Centrality

Steve Borgatti

Copyright © 2006 by Steve Borgatti. All rights reserved.

Life in the Military

A case by David Krackhardt

Roger was in charge of a prestigious Advisory Team, which made recommendations to the Joint Chiefs of Staff. His experience was considerable, and he was a well-respected authority in the area. Of the 16 people who worked for him, he trusted those who also had a considerable amount of wartime experience, either in Vietnam or in other combat operations. He found their counsel to be particularly valuable.

Roger and Rick each had a PhD, and the remaining people all had graduate professional degrees in a variety of areas. Bob, Pete, Red and Sally were the newest members of the Team (they had been there for almost a year), and were fresh out of training in advanced weapons technology. Pete was the youngest member of the team. His background was computer science, and he had worked at MIT in their Draper Labs on simulations of war strategies using various weaponry.

Copyright © 2006 by Steve Borgatti. All rights reserved.

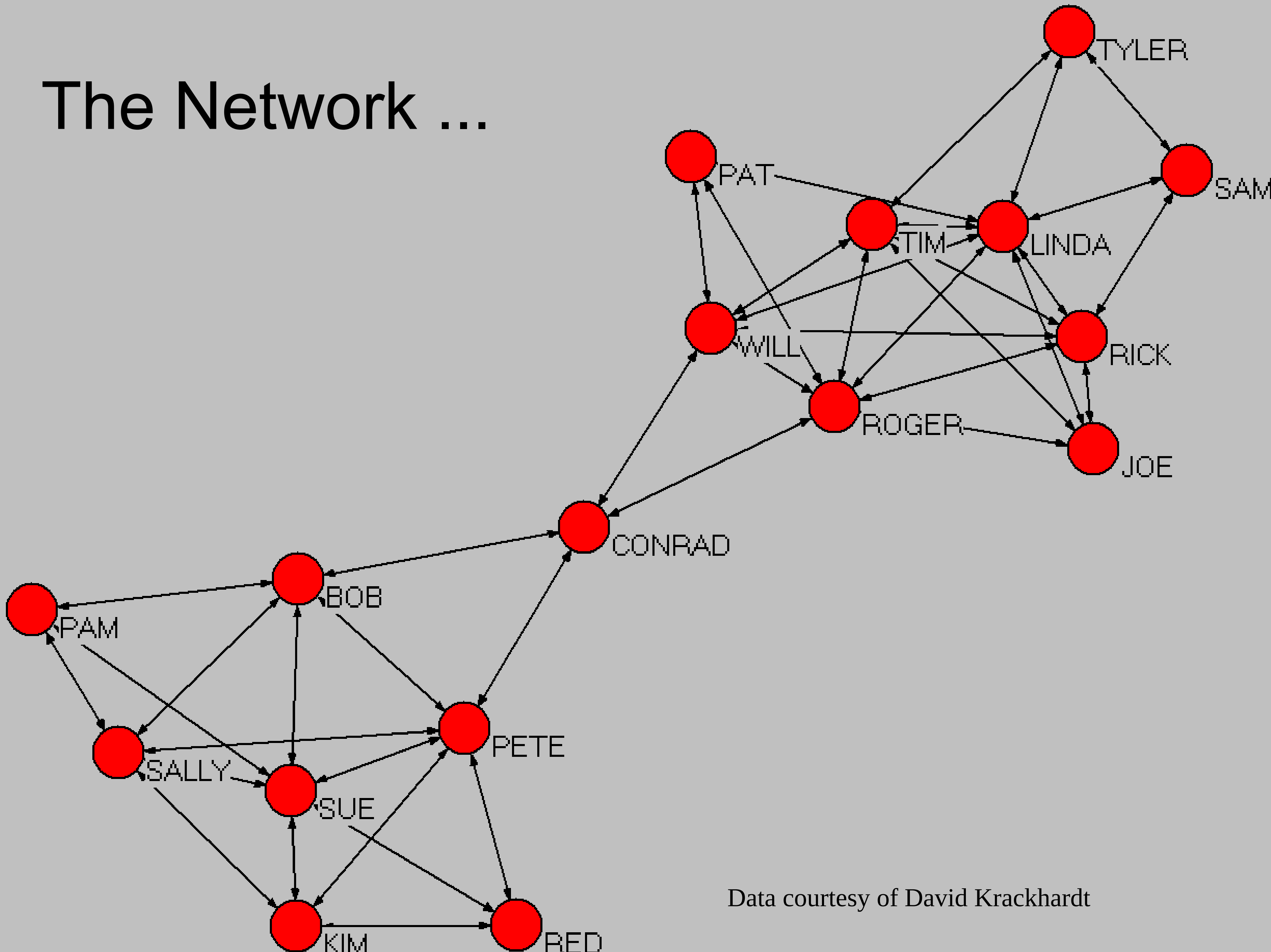
Life in the Military ... cont.

Linda was a senior member of the team and also one of the most approachable. She saw it as part of her responsibility to make sure people were getting along with each other, since cooperation across this disparate group was critical to its effectiveness. She and Rick would frequently hold social events to help solidify the group. Linda had been with the group the longest (almost 12 years) and had seen it grow in stature and respect over that time.

Roger had been criticized recently for his management style, which was admittedly authoritarian. At the request of some of his colleagues, he had called in an organizations consultant to advise him and the Team how to best proceed with teamwork and other managerial issues. The consultant ran team-building workshops. Roger felt that the consultant was a “touchy-feely” type and that the experience had been a total waste of time. He refused to bring in any more consultants. Some of the Team members were talking behind the scenes about resigning or requesting a transfer.

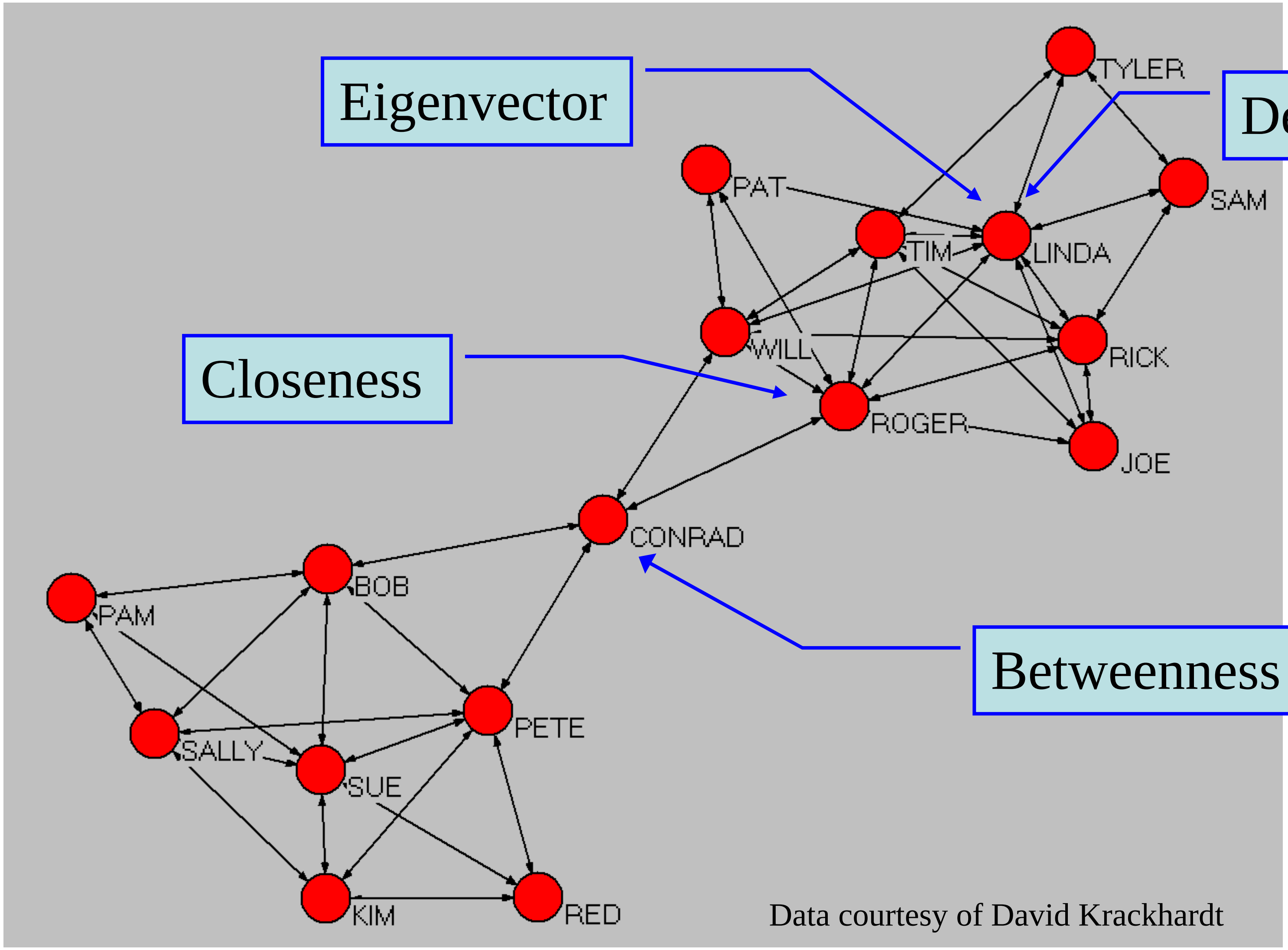
Copyright © 2006 by Steve Borgatti. All rights reserved.

