Slide 1

Centrality

Wednesday Morning

Centrality is a measure of how network structure and position contributes to a node's importance. However, there are multiple measures of centrality that capture different forms of importance such as power, influence, popularity, risk, etc. In this session we will discuss how we can identify central nodes.

Objectives:

After this section, you should be able to:

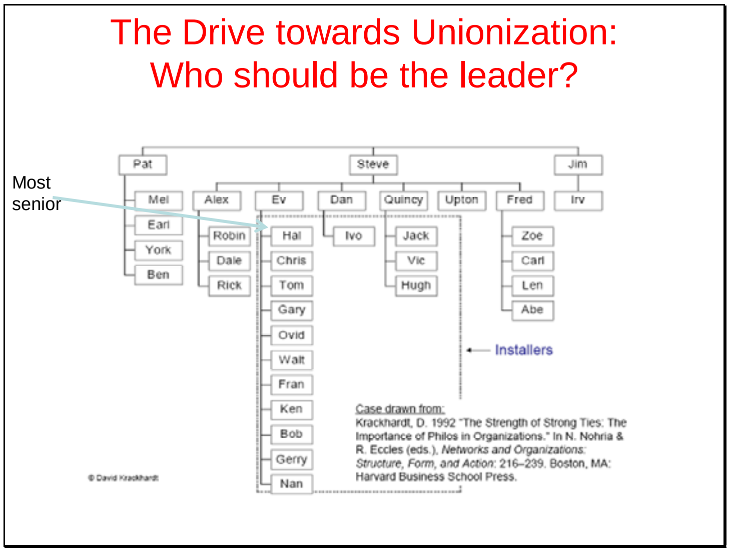
- Describe at least four different types of Centrality and the implications of each measure
- Use UCINET to calculate different measures of Centrality
- Use NetDraw to calculate and visualize different measures of Centrality

Slide 2

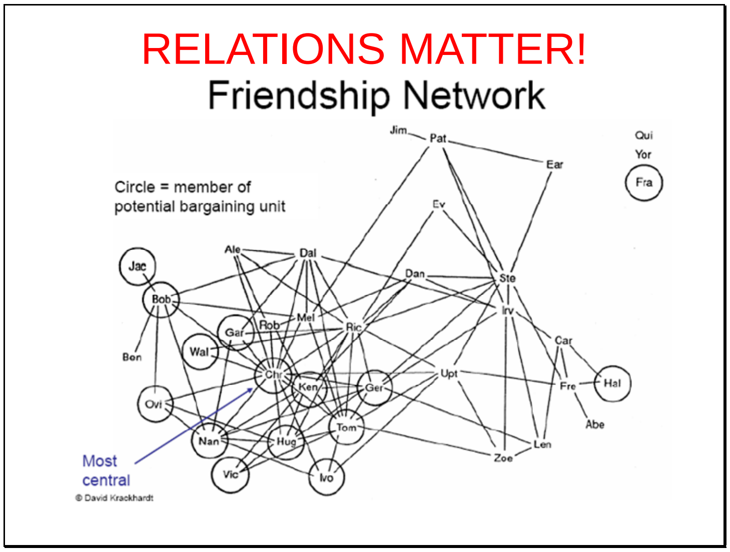
Algazi, Alvarez, Alpern, Ametrano, Andrews, Aran, Arnstein, Ashford,
Bailey Ballout, Bamberger, Baptista, Barr, Barrows, Baskerville,
Bassiri, Bell, Bokgese, Brandao, Bravo, Brooke, Brightman, Billy,
Blau, Bohen, Bohn, Borsuk, Brendle, Butler, Calle, Cantwell, Carrell,
Chinlund, Cirker, Cohen, Collas, Couch, Callegher, Calcaterra, Cook,
Carey, Cassell, Chen, Chung, Clarke, Cohn, Carton, Crowley, Curbelo,
Dellamanna, Diaz, Dirar, Duncan, Dagostino, Delakas, Dillon, Donaghey,
Daly, Dawson, Edery, Ellis, Elliott, Eastman, Easton, Famous, Fermin,
Fialco, Finklestein, Farber, Falkin, Feinman, Friedman, Gardner,
Gelpi, Glascock, Grandfield, Greenbaum Greenwood, Gruber, Garil, Goff,
Gladwell, Greenup, Gannon, Ganshaw, Garcia, Gennis, Gerard, Gericke,
Gilbert, Glassman, Glazer, Gomendio, Gonzalez, Greenstein, Guglielmo,
Gurman, Haberkorn, Hoskins, Hussein, Hamm, Hardwick, Harrell,
Hauptman, Hawkins, Henderson, Hayman, Hibara, Hehmann, Herbst, Hedges,
Hogan, Hoffman, Horowitz, Hsu, Huber, Ikiz, Jaroschy, Johann, Jacobs,
Jara, Johnson, Kassel, Keegan, Kuroda, Kavanau, Keller, Kevill, Kiew,
Kimbrough, Kline, Kossoff, Kotzitzky, Kahn, Kiesler, Kosser, Korte,
Leibowitz, Lin, Liu, Lowrance, Lundh, Laux, Leifer, Leung, Levine,
Leiw, Lockwood, Logrono, Lohnes, Lowet, Laber, Leonardi, Marten,
McLean, Michaels, Miranda, Moy, Marin, Muir, Murphy, Marodon, Matos,
Mendoza, Muraki, Neck, Needham, Noboa, Null, O'Flynn, O'Neill,
Orlowski, Perkins, Pieper, Pierre, Pons, Pruska, Paulino, Popper,
Potter, Purpura, Palma, Perez, Portocarrero, Punwasi, Rader, Rankin,
Ray, Reyes, Richardson, Ritter, Roos, Rose, Rosenfeld, Roth,
Rutherford, Rustin, Ramos, Regan, Reisman, Renkert, Roberts, Rowan,
Rene, Rosario, Rothbart, Saperstein, Schoenbrod, Schwed, Sears,
Statosky, Sutphen, Sheehy, Silverton, Silverman, Silverstein, Sklar,
Slotkin, Speros, Stollman, Sadowski, Schles, Shapiro, Sigdel, Snow,
Spencer, Steinkol, Stewart, Stires, Stopnik, Stonehill, Tayss, Tilney,
Temple, Torfield, Townsend, Trimpin, Turchin, Villa, Vasillov, Voda,
Waring, Weber, Weinstein, Wang, Wegimont, Weed, Weishaus

