GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Centrality

Social Networks Theories and Methods

James Hollway

Preliminary business

- TA
- Material
- Tutorials
- Slack
- Github

What is ‘centrality’?

- Being *central* in a network is a central concept in networks research…

- Why could ‘being central’ be important…
  - In your friendship network?
  - At work/uni?
  - In your empirical domain of interest?

Centrality questions

- Does a node have access to lots of resources?
- Does a node have access to different resources?
- Does a node connect different parts of the network?
- Does a node control the interaction between other nodes?
- Would the failure of a node cause the system to collapse?
- Is a node considered popular by others?
- Which nodes' esteem matters?

Warning

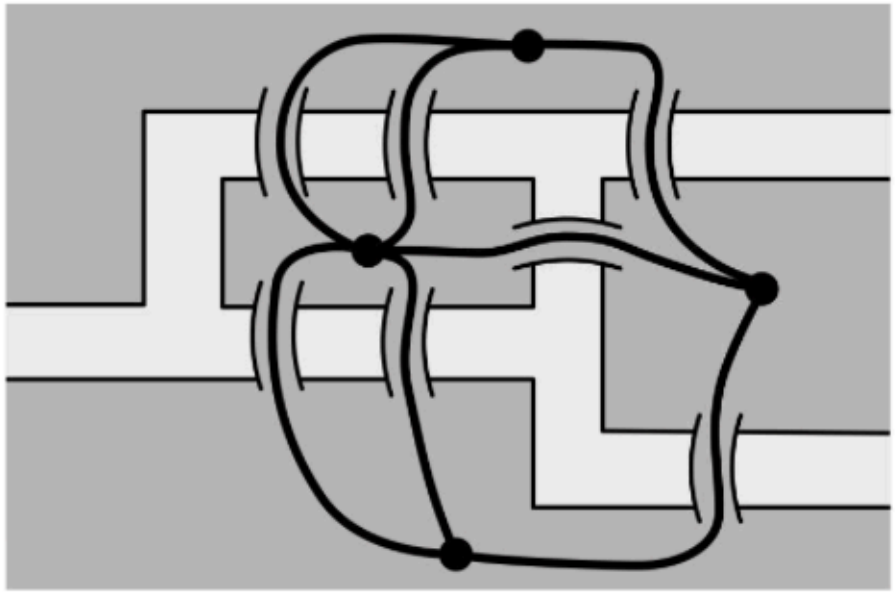
Not *one* kind of centrality

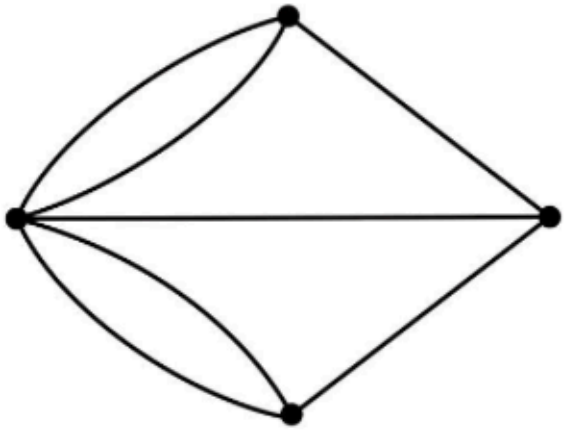
Many, with several in regular use:

- Degree
- Closeness
- Betweenness
- Eigenvector
- ...

They need to be deployed with care, *theoretically*...

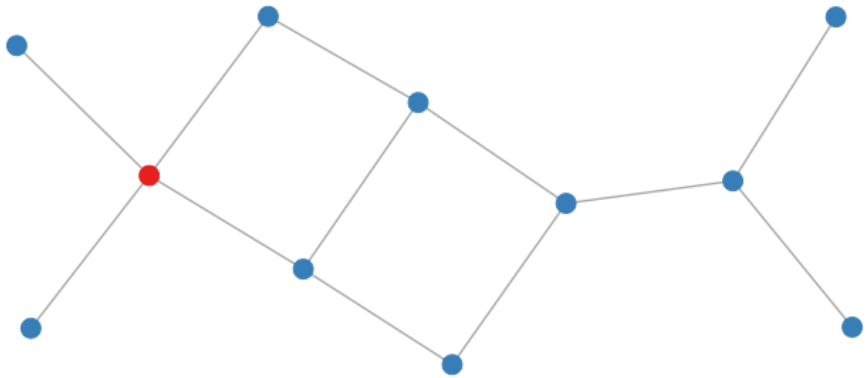
Recap: handshaking lemma





Degree centrality

- Question: how many ties does a node have?
- Theory: about e.g. activity, popularity, attention
- Method: degree, outdegree, and indegree



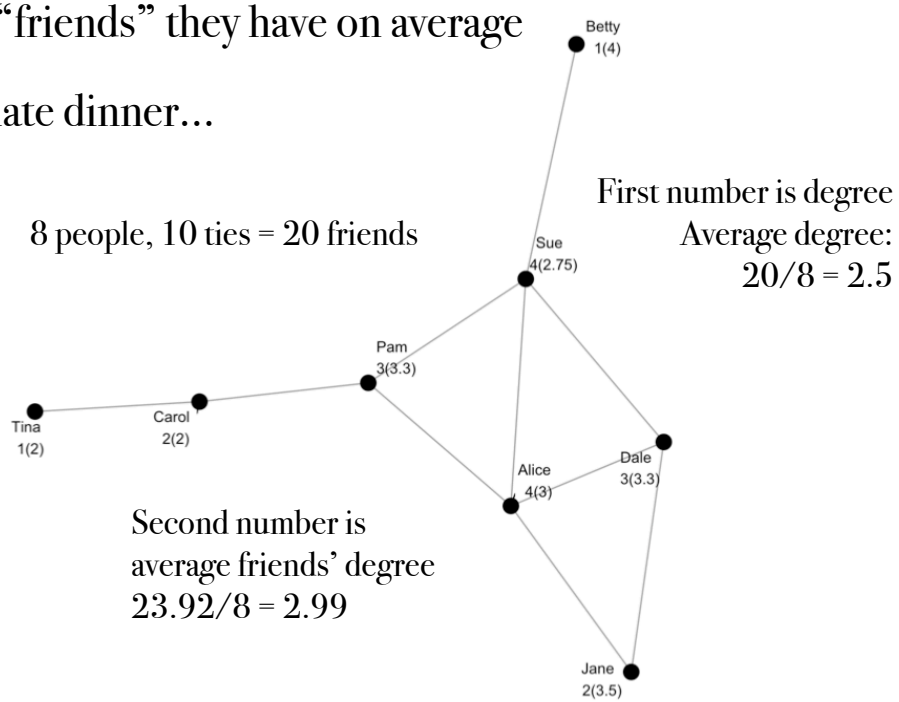
$$c_D(i) = \sum_{j \in V} x_{ij} \qquad c_{D+}(i) = \sum_{j \in V} x_{ij} \qquad c_{D-}(i) = \sum_{j \in V} x_{ji}$$

See also strength and weighted centrality extensions, e.g. Opsahl et al. 2010

Do you have +/- friends than your friends?

Do you have +/- friends than your friends?

- Option 1: data collection
  - Go on Facebook, count how many "friends" you have
  - Click on every one of those "friends" and see how many "friends" they have on average
  - Get a late dinner…



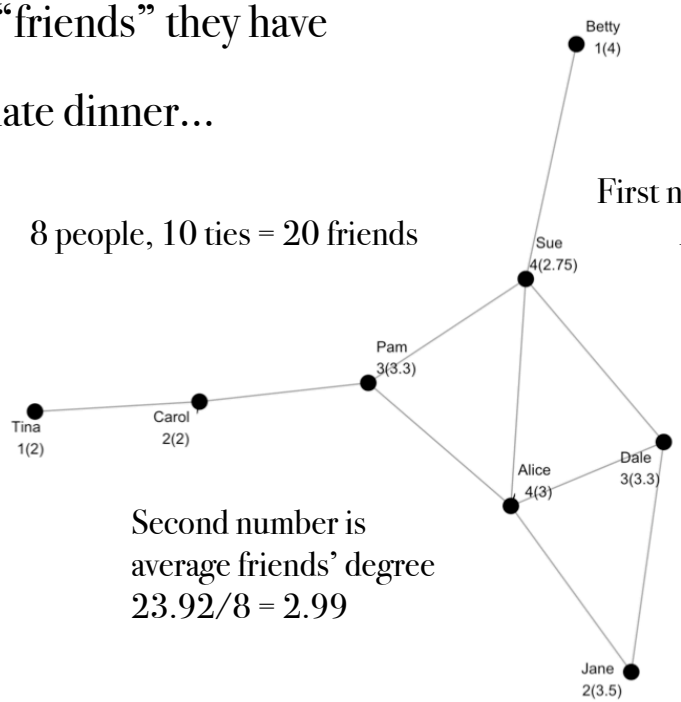
8 people, 10 ties = 20 friends

First number is degree
Average degree:
20/8 = 2.5

Second number is average friends' degree
23.92/8 = 2.99

Do you have +/- friends than your friends?

