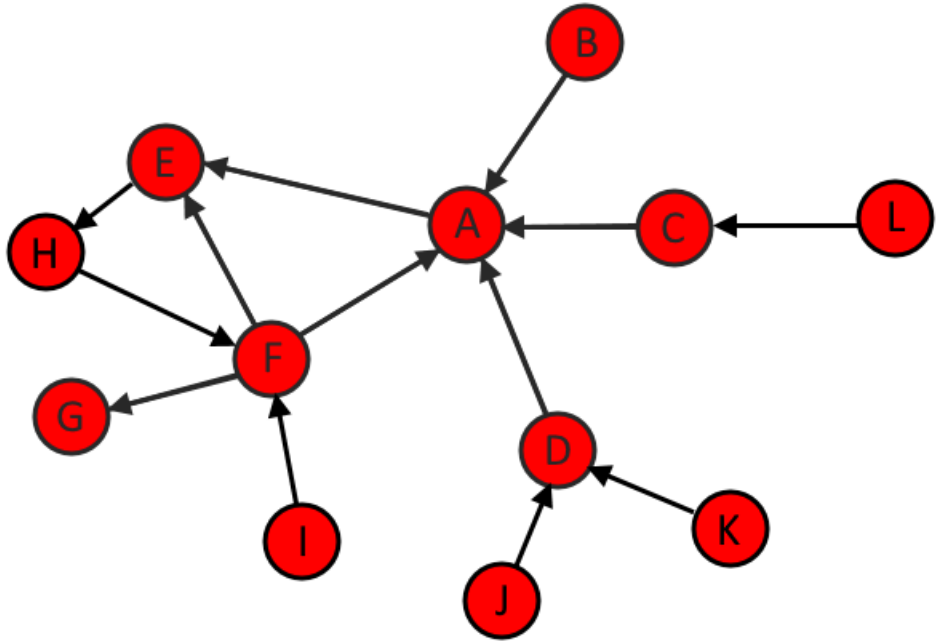


	A	E	F	H	B	C	D	L	J	K	G	I
A	1	1	1	1	0	0	0	0	0	0	1	0
E	1	1	1	1	0	0	0	0	0	0	1	0
F	1	1	1	1	0	0	0	0	0	0	1	0
H	1	1	1	1	0	0	0	0	0	0	1	0
B	1	1	1	1	1	0	0	0	0	0	1	0
C	1	1	1	1	0	1	0	0	0	0	1	0
D	1	1	1	1	0	0	1	0	0	0	1	0
L	1	1	1	1	0	1	0	1	0	0	1	0
J	1	1	1	1	0	0	1	0	1	0	1	0
K	1	1	1	1	0	0	1	0	0	1	1	0
G	0	0	0	0	0	0	0	0	0	0	1	0
I	1	1	1	1	0	0	0	0	0	0	1	1

Can row node reach column node ?

Directed Connectedness = 0.41

Directed Connectedness

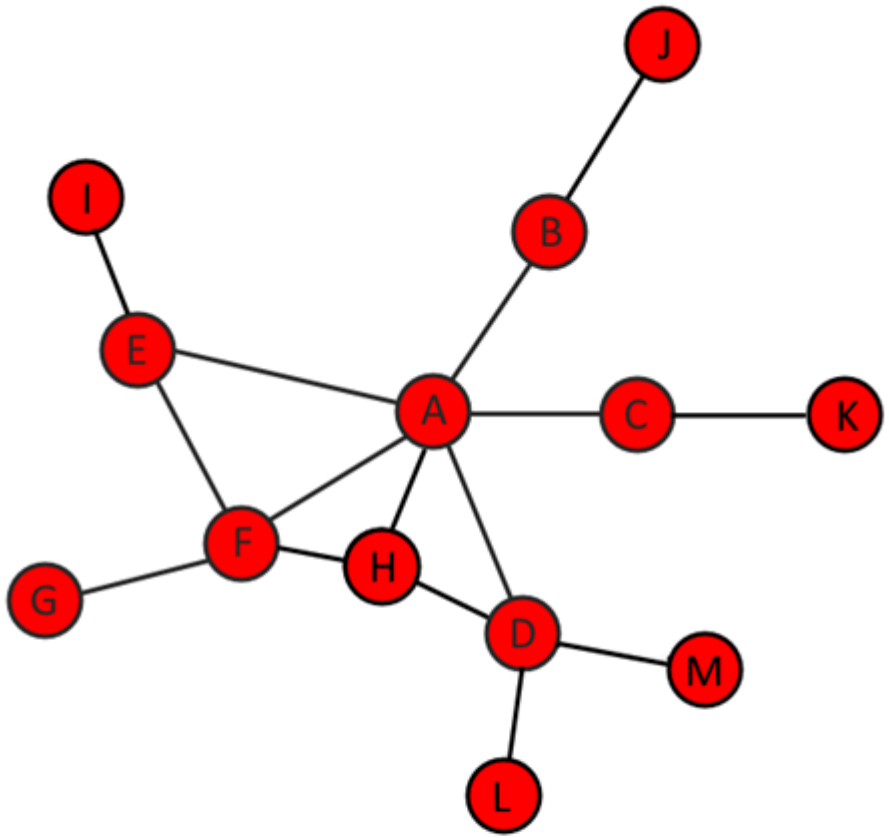


Geodesic distance matrix

	A	E	F	H	B	C	D	L	J	K	G	I
A	0	1	3	2	.	.	.	.	.	.	4	.
E	3	0	2	1	.	.	.	.	.	.	3	.
F	1	1	0	2	.	.	.	.	.	.	1	.
H	2	2	1	0	.	.	.	.	.	.	2	.
B	1	2	4	3	0	.	.	.	.	.	5	.
C	1	2	4	3	.	0	.	.	.	.	5	.
D	1	2	4	3	.	.	0	.	.	.	5	.
L	2	3	5	4	.	1	.	0	.	.	6	.
J	2	3	5	4	.	.	1	.	0	.	6	.
K	2	3	5	4	.	.	1	.	.	0	6	.
G	.	.	.	.	.	.	.	.	.	.	0	.
I	2	2	1	3	.	.	.	.	.	.	2	0

Directed Connectedness = 0.41

Neighbors



Neighbors are nodes that are connected

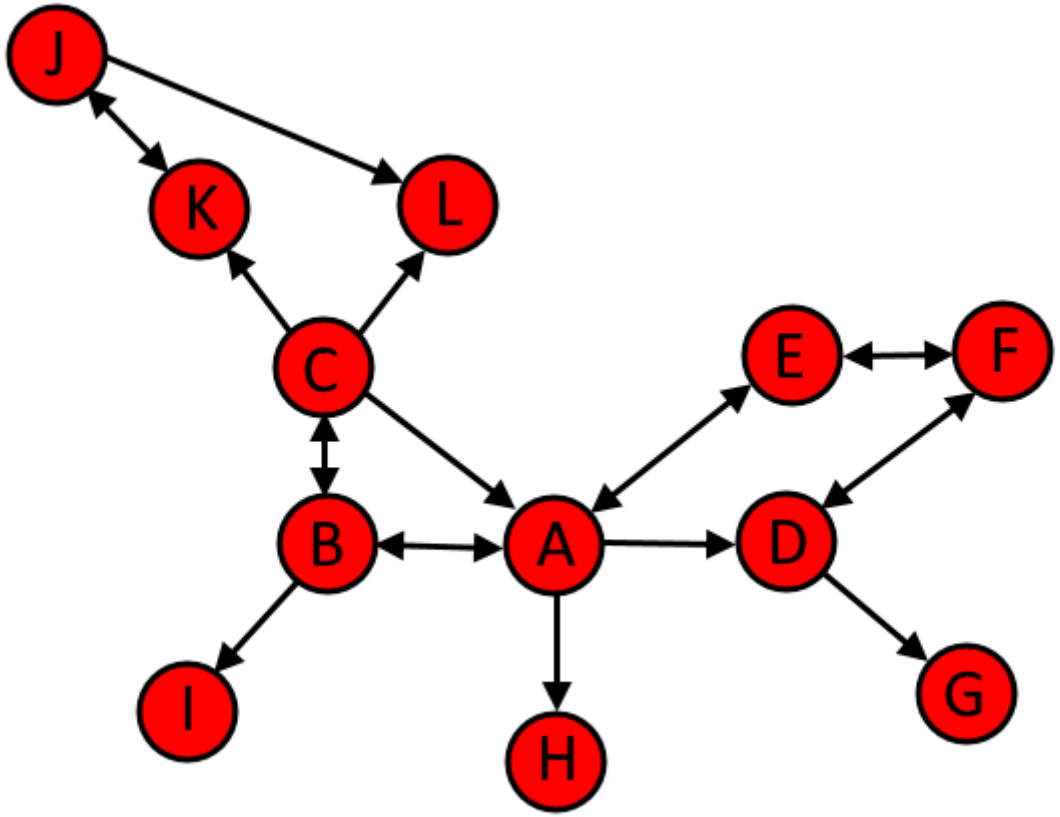
Co-citation is a method for identifying how many neighbors in common nodes have and is a measure of neighborhoods