From Gladwell(2000)

Slide 3


The Drive towards Unionization:
Who should be the leader?
Most senior
Pat
Steve
Jim
Mel
Earl
York
Ben
Alex
Robin
Dale
Rick
Ev
Hal
Chris
Tom
Gary
Ovid
Walt
Fran
Ken
Bob
Gerry
Nan
Dan
Ivo
Quincy
Jack
Vic
Hugh
Upton
Fred
Zoe
Carl
Len
Abe
Irv
Installers
Case drawn from:
Krackhardt, D. 1992 "The Strength of Strong Ties: The Importance of Philos in Organizations." In N. Nohria & R. Eccles (eds.), Networks and Organizations: Structure, Form, and Action: 216–239. Boston, MA: Harvard Business School Press.
© David Krackhardt

Slide 4


RELATIONS MATTER!
Friendship Network
Circle = member of potential bargaining unit
Jim
Pat
Ear
Qui
Yor
Fra
Ev
Ale
Dal
Dan
Ste
Irv
Jac
Bob
Mel
Rob
Gar
Ric
Wal
Ben
Car
Chr
Ken
Ger
Upt
Fre
Hal
Ovi
Abe
Tom
Nan
Hug
Len
Zoe
Vic
Ivo
Most central
© David Krackhardt

____

____

____

____

____

____

____

____

____

____

____

Slide 5

RELATIONS MATTER

Slide 6

Centrality

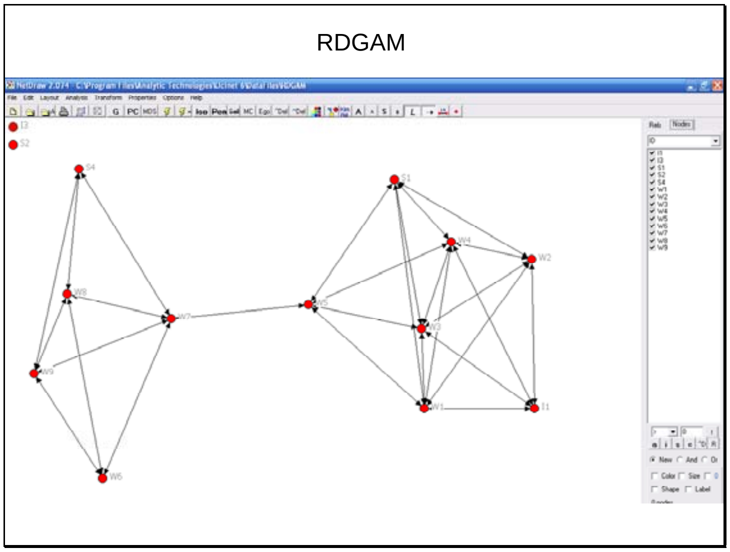
- A measure of how network structure and position contributes to a node’s importance
- Value associated with every node
- Many different measures which capture different aspects
- Can be characterized by the nature of the flow