- Option 1: data collection
  - Go on Facebook, count how many "friends" you have
  - Click on every one of those "friends" and see how many "friends" they have
  - Get a late dinner…



8 people, 10 ties = 20 friends

First number is degree
Average degree:
20/8 = 2.5

Second number is
average friends' degree
23.92/8 = 2.99

- Option 2: graph theory

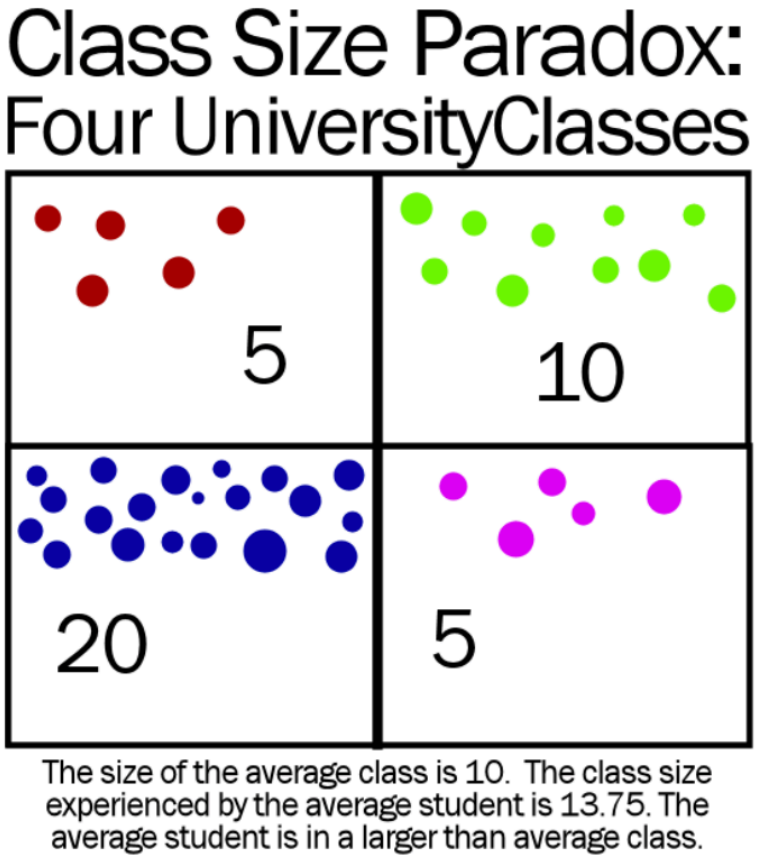
$$\frac{\sum d_v^2}{\sum d_v} \geq \frac{\sum d_v}{n}$$

- q.e.d

That is, you are unlikely to be most popular (1/n) but likely to be connected to most popular (d*/n)

See Feld (1991) "Why your friends have more friends than you do" *American Journal of Sociology* 96(6): 1464-1477.

Similar situations



- Class paradox: larger class than average
- Mothers paradox: more children than average
- Lovers paradox: more lovers than you
- Disease paradox: friends get flu first

Lesson #1

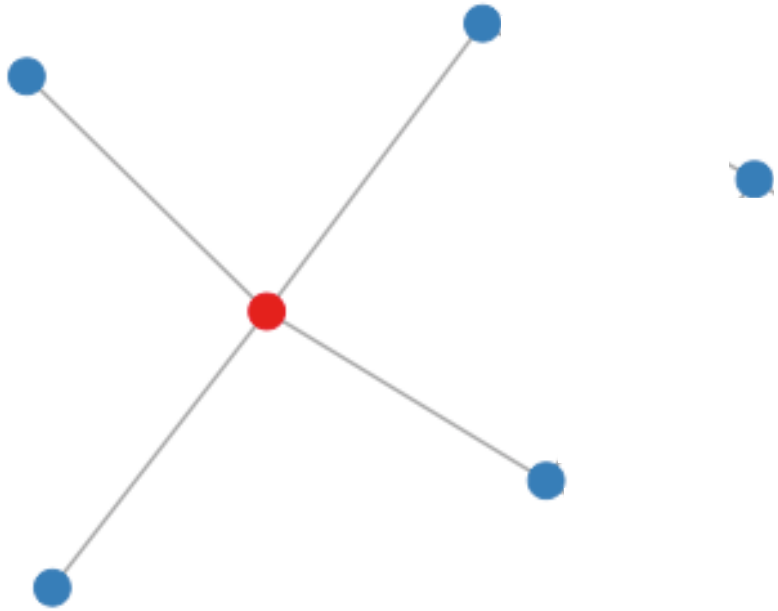
- Networks can fool people: just like how the earth seems flat because we're on it, people may get incorrect ideas about society because they're *in* it

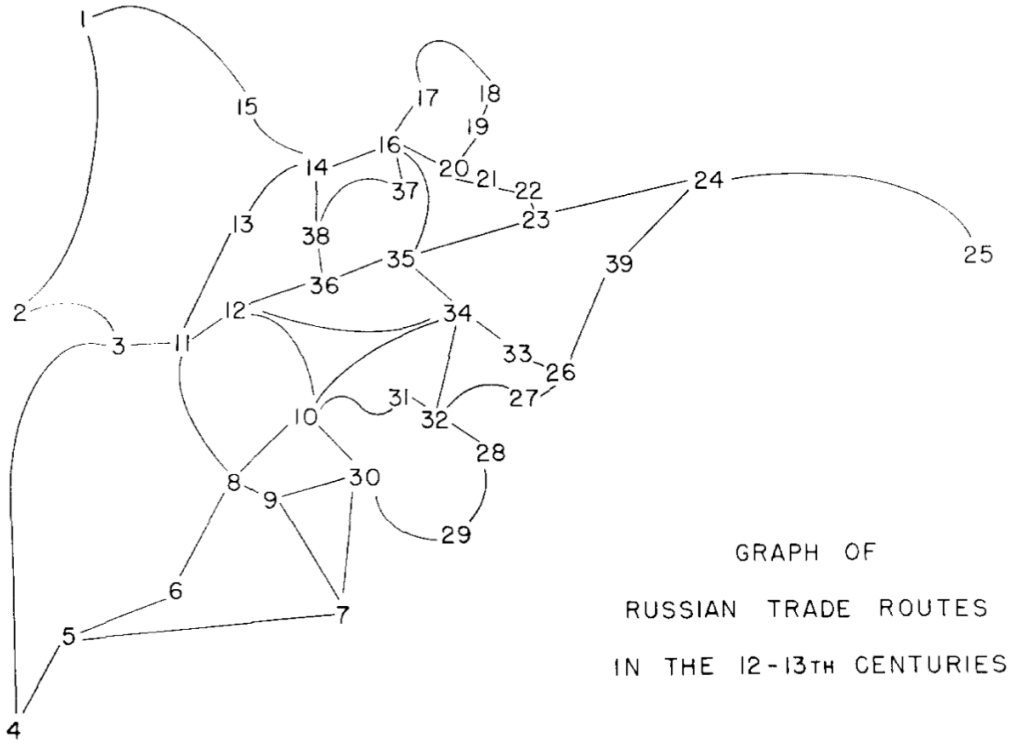
Tasks

- EASY: *Can you think of a network in which nobody has fewer friends than their friends?*
- MEDIUM: *Can you think of a network in which almost everybody has fewer friends than their friends?*
- HARD: *Can you think of a network where most have more friends than their friends?*