The Network ...



Data courtesy of David Krackhardt

Four Aspects of Centrality



Copyright © 2006 by Steve Borgatti. All rights reserved.

Degree Centrality

- Number of ties that involve a given node
  - Marginals of symmetric adjacency matrix
- Index of exposure to what is flowing through the network
  - Gossip network: central actor more likely to hear a given bit of gossip
- Interpreted as opportunity to influence & be influenced directly
- Predicts variety of outcomes from virus resistance to power & leadership to job satisfaction to knowledge

Copyright © 2006 by Steve Borgatti. All rights reserved.

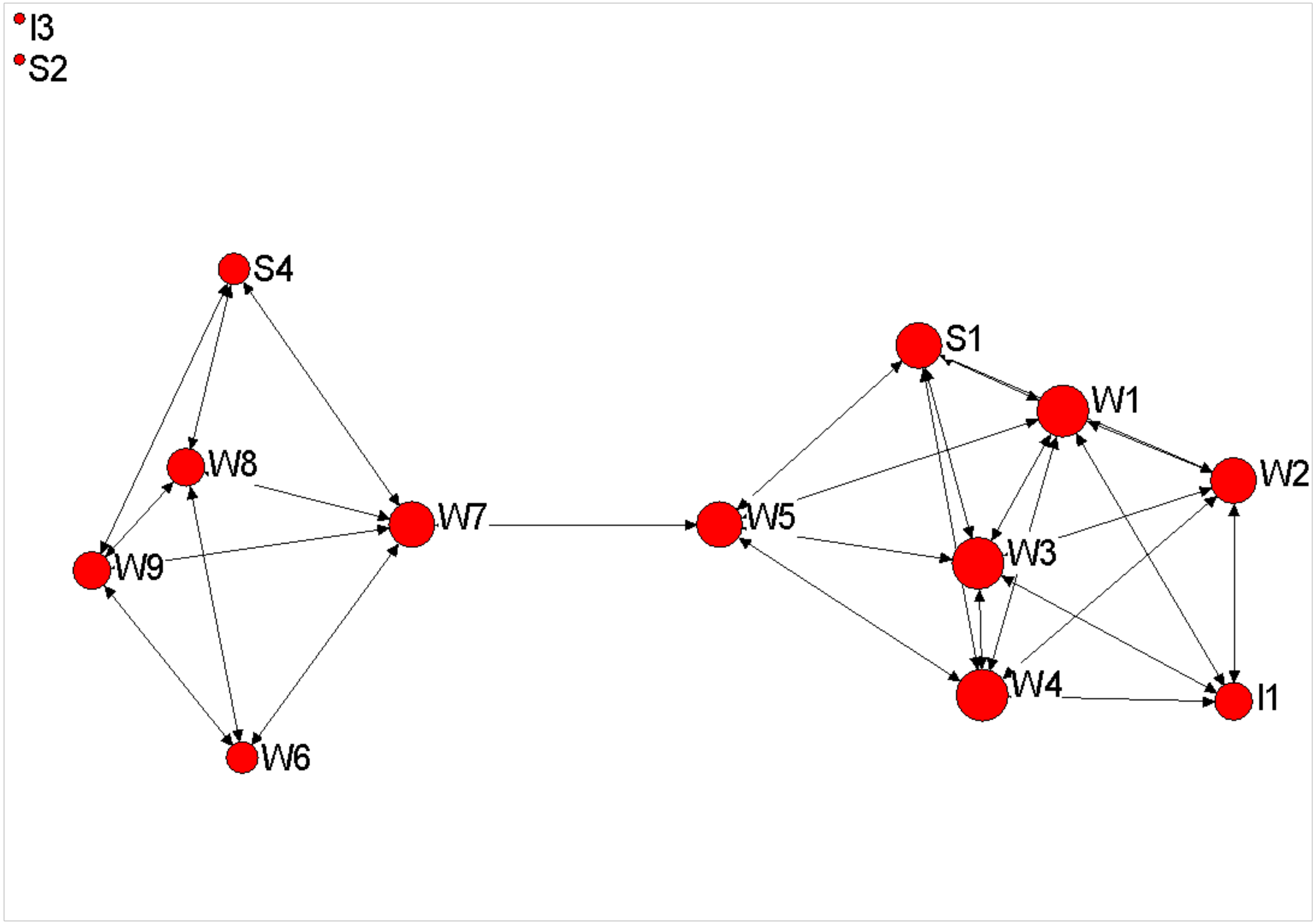
Degree Centrality

- Number of ties that involve a given node
  - Marginals of symmetric adjacency matrix

	I1	I3	W1	W2	W3	W4	W5	W6	W7	W8	W9	S1	S2	S4	Deg
I1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	4
I3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W1	1	0	0	1	1	1	1	0	0	0	0	1	0	0	6
W2	1	0	1	0	1	1	0	0	0	0	0	1	0	0	5
W3	1	0	1	1	0	1	1	0	0	0	0	1	0	0	6
W4	1	0	1	1	1	0	1	0	0	0	0	1	0	0	6
W5	0	0	1	0	1	1	0	0	1	0	0	1	0	0	5
W6	0	0	0	0	0	0	0	0	1	1	1	0	0	0	3
W7	0	0	0	0	0	0	1	1	0	1	1	0	0	1	5
W8	0	0	0	0	0	0	0	1	1	0	1	0	0	1	4
W9	0	0	0	0	0	0	0	1	1	1	0	0	0	1	4
S1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	5
S2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S4	0	0	0	0	0	0	0	0	1	1	1	0	0	0	3

Copyright © 2006 by Steve Borgatti. All rights reserved.

Wiring/Games Degree



	Deg
I1	4
I3	0
W1	6
W2	5
W3	6
W4	6
W5	5
W6	3
W7	5
W8	4
W9	4
S1	5
S2	0
S4	3

Copyright © 2006 by Steve Borgatti. All rights reserved.

Closeness Centrality

- Sum of distances to all other nodes
  - Is an inverse measure of centrality
- Index of expected time until arrival for given node of whatever is flowing through the network
  - Gossip network: central player hears things first

Copyright © 2006 by Steve Borgatti. All rights reserved.