	A	B	C	E	D	F	H	J	K	M	L	I	G
A	0	0	0	1	1	2	2	1	1	1	1	1	1
B	0	0	1	1	1	1	1	0	0	0	0	0	0
C	0	1	0	1	1	1	1	0	0	0	0	0	0
E	1	1	1	0	1	1	2	0	0	0	0	0	1
D	1	1	1	1	0	2	1	0	0	0	0	0	0
F	2	1	1	1	2	0	1	0	0	0	0	1	0
H	2	1	1	2	1	1	0	0	0	1	1	0	1
J	1	0	0	0	0	0	0	0	0	0	0	0	0
K	1	0	0	0	0	0	0	0	0	0	0	0	0
M	1	0	0	0	0	0	1	0	0	0	1	0	0
L	1	0	0	0	0	0	1	0	0	1	0	0	0
I	1	0	0	0	0	1	0	0	0	0	0	0	0
G	1	0	0	1	0	0	1	0	0	0	0	0	0

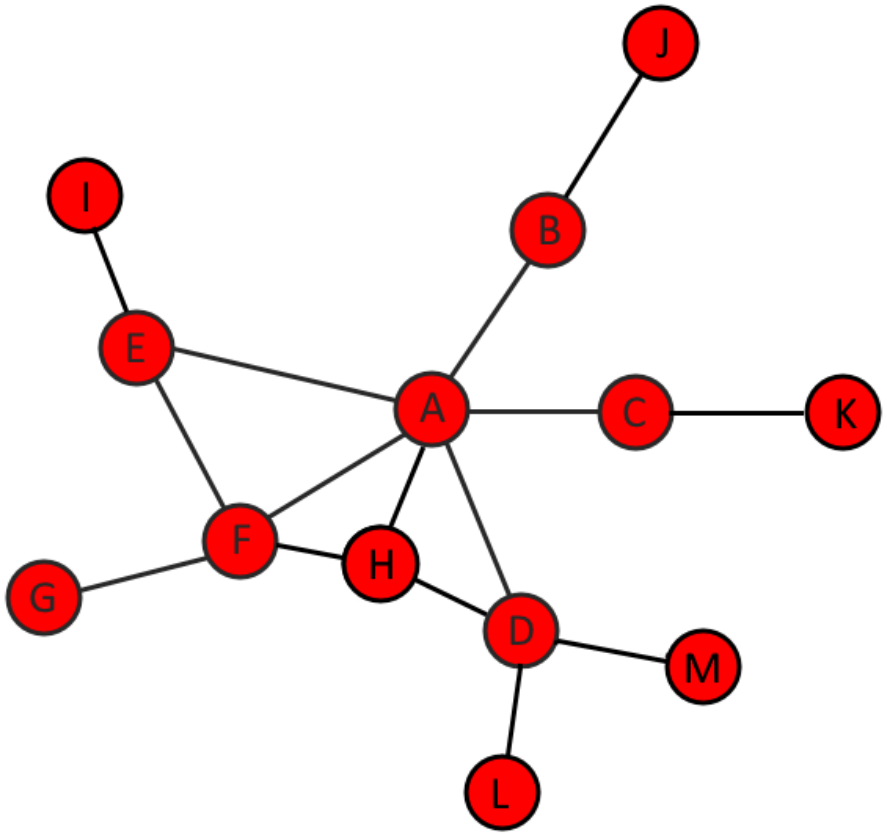
D and F have 2 neighbors in common

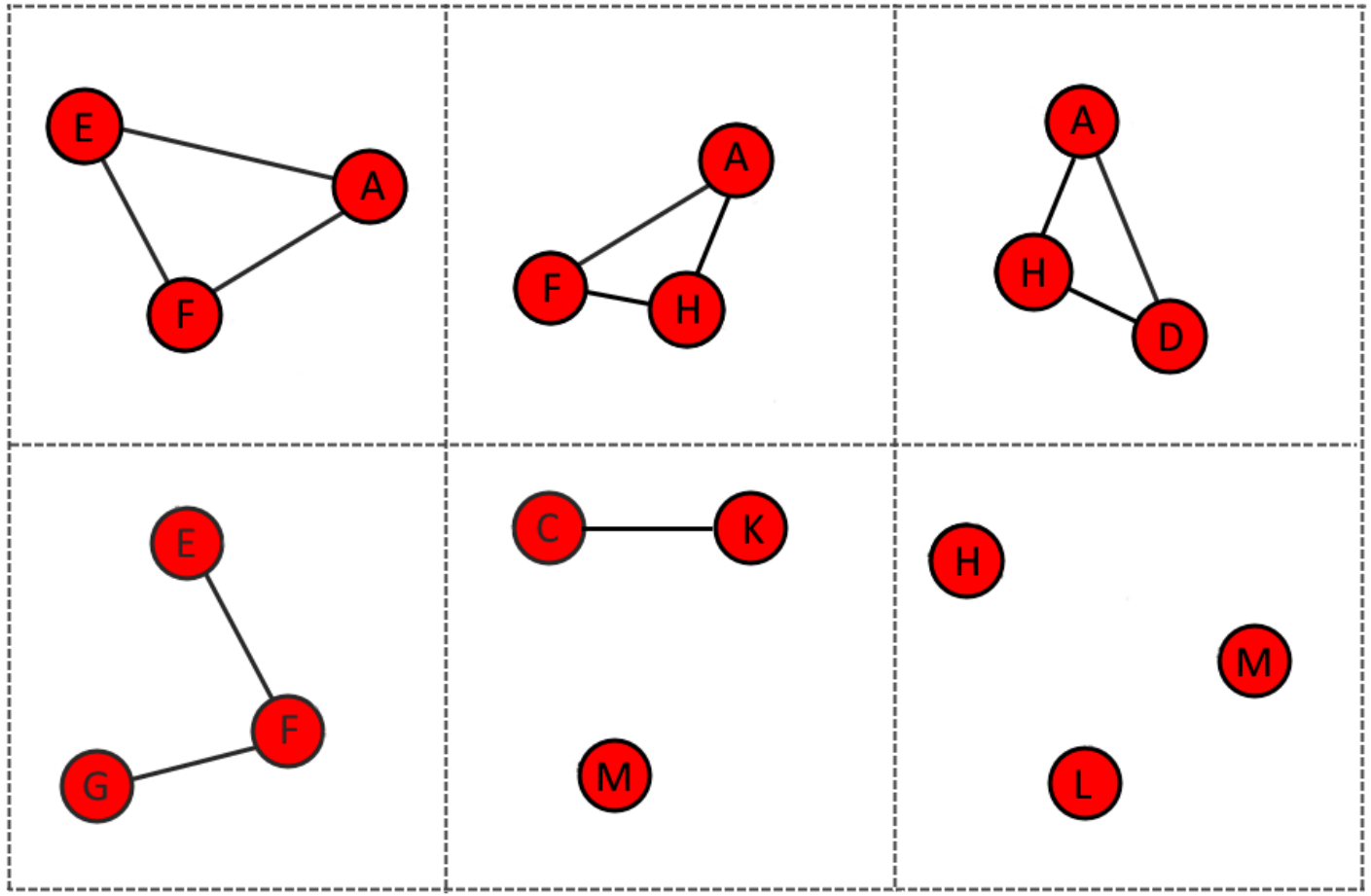
Reciprocity

What proportion of dyads have mutual/ bidirectional/ symmetrical/ reciprocated edges?



