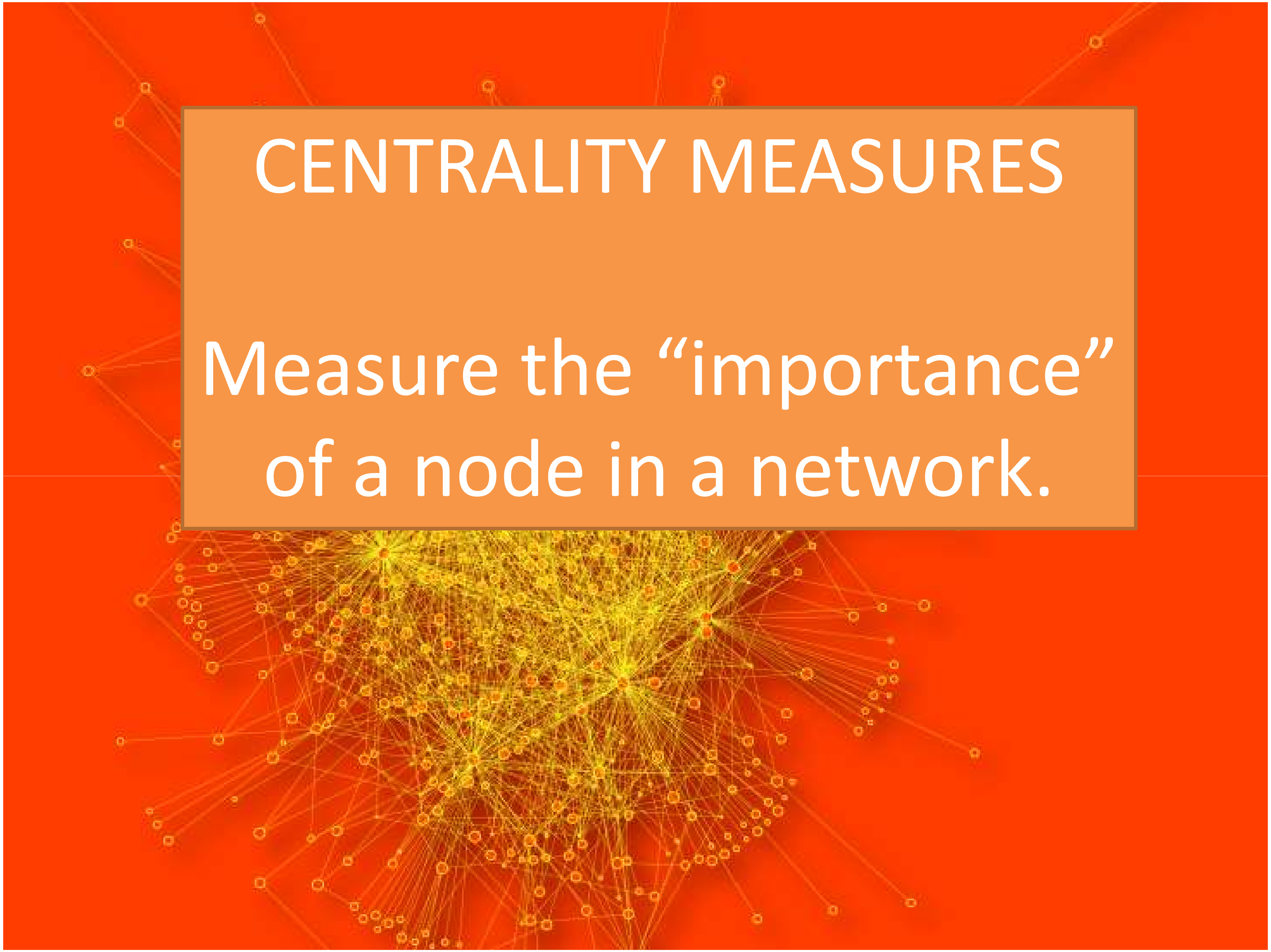
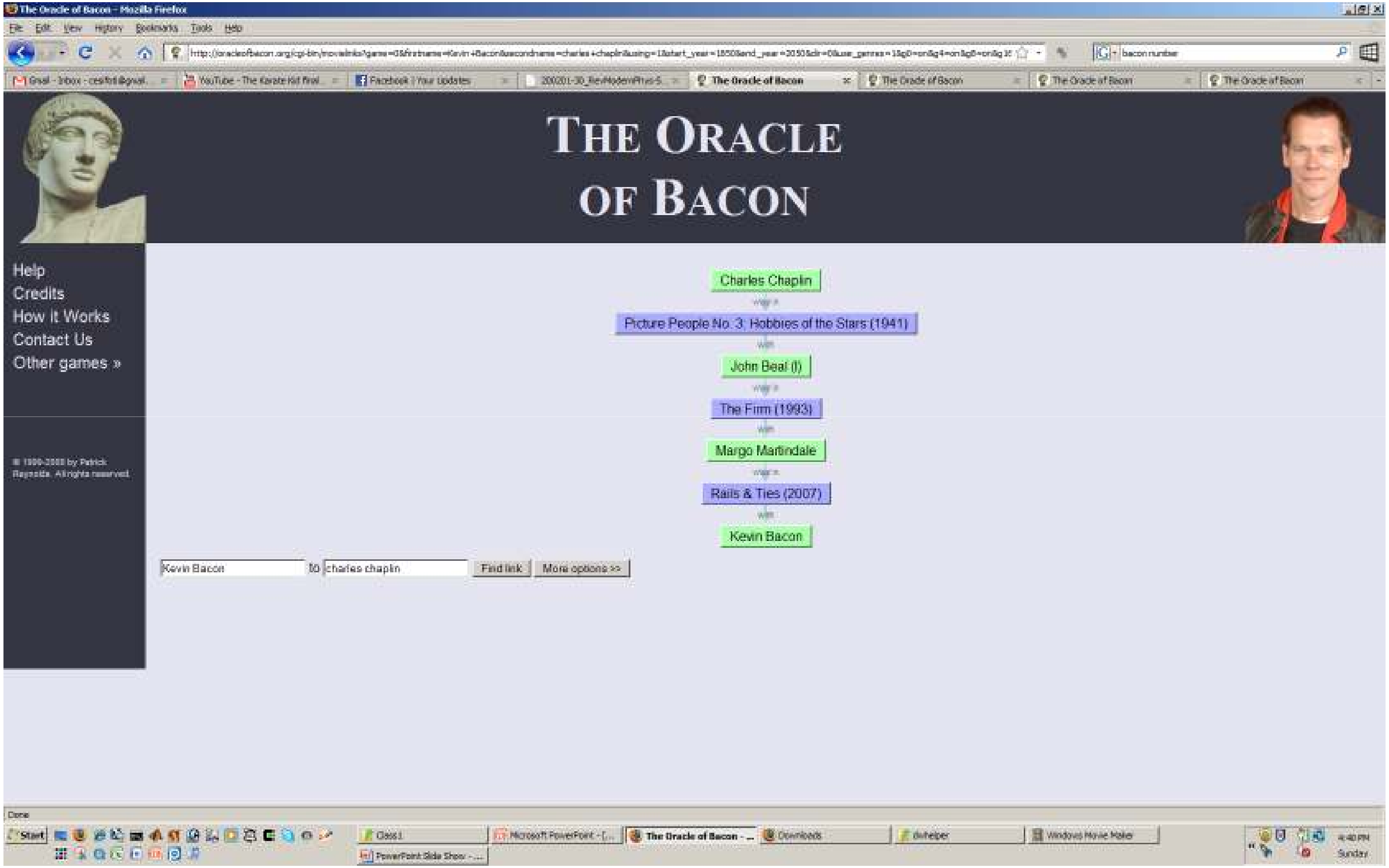
CENTRALITY MEASURES

Measure the “importance” of a node in a network.




The Oracle of Bacon - Mozilla Firefox
THE ORACLE
OF BACON
Help
Credits
How it Works
Contact Us
Other games »
Charles Chaplin
was in
Picture People No. 3: Hobbies of the Stars (1941)
with
John Beal (I)
was in
The Firm (1993)
with
Margo Martindale
was in
Rails & Ties (2007)
with
Kevin Bacon
Kevin Bacon
to
charles chaplin
Find link
More options >>

Hollywood Revolves Around

Click on a name to see that person's table.