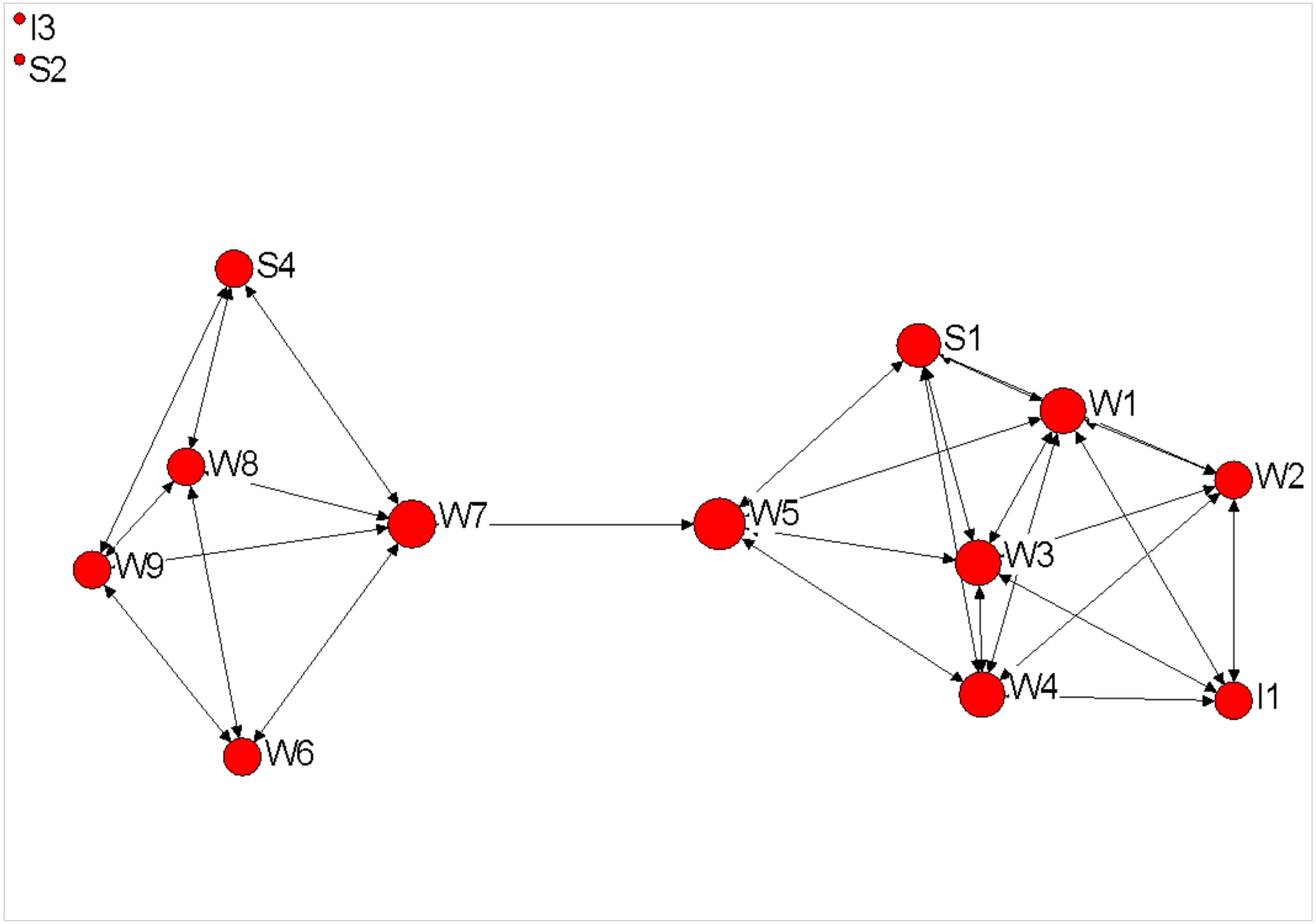
Closeness Centrality

- Computed as marginals of symmetric geodesic distance matrix

	I1	I3	W1	W2	W3	W4	W5	W6	W7	W8	W9	S1	S2	S4	Clo
I1	0	?	1	1	1	1	2	4	3	4	4	2	?	4	27
I3	?	0	?	?	?	?	?	?	?	?	?	?	?	?	0
W1	1	?	0	1	1	1	1	3	2	3	3	1	?	3	20
W2	1	?	1	0	1	1	2	4	3	4	4	1	?	4	26
W3	1	?	1	1	0	1	1	3	2	3	3	1	?	3	20
W4	1	?	1	1	1	0	1	3	2	3	3	1	?	3	20
W5	2	?	1	2	1	1	0	2	1	2	2	1	?	2	17
W6	4	?	3	4	3	3	2	0	1	1	1	3	?	2	27
W7	3	?	2	3	2	2	1	1	0	1	1	2	?	1	19
W8	4	?	3	4	3	3	2	1	1	0	1	3	?	1	26
W9	4	?	3	4	3	3	2	1	1	1	0	3	?	1	26
S1	2	?	1	1	1	1	1	3	2	3	3	0	?	3	21
S2	?	?	?	?	?	?	?	?	?	?	?	?	0	?	0
S4	4	?	3	4	3	3	2	2	1	1	1	3	?	0	27

Copyright © 2006 by Steve Borgatti. All rights reserved.

Wiring/Games Closeness



	Clo
I1	27
I3	
W1	20
W2	26
W3	20
W4	20
W5	17
W6	27
W7	19
W8	26
W9	26
S1	21
S2	
S4	27

Copyright © 2006 by Steve Borgatti. All rights reserved.

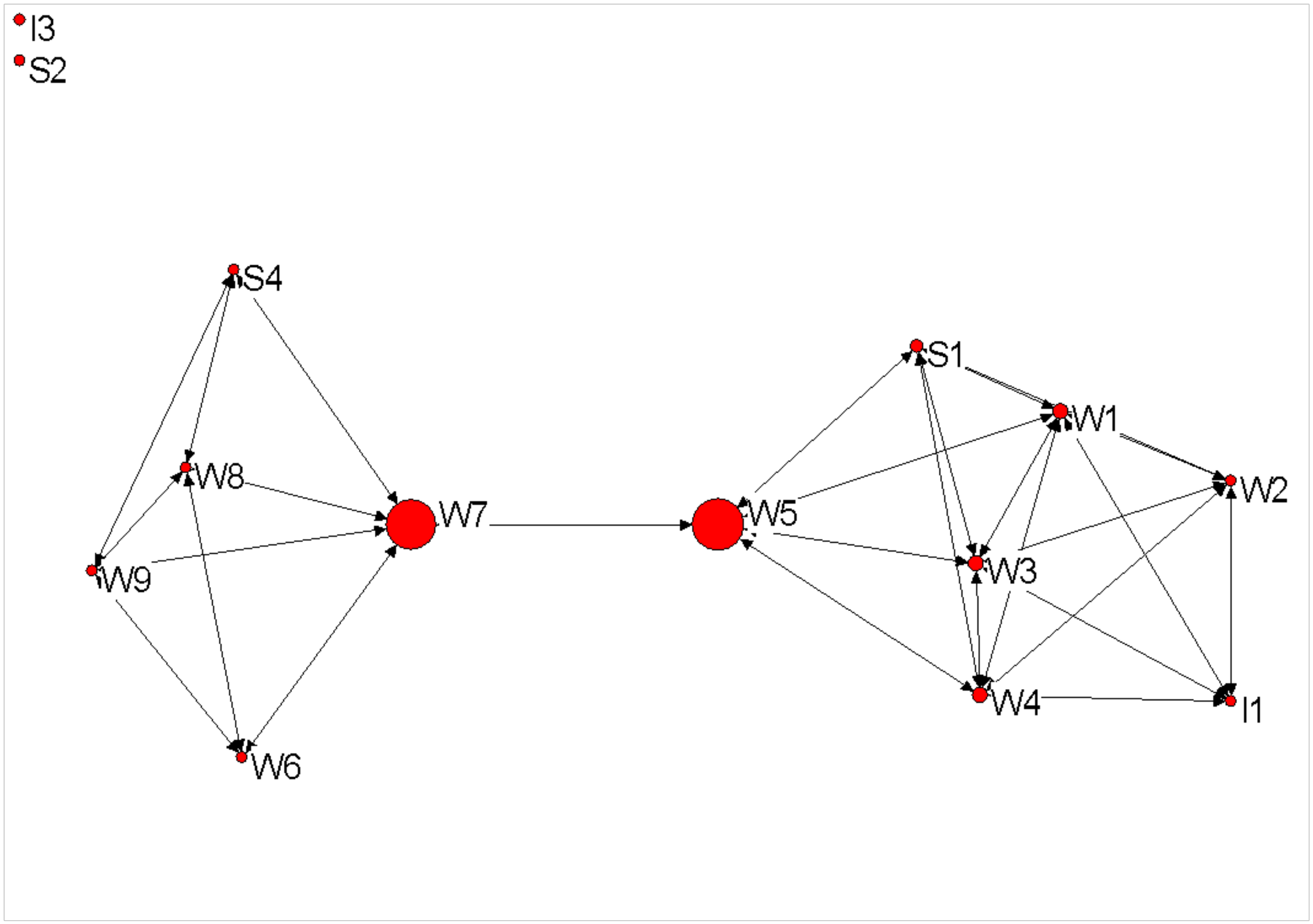
Betweenness Centrality

- How often a node lies along the shortest path between two other nodes
  - Computed as: $$b_k = \sum_{i,j} \frac{g_{ikj}}{g_{ij}}$$

    where gij is number of geodesic paths from i to j and gikj is number of those paths that pass through k
- Index of potential for gatekeeping, brokering, controlling the flow, and also of liaising otherwise separate parts of the network;
- Interpreted as indicating power and access to diversity of what flows; potential for synthesizing

Copyright © 2006 by Steve Borgatti. All rights reserved.