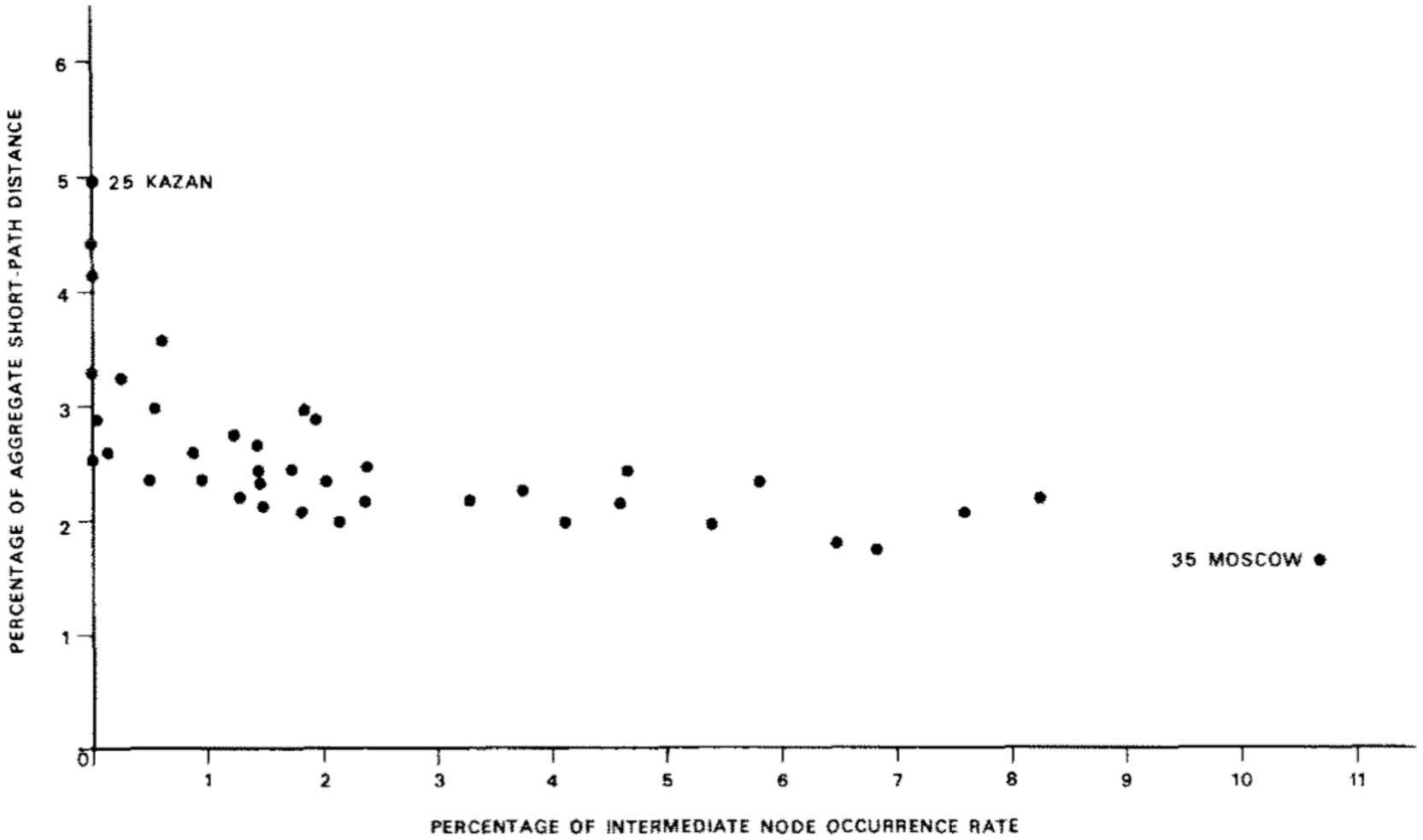
Special terms

- Isolate: node with *degree = 0*
- Leaf or pendant: node with *degree = 1*
- Pendant tie: tie adjacent to node with *degree = 1*
- Dominating: node with *degree = n-1*
- Star graph: network where *n-1* ties adjacent to single node





See Pitts (1978) "The medieval river trade network of Russia" *Social Networks* 1: 285–292.

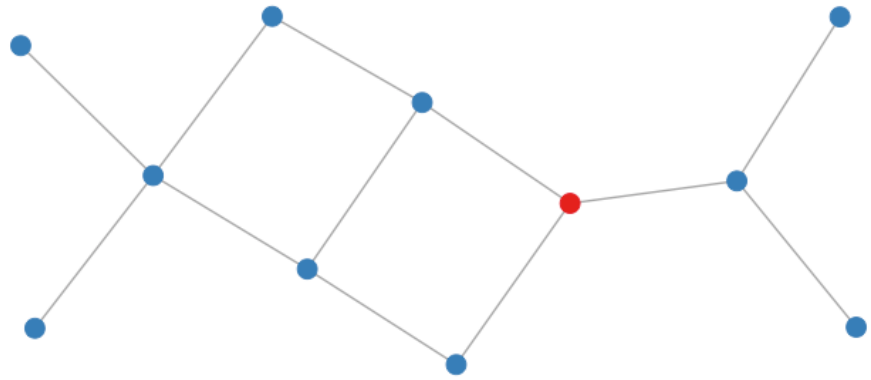

PERCENTAGE OF AGGREGATE SHORT-PATH DISTANCE
6
5
4
3
2
1
0
25 KAZAN
35 MOSCOW
1
2
3
4
5
6
7
8
9
10
11
PERCENTAGE OF INTERMEDIATE NODE OCCURRENCE RATE

Betweenness centrality

- AKA *medial centrality*
- Question: how many shortest paths go through the node?
- Theory: about e.g. brokering, transmission, contagion, innovation
- Method:

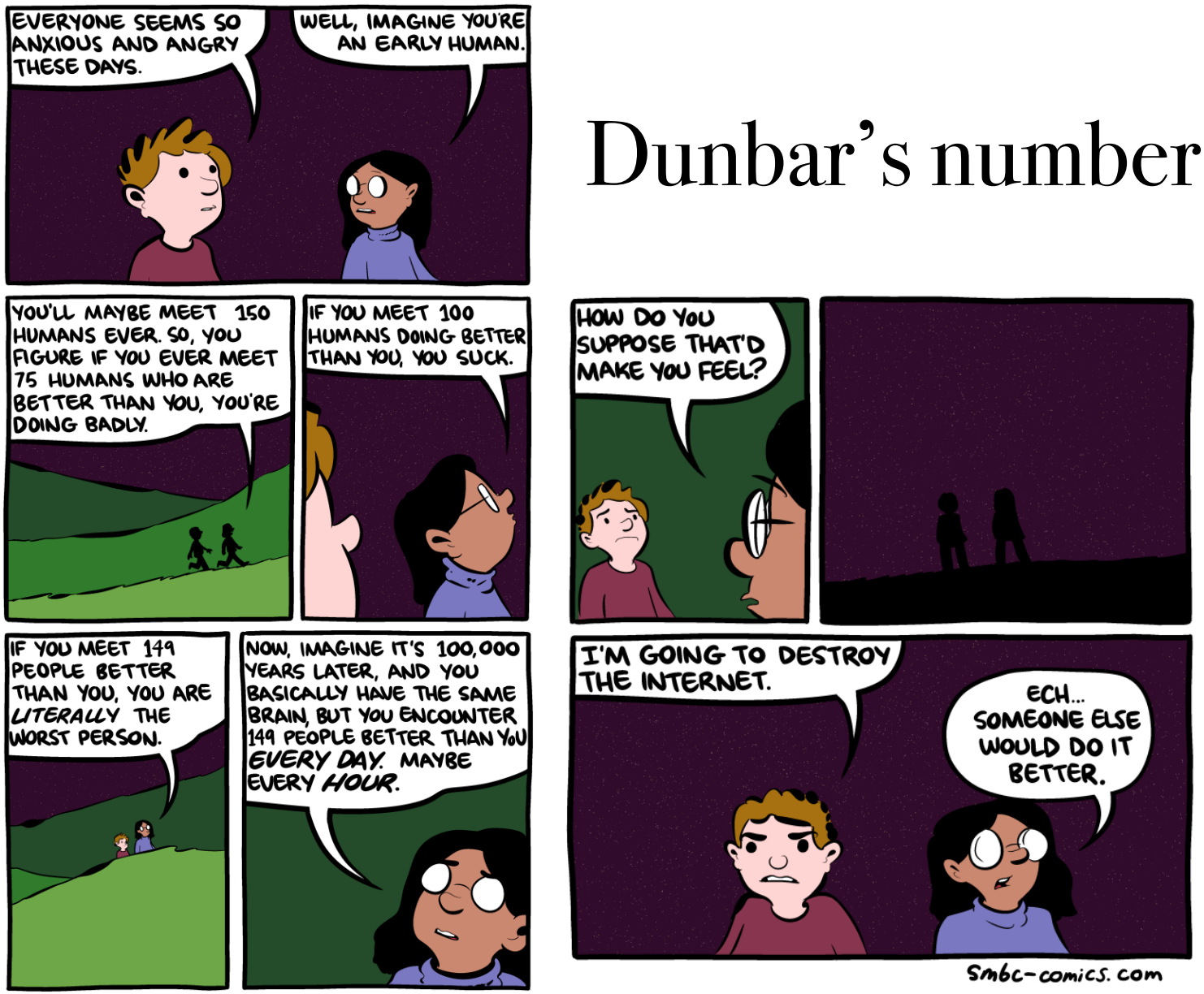
$$c_B(i) = \sum_{i,j,k \in N} \frac{\sigma_{jk\ i}}{\sigma_{jk}}$$