Transitivity





Triad census

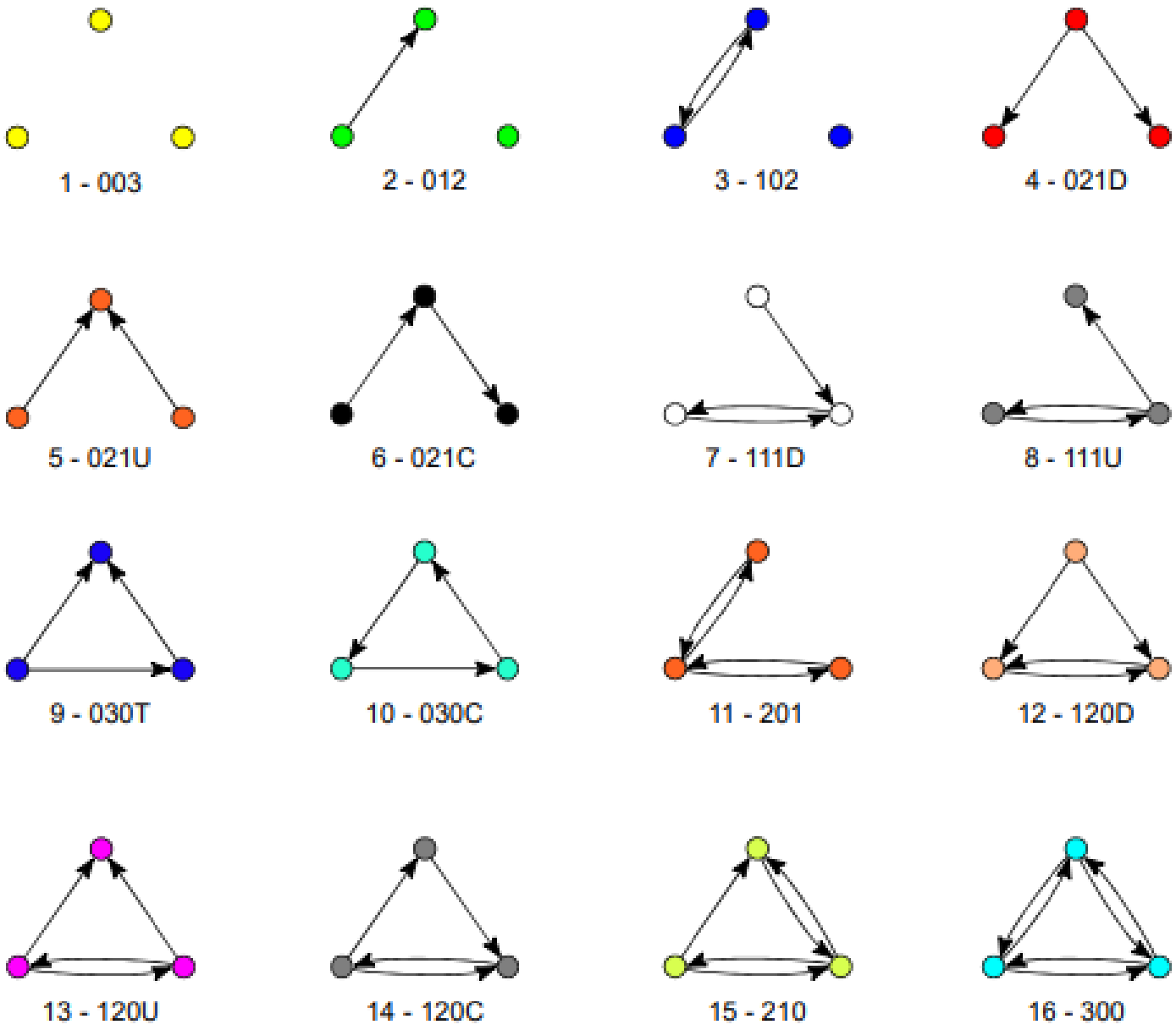
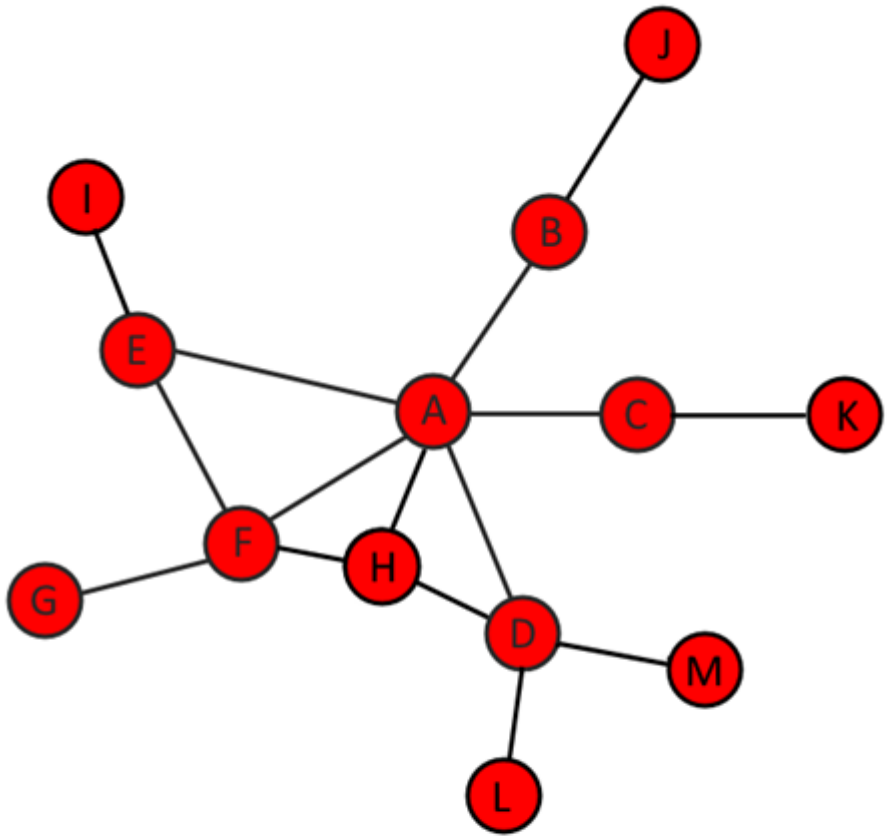


Figure 3 - Triad types with their sequential numbers in Pajek.

Clustering Coefficient



(Watts&Strogatz 1998)

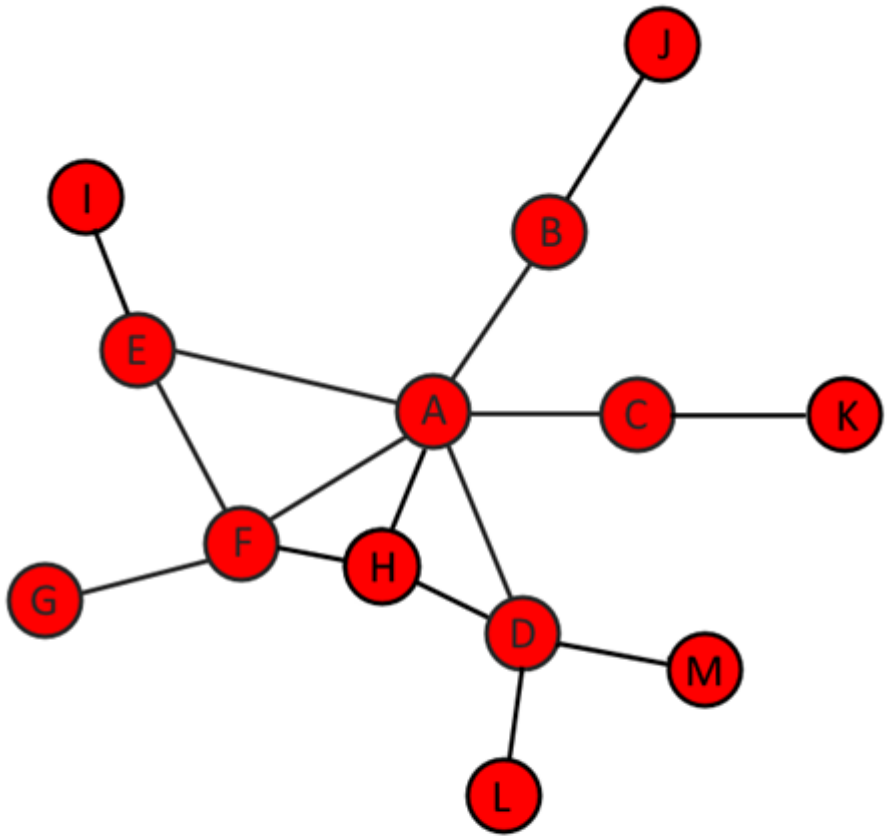
Transitivity relates to Watts & Strogatz clustering coefficient for undirected networks

Get the individual clustering of each node (density of ties in each ego network). Then the clustering coefficient is a weighted average of individual clustering coefficients across all nodes.

High clustering may occur in networks that aren't globally cohesive (local vs global cohesion)

$$C = \frac{3 \times \text{number of triangles in the graph}}{\text{number of connected triples}}$$

Local Clustering Coefficient



(Watts&Strogatz 1998)

For each node, calculate the number of its neighbors who share a connection. This can be used to calculate the local clustering coefficient (local transitivity) for each node.

$$\text{C.directed} = \frac{\text{No. directed connections between i's neighbors}}{n_i * (n_i - 1)}$$

$$\text{C.undirected} = \frac{\text{No. undirected connections between i's neighbors}}{n_i * (n_i - 1)/2}$$

n_i = number of neighbors of i

Comparing networks with correspondence analysis

e.g. comparison of triads between networks (K. Faust - https://www.stat-d.si/mz/mz3.1/faust.pdf)

Table 2: Triad distributions for 51 Networks, expressed as percentages.