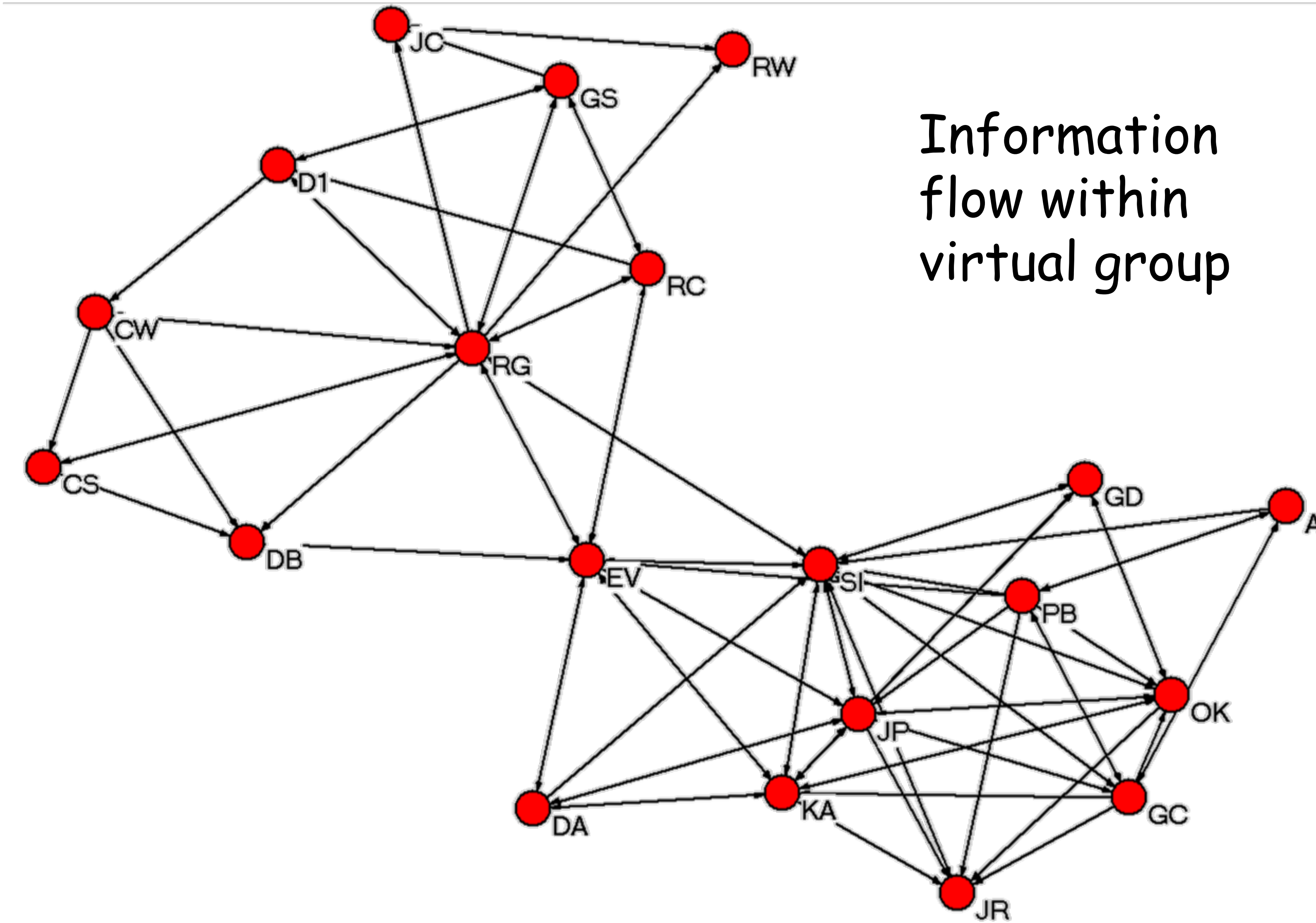
Wiring/Games Betweenness



	Bet
I1	0.00
I3	
W1	6.82
W2	0.45
W3	6.82
W4	6.82
W5	54.55
W6	0.00
W7	51.52
W8	0.61
W9	0.61
S1	2.73
S2	
S4	0.00

Copyright © 2006 by Steve Borgatti. All rights reserved.

Local Gain is Global Pain



Information flow within virtual group

Copyright © 2006 by Steve Borgatti. All rights reserved.

Eigenvector Centrality

- Node has high score if connected to many nodes are themselves well connected
  - Definition: $$\lambda v = Av$$

    where A is adjacency matrix and V is eigenvector centrality. V is the principal eigenvector of A.
- Indicator of popularity, “in the know”
- Like degree, is index of exposure, risk
- Tends to identify centers of large cliques

Copyright © 2006 by Steve Borgatti. All rights reserved.

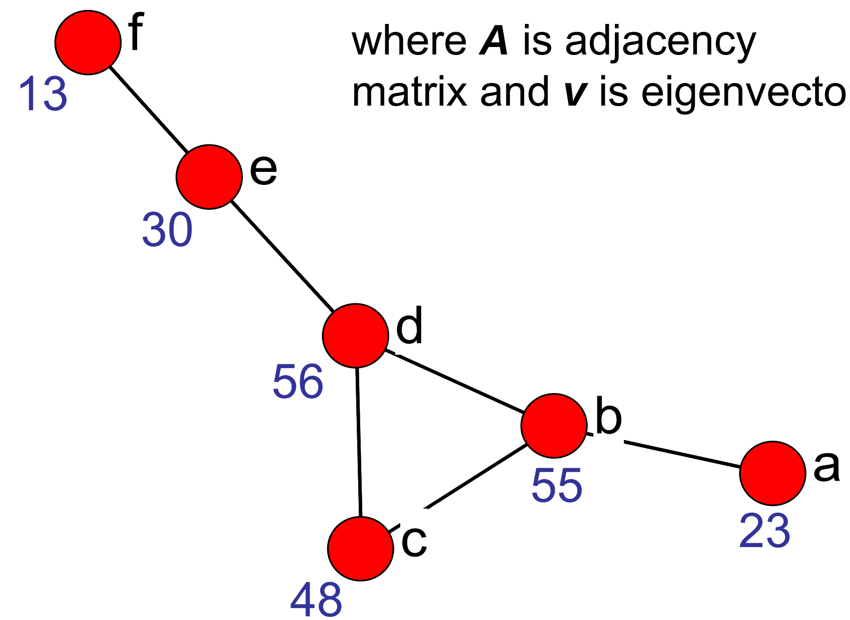
Eigenvector as turbo-charged degree

- Node is well-connected if connected to many well-connected nodes
  - Each person's score is proportional to the sum of the scores of those connected to them
  - This is an eigenvector
- Can (sometimes) compute using iterated weighted degree
  - Eigenvector as nuanced degree