Steiger, Rod (2.678695)
Lee, Christopher (I) (2.684104)
Hopper, Dennis (2.698471)
Sutherland, Donald (I) (2.701850)
Keitel, Harvey (2.705573)
Pleasence, Donald (2.707490)
von Sydow, Max (2.708420)
Caine, Michael (I) (2.720621)
Sheen, Martin (2.721361)
Quinn, Anthony (2.722720)
Heston, Charlton (2.722904)
Hackman, Gene (2.725215)
Connery, Sean (2.730801)
Stanton, Harry Dean (2.737575)
Welles, Orson (2.744593)
Mitchum, Robert (2.745206)
Gould, Elliott (2.746082)
Plummer, Christopher (I) (2.746427)
Coburn, James (2.746822)
Borgnine, Ernest (2.747229)

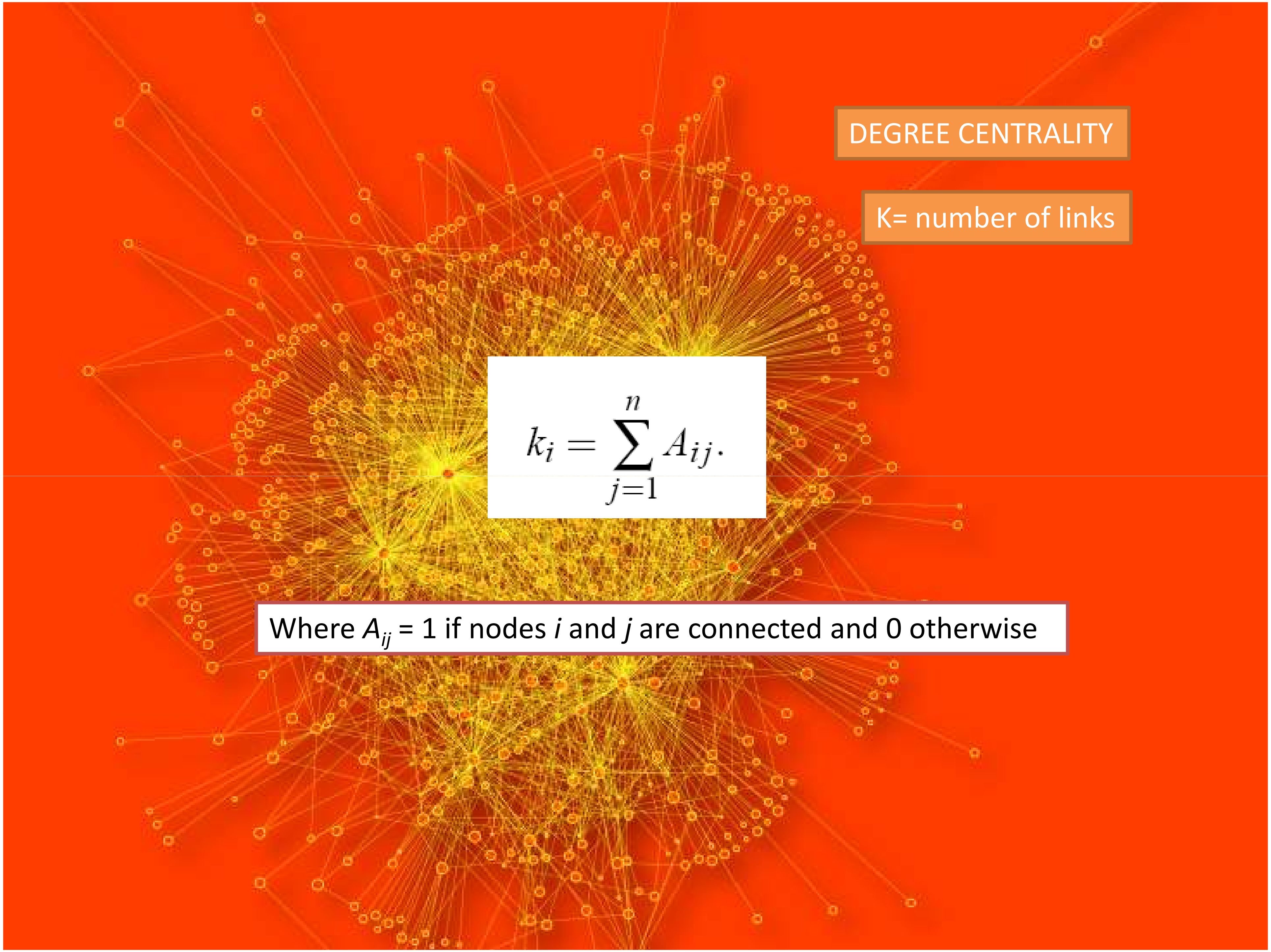
Most Connected Actors in Hollywood

(measured in the late 90's)

Mel Blanc	759
Tom Byron	679
Marc Wallice	535
Ron Jeremy	500
Peter North	491
TT Boy	449
Tom London	436
Randy West	425
Mike Horner	418
Joey Silvera	410

XXX

A-L Barabasi, "Linked", 2002



DEGREE CENTRALITY

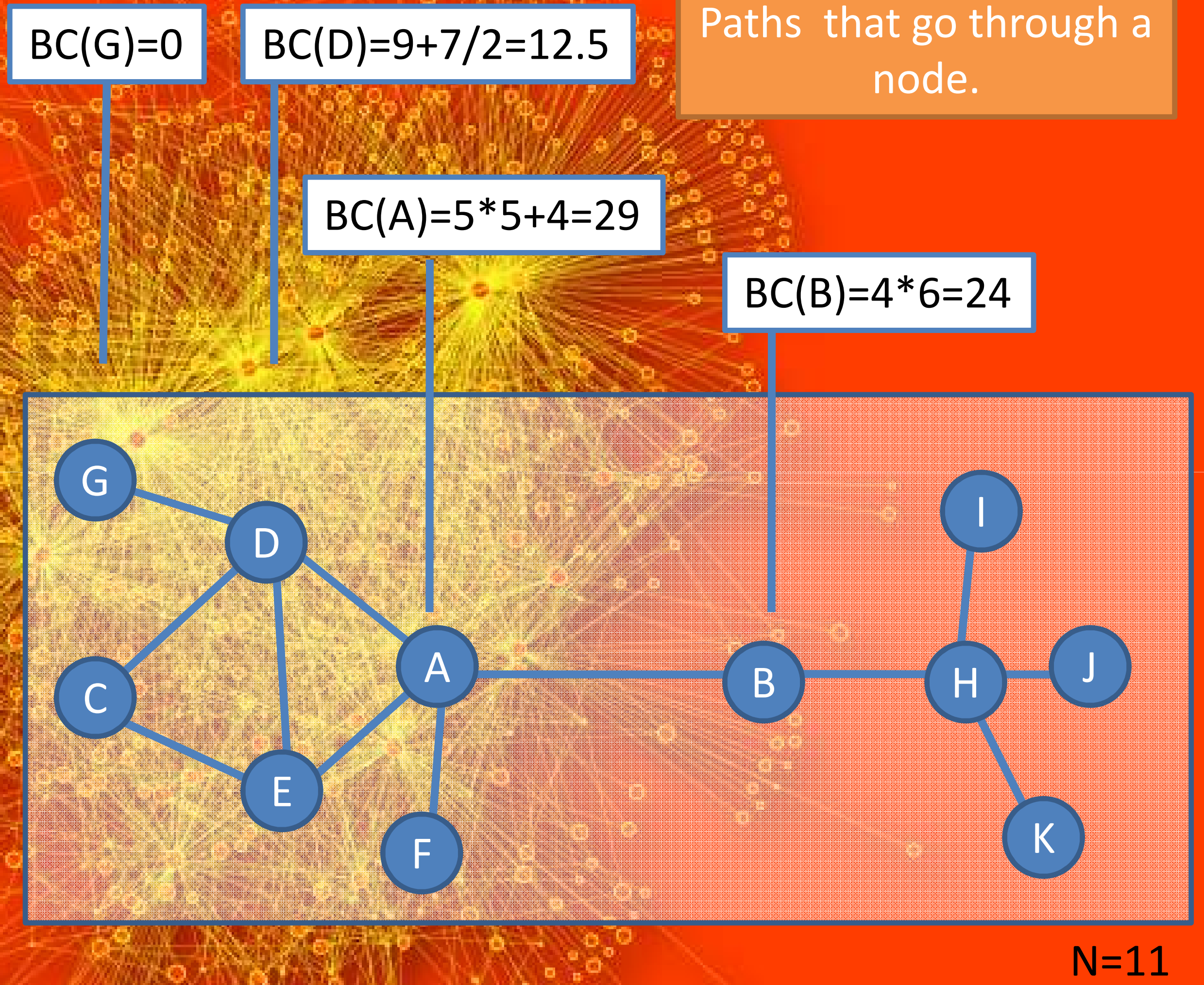
K= number of links

$$k_i = \sum_{j=1}^{n} A_{ij}.$$

Where A_{ij} = 1 if nodes *i* and *j* are connected and 0 otherwise

BETWENNESS CENTRALITY

BC= number of shortest Paths that go through a node.



