Working with Node Attributes

Steve Borgatti

2009 LINKS Center Workshop

ADVANCED Session

Egonet Composition

- Concept
  - Characterize the alters that an ego has ties a given type of ties with
    - What proportion of friends are men, women, gay, etc.
    - Average income of a person's friends
- Can define ego network in terms of outgoing ties or incoming ties
  - What kinds of actors are nominating a given node
  - What kinds of actors is a given actor nominating?
- Categorical & continuous alter attributes
  - Gender, department, religion, etc.
  - Age, income, centrality, structural holes, etc. of the alter

Egonet Strength

- Concept
  - Average value of an alter attribute, such as wealth or power
  - Social capital ala social resource theory (Lin)
- Measure
  - Given adjacency matrix A and attribute vector v, the matrix product Av gives the sum of attribute values for the alters of each node
  - Can also compute average, maximum, minimum etc

$$s_i = \frac{\sum_j a_{ij} v_j}{\sum_j a_{ij}}$$

Egonet Heterogeneity

- Concept
  - Measure the diversity of an actor's contacts
- For categorical attributes, e.g., gender
  - use Blau|Herfindahl index $H = \sum_k p_k^2$ where pk gives the proportion of alters that fall into category k
- For continuous attributes, e.g., wealth of alters
  - Calculate standard deviation

Homophily

- Concept
  - Extent to which actors tend to have ties with actors who are similar to themselves
    - E.g., girls confide mostly to girls, boys to boys
- Dozens of Measures
  - Pct homophilous matches
  - E-I
  - Correlation

“converting” attributes to matrices

- Problem
  - Given vector **v** representing a node attribute, construct a matrix X
- Categorical attributes
  - $x_{ij} = 1$ if $v_i = v_j$, and $x_{ij} = 0$ otherwise
- Continuous attributes
  - $x_{ij} = v_i - v_j$
  - $x_{ij} = (v_i - v_j)^2$

“converting” attributes to matrices

Categorical example

	V
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	2
10	2
11	2
12	2
13	2
14	2
15	2
16	2
17	2
18	2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
10	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
11	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
12	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
13	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
14	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
15	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
16	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
17	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
18	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1

$x_{ij} = 1$ if $v_i = v_j$, and $x_{ij} = 0$ otherwise

Set-up for homophily measures

- Given
  - A social relation R
  - A categorical attribute vector **a**
- Construct
  - Similarity relation S in which $s_{ij} = 1$ if $a_i = a_j$, and $s_{ij} = 0$ otherwise

R

	f	f	f	f	f	f	f	f	m	m	m	m	m	m	m	m	m	m
f	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
f	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
f	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
f	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
m	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
m	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0
m	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
m	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
m	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1
m	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1
m	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
m	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1
m	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0

a

1
1
1
1
1
1
1
1
2
2
2
2
2
2
2
2
2
2

S

	f	f	f	f	f	f	f	f	m	m	m	m	m	m	m	m	m	m
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1

Homophily set-up - cont

- Construct "relational contingency table"

		S 1	S 0
R	1	a	b
R	0	c	d

R = the data – the social relation
S = similarity – is 1 if same attrib value

- Campnet dataset:

		S 1	S 0
R	1	45	9
R	0	101	151

Pct homophilous matches (H%)

- H% = a/(a+b)

S

R		1	0
	1	a	b
	0	c	d

- Campnet dataset

S

R		1	0
	1	45	9
	0	101	151

H% = 0.83

E-I index

- Krackhardt & Stern
  - Number external ties minus number of internal ties as a proportion of all ties

		S 1	S 0
R	1	a	b
R	0	c	d

$$EI = \frac{b-a}{b+a}$$

Negative values indicated greater homophily

- Campnet

		S 1	S 0
R	1	45	9
R	0	101	151

EI = -0.667

Point bi-serial correlation (pbsc) approach

- Take into account non-choices as well:

R \ S	1	0
1	a	b
0	c	d

$$r(R,S) = \frac{ad - bc}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

- Campnet dataset:

R \ S	1	0
1	45	9
0	101	151

r(R,S) = 0.33

H% = 0.83

R \ S	1	0
1	45	9
0	90	18

r(R,S) = 0.00

H% = 0.83

Correlation

- The pbsc measure is the same as a QAP correlation of the two dyadic variables R and S

r(R,S) = 0.33

	f	f	f	f	f	f	f	f	m	m	m	m	m	m	m	m	m	m
f	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
f	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0
f	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
f	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
f	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
m	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
m	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0
m	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
m	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
m	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1
m	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1
m	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
m	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1
m	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0

R

	f	f	f	f	f	f	f	f	m	m	m	m	m	m	m	m	m	m
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
f	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
m	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1

S

Egonet Homophily

- Concept
  - To what extent an ego's alters are like ego on a given attribute
- Approach
  - Construct relational contingency table for each node
- Measures
  - Pct homophilous (%H) = 0.67
  - E-I index = -0.333
  - PBSC = 0.24

HOLLY

		S: 1	S: 0
R	1	2	1
	0	5	9

Density Tables

- Concept
  - Number of ties within and between groups
  - Called density when expressed as a function of the number possible

	BHS	CCG	DCL	ES	HEW	IS	MS	SRG	STAT	TAS	N
BHS	1613	356	239	1601	717	74	862	231	576	239	178
CCG	356	42	272	329	206	59	228	32	58	231	15
DCL	239	272	2117	521	616	844	1005	61	79	1541	177
ES	1601	329	521	968	445	117	712	119	326	416	146
HEW	717	206	616	445	374	64	380	67	161	245	89
IS	74	59	844	117	64	188	159	16	15	397	52
MS	862	228	1005	712	380	159	666	84	263	845	134
SRG	231	32	61	119	67	16	84	15	49	32	20
STAT	576	58	79	326	161	15	263	49	62	88	19
TAS	239	231	1541	416	245	397	845	32	88	1397	130
N	178	15	177	146	89	52	134	20	19	130	960