  - Naïve calculation complex

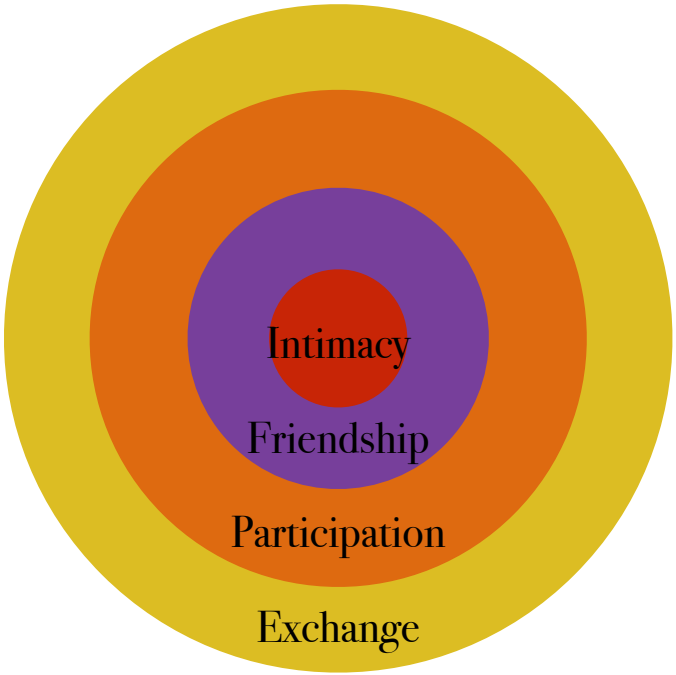


See also weighted and percolation centrality extensions, e.g. Piraveenan et al. 2013

Dunbar's number



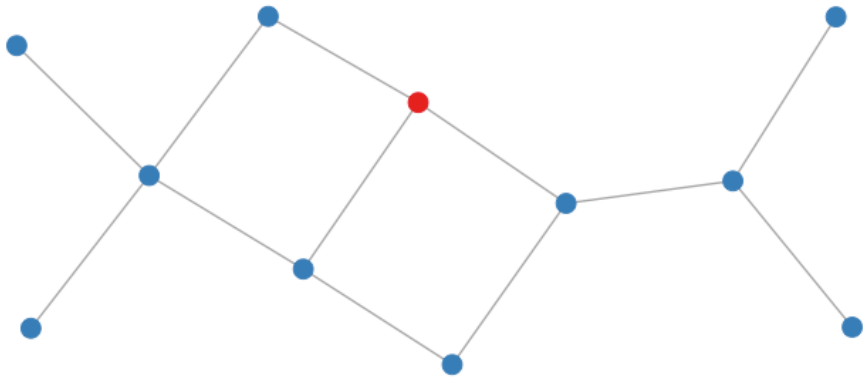
How many people do we need to be close to to be happy?



- Dunbar's number (150) can be more differentiated:
  - 3-5 closest BFFs
  - 9-15 friends and family who are your go-to comrades
  - 30-45 allies and colleagues
  - 90-135 wider acquaintances

Closeness centrality

- AKA *radial centrality*
- Question: how short are the distances to all other nodes?
- Theory: about capacity to e.g. communicate, diffuse, draw on all networked resources
- Method: normalised inverse of the sum over all dyadic distances, *d(i,j)*, i.e. the farness

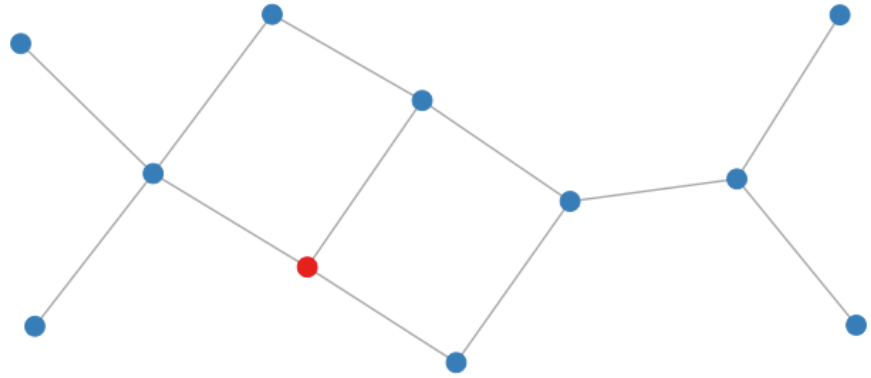


$$c_C(i) = \frac{N-1}{\sum_{j \in V} d(i,j)}$$

Bavelas (1950). See also harmonic centrality (Marchiori and Latora 2000), information centrality (Stephenson and Zelen 1989) and reach centrality extensions

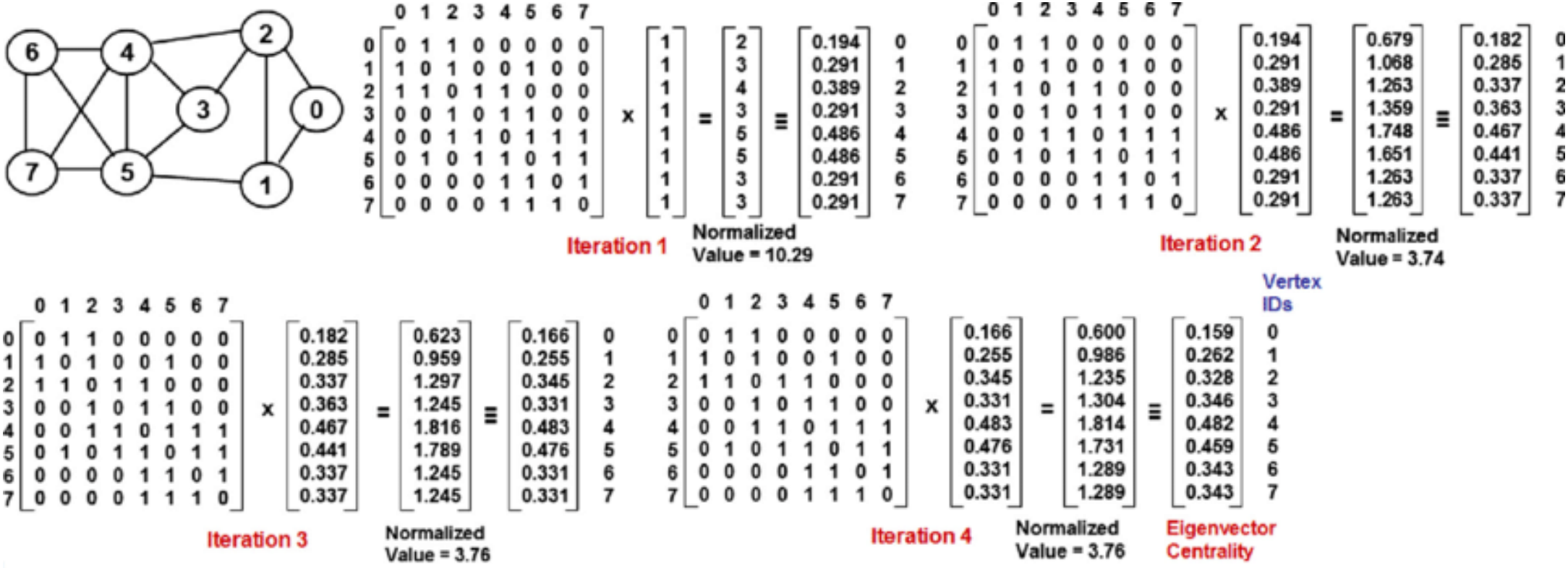
Eigenvector centrality

- AKA *feedback centrality*
- Question: how connected is the node to important (central) neighbours
- Theory: about e.g. power, influence, support
- Method: leading eigenvector of sociomatrix $c_E(i) = \frac{1}{\lambda} \sum_{j \in N(i)} a_{i,j} x_i$
  - With $\delta = 1/\alpha$ being an eigenvalue of the adjacency matrix and $N(i)$ the neighbour nodes of i
  - In undirected networks, $c_E(i)$ with $\alpha = 1$ equals the degree centrality (stationary distribution of a random walk)
  - Perron-Frobenius theorem: non-negative irreducible matrices have leading eigenvalues, leading eigenvectors, and symmetric eigenvalues