Slide 7

Centrality

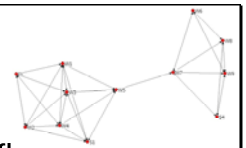
- Degree
  - how well connected; direct influence
- Closeness
  - how far from all others
  - how long information takes to arrive
- Eigenvector
  - being connected to the well connected (a popularity & power measure)
- Betweenness
  - brokerage, gatekeeping, control of info

Slide 8



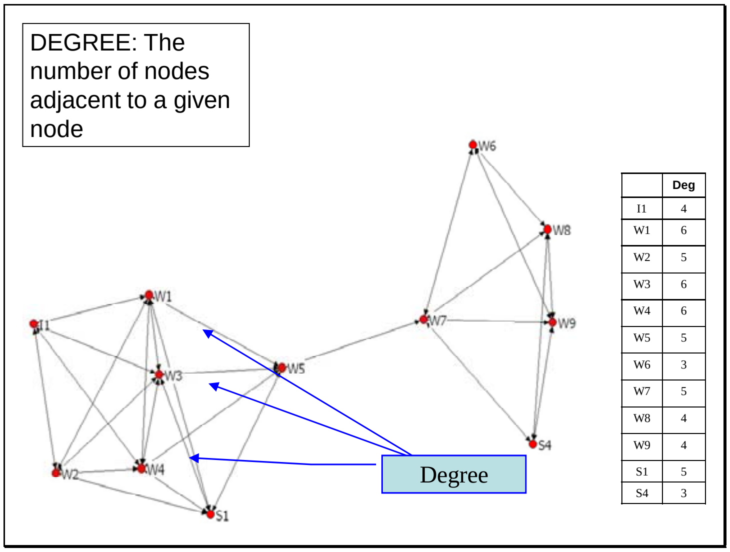
Slide 9

Degree Centrality



- 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

Slide 10



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

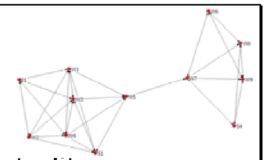
____
____
____
____
____
____
____
____
____
____
____

Slide 11

	I1	W1	W2	W3	W4	W5	W6	W7	W8	W9	S1	S4	SUM
I1	0	1	1	1	1	0	0	0	0	0	0	0	**4**
W1	1	0	1	1	1	1	0	0	0	0	1	0	**6**
W2	1	1	0	1	1	0	0	0	0	0	1	0	**5**
W3	1	1	1	0	1	1	0	0	0	0	1	0	**6**
W4	1	1	1	1	0	1	0	0	0	0	1	0	**6**
W5	0	1	0	1	1	0	0	1	0	0	1	0	**5**
W6	0	0	0	0	0	0	0	1	1	1	0	0	**3**
W7	0	0	0	0	0	1	1	0	1	1	0	1	**5**
W8	0	0	0	0	0	0	1	1	0	1	0	1	**4**
W9	0	0	0	0	0	0	1	1	1	0	0	1	**4**
S1	0	1	1	1	1	1	0	0	0	0	0	0	**5**
S4	0	0	0	0	0	0	0	1	1	1	0	0	**3**

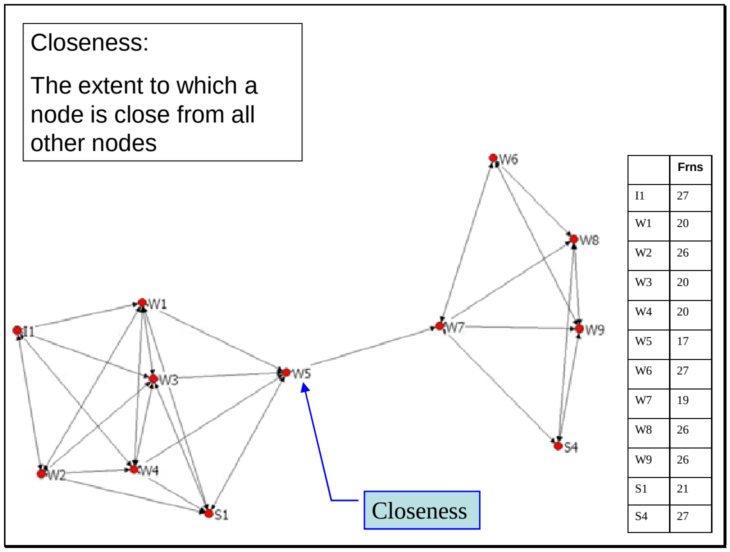
Slide 12

Closeness Centrality



- Is an inverse measure of centrality
- The extent to which a node is close (or “far”) from all other nodes
- Index of expected time until arrival for given node of whatever is flowing through the network
  - –Gossip network: central player hears things first