	Tryad type															
Network	003	012	102	021D	021U	021C	111D	111U	030T	030C	201	120D	120U	120C	210	300
baboonf	12	35	02	10	09	07	05	01	14	01	00	03	00	01	01	00
baboonm1	00	00	00	00	00	00	00	00	15	00	00	00	50	00	15	20
baboonm2	00	05	00	10	10	00	00	10	40	00	00	00	20	00	05	00
baboonm3	02	18	02	02	00	08	01	02	51	00	00	03	10	00	01	00
banka	13	36	02	12	08	05	03	01	15	00	01	03	00	01	00	00
bankc	21	33	08	06	12	04	05	04	04	00	01	02	02	00	00	00
bankf	38	29	11	05	04	03	04	02	00	00	00	02	01	00	00	01
banks	07	13	01	19	01	04	00	13	09	00	02	01	14	03	10	03
cattle	12	31	02	13	07	10	02	02	17	00	00	02	01	01	00	00
chimp1	00	08	01	10	19	06	01	04	42	00	00	01	05	02	01	00
chimp2	00	00	00	04	00	00	02	08	06	00	02	12	19	08	27	11
chimp3	00	00	02	00	00	00	04	00	00	00	17	00	01	00	27	49
cole1	81	12	06	00	00	00	00	00	00	00	00	00	00	00	00	00
cole2	79	14	06	00	00	00	00	00	00	00	00	00	00	00	00	00
colobus1	00	00	00	25	00	00	00	25	00	00	00	00	25	00	25	00
colobus2	00	00	10	20	00	00	20	00	00	00	00	10	00	10	20	10
colobus3	02	08	19	00	00	02	10	13	00	00	21	02	02	02	08	08
cowg	93	02	04	00	00	00	00	00	00	00	00	00	00	00	00	00
cowl	87	12	01	00	00	00	00	00	00	00	00	00	00	00	00	00
eiesk1	49	23	15	01	02	01	02	03	00	00	01	00	01	00	01	00
eiesk2	39	23	17	02	03	01	05	04	01	00	02	01	01	00	01	01
eiesm	20	09	12	06	00	01	01	13	01	00	14	01	02	00	08	10
fifth	29	32	21	02	02	01	02	03	01	00	01	02	01	01	01	01
fourth	19	25	28	04	02	02	05	05	01	00	03	02	01	01	03	01
hyenaf	55	31	01	06	01	02	00	01	02	00	00	00	01	00	00	00
hyenam	28	34	02	15	04	05	00	04	07	00	00	00	01	00	00	00
kids1	06	11	13	05	01	04	08	14	03	00	06	03	04	03	12	07
kids2	60	28	04	02	02	02	01	01	01	00	00	00	00	00	00	00
macaca	09	13	21	03	02	02	09	09	00	00	10	02	02	01	10	07
nfponies	21	41	00	12	06	08	00	00	13	00	00	00	00	00	00	00
patasf	00	00	00	03	03	05	00	00	81	01	00	05	02	00	00	00
patasg	38	29	14	02	01	03	03	04	00	00	02	00	00	01	01	01
ponies	00	01	00	02	04	04	01	02	52	00	00	17	12	01	04	00
prison	82	12	05	00	00	00	00	00	00	00	00	00	00	00	00	00
reddeer	06	20	09	06	00	00	00	26	00	00	09	00	11	00	06	09
rhesus1	00	00	00	03	00	09	00	03	74	00	00	03	09	00	00	00
rhesus2	00	10	00	00	00	40	00	00	50	00	00	00	00	00	00	00
rhesus4	01	07	00	10	10	16	01	01	45	01	00	04	01	02	01	00
rhesus5	00	25	00	10	08	16	00	01	37	00	00	01	02	00	00	00
rhesus6	00	03	01	13	01	05	00	02	61	01	00	10	02	01	03	00
silver	00	00	03	06	01	01	02	06	19	01	00	21	09	03	17	12
sparrow	01	08	00	09	05	09	01	01	54	01	00	05	03	01	00	00
third	10	21	19	05	04	04	06	08	03	00	04	05	02	01	06	03
tits	00	00	02	09	00	04	00	05	21	00	00	14	16	04	20	05
vcbf	35	22	20	03	01	02	03	06	01	00	01	02	01	00	01	01
vcg	09	14	17	05	02	02	06	12	03	00	06	05	03	02	08	05
vcw	28	27	14	05	03	04	03	05	02	00	00	03	02	01	01	01
vervet1a	01	11	00	16	08	08	03	03	37	00	00	05	05	02	01	00
vervet1m	02	07	01	15	12	02	04	04	34	00	01	08	05	02	03	00
vervet2a	00	00	01	08	01	00	01	01	71	00	00	14	04	00	01	00

Table 3: Correspondence analysis of triad censuses, scores for networks and triads.

	Dimension			
Network	1	2	3	4
baboonf	0.078	-0.360	-0.448	0.190
baboonm1	-0.980	1.141	1.395	0.307
baboonm2	-0.803	-0.187	0.378	0.473
baboonm3	-0.656	-0.573	0.110	-0.131
banka	0.133	-0.371	-0.420	0.191
bankc	0.402	-0.166	-0.422	0.283
bankf	0.756	-0.107	-0.175	0.117
banks	-0.341	0.388	0.126	0.597
cattle	0.024	-0.411	-0.361	0.172
chimp1	-0.667	-0.563	-0.171	0.105
chimp2	-0.767	1.182	0.542	0.346
chimp3	-0.548	2.390	0.080	-1.675
cole1	1.358	-0.206	0.620	-0.222
cole2	1.336	-0.209	0.582	-0.205
colobus1	-0.641	0.906	0.516	1.320
colobus2	-0.406	1.013	-0.392	0.223
colobus3	-0.014	1.214	-0.841	-0.388
cowg	1.478	-0.225	0.869	-0.314
cowl	1.414	-0.267	0.772	-0.252
eiesk1	0.955	-0.026	-0.018	0.002
eiesk2	0.784	0.077	-0.213	0.037
eiesm	0.208	0.904	-0.289	-0.298
fifth	0.700	0.072	-0.388	0.124
fourth	0.517	0.274	-0.605	0.147
hyenaf	0.971	-0.275	0.242	0.036
hyenam	0.456	-0.286	-0.174	0.273
kids1	-0.102	0.782	-0.397	0.142
kids2	1.092	-0.229	0.250	-0.041
macaca	0.117	0.848	-0.602	-0.084
nfponies	0.304	-0.462	-0.293	0.153
patasf	-1.081	-0.965	0.181	-0.570
patasg	0.797	0.011	-0.206	0.078
ponies	-0.985	-0.427	0.354	-0.099
prison	1.369	-0.222	0.651	-0.225
reddeer80	0.768	-0.347	0.019	0.023
rhesus1	-1.071	-0.829	0.302	-0.393
rhesus2	-0.750	-1.008	-0.344	-0.644
rhesus4	-0.729	-0.721	-0.217	-0.172
rhesus5	-0.463	-0.695	-0.349	-0.060
rhesus6	-0.938	-0.675	0.094	-0.243
silver	-0.799	0.642	0.341	-0.015
sparrow	-0.783	-0.710	-0.058	-0.206
third	0.154	0.370	-0.550	0.188
tits	-0.837	0.493	0.480	0.299
vcbf	0.725	0.112	-0.235	0.091
vcg	0.045	0.632	-0.461	0.122
vcw	0.539	0.003	-0.266	0.187
vervet1a	-0.634	-0.477	-0.136	0.138
vervet1m	-0.622	-0.291	-0.100	0.200
vervet2a	-1.059	-0.739	0.265	-0.343
vervet2m	-0.802	-0.426	-0.041	0.015

Triad	1	2	3	4
003	1.209	-0.179	0.452	-0.139
012	0.489	-0.207	-0.394	0.183
102	0.542	0.450	-0.657	0.053
021D	-0.360	-0.066	-0.166	0.520
021U	-0.243	-0.485	-0.399	0.258
021C	-0.426	-0.693	-0.437	-0.295
111D	-0.011	0.597	-0.738	0.066
111U	-0.189	0.606	-0.220	0.623
030T	-0.933	-0.760	0.126	-0.302
030C	-0.681	-0.460	-0.195	-0.263
201	-0.011	1.538	-0.843	-0.940
120D	-0.737	0.017	0.093	0.033
120U	-0.859	0.589	0.999	0.732
120C	-0.501	0.708	-0.199	0.390
210	-0.603	1.311	0.295	0.181
300	-0.543	1.955	0.256	-1.142