$$v_i = \tfrac{1}{\lambda} \sum_j a_{ij} v_j$$

$$\lambda v = A v$$

where ***A*** is adjacency matrix and ***v*** is eigenvector



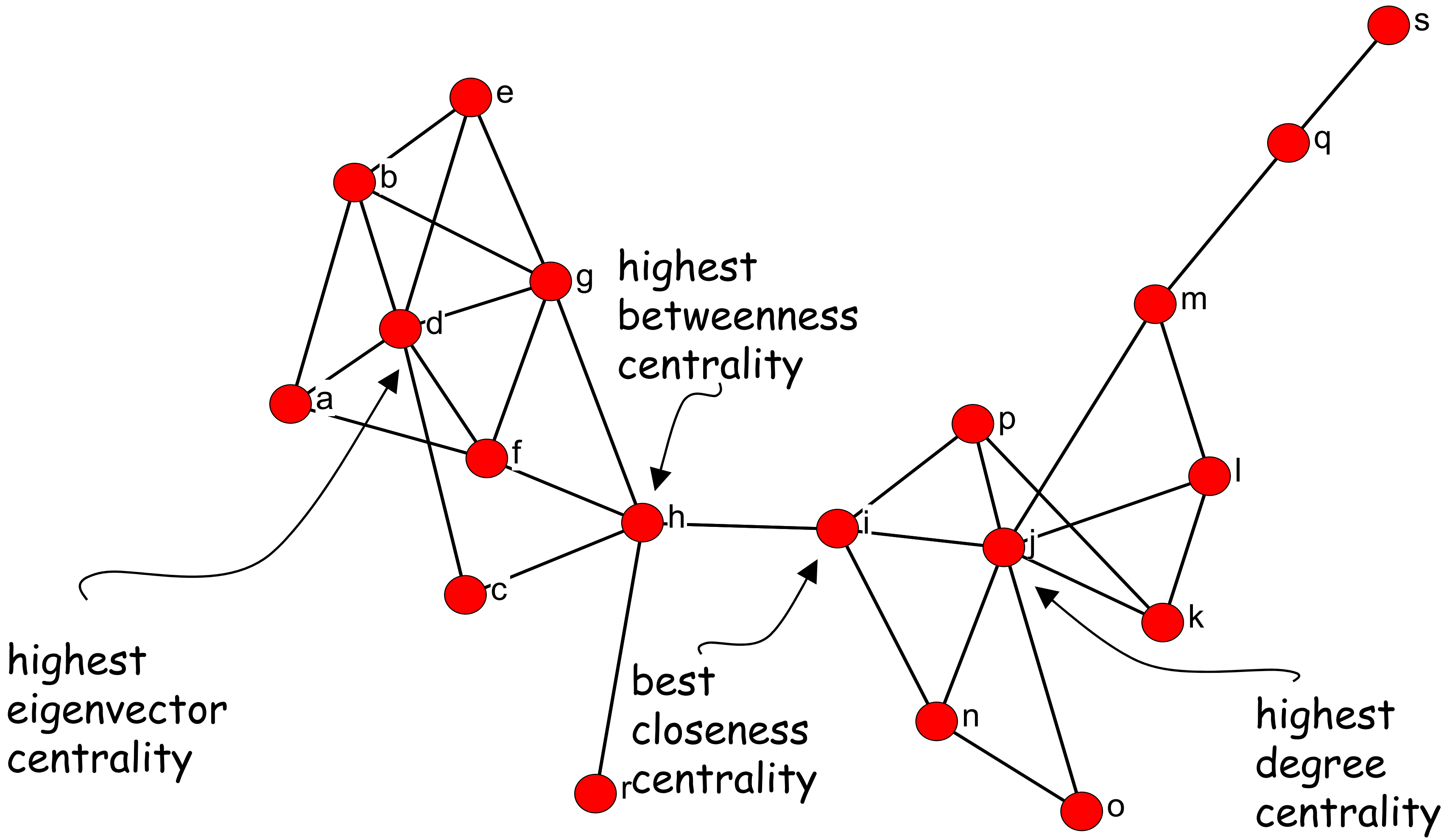
Eigenvector as Iterated Weighted Degree

	a	b	c	d	e	f	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
a	0	1	0	0	0	0	1	3	6	16	35	86	195	465	1071	2524
b	1	0	1	1	0	0	3	6	16	35	86	195	465	1071	2524	5854
c	0	1	0	1	0	0	2	6	13	32	73	173	401	940	2190	5117
d	0	1	1	0	1	0	3	7	16	38	87	206	475	1119	2593	6086
e	0	0	0	1	0	1	2	4	9	20	47	107	253	582	1372	3175
f	0	0	0	0	1	0	1	2	4	9	20	47	107	253	582	1372

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
A	8.3	10.7	9.4	10.7	10.1	10.6	10.3	10.5	10.4	10.5
B	25.0	21.4	25.0	23.3	24.7	24.0	24.5	24.2	24.4	24.3
C	16.7	21.4	20.3	21.3	21.0	21.3	21.1	21.2	21.2	21.2
D	25.0	25.0	25.0	25.3	25.0	25.3	25.1	25.3	25.1	25.2
E	16.7	14.3	14.1	13.3	13.5	13.1	13.3	13.1	13.3	13.2
F	8.3	7.1	6.3	6.0	5.7	5.8	5.6	5.7	5.6	5.7

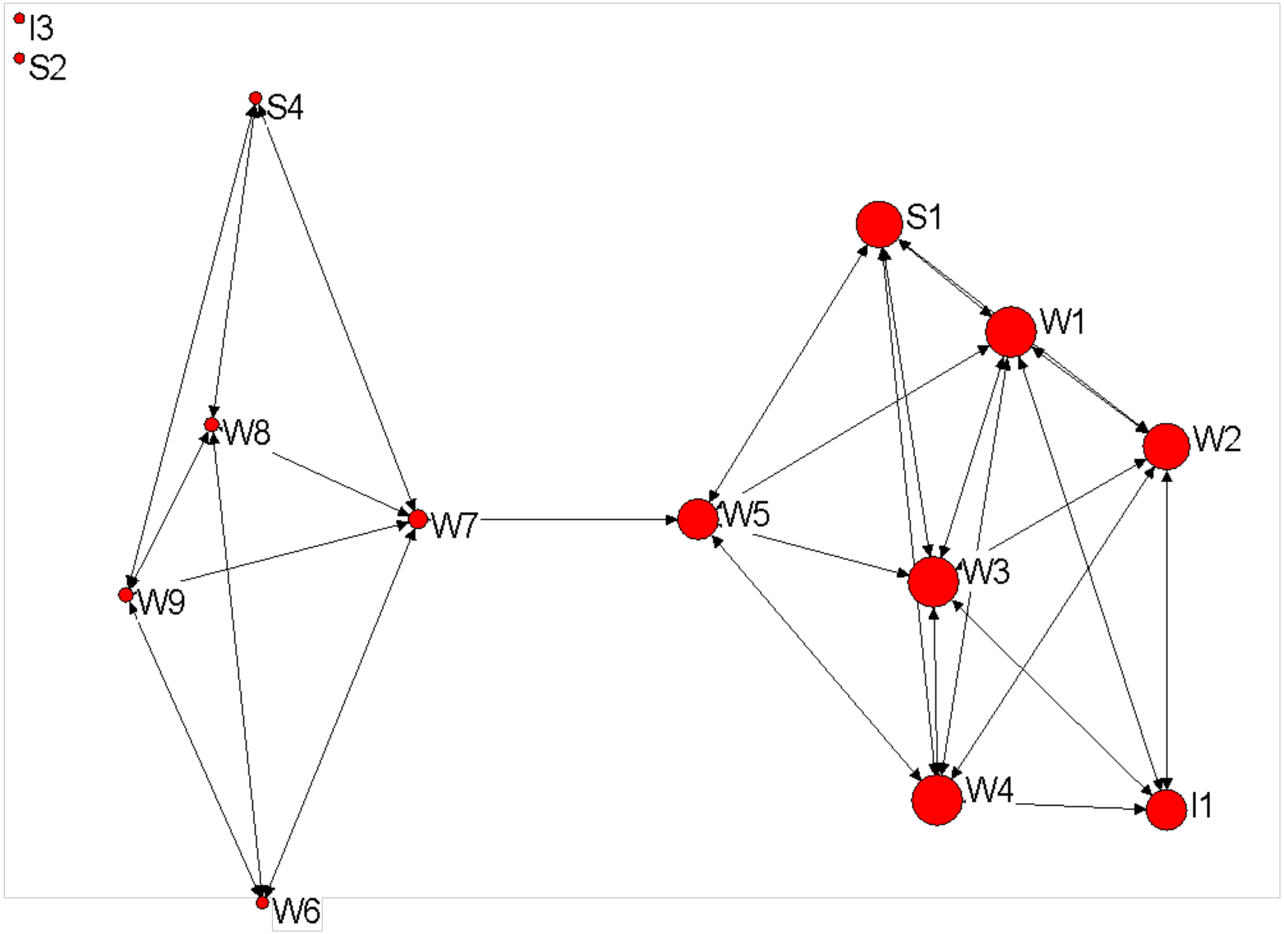
Copyright © 2006 by Steve Borgatti. All rights reserved.

Four Centralities



Copyright © 2006 by Steve Borgatti. All rights reserved.

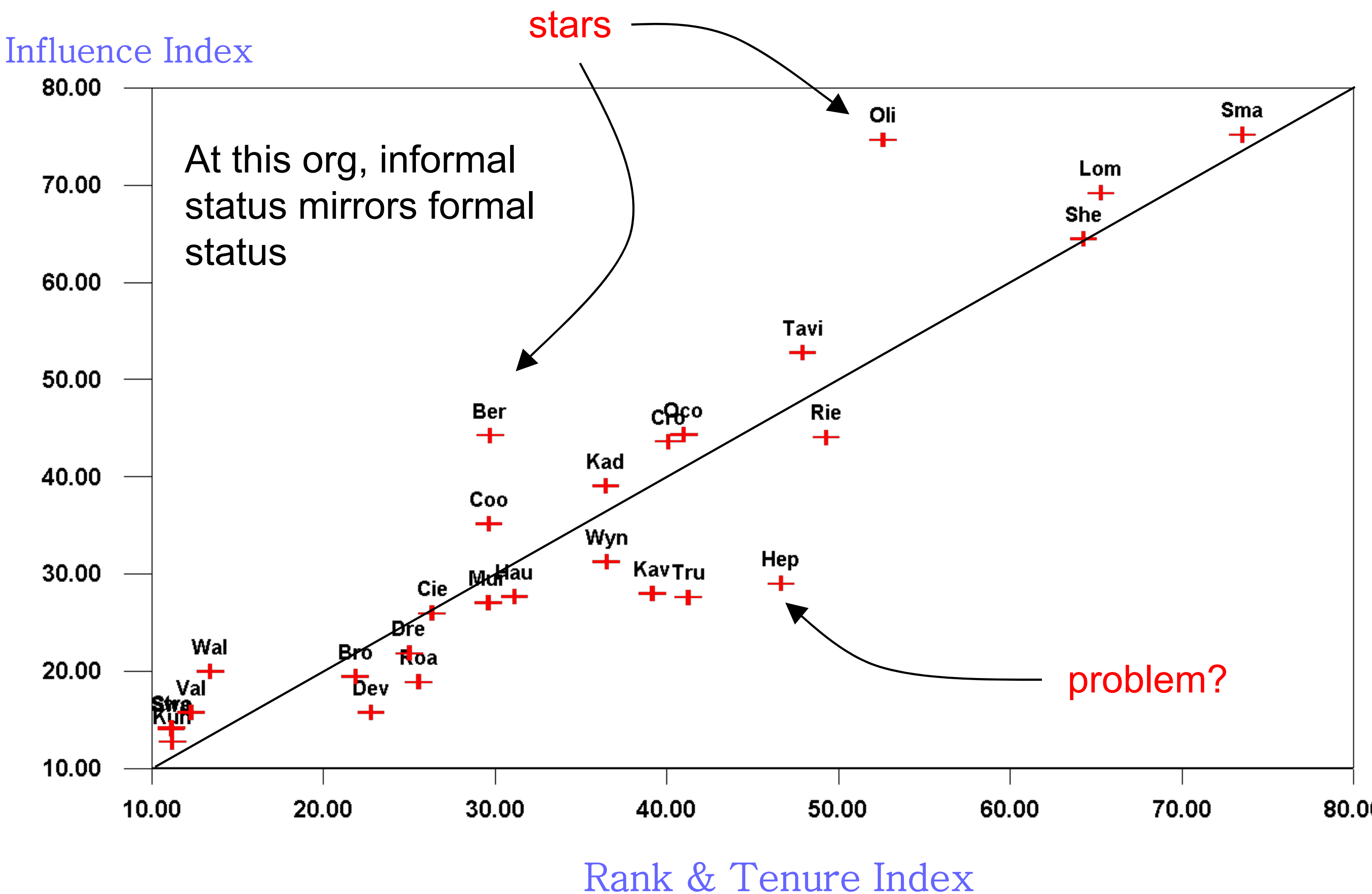
Wiring/Games Eigenvector



	Eig
I1	43
W1	59
W2	52
W3	59
W4	59
W5	46
W6	4
W7	12
W8	5
W9	5
S1	52
S4	4

Copyright © 2006 by Steve Borgatti. All rights reserved.