Slide 13



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

Slide 14

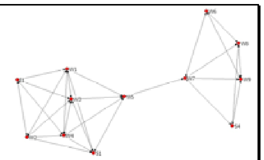
Closeness Centrality

Sum of distances to all other nodes

	I1	W1	W2	W3	W4	W5	W6	W7	W8	W9	S1	S4	Sum
I1	0	1	1	1	1	2	4	3	4	4	2	4	**27**
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**
S4	4	3	4	3	3	2	2	1	1	1	3	0	**27**

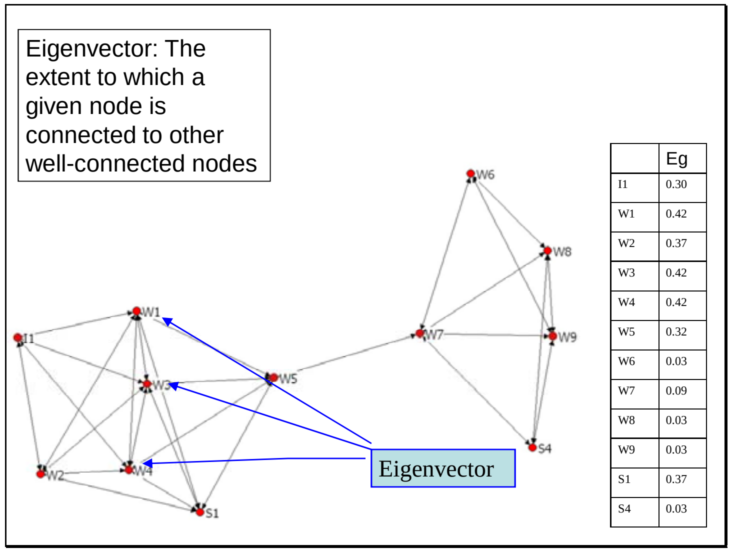
Slide 15

Eigenvector Centrality

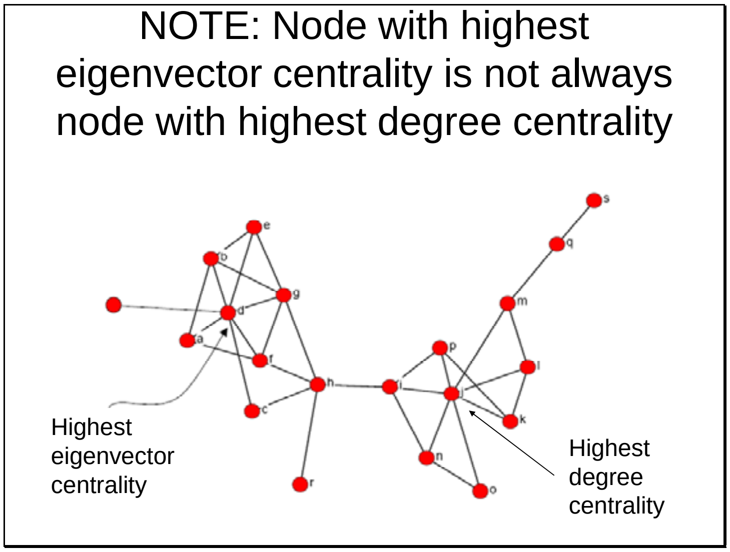


- Node has high score if connected to many nodes that are themselves well connected
- Indicator of popularity,
  - "in the thick of things"
- Like degree, is index of exposure, risk
- Tends to identify centers of large cliques

Slide 16

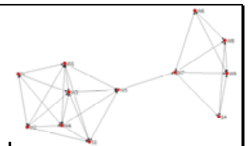

Eigenvector: The extent to which a given node is connected to other well-connected nodes
W6
W8
W1
I1
W7
W9
W5
W3
S4
W2
W4
Eigenvector
S1
Eg
I1 0.30
W1 0.42
W2 0.37
W3 0.42
W4 0.42
W5 0.32
W6 0.03
W7 0.09
W8 0.03
W9 0.03
S1 0.37
S4 0.03