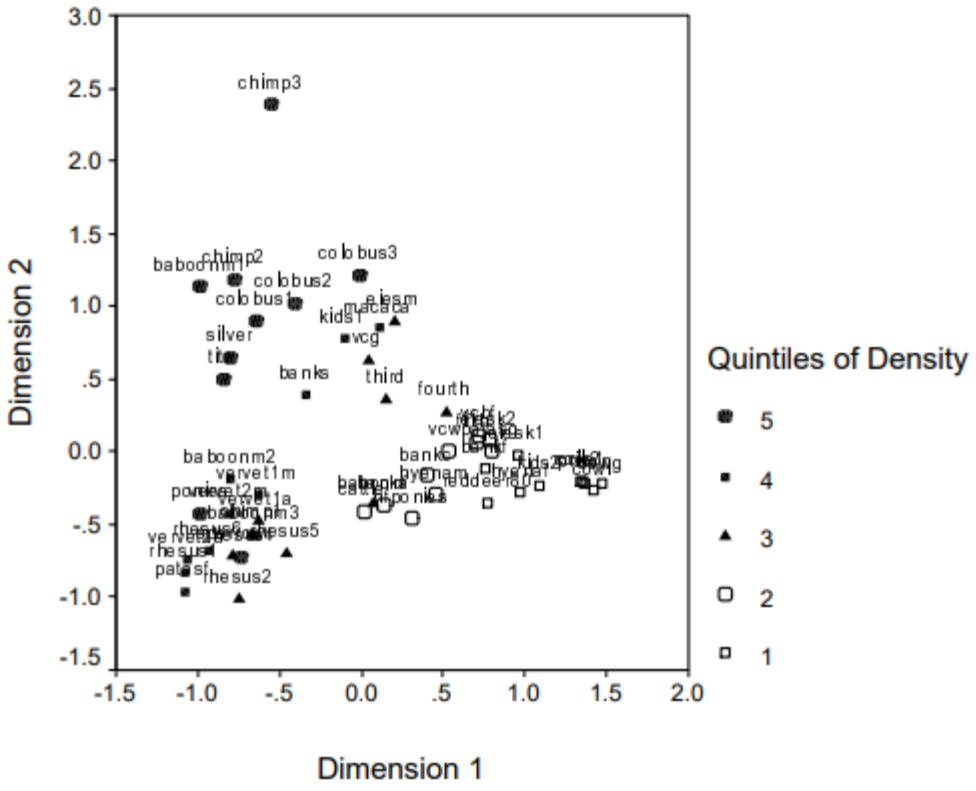


Figure 2a: Correspondence analysis of triad censuses, network space.

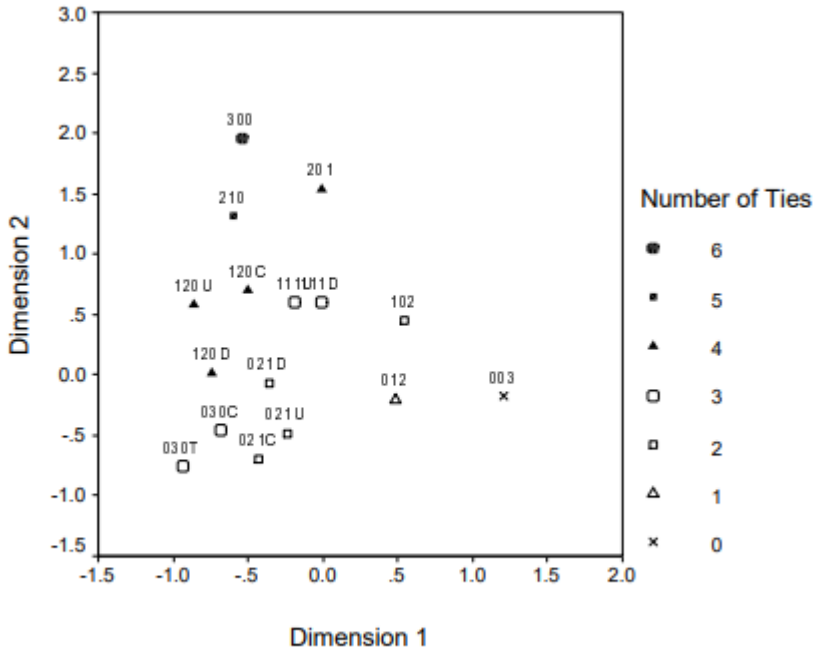


Figure 2b: Correspondence analysis of triad censuses, triad space.

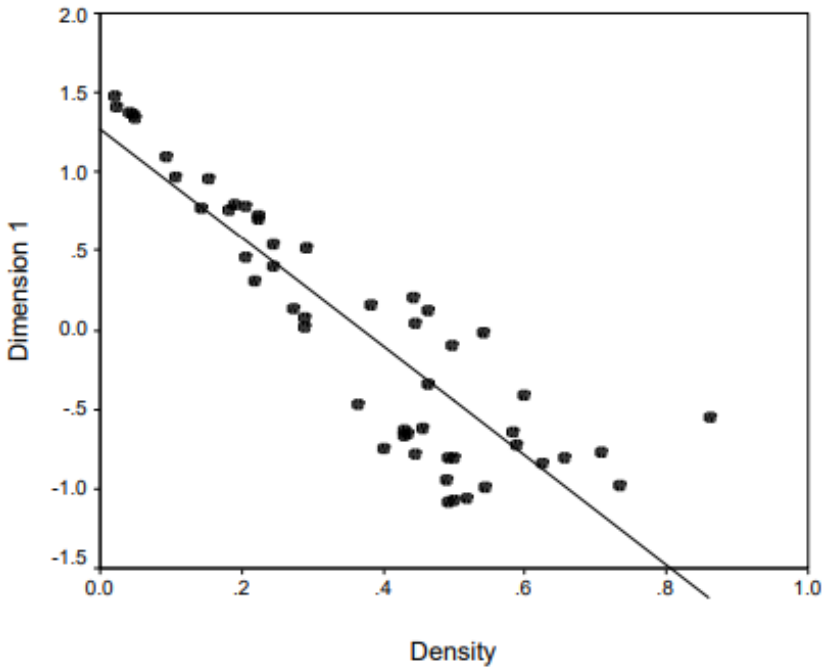


Figure 3: Scatterplot of correspondence analysis network space dimension 1 and the network density, N = 51 networks.

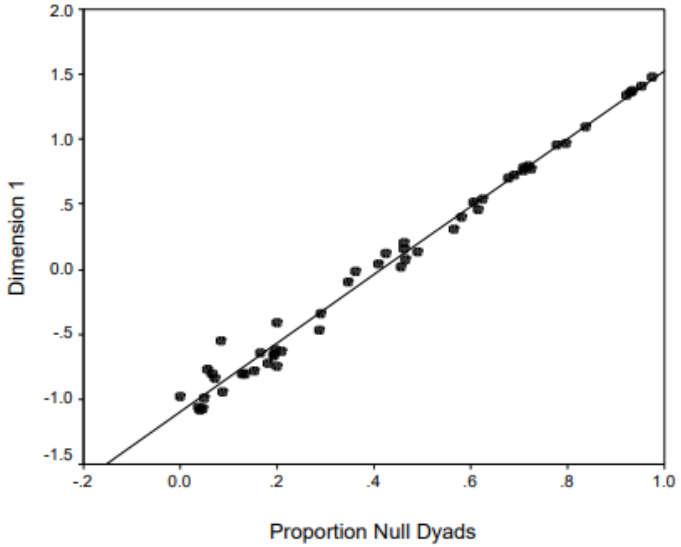


Figure 4: Scatterplot of correspondence analysis network space dimension 1 and the proportion of null dyads, N = 51 networks.

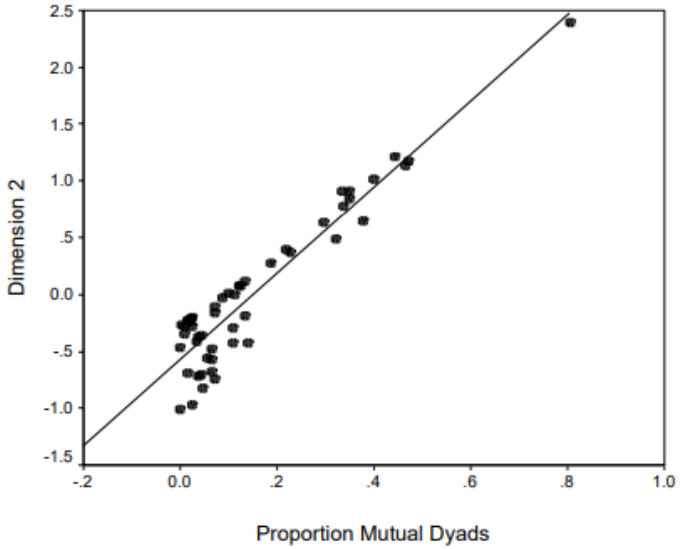


Figure 5: Scatterplot of correspondence analysis network space dimension 2 and the proportion of mutual dyads, N = 51 networks.