Influence Network



Copyright © 2006 by Steve Borgatti. All rights reserved.

Walk-Based Measures

- Multiple motivations
  - actor's status is function of not only the number of people who choose them, but their status
  - in an influence process, an actor's impact on another is function of all sequences (walks) linking them
- Resulting measures are similar / related

Copyright © 2006 by Steve Borgatti. All rights reserved.

Influence Approach

- Variations by Katz, Friedkin, Taylor, etc.
- Generic approach
  - R is network matrix, α is attenuation parameter
  - $\mathbf{Q} = \alpha^0\mathbf{R}^0 + \alpha^1\mathbf{R}^1 + \alpha^2\mathbf{R}^2 + \alpha^3\mathbf{R}^3 + \ldots \alpha^\infty\mathbf{R}^\infty$
  - $\mathbf{Q} = (\mathbf{I}-\alpha\mathbf{R})^{-1}$, assuming $\alpha^{-1} > \lambda_1$
  - $\mathbf{s} = (\mathbf{I}-\alpha\mathbf{R})^{-1}\mathbf{1} = \mathbf{Q1}$ (row sums of **Q**)

Copyright © 2006 by Steve Borgatti. All rights reserved.

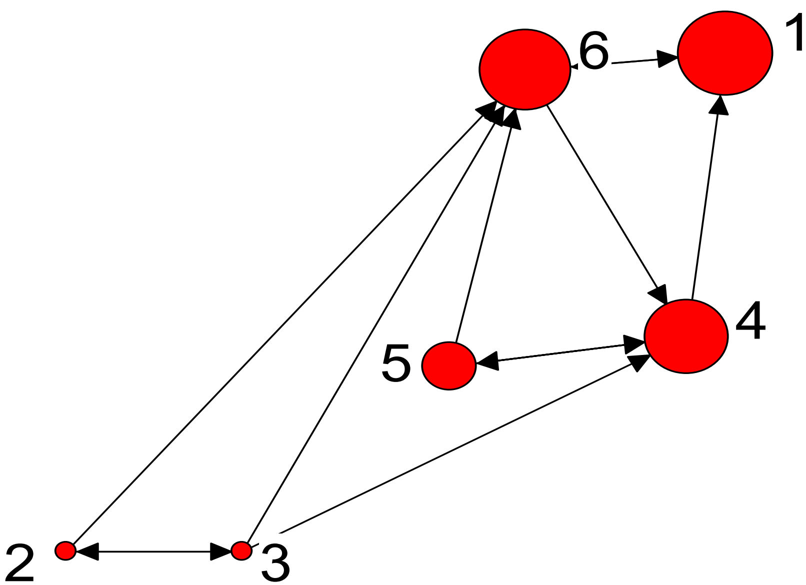
Recursive Status Approach

- Hubbell
  - **s** = **Ws** + **e**, where **W** is adj matrix w/ equal col sums < 1, **s** is vector representing status, **e** is vector of exogeneous inputs (usually 1s)
  - $\mathbf{s} = (\mathbf{I}-\mathbf{W})^{-1}\mathbf{e}$
  - Compare with Katz: $\mathbf{s} = (\mathbf{I}-\alpha\mathbf{R})^{-1}\mathbf{1} = \mathbf{Q1}$ (row sums of **Q**)
- Bonacich, Coleman, Burt, etc.
  - Principal eigenvector of **W**
  - $\lambda\mathbf{c} = \mathbf{Wc}$ (or **W'c** if appropriate)
  - $\mathbf{c} = (\mathbf{I}-\mathbf{W})^{-1}\mathbf{c}$

Copyright © 2006 by Steve Borgatti. All rights reserved.

Katz example

Who really knows what's going on?



Node	Out	In	Katz
1	1	2	13.0
2	2	1	1.0
3	3	1	1.0
4	2	3	11.4
5	2	1	6.2
6	2	4	12.6

Indegree gives same score to 5 as to 2 and 3. But 5 is chosen by 4, who is chosen by popular nodes like 6. Katz score gives 5 much higher score than 2 or 3. Similarly node 1 has only two incoming choices, but they are from the most sought-after players, so 1 must be even more knowledgeable than they.

Copyright © 2006 by Steve Borgatti. All rights reserved.

Bonacich Power

$$c_i(\alpha, \beta) = \sum_j (\alpha + \beta c_j) R_{ij}.$$

$$c(\alpha, \beta) = \alpha(I - \beta R)^{-1} R 1,$$

where "1" is a column vector of ones and I is an identity matrix.[6]

[6] The vector $c(\alpha, \beta)$ approaches e as a limit as β approaches the reciprocal of the largest eigenvalue of R.

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

Copyright © 2006 by Steve Borgatti. All rights reserved.

Typology of Centrality Measures

- Measures of centrality assess involvement of nodes in walk structure of the network

Paths
Trails
Walks

	Radial (emanating from node)	**Medial (passing thru node)**
Length	Closeness Information	< New >
Frequency	Degree Eigenvector	Betweenness Random Walk Bet

Borgatti, S. P. and Everett, M.G. (2007). A graph-theoretic framework for classifying centrality measures. *Social Networks*

Copyright © 2006 by Steve Borgatti. All rights reserved.

- **Objective**: How does a node's centrality vary with the way things flow through a network?
  - Different kinds of flows: guns, information, infection, etc.
- Freeman's Closeness
  - Sum of distances from a given node to all others
  - Time until arrival of something flowing through net
- Freeman's Betweenness
  - Share of shortest paths that pass through a given node
  - Gatekeeper control of flows
- What if things don't flow along shortest paths?
- Or copy rather than move?

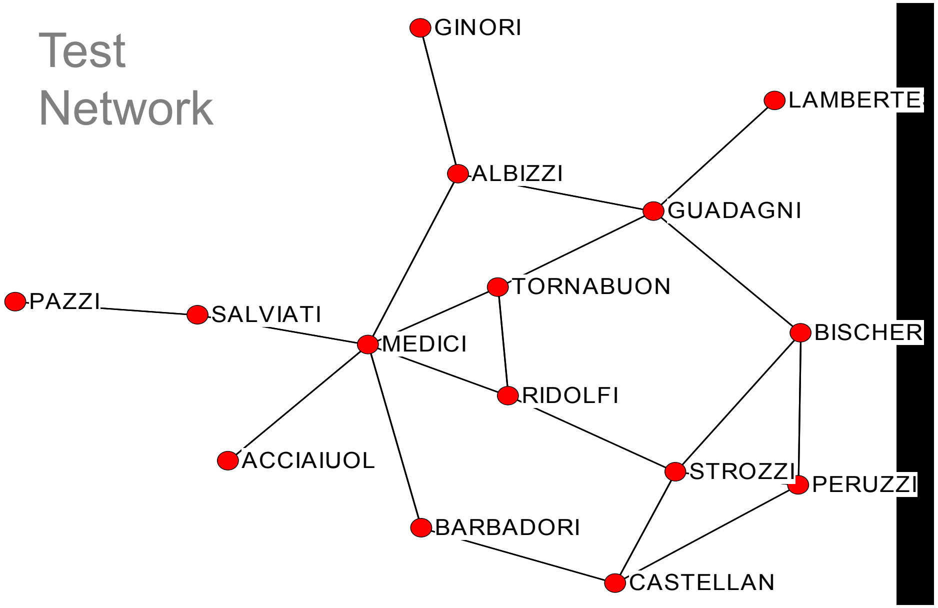
Borgatti, S.P. 2005. Centrality and network flow. *Social Networks*. 27(1): 55-71.