A set of measures of centrality based on betweenness

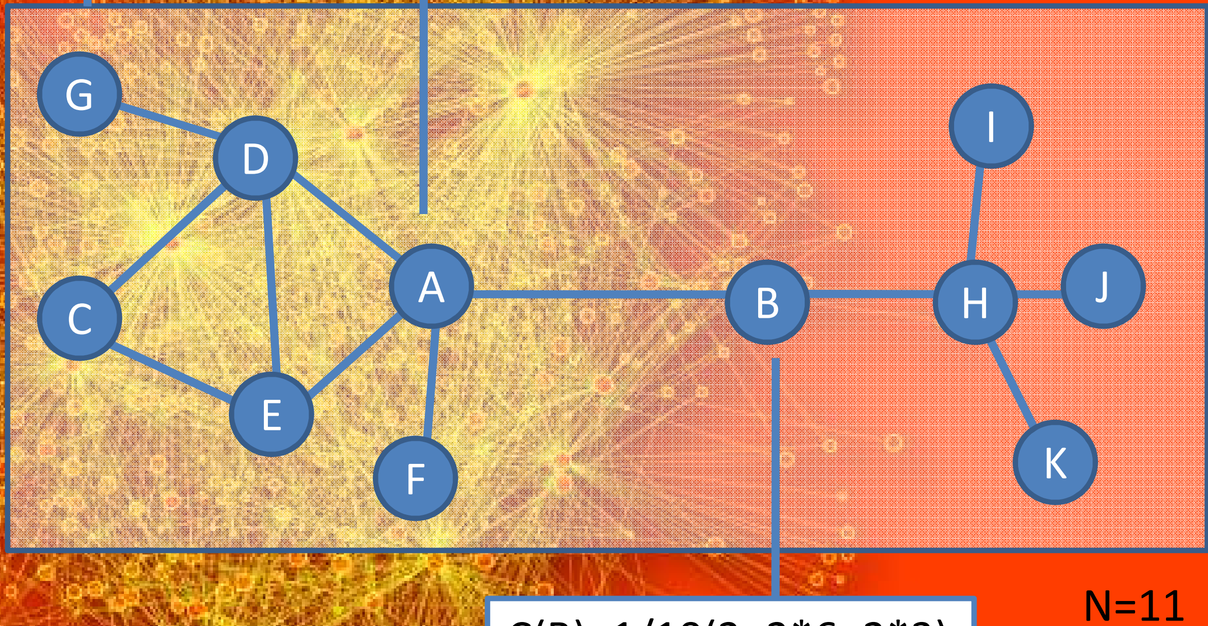
LC **Freeman** - Sociometry, 1977 - jstor.org

CLOSENESS CENTRALITY

C= Average Distance to neighbors

C(G)=1/10(1+2*3+2*3+4+3*5)
C(G)=3.2

C(A)=1/10(4+2*3+3*3)
C(A)=1.9



C(B)=1/10(2+2*6+2*3)
C(B)=2

EIGENVECTOR CENTRALITY

Consider the Adjacency Matrix A_{ij} = 1 if node *i* is connected to node *j* and 0 otherwise. Now, measure the centrality of a node, as the sum over the centralities of all nodes....

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j = \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j} x_j$$

This is equivalent to eigenvalue problem:

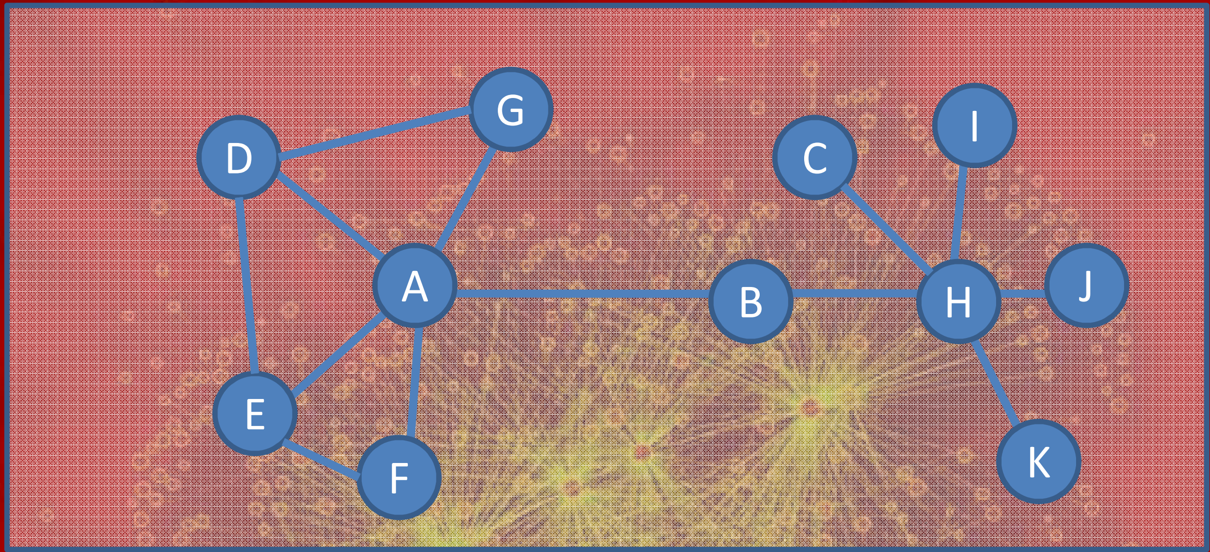
$$A\mathbf{x} = \lambda \mathbf{x}$$

Then the eigenvector centrality of node (i) is defined as:

$$x_i$$

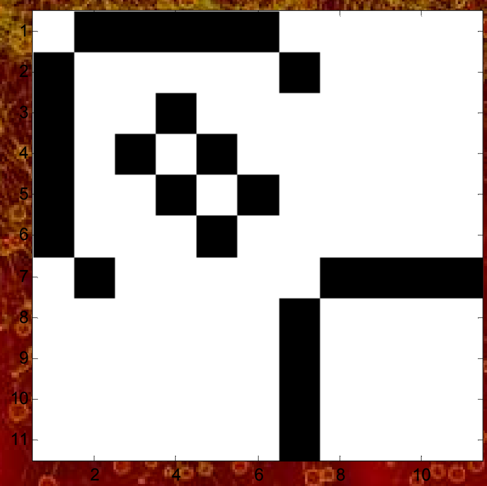
where λ is the largest eigenvalue associated with **A** and x is its associated eigenvector.

EIGENVECTOR CENTRALITY



Adjacency Matrix

-0.5727	A
-0.2289	B
-0.3286	G
-0.4357	D
-0.4357	E
-0.3286	F
-0.1297	H
-0.0423	C
-0.0423	I
-0.0423	J
-0.0423	K

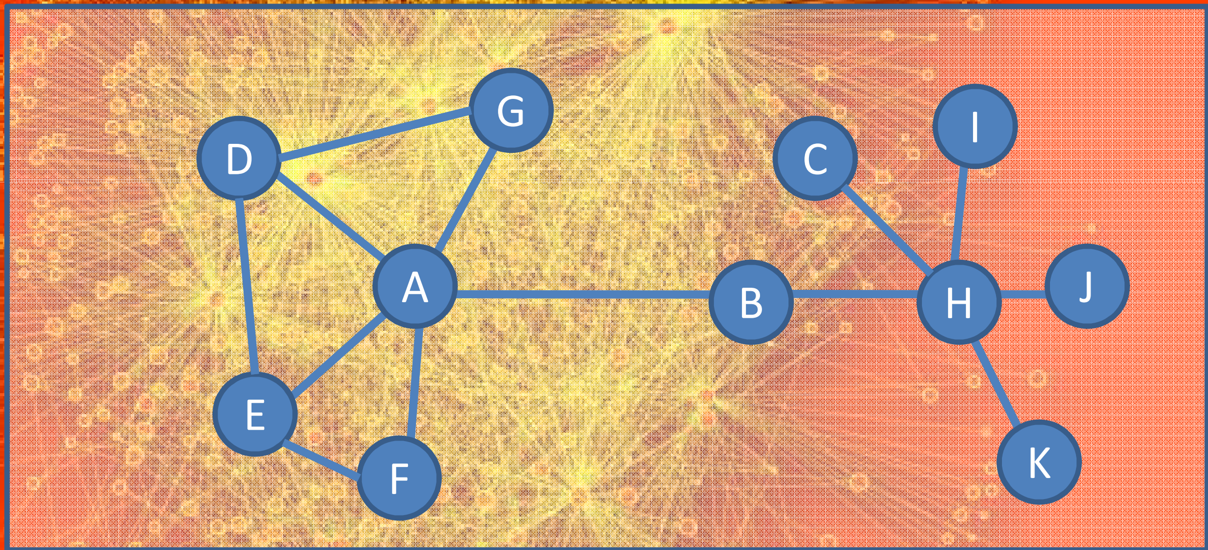


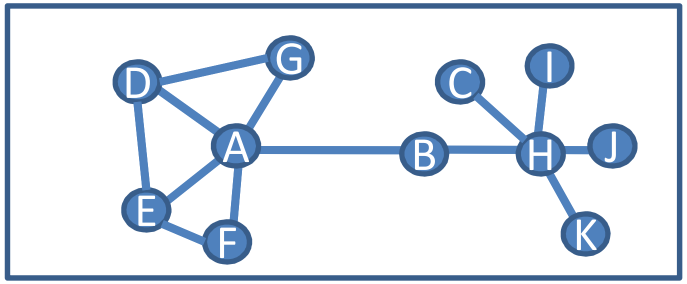
Power and Centrality: A Family of Measures