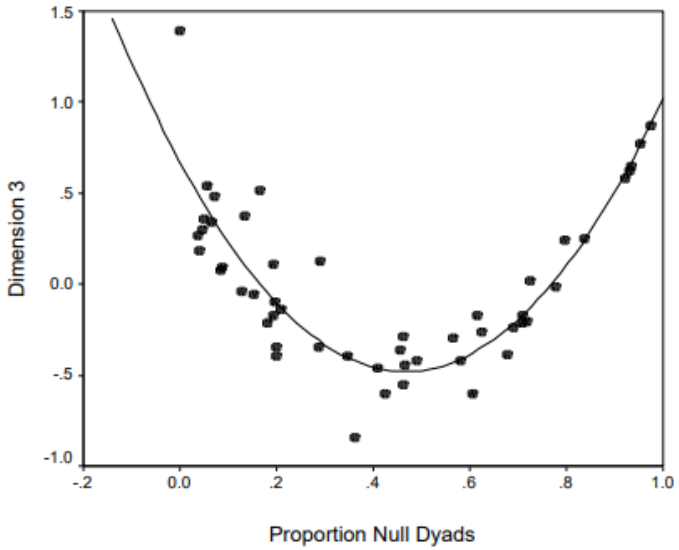


Figure 6: Scatterplot of correspondence analysis network space dimension 3 and the proportion of null dyads, N = 51 networks $Y = 0.67 - 4.9X + 5.2X^2$.

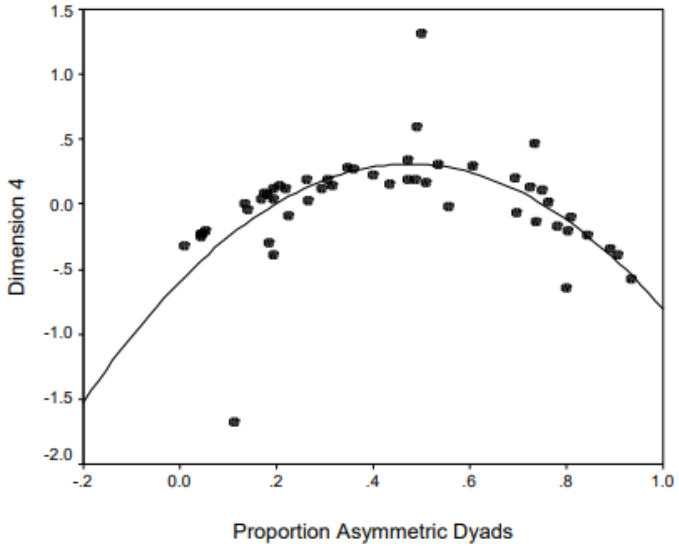
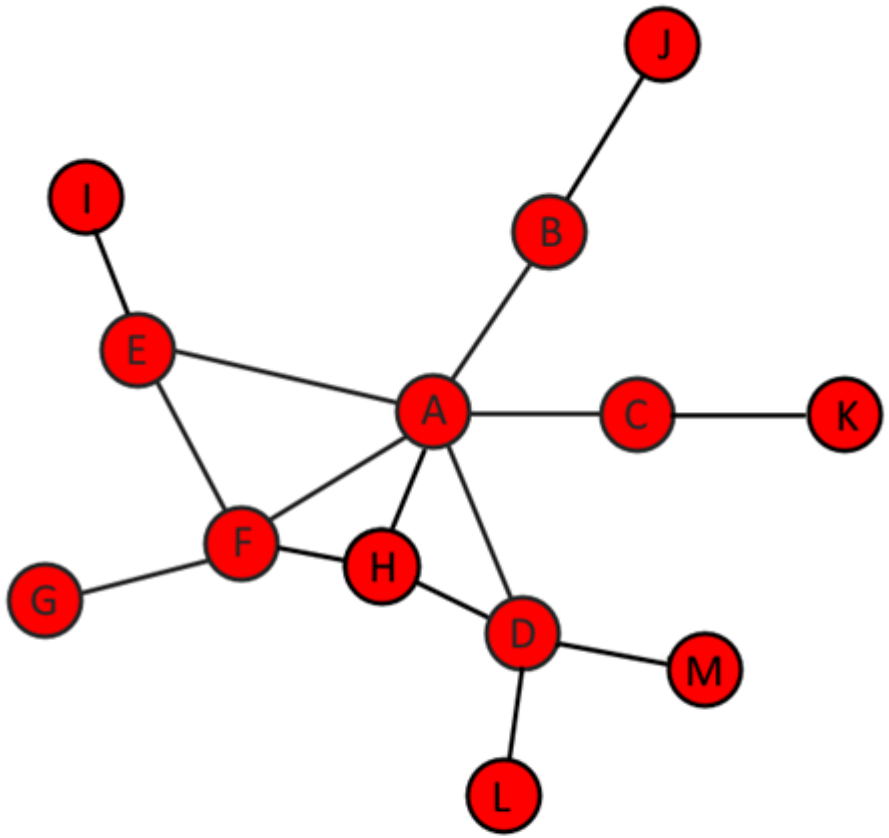


Figure 7: Scatterplot of correspondence analysis network space dimension 4 and the proportion of asymmetric dyads, N = 51 networks $Y = -0.59 + 3.82X - 4.03X^2$.