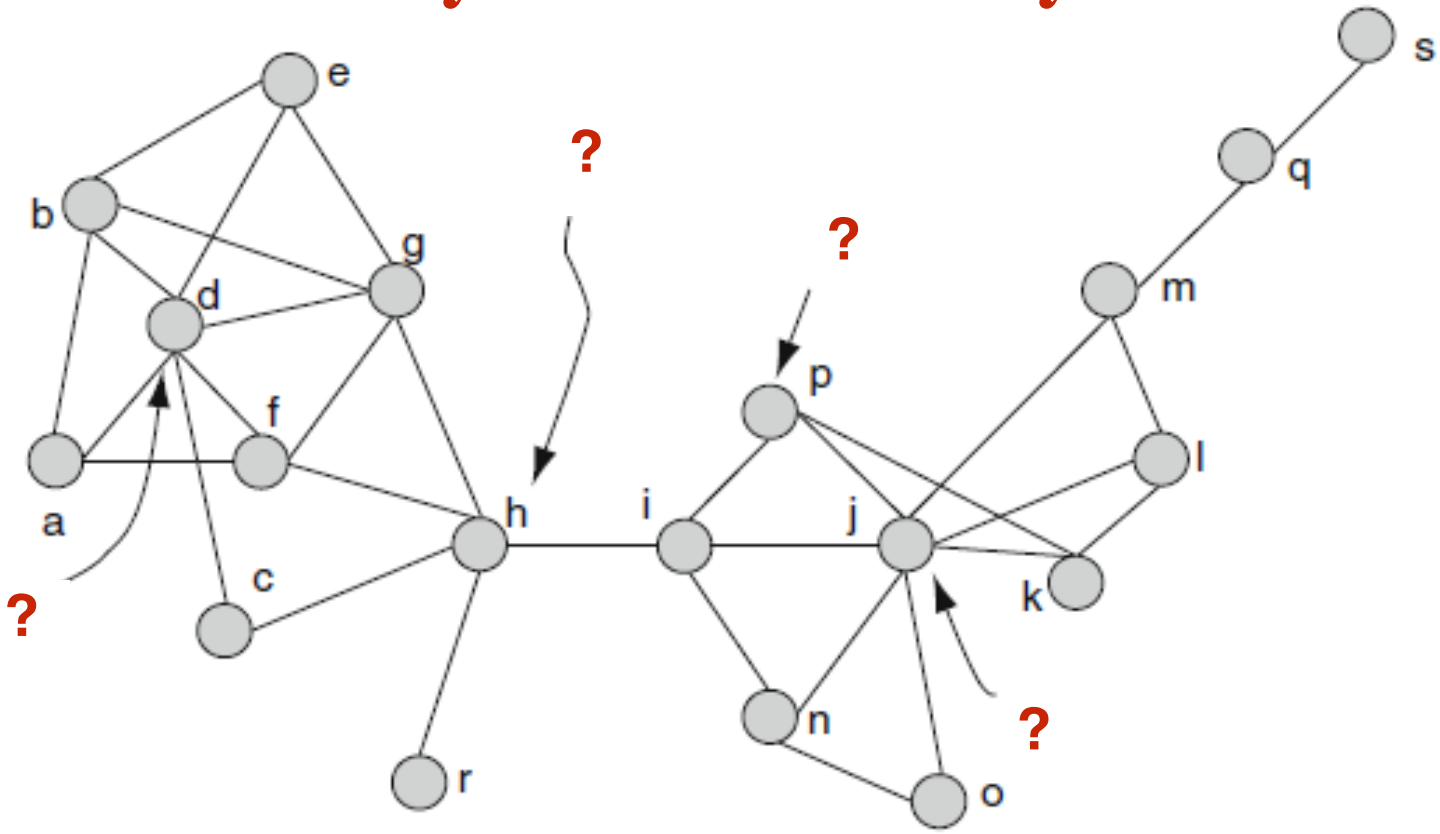


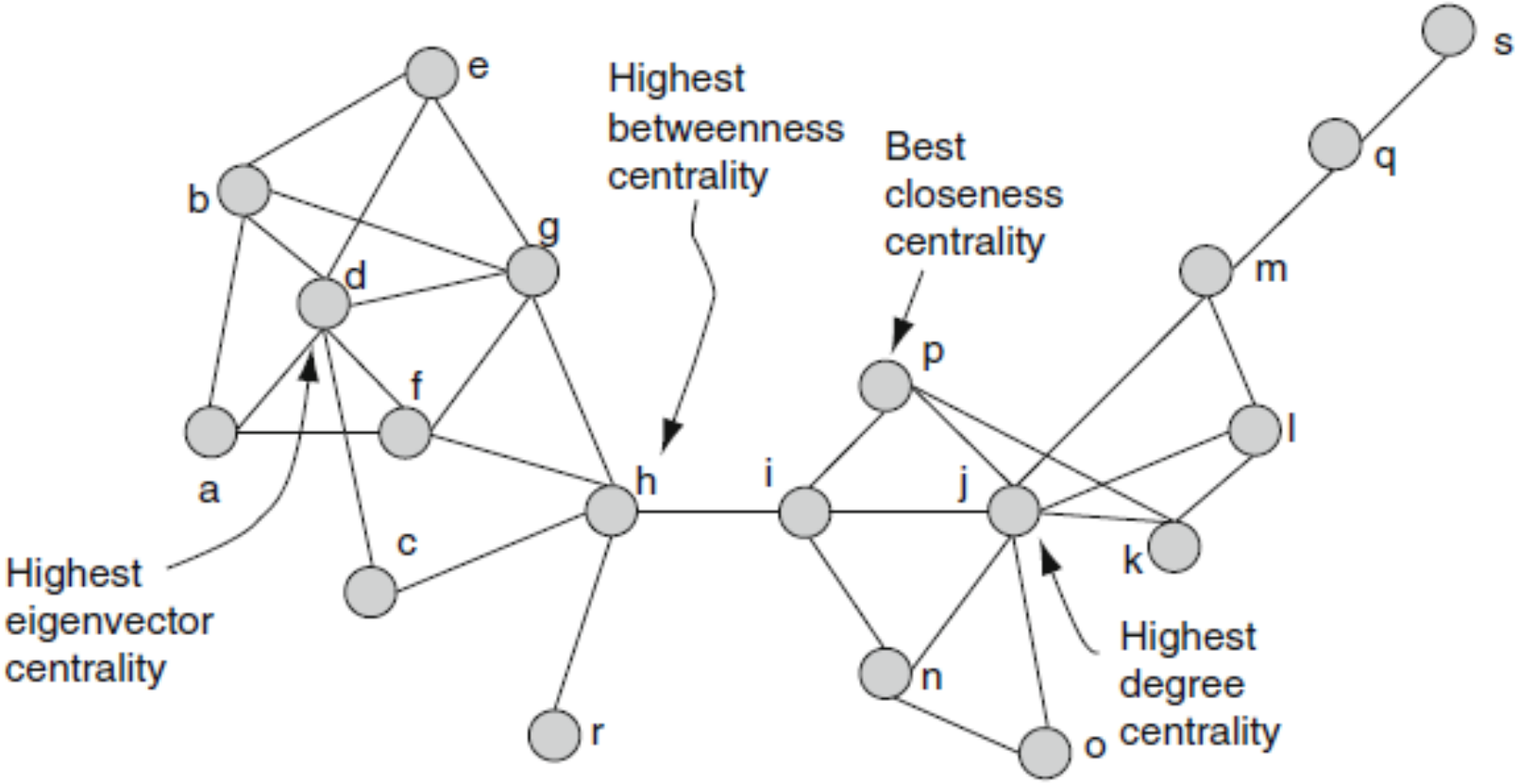
Remarkably old method: Landau 1895, Leontief 1941, Seeley 1949.
See also Bonacich power and alpha centrality,
HITS (Hubs and Authorities), PageRank (Brin and Page 1998) and Katz centrality extensions