Phillip Bonacich

The American Journal of Sociology, Vol. 92, No. 5 (Mar., 1987), 1170-1182.

$$c(\alpha, \beta) = \alpha \sum_{k=0}^{\infty} \beta^k R^{k+1} 1 = \alpha(R1 + \beta R^2 1 + \beta^2 R^3 1 + \ldots). \quad (5)$$





$$c(\alpha, \beta) = \alpha \sum_{k=0}^{\infty} \beta^k R^{k+1} 1 = \alpha(R1 + \beta R^2 1 + \beta^2 R^3 1 + \ldots). \quad (5)$$

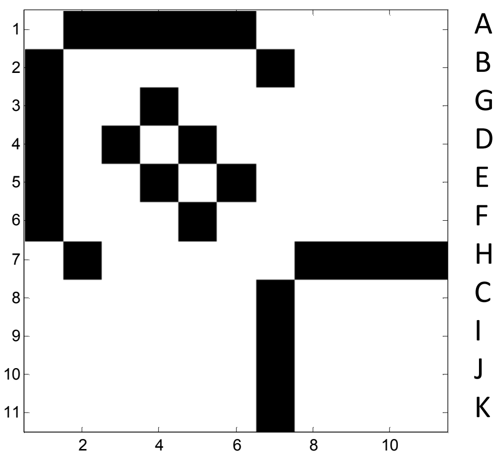
degree	# of nodes two hops away	#of nodes Three hops away
A=5	A=12	A=46
B=2	B=10	B=18
H=5	H=6	H=30
K=1	K=5	K=6
E=3	E=10	E=30

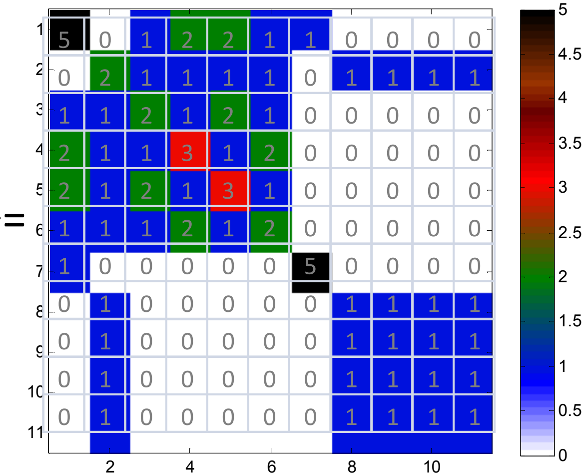
β=0.5 (after 3 iterations)

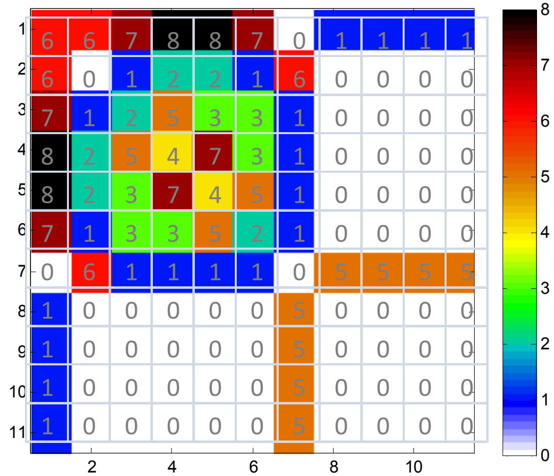
A-> 22.5 K-> 5
B-> 11.5 E->15.5
H-> 15.5

β=-0.5 (after 3 iterations)

A-> 10.5 K->0
B-> 1.5 E->5.5
H-> 9.5

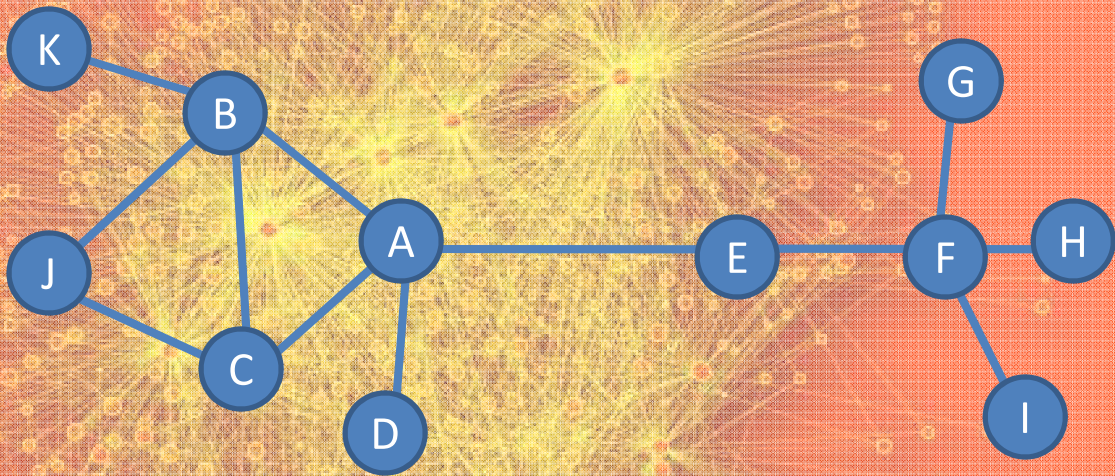
R=


R^2=


R^3=


PAGE RANK

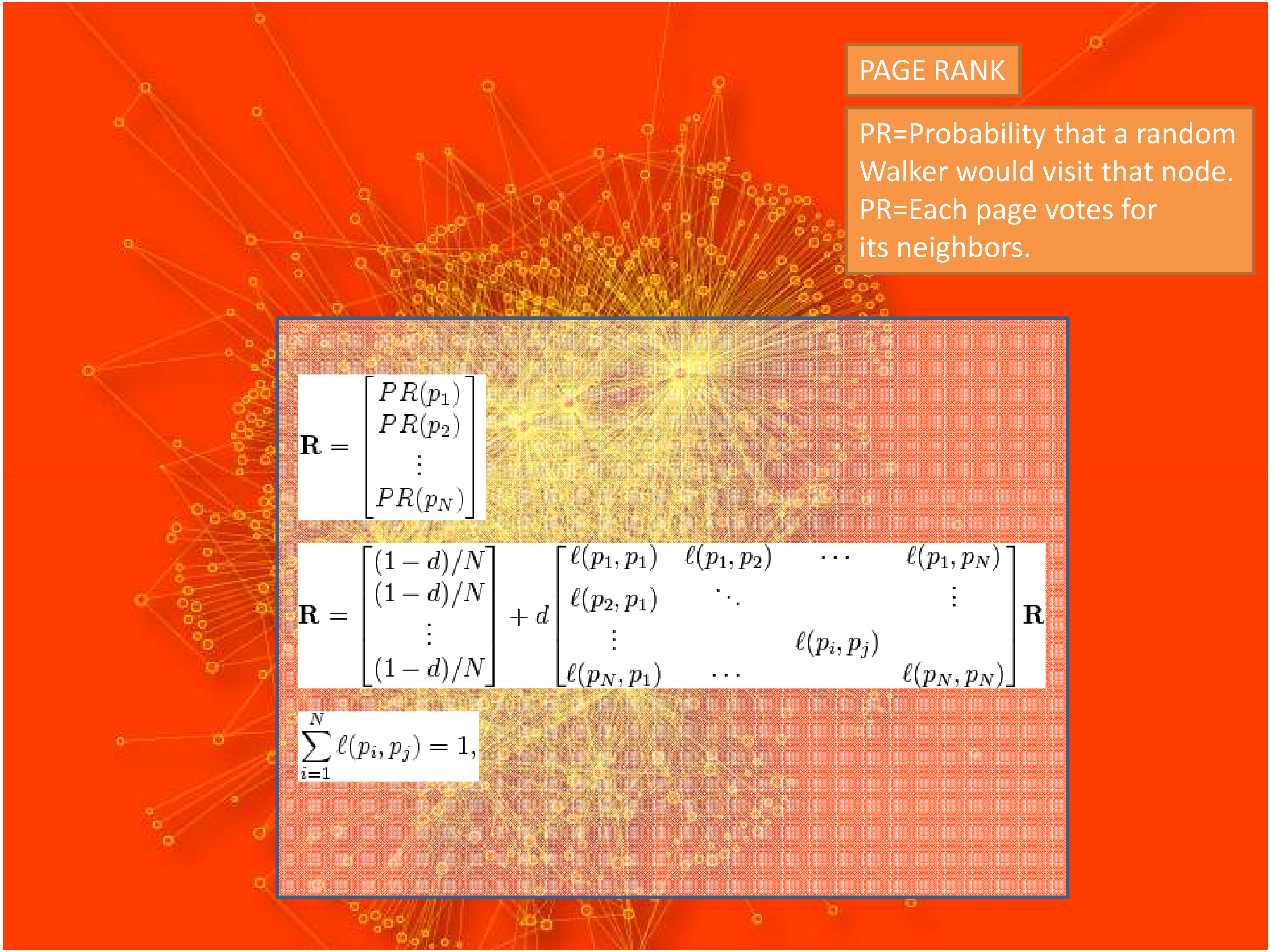
PR=Probability that a random walker with interspersed Jumps would visit that node.
PR=Each page votes for its neighbors.