Slide 17


NOTE: Node with highest eigenvector centrality is not always node with highest degree centrality
Highest eigenvector centrality
Highest degree centrality

Slide 18

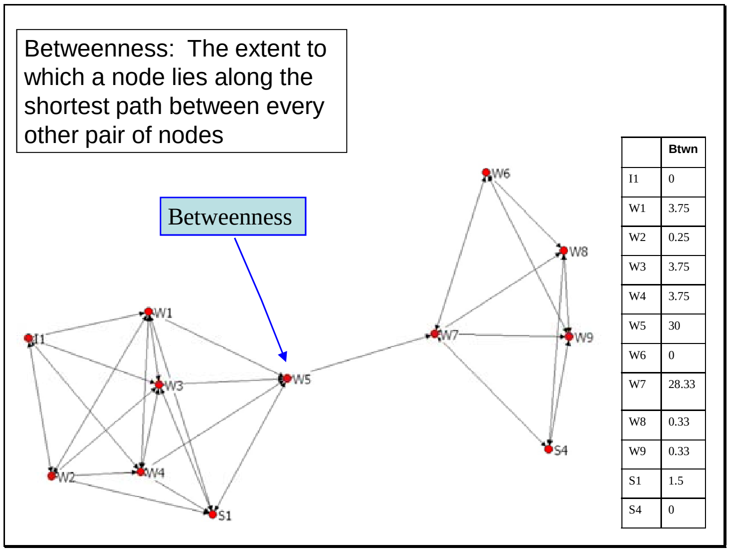
Betweenness Centrality



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

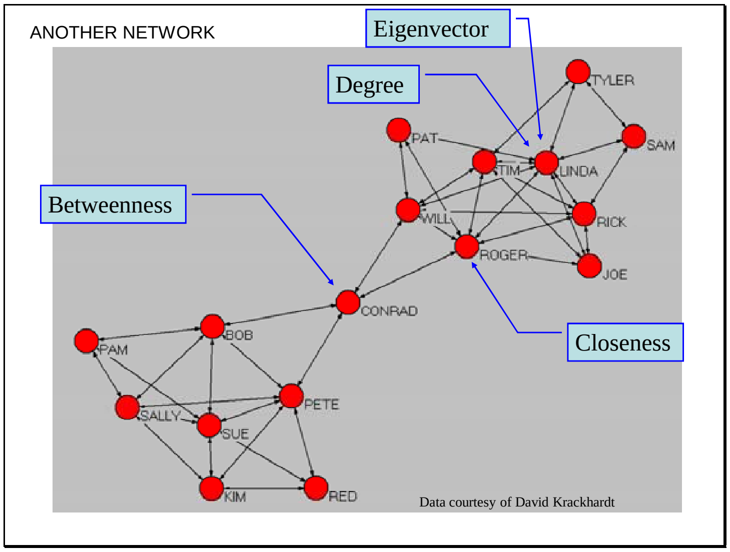
    where g_{ij} is number of geodesic paths from i to j and g_{ikj} is number of those paths that pass through k
- Index of potential for gatekeeping, brokering, controlling the flow, and also of liaising otherwise separate parts of the network
- Interpreted as indicating power and access to diversity of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

Slide 19


Betweenness: The extent to which a node lies along the shortest path between every other pair of nodes
Betweenness
W6
W8
W1
I1
W7
W9
W3
W5
W2
W4
S4
S1
Btwn
I1 0
W1 3.75
W2 0.25
W3 3.75
W4 3.75
W5 30
W6 0
W7 28.33
W8 0.33
W9 0.33
S1 1.5
S4 0

____
____
____
____
____
____
____
____
____
____
____

Slide 20


ANOTHER NETWORK
Eigenvector
Degree
Betweenness
Closeness
TYLER
SAM
PAT
TIM
LINDA
WILL
RICK
ROGER
JOE
CONRAD
BOB
PAM
SALLY
PETE
SUE
KIM
RED
Data courtesy of David Krackhardt