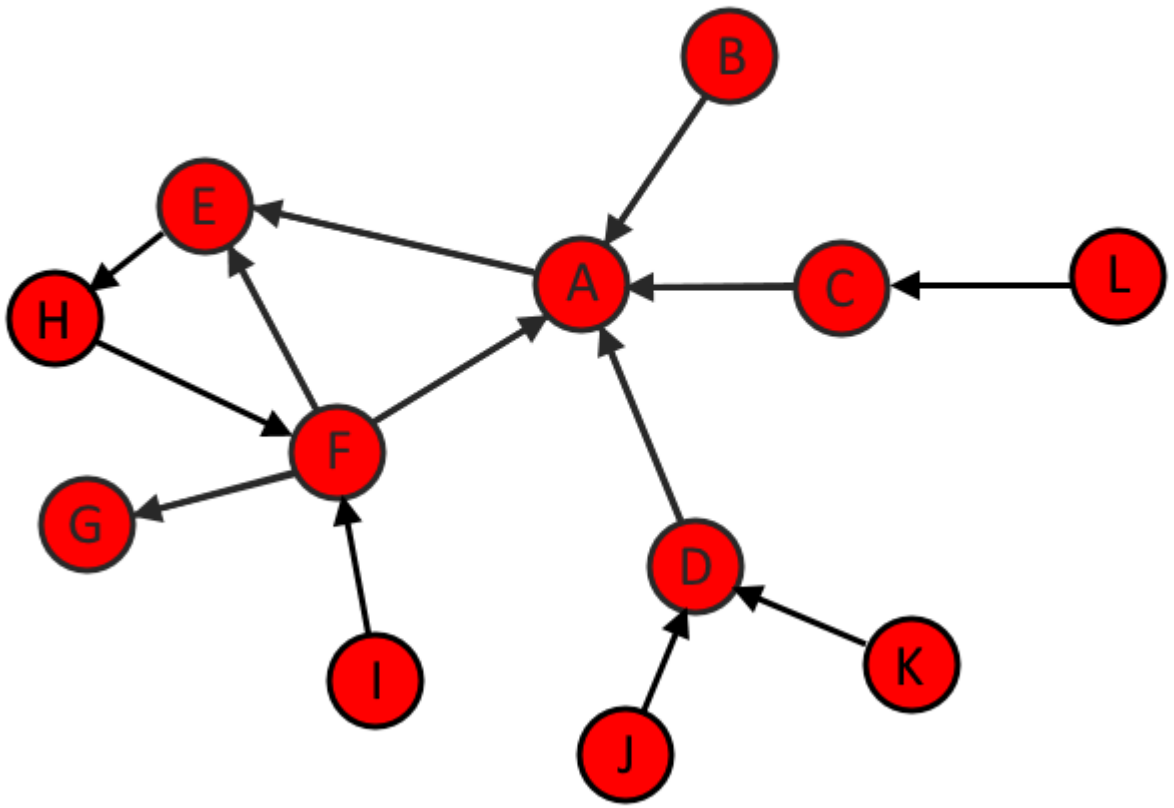
Node Metrics

Degree Centrality



A	B	C	E	D	F	H	J	K	M	L	I	G
6	2	2	3	4	4	3	1	1	1	1	1	1

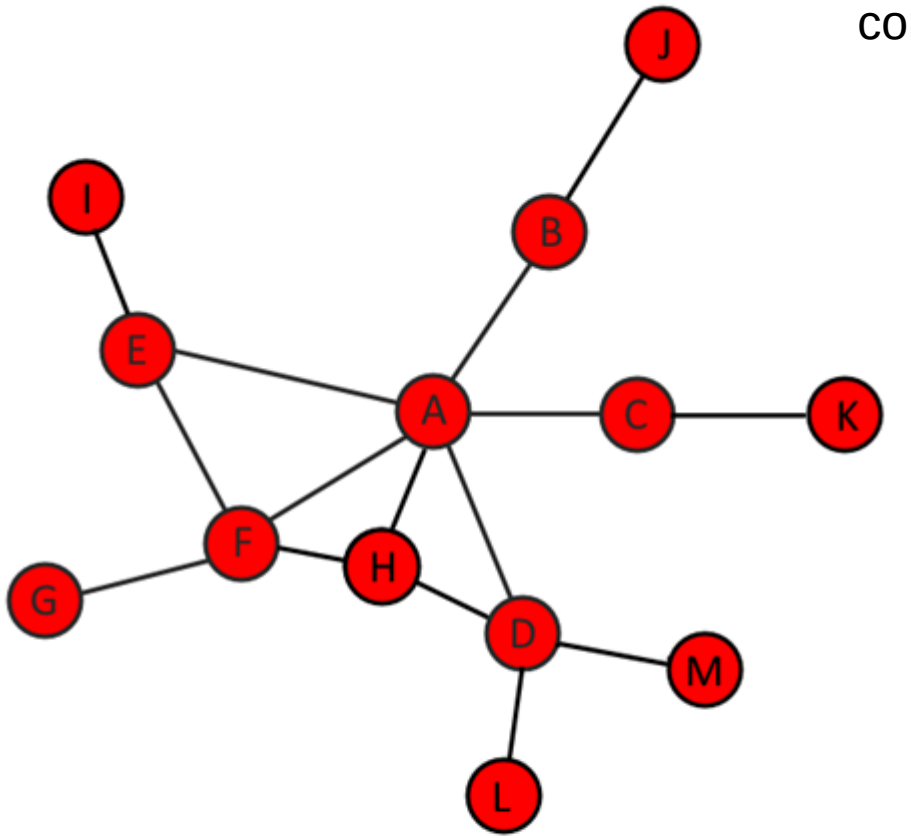
Degree Centrality



	out-degree	in-degree
A	1	4
B	1	0
C	1	1
D	1	2
E	1	2
F	3	2
G	0	1
H	1	1
I	0	1
J	1	0
K	1	0

Betweenness Centrality

How many shortest paths each node lies on when examining all pairs of nodes. If multiple shortest paths between two nodes – shared credit for each node. Individuals who have high betweenness are bridges connecting nodes from different parts of the network.

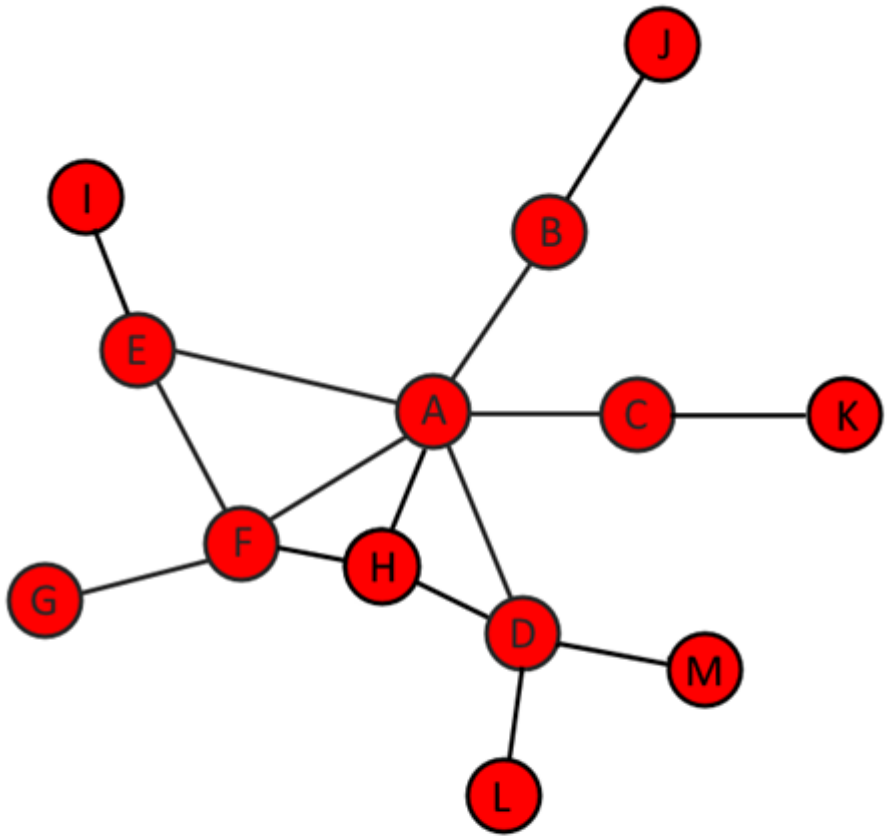


	Norm	Total
A	0.70	46
B	0.17	11
C	0.17	11
E	0.17	11
D	0.32	21
F	0.18	12
H	0.05	3
J	0.00	0
K	0.00	0
M	0.00	0
L	0.00	0
I	0.00	0
G	0.00	0

Freeman 1979

Eigenvector Centrality

Higher in individuals who are well-connected and connected to others who are also well-connected. Related to sum of centralitites of neighbors

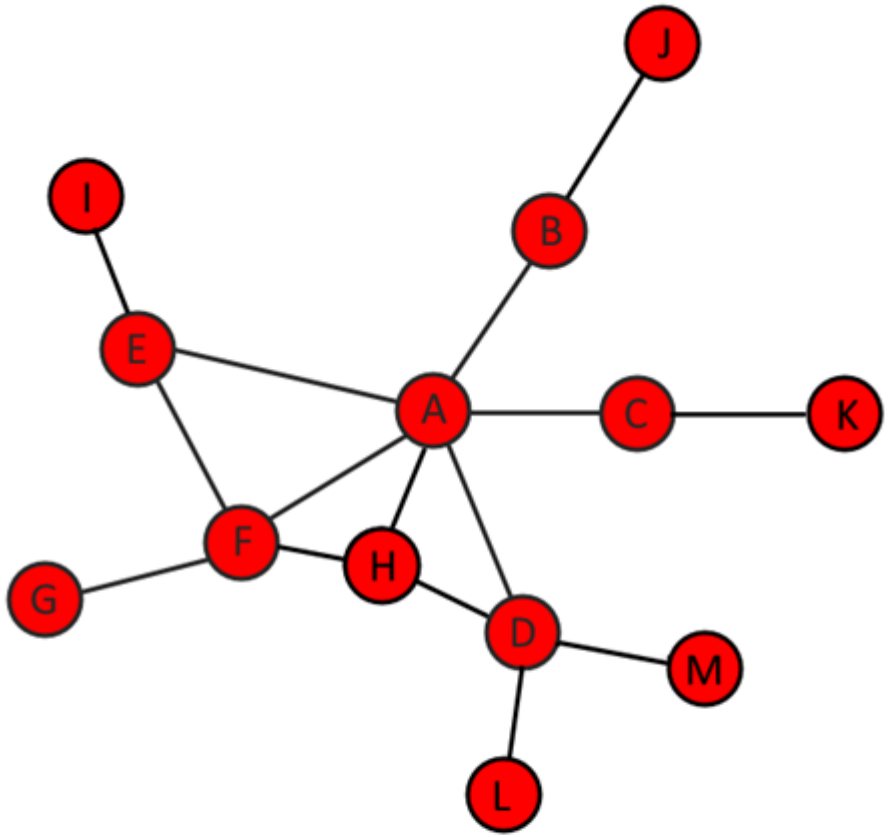


A	1.00
B	0.33
C	0.33
E	0.58
D	0.63
F	0.76
H	0.71
J	0.10
K	0.10
M	0.19
L	0.19
I	0.17
G	0.23

Bonacich 1972

Closeness Centrality

Related to the sum of geodesic distances from a node to all others. Essentially the inverse of a node's average geodesic distance to others in the network.