PR(A)=PR(B)/4 + PR(C)/3 + PR(D)+PR(E)/2

A random surfer eventually stops clicking

$PR(X)=(1-d)/N + d(\sum PR(y)/k(y))$



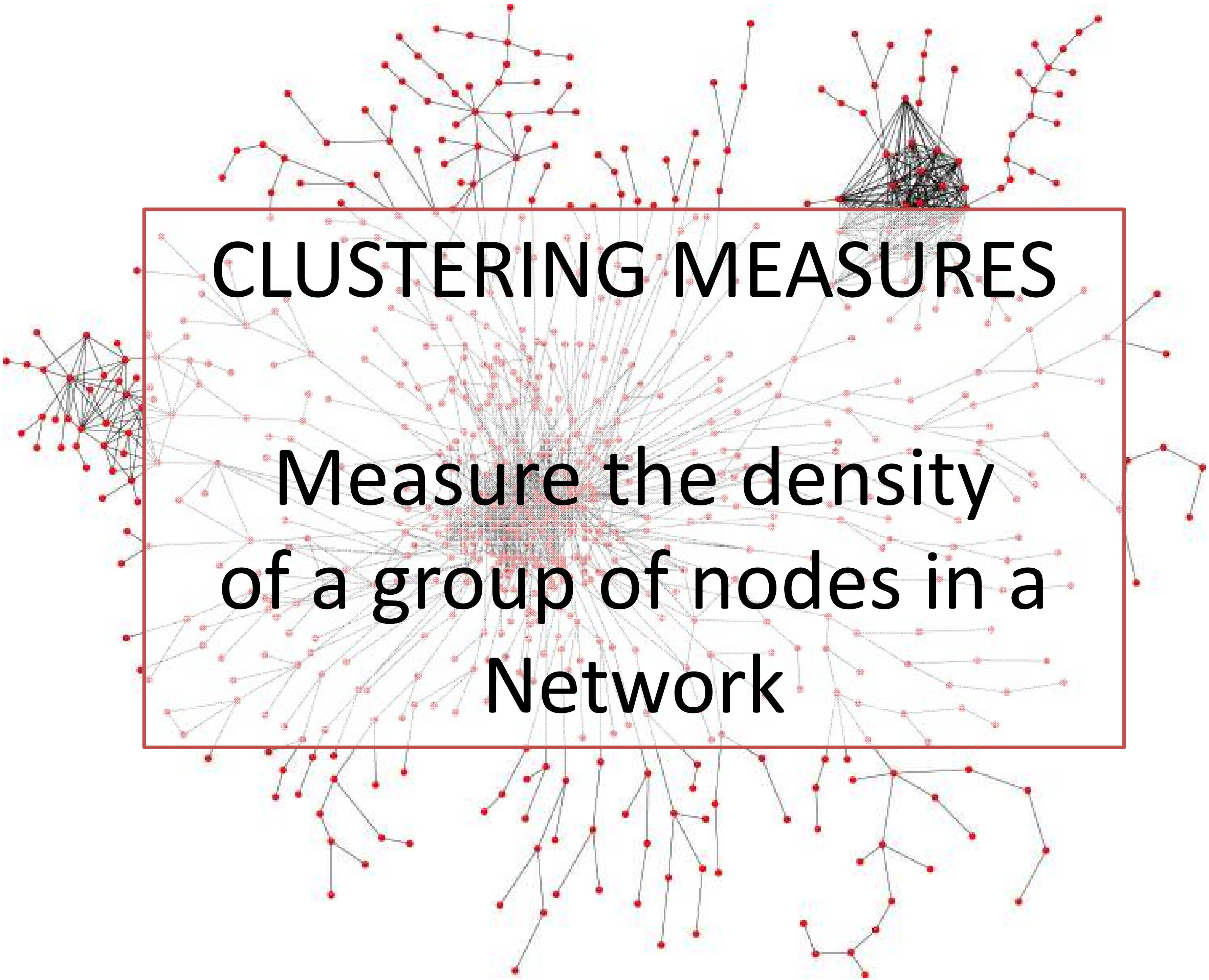
PAGE RANK

PR=Probability that a random Walker would visit that node.
PR=Each page votes for its neighbors.

$$\mathbf{R} = \begin{bmatrix} PR(p_1) \\ PR(p_2) \\ \vdots \\ PR(p_N) \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} (1-d)/N \\ (1-d)/N \\ \vdots \\ (1-d)/N \end{bmatrix} + d \begin{bmatrix} \ell(p_1,p_1) & \ell(p_1,p_2) & \cdots & \ell(p_1,p_N) \\ \ell(p_2,p_1) & \ddots & & \vdots \\ \vdots & & \ell(p_i,p_j) & \\ \ell(p_N,p_1) & \cdots & & \ell(p_N,p_N) \end{bmatrix} \mathbf{R}$$