Typology of Flows

	REPLICATION	TRANSFER
SHORT PATHS		<u>Mail Package</u> - One place a time - Travels along shortest paths
PATHS	<u>Viral Infection</u> -Multiple copies - Doesn't return to same node twice	<u>Mooch</u> -One place a time -Burns bridges, can't return
TRAILS	<u>Gossip</u> -Multiple copies -Can revisit node -Doesn't travel between same pairs twice	<u>Used Goods</u> -One place at a time -Can revisit node -Doesn't travel between same pairs twice
WALKS		<u>Money</u> - One place at a time - Unrestricted travel

Copyright © 2006 by Steve Borgatti. All rights reserved.

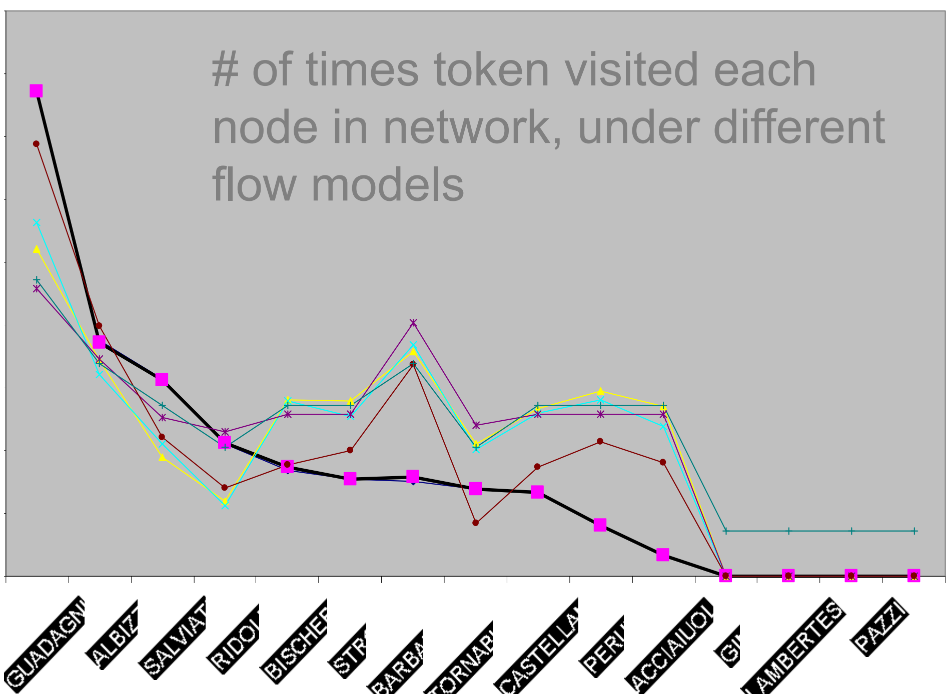
Test Network



Betweenness Results

Number of times token passed thru node

Node	BET	Pkg	Mooch	Used Good	Gossip	Infect	Cash
MEDICI	47.5	47.5	113.7	129.8	334	887.0	1155
GUADAGNI	23.2	22.8	74.9	73.8	252	513.5	828
ALBIZZI	19.3	19.2	41.5	48.5	185	285.5	666
SALVIATI	13.0	13.0	26.0	26.0	168	182.0	503
RIDOLFI	10.3	10.7	61.3	64.2	189	227.9	665
BISCHERI	9.5	9.5	60.9	58.6	189	257.2	665
STROZZI	9.3	9.7	78.1	84.8	295	435.1	828
BARBADORI	8.5	8.5	45.8	46.5	176	107.7	504
TORNABUO	8.3	8.2	58.2	59.8	189	223.0	666
CASTELLAN	5.0	5.0	64.5	64.7	188	277.2	665
PERUZZI	2.0	2.0	59.1	55.1	189	232.3	665
ACCIAIUOL	0.0	0.0	0.0	0.0	0	0.0	177
GINORI	0.0	0.0	0.0	0.0	0	0.0	177
LAMBERTE	0.0	0.0	0.0	0.0	0	0.0	177
PAZZI	0.0	0.0	0.0	0.0	0	0.0	177



- Centrality measures as expected values for nodal outcomes under specific flow models

Copyright © 2006 by Steve Borgatti. All rights reserved.

Sponsored by:

KeyPlayer Project

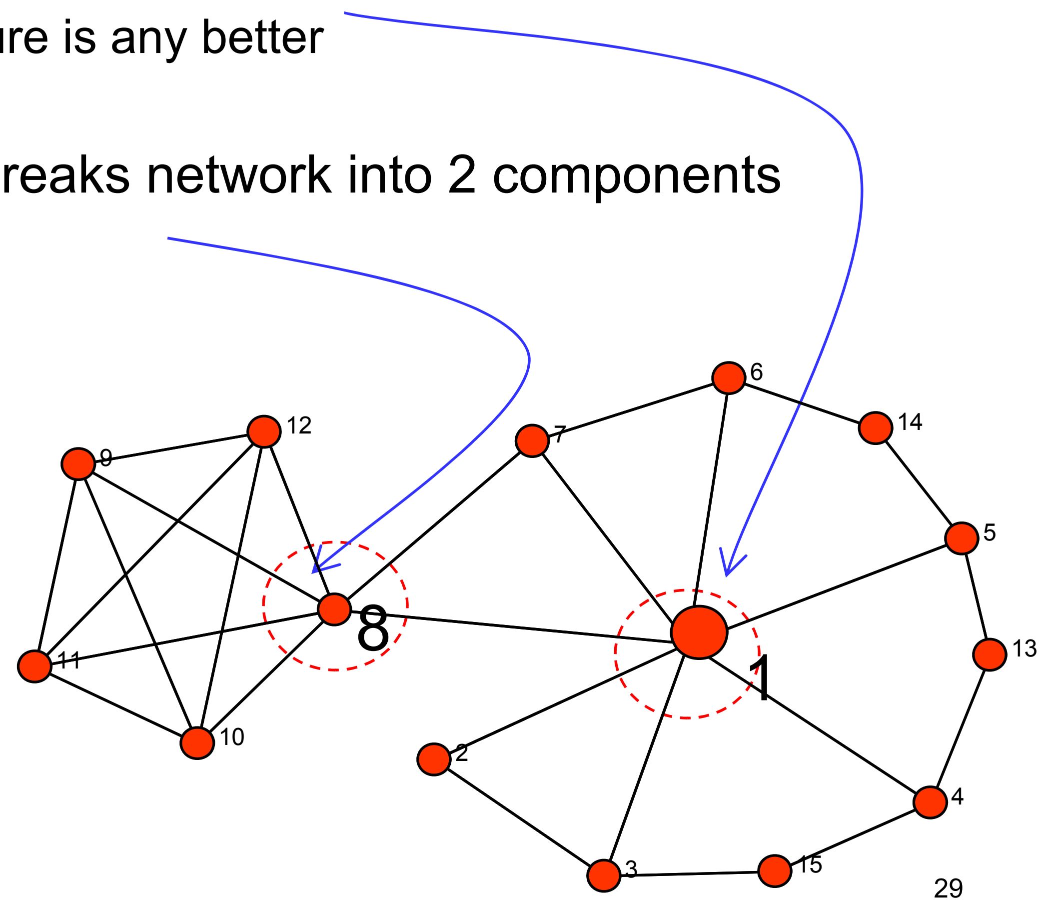
- Identifying key players in network.
- Three broad classes of purposes
  - Disruption
  - Diffusion
  - Debriefing
- Off-the-shelf centrality measures not optimal for specific purposes
- Sometimes we need to select an optimal set of key players
  - Best ensemble not usually composed of top individuals
- Research issues:
  - Metrics
  - Algorithms

DISRUPT		DIFFUSE
Who/how many to **immunize or quarantine** in order to slow spread of infectious disease	← **Health Care** →	Selecting peer health advocates for **diffusing** safe practices (e.g. bleaching) and material
Who to **arrest or discredit** to disrupt criminal networks	← **Security** →	Who to “turn” or feed **false information** to
Where is an organization most vulnerable to **turnover**?	← **Mgmt** →	Select subset of employees for **intervention** prior to change initiative
Which opinion leaders do I need to take care of?	← **Marketing** →	Select efficient opinion leaders

Copyright © 2006 by Steve Borgatti. All rights reserved.