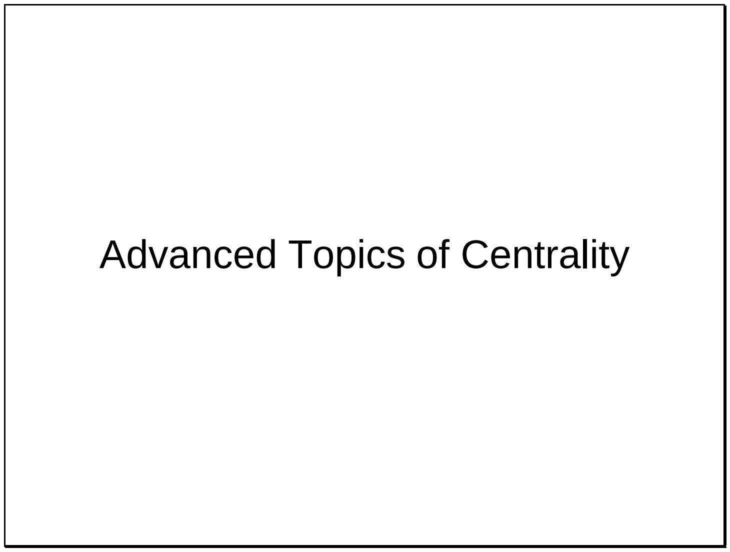
____
____
____
____
____
____
____
____
____
____
____

Slide 21

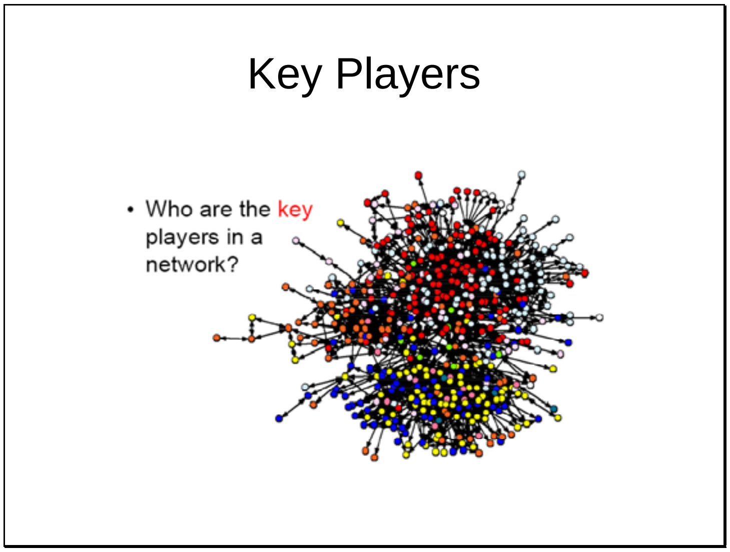
Data Types: Centrality

	Disconnected or Connected	Binary or Valued	Directed or Undirected
Degree	Both	Both	Both
Closeness	Strongly Connected	Binary	Both
Betweenness	Both	Binary (see Brandes, 2001 for discussion of Valued)	Both
Eigenvector	Connected	Both	Undirected*