$$\sum_{i=1}^{N} \ell(p_i,p_j) = 1,$$



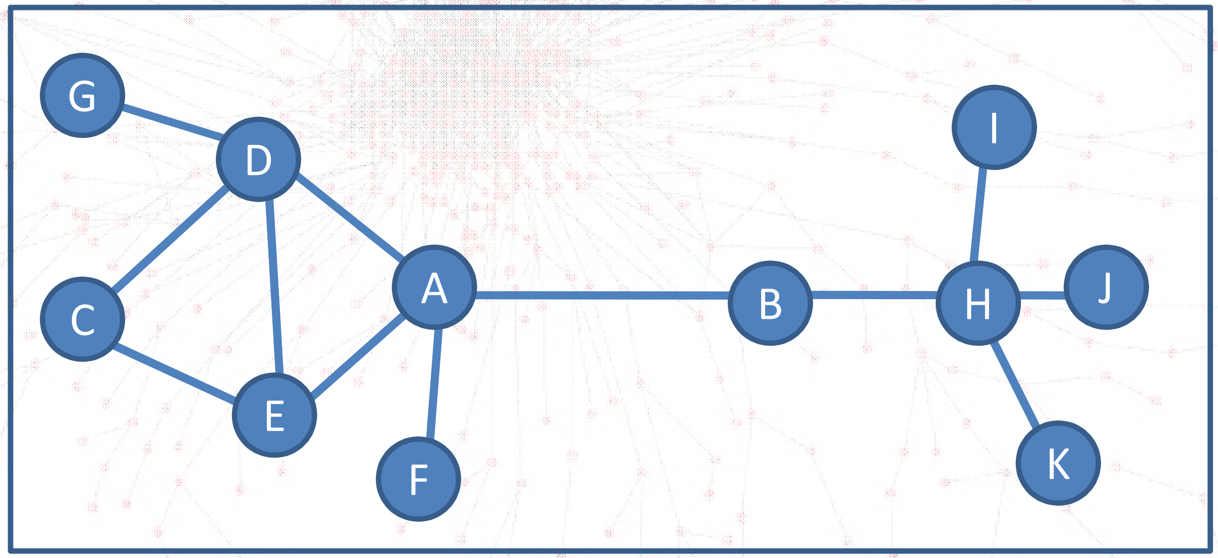
CLUSTERING MEASURES

Measure the density of a group of nodes in a Network

Clustering Coefficient, Transitivity

$$C_i = 2\Delta / k(k-1)$$

$C_A = 2/12 = 1/6$ $\quad$ $C_C = 2/2 = 1$ $\quad$ $C_E = 4/6 = 2/3$



Topological Overlap
Mutual Clustering

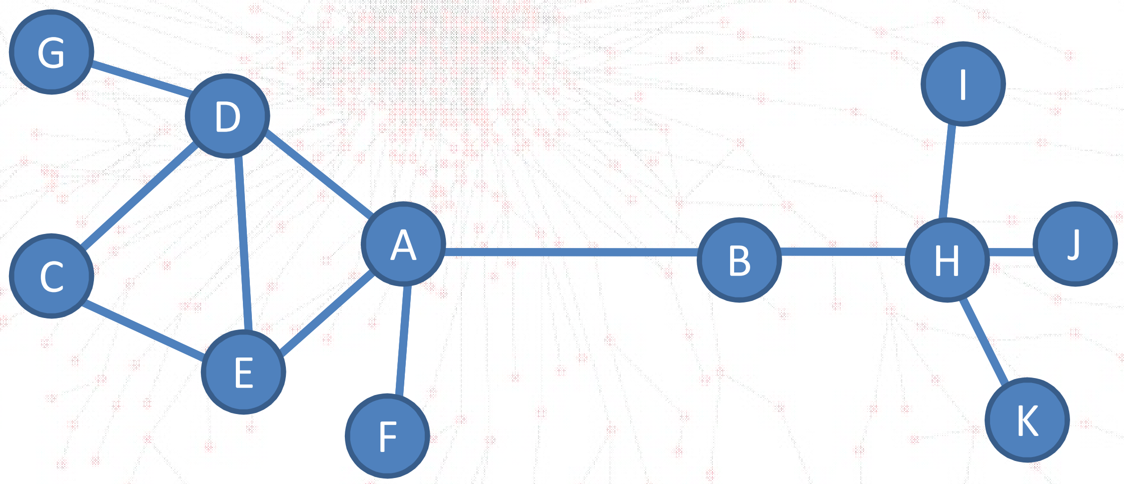
TO(A,B)=Overlap(A,B)/NormalizingFactor(A,B)

$TO(A,B)=N(A,B)/\max(k(A),k(B))$

$TO(A,B)=N(A,B)/\min(k(A),k(B))$

$TO(A,B)=N(A,B)/(k(A) \times k(B))^{1/2}$

$TO(A,B)=N(A,B)/(k(A)+k(B))$



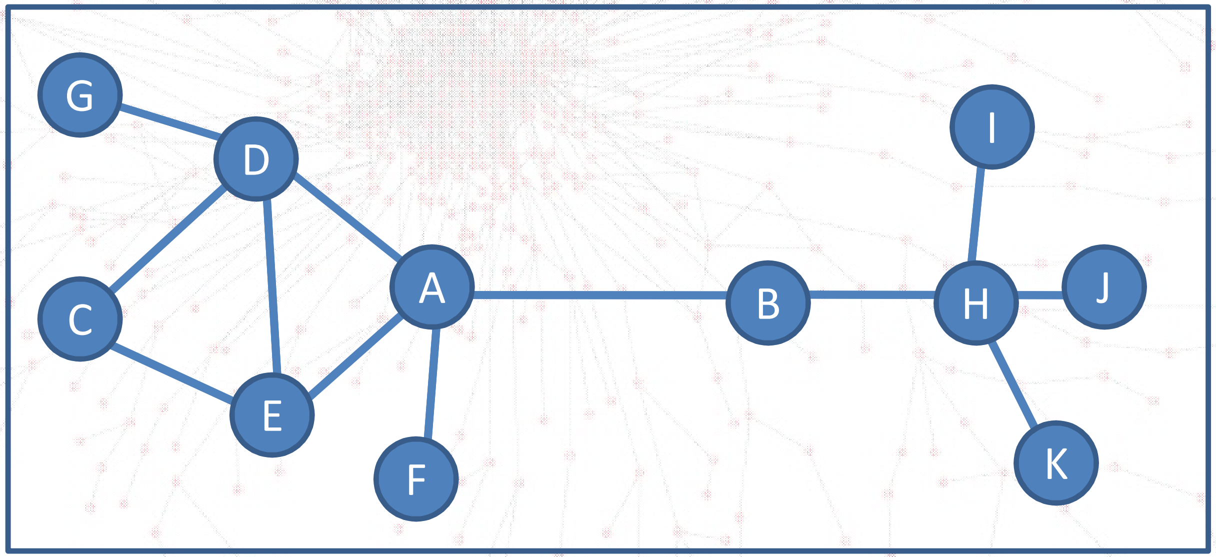
Topological Overlap
Mutual Clustering

TO(A,B)=N(A,B)/max(k(A),k(B))