The Design Issue

- Node 1 has highest betweenness centrality, but deleting it …
  - does not disconnect the network
  - And no other existing measure is any better
- In contrast, deleting node 8 breaks network into 2 components
  - Yet node 8 is not highest in centrality
  - So centrality is not optimal

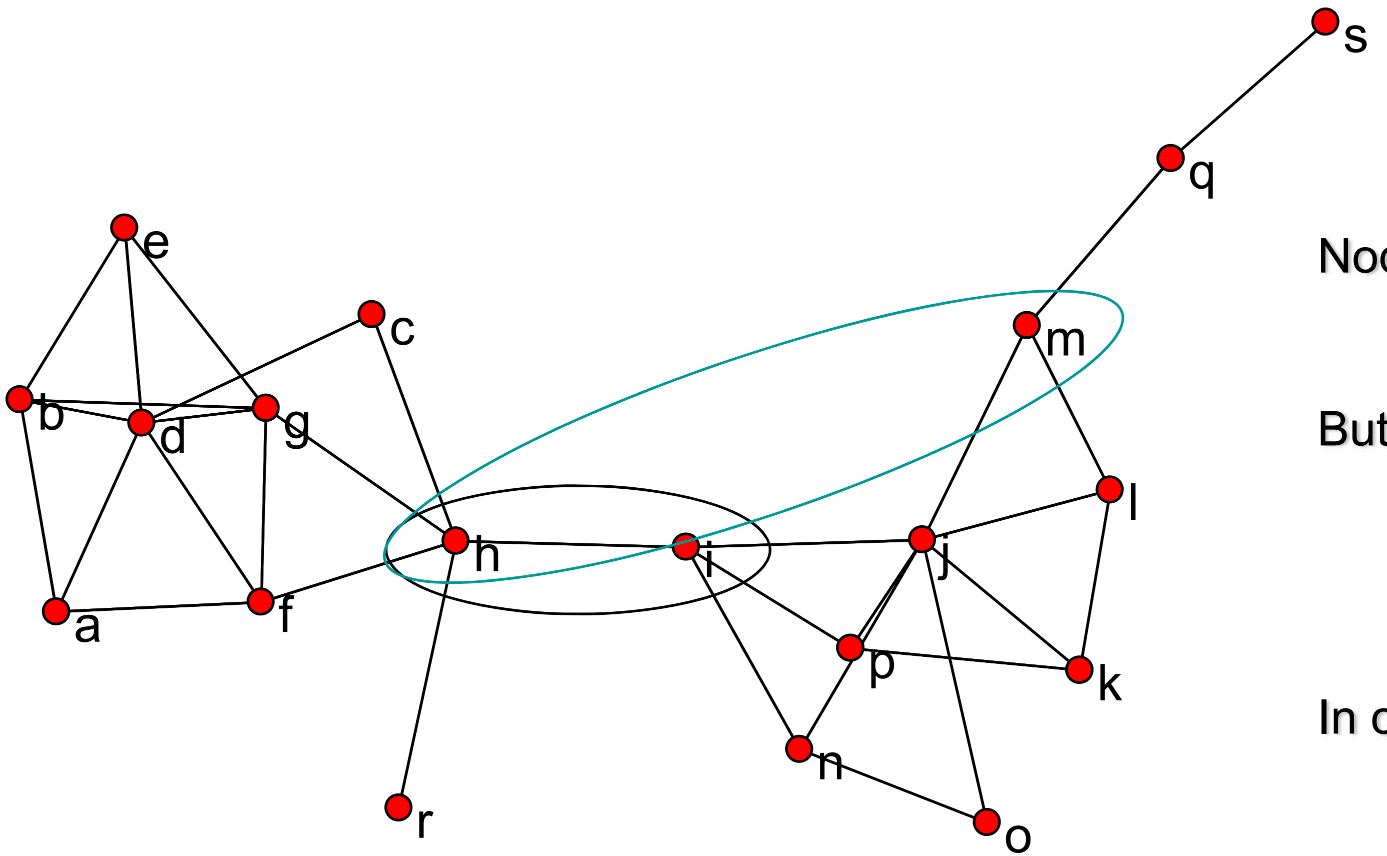


Copyright © 2006 by Steve Borgatti. All rights reserved.

DISRUPTION

The Ensemble Issue

Node Redundancy Requires Choosing Complementary Nodes



Nodes *h* and *i* are individually optimal

But deleting both is no better than deleting *h* alone -- *h* and *i* are redundant

In contrast, {*h*,*m*} splits graph into 4 fragments (is optimal)

Copyright © 2006 by Steve Borgatti. All rights reserved.

KeyPlayer DISRUPTION Metrics

- Strategy
  - Measure cohesiveness of network
  - Remove key players
  - Measure cohesiveness of new network
  - Proportion reduction in cohesion (PRC) measures fragmentation potential of the key player set
- Optimization
  - Find sets of keyplayers with maximum fragmentation potential

- Cohesion measures
  - Fragmentation

$$F = 1 - \frac{2\sum_{j<i} r_{ij}}{n(n-1)}$$

  - Breadth

$$S = 1 - \frac{2\sum_{i>j} \frac{1}{d_{ij}}}{n(n-1)}$$

Medial measures of centrality (e.g., betweenness) can be rewritten as PRC measures

Copyright © 2006 by Steve Borgatti. All rights reserved.

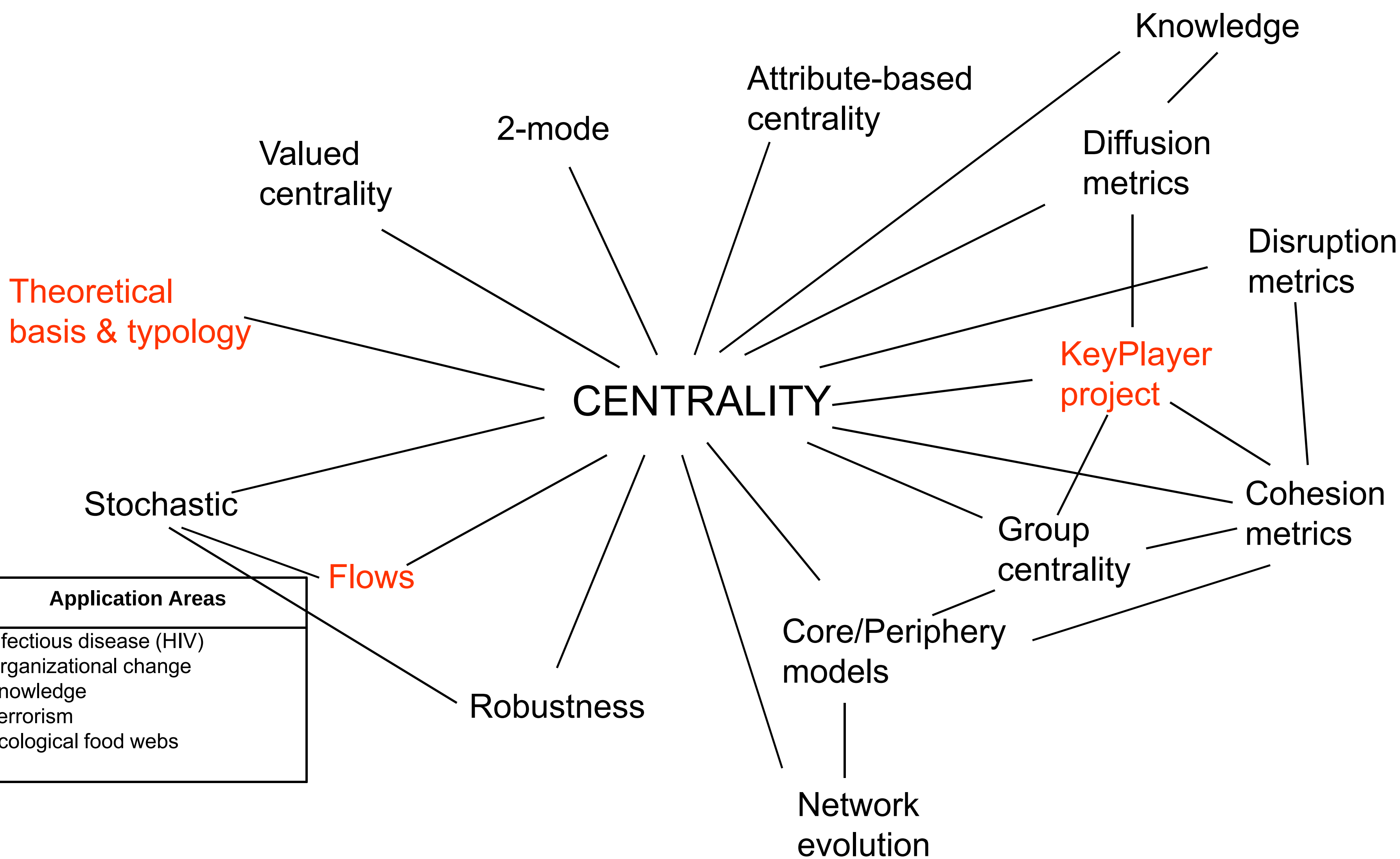
Centrality

- Structural importance
- Many measures
  - very different assumptions about data, processes & objectives
- Basically count paths or walks
  - emanating from / terminating at given node
  - passing through a given node

Copyright © 2006 by Steve Borgatti. All rights reserved.

My own …

Centrality Stream of Research



Application Areas
Infectious disease (HIV) Organizational change Knowledge Terrorism Ecological food webs

Copyright © 2006 by Steve Borgatti. All rights reserved.