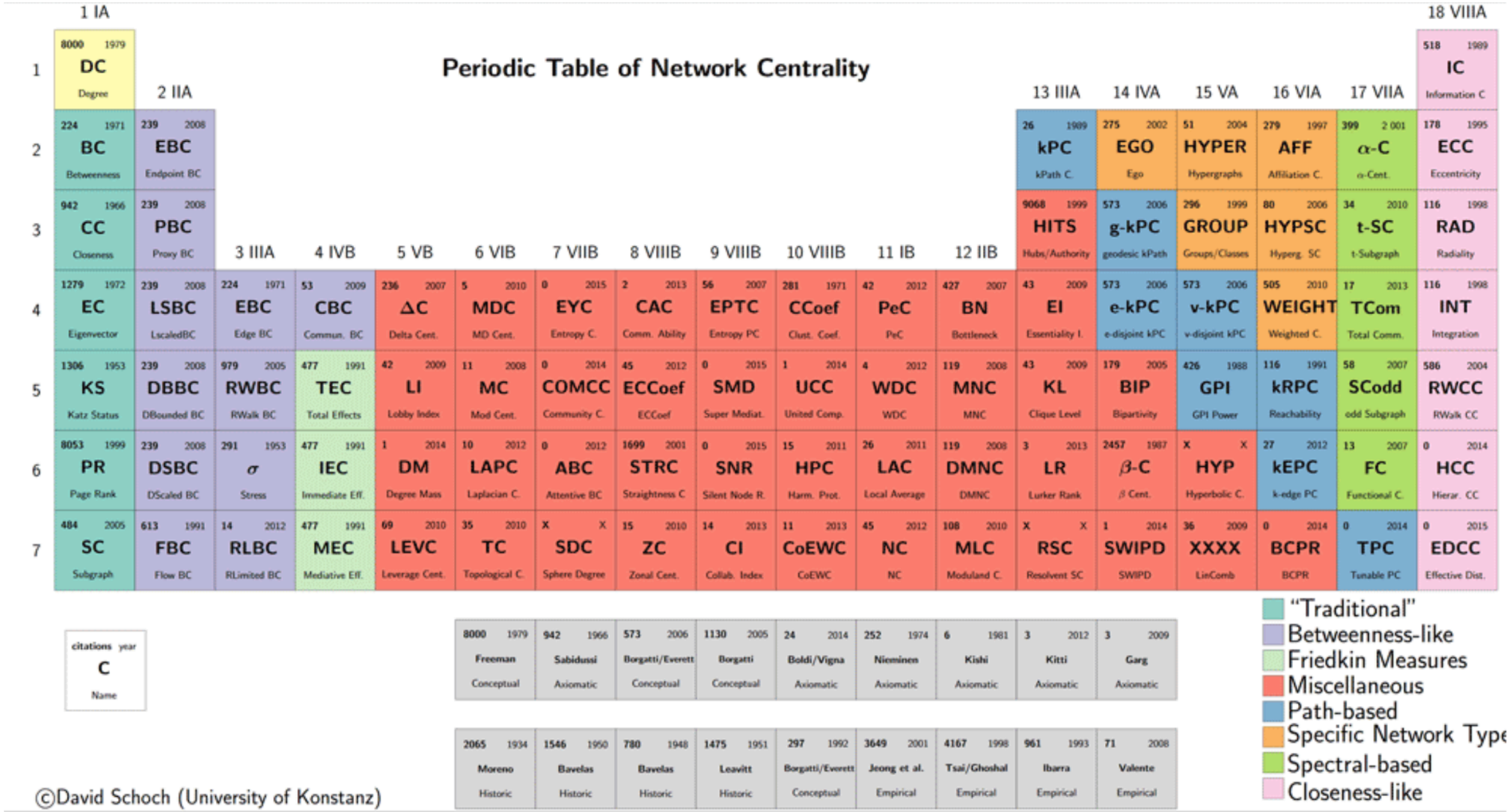
How the sausage is made



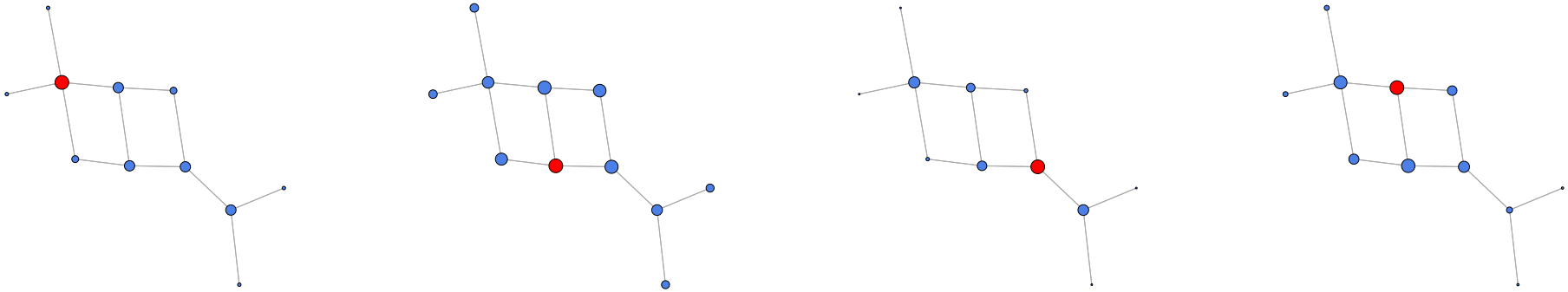
Test your memory




e
b
g
d
f
a
c
h
r
i
p
j
n
o
k
l
m
q
s
Highest
betweenness
centrality
Best
closeness
centrality
Highest
eigenvector
centrality
Highest
degree
centrality

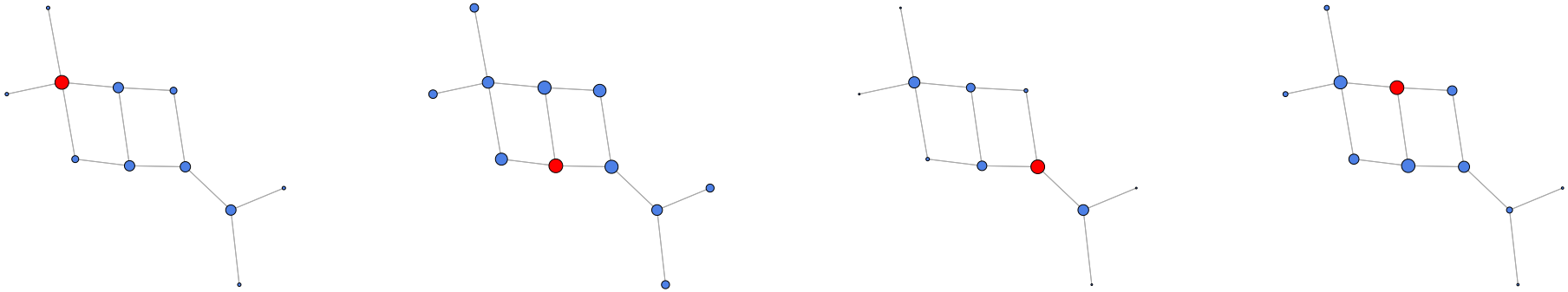

Periodic Table of Network Centrality
1 IA
2 IIA
3 IIIA
4 IVB
5 VB
6 VIB
7 VIIB
8 VIIIB
9 VIIIB
10 VIIIB
11 IB
12 IIB
13 IIIA
14 IVA
15 VA
16 VIA
17 VIIA
18 VIIIA
1
2
3
4
5
6
7
8000 1979 DC Degree
518 1989 IC Information C
224 1971 BC Betweenness
239 2008 EBC Endpoint BC
26 1989 kPC kPath C.
275 2002 EGO Ego
51 2004 HYPER Hypergraphs
279 1997 AFF Affiliation C.
399 2 001 α-C α-Cent.
178 1995 ECC Eccentricity
942 1966 CC Closeness
239 2008 PBC Proxy BC
9068 1999 HITS Hubs/Authority
573 2006 g-kPC geodesic kPath
296 1999 GROUP Groups/Classes
80 2006 HYPSC Hyperg. SC
34 2010 t-SC t-Subgraph
116 1998 RAD Radiality
1279 1972 EC Eigenvector
239 2008 LSBC LscaledBC
224 1971 EBC Edge BC
53 2009 CBC Commun. BC
236 2007 ΔC Delta Cent.
5 2010 MDC MD Cent.
0 2015 EYC Entropy C.
2 2013 CAC Comm. Ability
56 2007 EPTC Entropy PC
281 1971 CCoef Clust. Coef.
42 2012 PeC PeC
427 2007 BN Bottleneck
43 2009 EI Essentiality I.
573 2006 e-kPC e-disjoint kPC
573 2006 v-kPC v-disjoint kPC
505 2010 WEIGHT Weighted C.
17 2013 TCom Total Comm.
116 1998 INT Integration
1306 1953 KS Katz Status
239 2008 DBBC DBounded BC
979 2005 RWBC RWalk BC
477 1991 TEC Total Effects
42 2009 LI Lobby Index
11 2008 MC Mod Cent.
0 2014 COMCC Community C.
45 2012 ECCoef ECCoef
0 2015 SMD Super Mediat.
1 2014 UCC United Comp.
4 2012 WDC WDC
119 2008 MNC MNC
43 2009 KL Clique Level
179 2005 BIP Bipartivity
426 1988 GPI GPI Power
116 1991 kRPC Reachability
58 2007 SCodd odd Subgraph
586 2004 RWCC RWalk CC
8053 1999 PR Page Rank
239 2008 DSBC DScaled BC
291 1953 σ Stress
477 1991 IEC Immediate Eff.
1 2014 DM Degree Mass
10 2012 LAPC Laplacian C.
0 2012 ABC Attentive BC
1699 2001 STRC Straightness C
0 2015 SNR Silent Node R.
15 2011 HPC Harm. Prot.
26 2011 LAC Local Average
119 2008 DMNC DMNC
3 2013 LR Lurker Rank
2457 1987 β-C β Cent.
X X HYP Hyperbolic C.
27 2012 kEPC k-edge PC
13 2007 FC Functional C.
0 2014 HCC Hierar. CC
484 2005 SC Subgraph
613 1991 FBC Flow BC
14 2012 RLBC RLimited BC
477 1991 MEC Mediative Eff.
69 2010 LEVC Leverage Cent.
35 2010 TC Topological C.
X X SDC Sphere Degree
15 2010 ZC Zonal Cent.
14 2013 CI Collab. Index
11 2013 CoEWC CoEWC
45 2012 NC NC
108 2010 MLC Moduland C.
X X RSC Resolvent SC
1 2014 SWIPD SWIPD
36 2009 XXXX LinComb
0 2014 BCPR BCPR
0 2014 TPC Tunable PC
0 2015 EDCC Effective Dist.
citations year
C
Name
8000 1979 Freeman Conceptual
942 1966 Sabidussi Axiomatic
573 2006 Borgatti/Everett Conceptual
1130 2005 Borgatti Conceptual
24 2014 Boldi/Vigna Axiomatic
252 1974 Nieminen Axiomatic
6 1981 Kishi Axiomatic
3 2012 Kitti Axiomatic
3 2009 Garg Axiomatic
2065 1934 Moreno Historic
1546 1950 Bavelas Historic
780 1948 Bavelas Historic
1475 1951 Leavitt Historic
297 1992 Borgatti/Everett Conceptual
3649 2001 Jeong et al. Empirical
4167 1998 Tsai/Ghoshal Empirical
961 1993 Ibarra Empirical
71 2008 Valente Empirical
©David Schoch (University of Konstanz)
"Traditional"
Betweenness-like
Friedkin Measures
Miscellaneous
Path-based
Specific Network Type
Spectral-based
Closeness-like