TO(A,B)=0

TO(A,D)=1/4

TO(E,D)=2/4



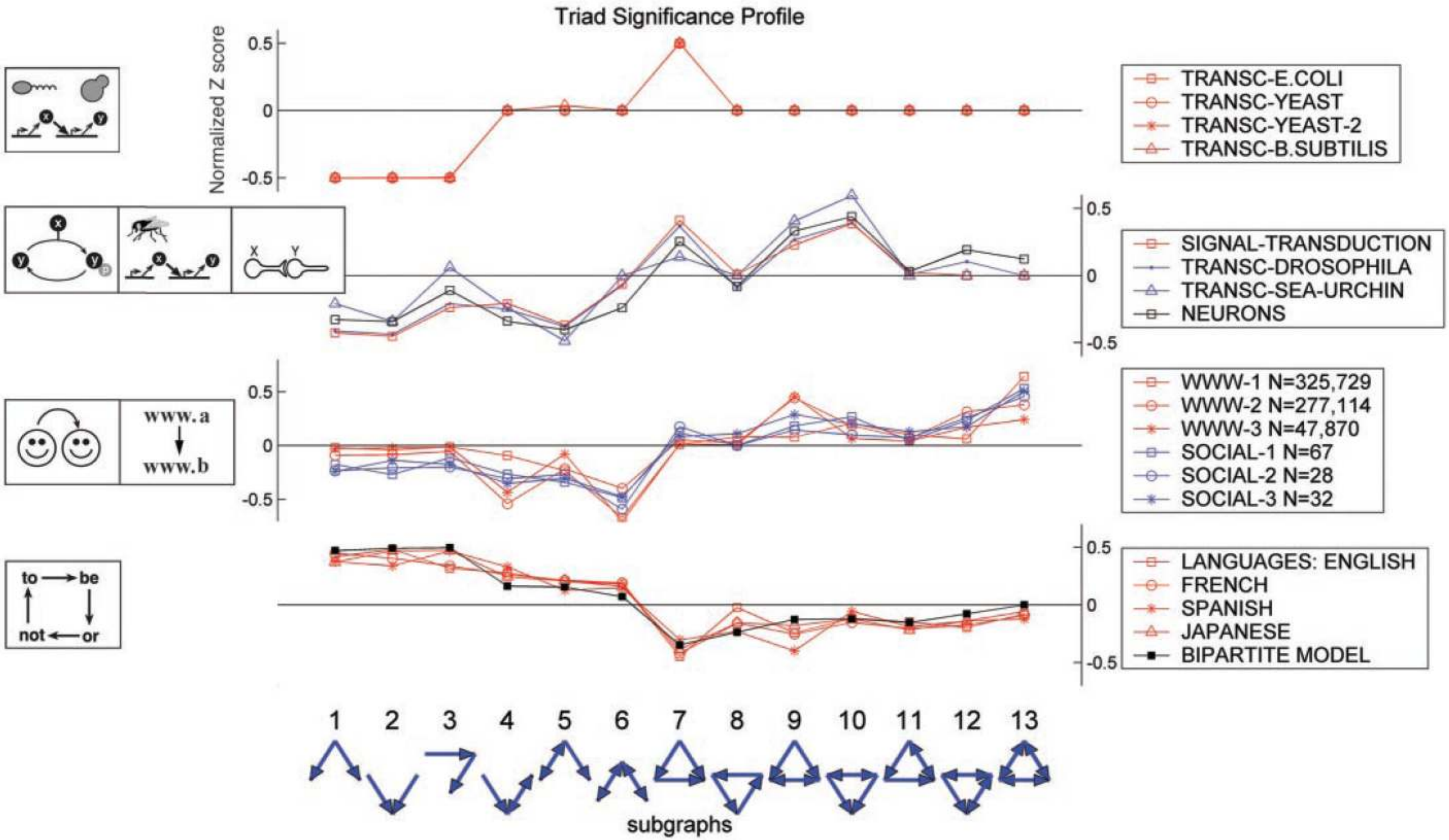
Global Measures

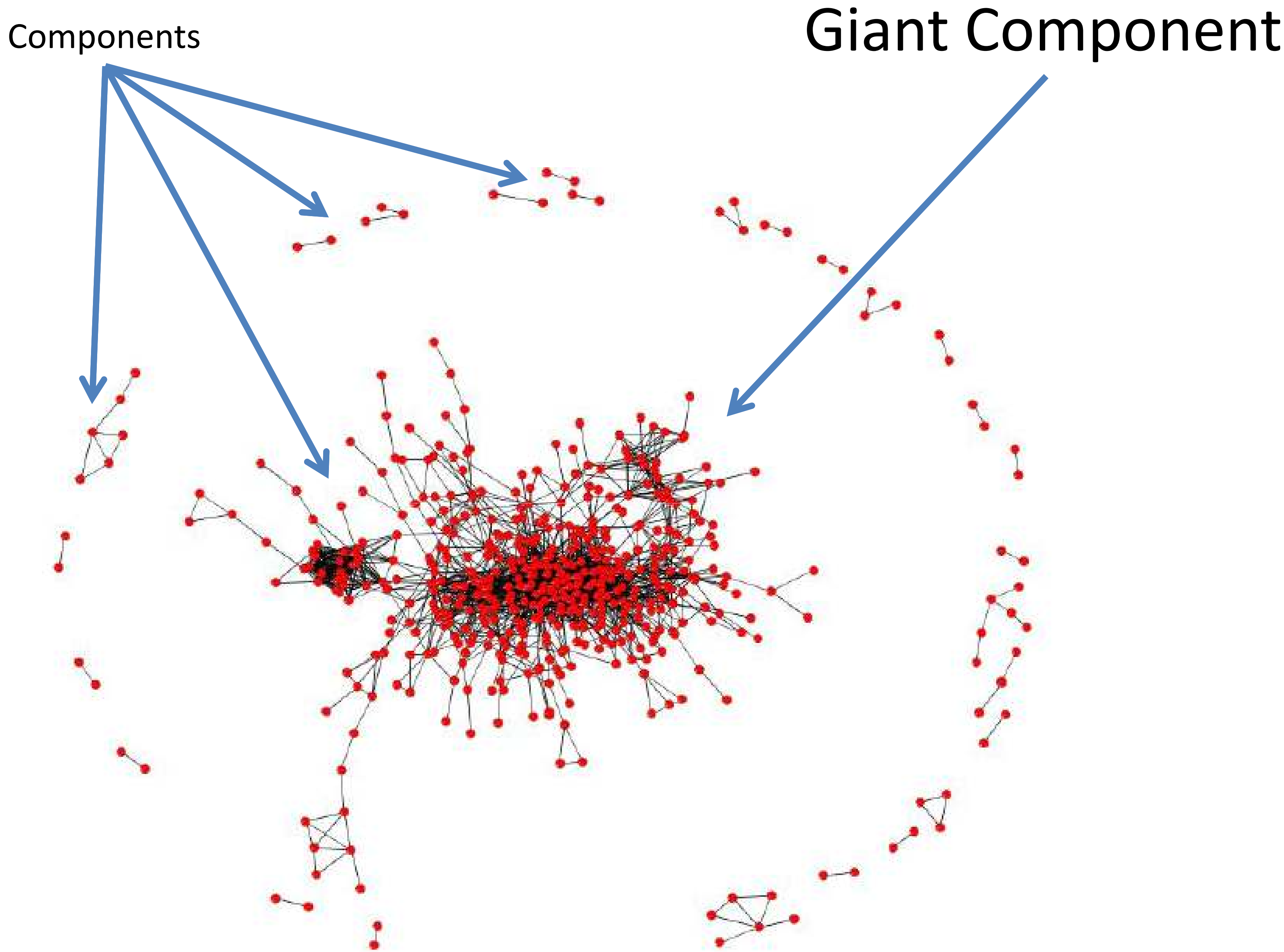
The Distribution of any of the previously introduced measures

5 MARCH 2004 VOL 303 SCIENCE

Superfamilies of Evolved and Designed Networks

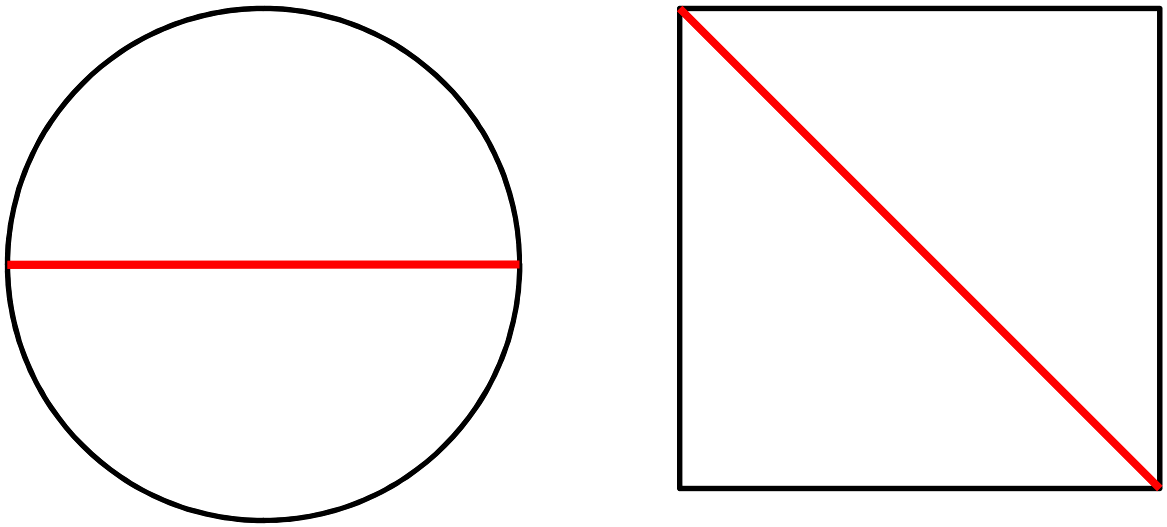
Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt,
Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, Uri Alon*





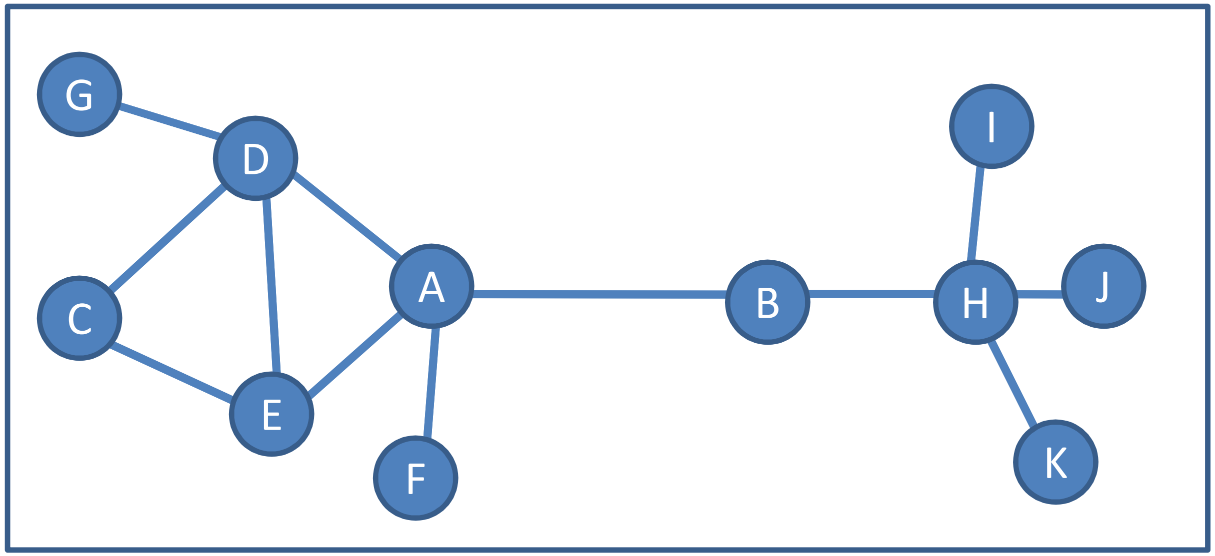
S=NumberOfNodesInGiantComponent/TotalNumberOfNodes

Diameter



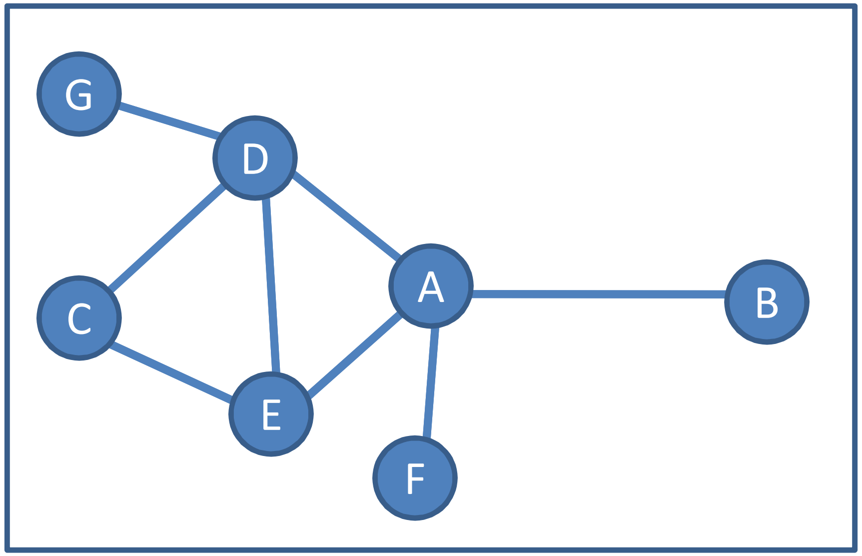
Diameter=Maximum Distance Between Elements in a Set