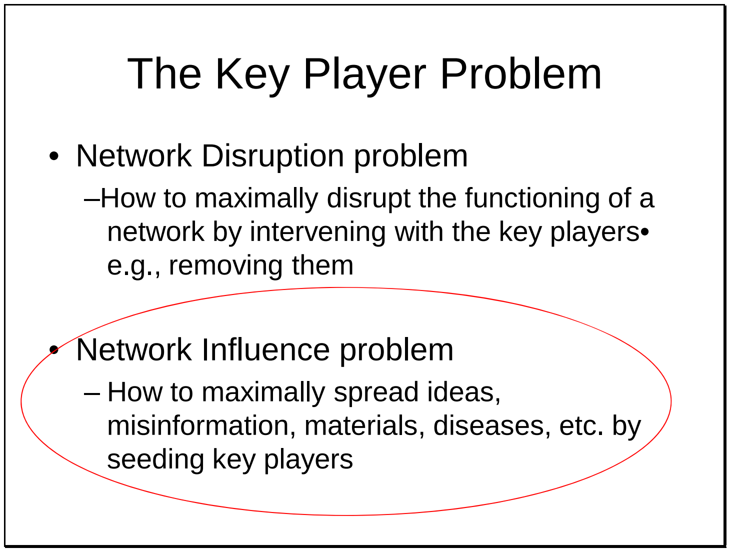
Slide 22



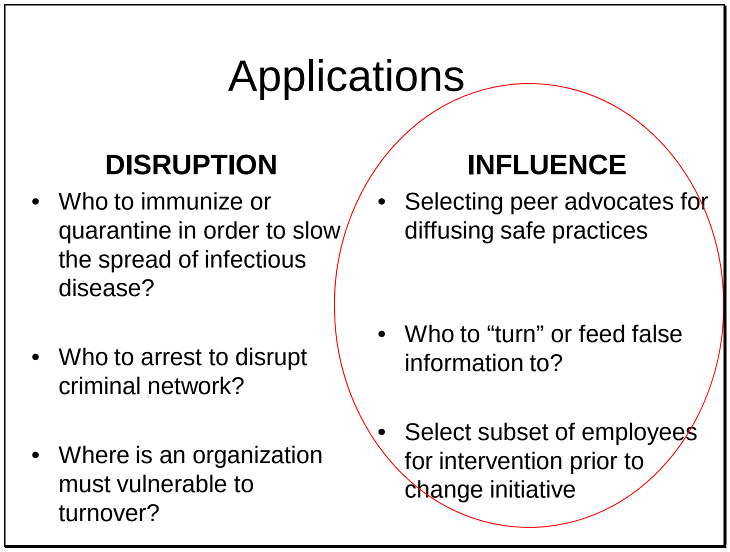
Slide 23



Slide 24


The Key Player Problem
• Network Disruption problem
–How to maximally disrupt the functioning of a network by intervening with the key players• e.g., removing them
• Network Influence problem
– How to maximally spread ideas, misinformation, materials, diseases, etc. by seeding key players

Slide 25


Applications
DISRUPTION
• Who to immunize or quarantine in order to slow the spread of infectious disease?
• Who to arrest to disrupt criminal network?
• Where is an organization must vulnerable to turnover?
INFLUENCE
• Selecting peer advocates for diffusing safe practices
• Who to “turn” or feed false information to?
• Select subset of employees for intervention prior to change initiative

Slide 26

The Naïve Approach

- Use UCINET to identify nodes that should be influenced

1. Calculate degree centrality
2. Identify top set of actors with highest degree centrality and use them to influence others

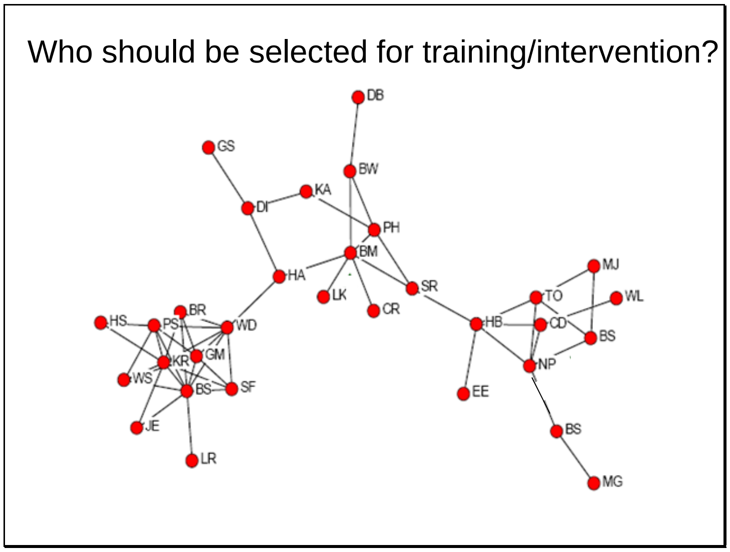
Slide 27

But this fails for multiple reasons

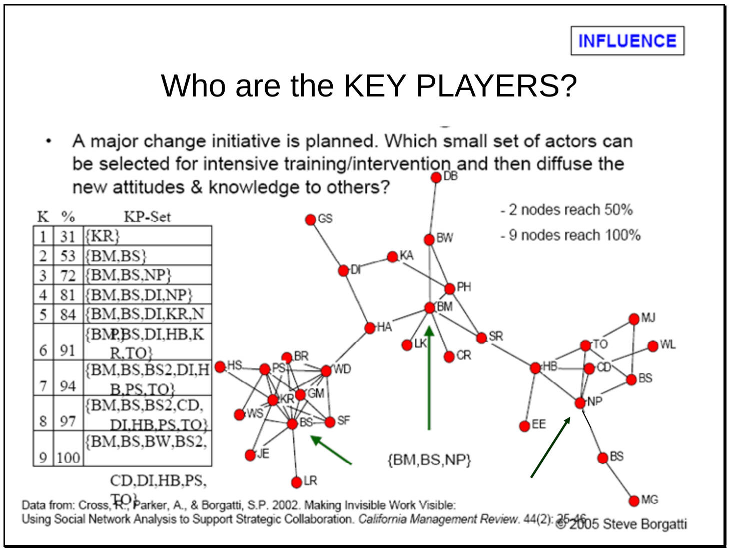
1. The **DESIGN** Issue
   - Centrality measures not specifically designed for our specific problems, so are sub-optimal

2. The **ENSEMBLE** issue
   - Centrality measures are node-level, not group-level concepts
   - The optimal SET of players is not the same as the set of players that are INDIVIDUALLY optimal
     - What if the people with the top two measures of degree centrality know ALL of the same people. What good will it do to influence BOTH of these individuals?