Which centrality should I use?



Option 1: Just try all and see?

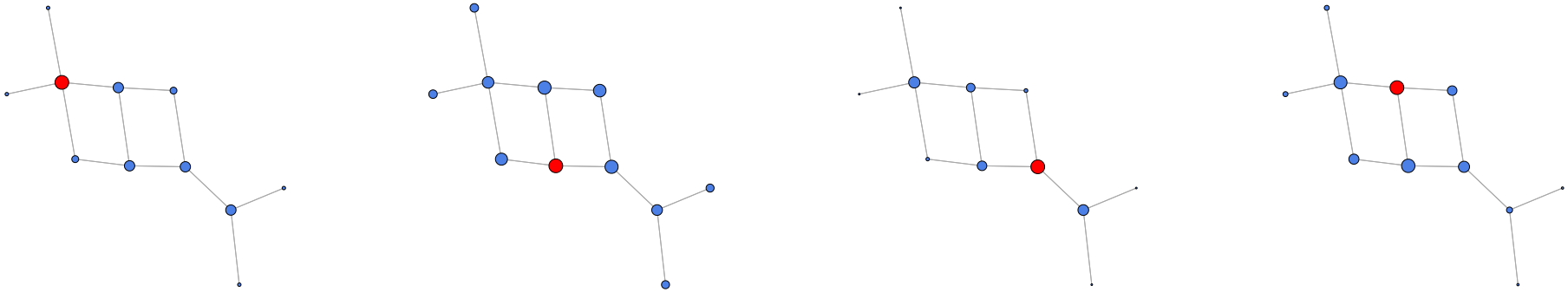
Which centrality should I use?



Option 1: Just try all and see?

Option 2: Create index of all?

Which centrality should I use?



Option 1: Just try all and see?

Option 2: Create index of all?

NO, option 3: derive centrality measure based on theoretical assumptions (see also Zweig 2016)

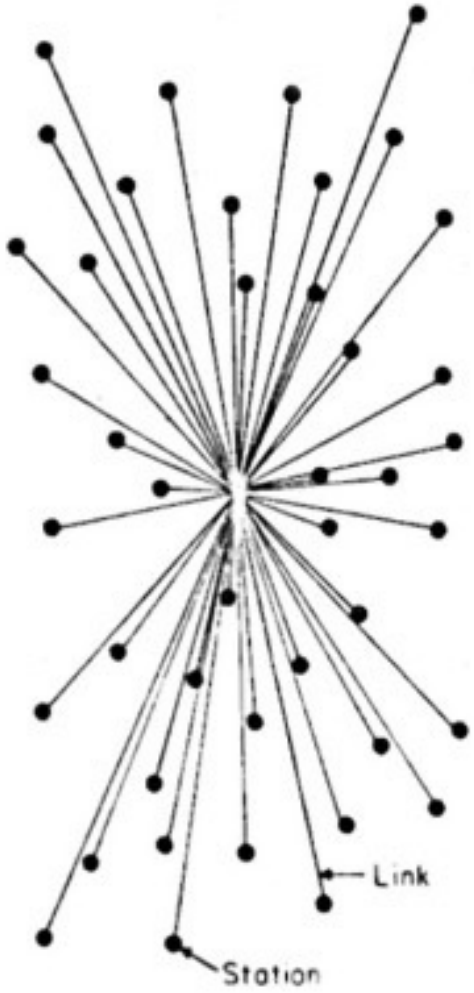
Depends on research question and meaning of the edges:

- Support?
- Flow?
- Power?
- Access?

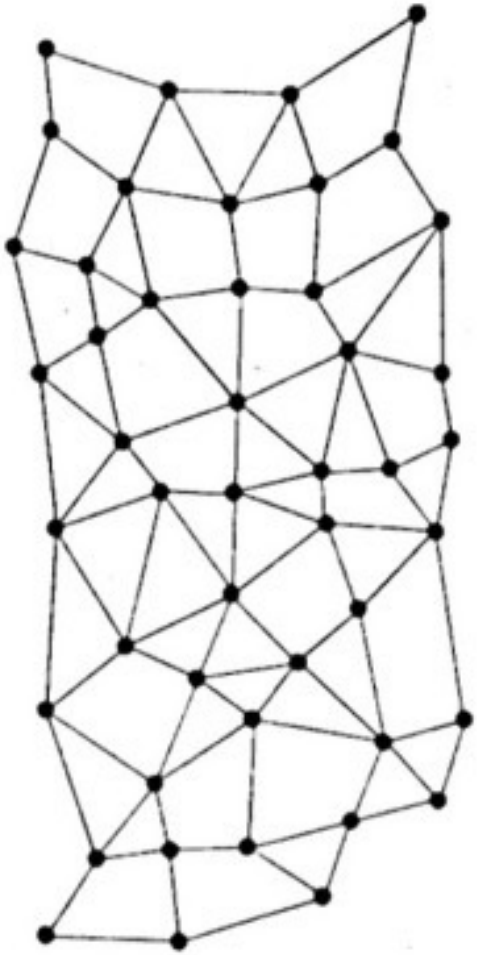
How should I use centrality?

- Ultimately ordinal, not interval, scale: there is no “unit of centrality”
  - Makes comparing across measures difficult
  - Makes comparing across networks difficult (even with normalisation)
- To identify nodes with the most (usually) or least central positions
  - Inequality of centrality thought to correlate with some other inequality among nodes
  - But you can also use differences in centrality to characterise the overall centralization of the network…

Lesson #2:
Choose centrality carefully



CENTRALIZED
(A)



DISTRIBUTED
(C).

Centralisation

Question: how central is a network's most central node compared to all others?

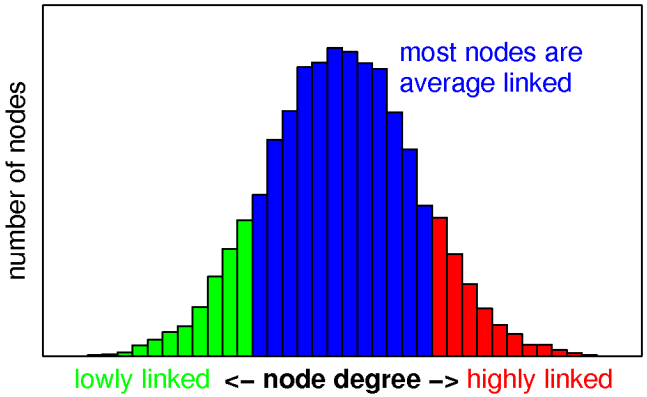
Theory: about inequality, coordination, efficiency