Diameter=D(G,J)=D(C,J)=D(G,I)=…=5



Average Path Length

	A	B	C	D	E	F	G
A		1	2	1	1	1	2
B			3	2	2	2	3
C				1	1	3	2
D					1	2	1
E						2	2
F							3
G							



D(1)=8
D(2)=9
D(3)=4

L=(8+2x9+3x4)/(8+9+4)   L=1.8

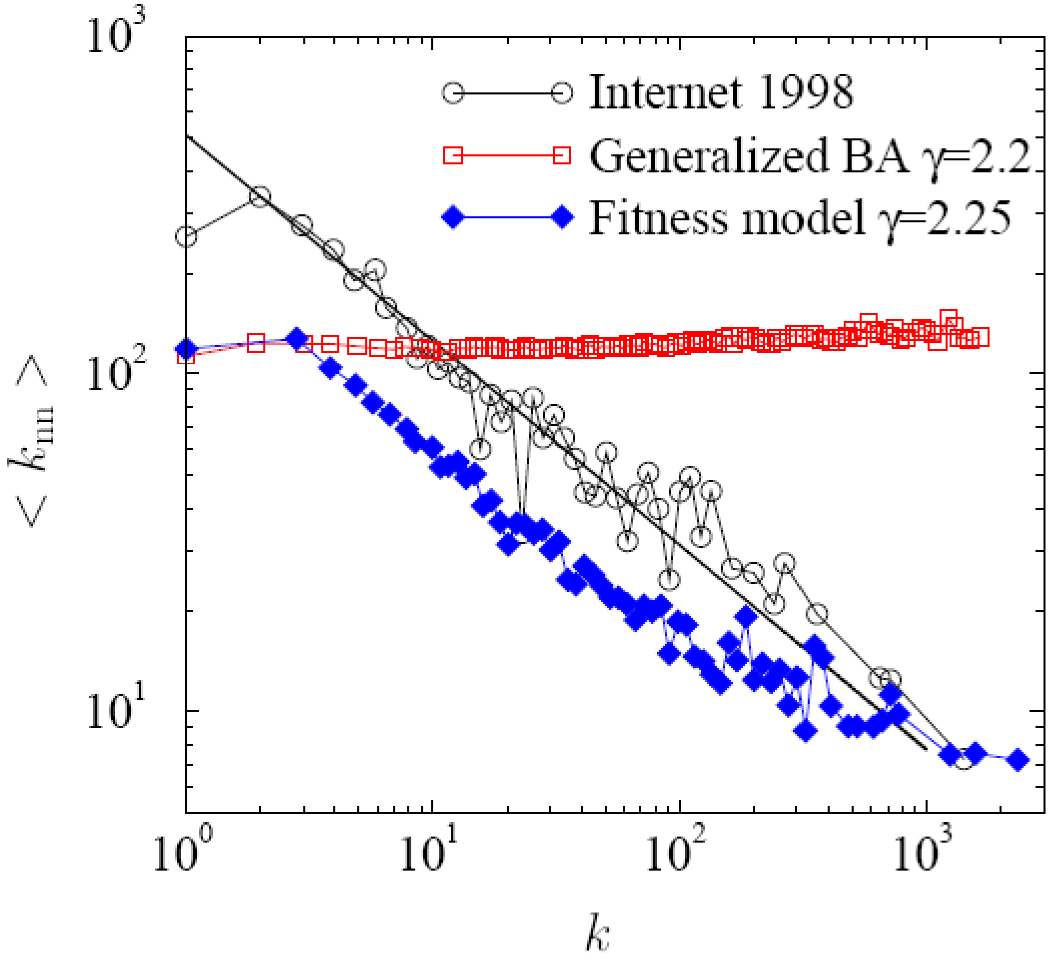
Degree Correlations

Are Hubs Connected to Hubs?

Dynamical and correlation properties of the Internet

Romualdo Pastor-Satorras,[1] Alexei Vázquez,[2] and Alessandro Vespignani[3]

Phys. Rev. Lett. 87, 258701 (2001)



Assortative Mixing in Networks

M. E. J. Newman

Department of Physics, University of Michigan, Ann Arbor, MI 48109–1120
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501
(Received 20 May 2002; published 28 October 2002)

A network is said to show assortative mixing if the nodes in the network that have many connections tend to be connected to other nodes with many connections. Here we measure mixing patterns in a variety of networks and find that social networks are mostly assortatively mixed, but that technological and biological networks tend to be disassortative. We propose a model of an assortatively mixed network, which we study both analytically and numerically. Within this model we find that networks percolate more easily if they are assortative and that they are also more robust to vertex removal.

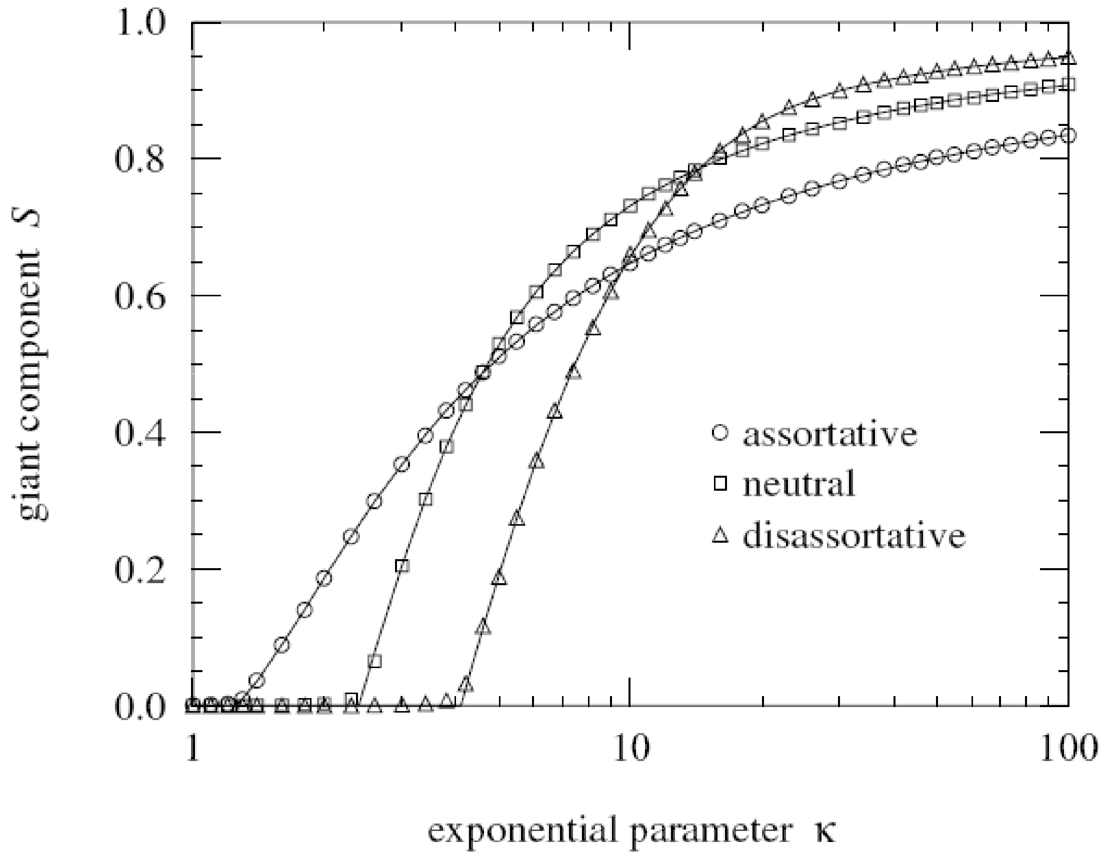
the normalized correlation function is

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - q_j q_k), \qquad (3)$$

which is simply the Pearson correlation coefficient of the degrees at either ends of an edge and lies in the range $-1 \leq r \leq 1$. For the practical purpose of evaluating r on

Network	n	r
Physics coauthorship (a)	52 909	0.363
Biology coauthorship (a)	1 520 251	0.127
Mathematics coauthorship (b)	253 339	0.120
Film actor collaborations (c)	449 913	0.208
Company directors (d)	7 673	0.276
Internet (e)	10 697	−0.189
World-Wide Web (f)	269 504	−0.065
Protein interactions (g)	2 115	−0.156
Neural network (h)	307	−0.163
Marine food web (i)	134	−0.247
Freshwater food web (j)	92	−0.276
Random graph (u)		0
Callaway *et al.* (v)		$\delta/(1+2\delta)$
Barabási and Albert (w)		0

What's a problem here?



1.0
0.8
0.6
0.4
0.2
0.0
giant component S
1
10
100
exponential parameter κ
assortative
neutral
disassortative