Slide 28


Who should be selected for training/intervention?
DB
GS
BW
KA
DI
PH
BM
HA
LK
CR
SR
MJ
TO
WL
HS
PS
BR
WD
HB
CD
BS
KR
GM
NP
WS
BS
SF
EE
JE
BS
LR
MG

Slide 29


INFLUENCE
Who are the KEY PLAYERS?
• A major change initiative is planned. Which small set of actors can be selected for intensive training/intervention and then diffuse the new attitudes & knowledge to others?
- 2 nodes reach 50%
- 9 nodes reach 100%
K % KP-Set
1 31 {KR}
2 53 {BM,BS}
3 72 {BM,BS,NP}
4 81 {BM,BS,DI,NP}
5 84 {BM,BS,DI,KR,NP}
6 91 {BM,BS,DI,HB,KR,TO}
7 94 {BM,BS,BS2,DI,HB,PS,TO}
8 97 {BM,BS,BS2,CD,DI,HB,PS,TO}
9 100 {BM,BS,BW,BS2,CD,DI,HB,PS,TO}
{BM,BS,NP}
Data from: Cross, R., Parker, A., & Borgatti, S.P. 2002. Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration. California Management Review. 44(2): 25-46.
©2005 Steve Borgatti

Centrality Lab

For this lab we will use 3 datasets:

WIRING:

This is a stacked dataset that includes many different files. We will be working with RDGAM. This is a dichotomous adjacency matrix of 14 employees of the bank wiring room of Western Electric. Ties are symmetric and represent participation in games during work breaks.

PRISON:

This is a dichotomous adjacency matrix of 67 prisoners. Ties are directed and represent each ego's friends. Each was free to choose as few or as many "friends" as he desired.

DRUGNET:

This is a dichotomous adjacency matrix of drug users in Hartford. Ties are directed and represent the lending of drug needles. We will also work with the attribute file **DRUGATTR**.

EXERCISES:

1) Centrality using UCINET and NetDraw with **RDGAM**

If you have not done so already use UCINET to unpack **WIRING**

a) Open **RDGAM** in Netdraw to familiarize yourself with the data
In UCINET calculate the following measures of cohesion using
Network | Centrality
- Degree
- Betweeness
- Closeness
- Eigenvector

b) Using your Netdraw visualization, compare your calculations of various Centrality measures

c) Now run Centrality multiple measures in UCINET using Network | Centrality | Multiple measures

d) Compare the profile of W1 with W5 across all measures. Note that W1 is stronger in eigenvector while W5 is stronger on betweenness. Interpret this result

e) Compare W5 with W7. They have same degree yet differ on eigenvector centrality. Why is W7 so much weaker on eigenvector centrality?

f) Remove isolates using Data | Remove Isolates on **RDGAM** and recalculate centrality measures

g) Compare the results for closeness centrality (especially the descriptive statistics) with those from the previous run.

2) Directed Centrality using UCINET with **PRISON**

a) Open **PRISON** in NetDraw to familiarize yourself with the data

b) Using UCINET calculate Centrality measures (first, using individual measures, and next using the “multiple measures” option) and compare the results. What is the program doing?

c) Identify which individuals have the most friends in this dataset

3) Directed Centrality using NetDraw with **PRISON**

a) Open **PRISON** in NetDraw

b) Using NetDraw calculate Centrality measures under Analysis | Centrality
c) Resize the nodes based on various Centrality measures
d) Identify which individuals list the most number of friends
e) Identify which individuals are listed as friends by the most number of others

4) Directed Centrality using UCINET with **DRUGNET**

a) Open **DRUGNET** in NetDraw to familiarize yourself with the data

b) Using UCINET identify which individuals are at highest risk of contracting a disease based on their needle sharing habits

c) In NetDraw, open **DRUGATTR** by clicking on the folder with the A

d) Calculate Centrality measures in NetDraw (remember that this is directed data)

e) Using NetDraw color the nodes based on different attributes and size the nodes based on different Centrality measures. Do you see any pattern?