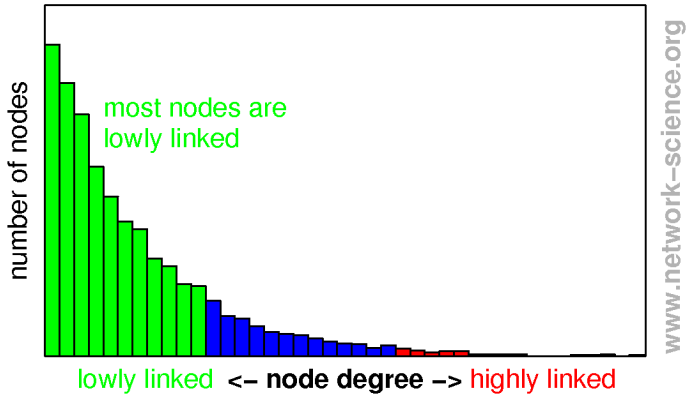
Method: every centrality measure has its own centralisation measure:

$$\max \sum_{i=1}^{N} C_x(p_*) - C_x(p_i)$$

1. obtain largest centrality
2. calculate sum of differences between each node and all other nodes
3. take the maximum of this
4. divide by theoretically largest such sum of differences in any network of the same size

Degree Distributions





www.network-science.org

Poisson:

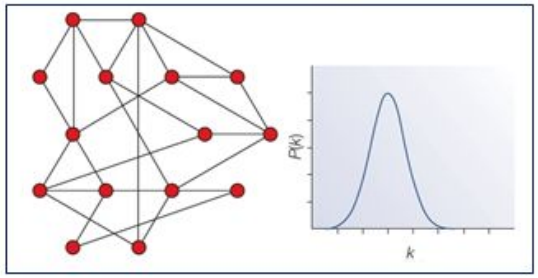
$$P(k) = \frac{e^{-d} d^{k}}{k!}$$

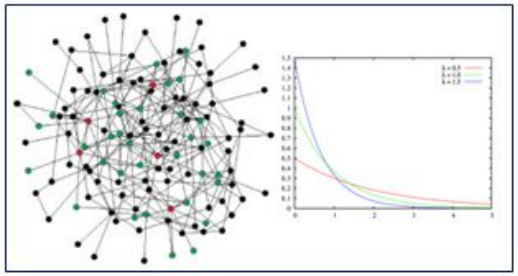
Exponential:

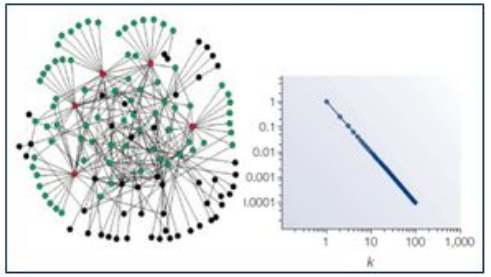
$$P(k) \propto e^{-k/d}$$

Power-law:

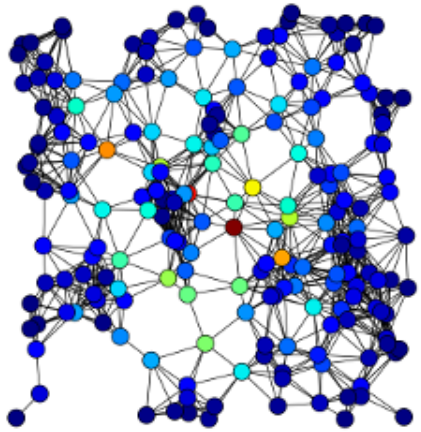
$$P(k) \propto k^{-c}, k \neq 0, c > 1$$



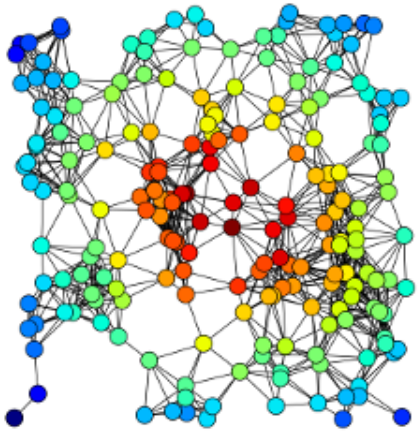




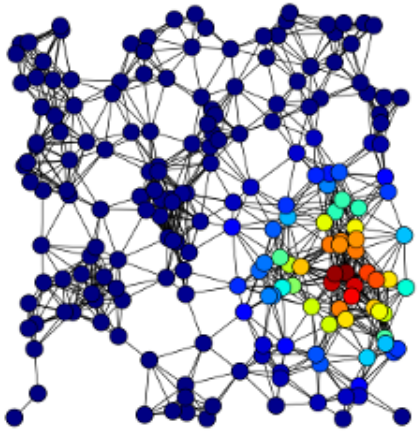
Jennings 1932



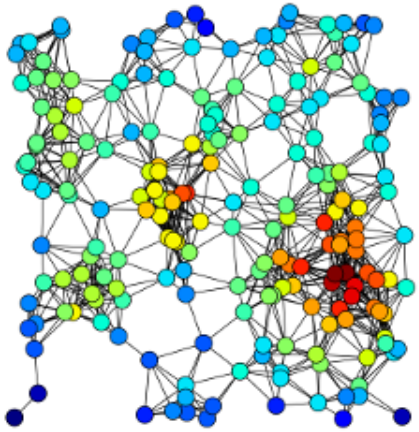
A



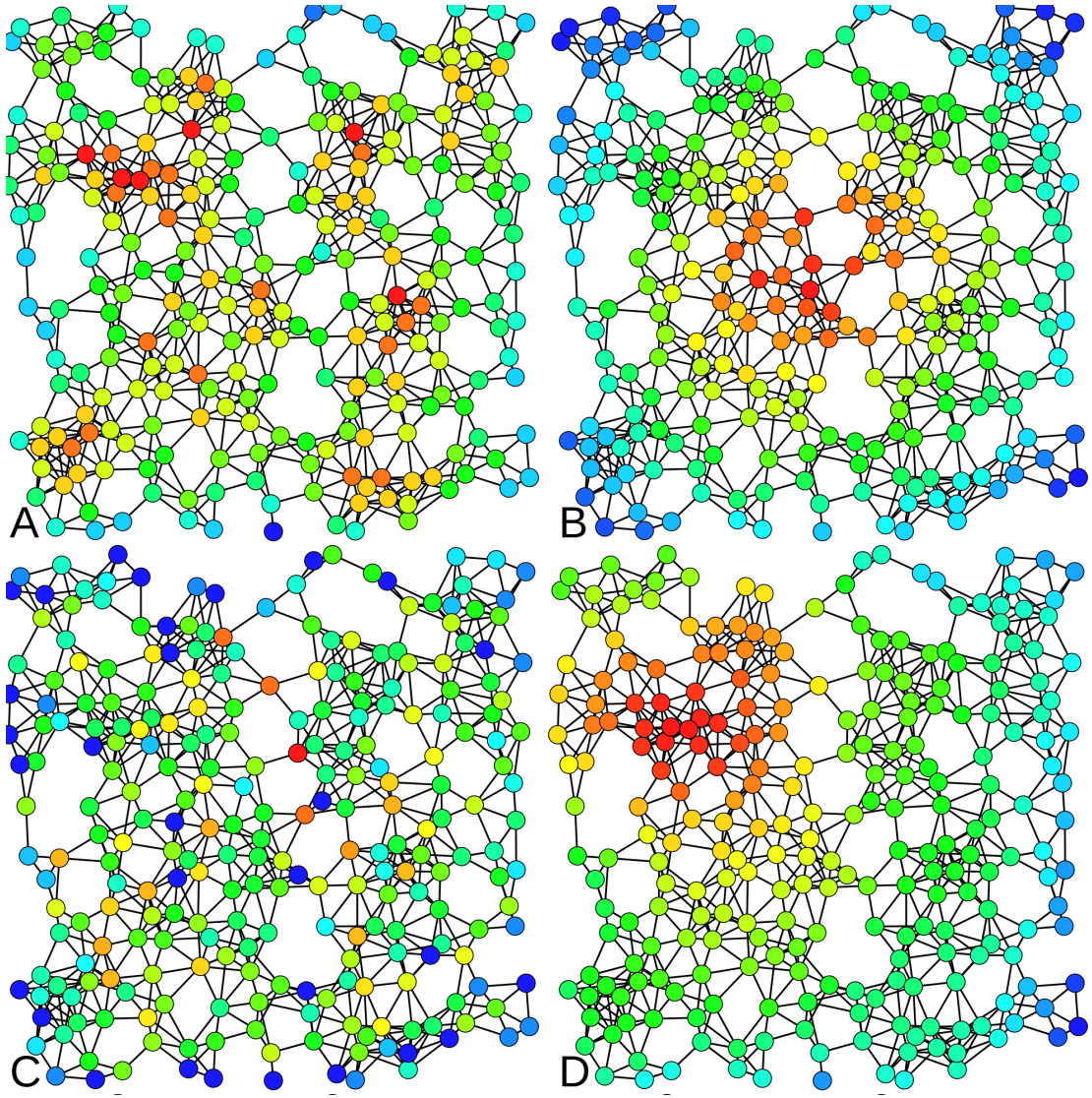
B



C



D


A
B
C
D