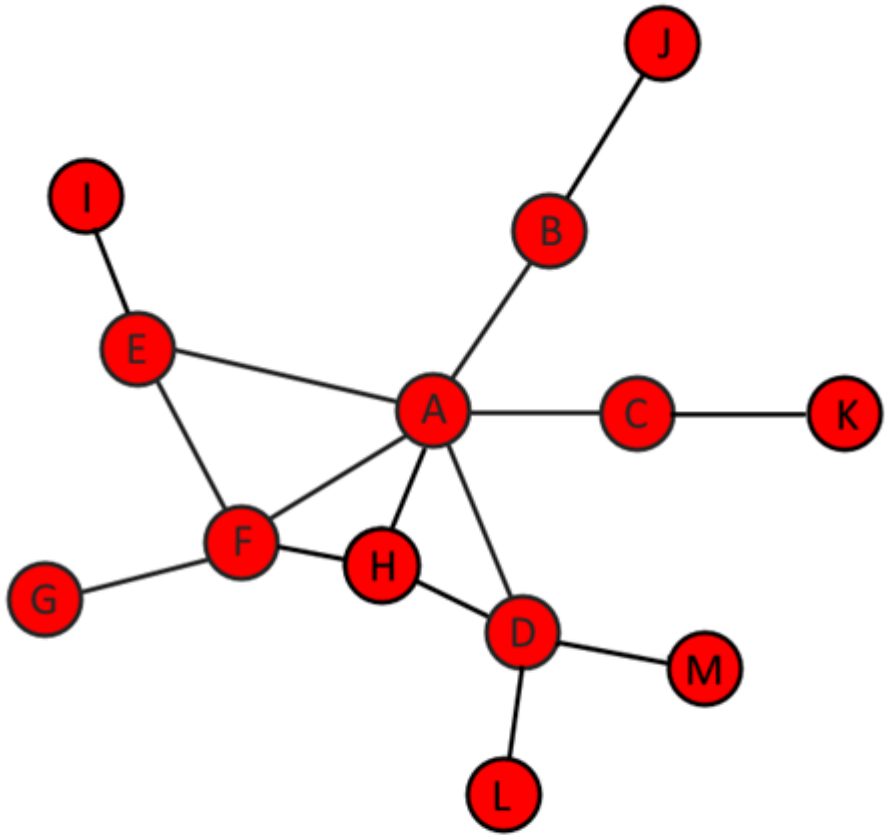
Freeman 1979

A	0.67
B	0.44
C	0.44
E	0.48
D	0.50
F	0.50
H	0.50
J	0.32
K	0.32
M	0.34
L	0.34
I	0.33
G	0.34

Closeness centrality is problematic in disconnected networks as no path exists between pairs of nodes. Also not particularly useful for directed networks due to fragmentation

K-reach Centrality

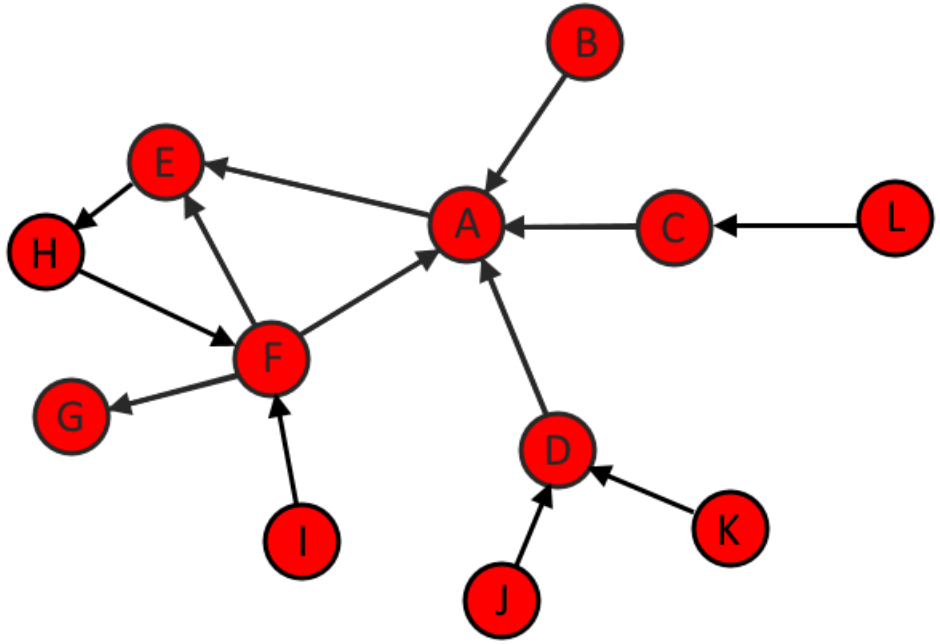
What proportion of nodes can be reached in K steps from each node ?



	K=2	K=3
A	1.00	1.00
B	0.62	1.00
C	0.62	1.00
E	0.69	1.00
D	0.69	1.00
F	0.69	1.00
H	0.77	1.00
J	0.23	0.62
K	0.23	0.62
M	0.38	0.69
L	0.38	0.69
I	0.31	0.69
G	0.38	0.69

Borgatti 2006

Hub Score

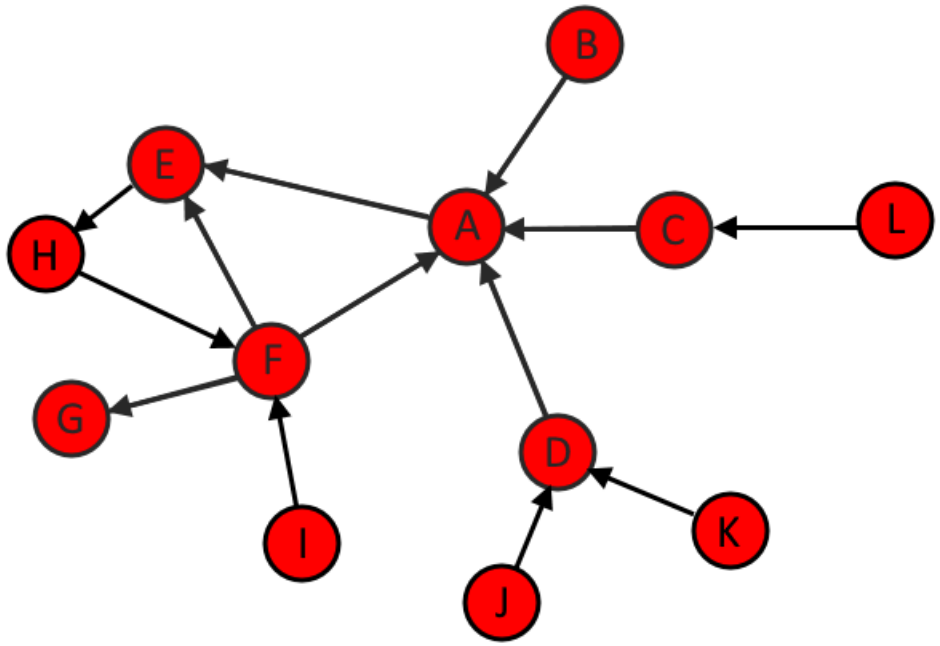


Hubs provide many links to other vertices (many outgoing ties) and are key drivers of flow

	raw	norm
A	0.19	0.26
E	0.00	0.00
F	0.72	1.00
H	0.00	0.00
B	0.39	0.54
C	0.39	0.54
D	0.39	0.54
L	0.00	0.00
J	0.00	0.00
K	0.00	0.00
G	0.00	0.00
I	0.00	0.00

Kleinberg

Authority Score

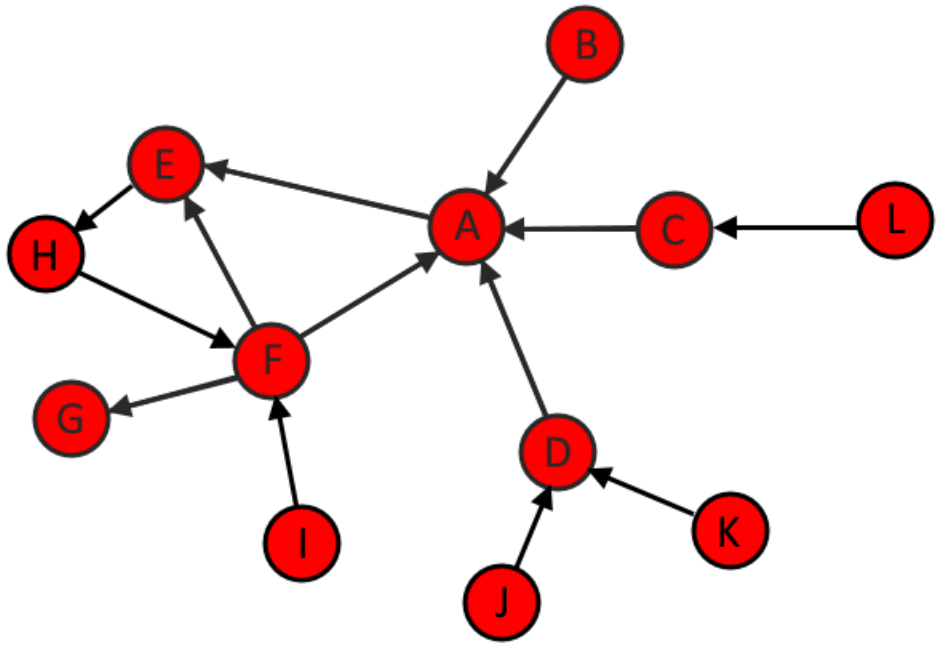


Authorities receive many incoming links particularly from hubs

	raw	norm
A	0.85	1.00
E	0.41	0.48
F	0.00	0.00
H	0.00	0.00
B	0.00	0.00
C	0.00	0.00
D	0.00	0.00
L	0.00	0.00
J	0.00	0.00
K	0.00	0.00
G	0.33	0.38
I	0.00	0.00

Kleinberg

Page Rank



A	0.16
E	0.21
F	0.20
H	0.19
B	0.02
C	0.03
D	0.05
L	0.02
J	0.02
K	0.02
G	0.07
I	0.02

Brin & Page 1998