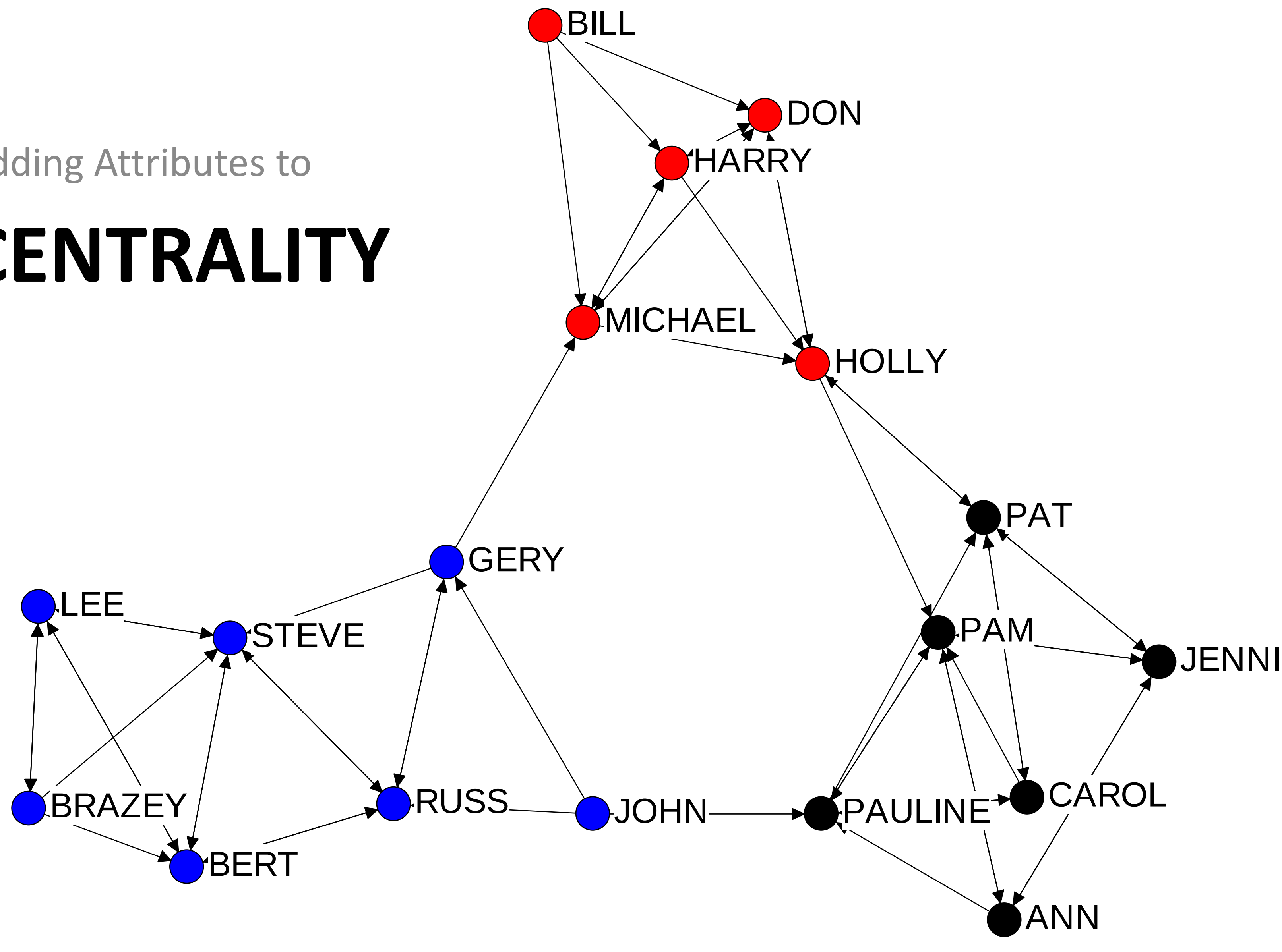
Tie Frequencies

	BHS	CCG	DCL	ES	HEW	IS	MS	SRG	STAT	TAS
BHS	0.10	0.13	0.01	0.06	0.05	0.01	0.04	0.06	0.17	0.01
CCG	0.13	0.40	0.10	0.15	0.15	0.08	0.11	0.11	0.20	0.12
DCL	0.01	0.10	0.14	0.02	0.04	0.09	0.04	0.02	0.02	0.07
ES	0.06	0.15	0.02	0.09	0.03	0.02	0.04	0.04	0.12	0.02
HEW	0.05	0.15	0.04	0.03	0.10	0.01	0.03	0.04	0.10	0.02
IS	0.01	0.08	0.09	0.02	0.01	0.14	0.02	0.02	0.02	0.06
MS	0.04	0.11	0.04	0.04	0.03	0.02	0.07	0.03	0.10	0.05
SRG	0.06	0.11	0.02	0.04	0.04	0.02	0.03	0.08	0.13	0.01
STAT	0.17	0.20	0.02	0.12	0.10	0.02	0.10	0.13	0.36	0.04
TAS	0.01	0.12	0.07	0.02	0.02	0.06	0.05	0.01	0.04	0.17

Densities

Adding Attributes to

CENTRALITY



Individual Level E-I Index

- We can partition degree centrality into inward and outward components
- Can we similarly partition other centrality measures the same way?

Degree centrality

		Intern	Extern	Total	E-I
1	HOLLY	4.000	5.000	9.000	0.111
2	BRAZEY	5.000	3.000	8.000	-0.250
3	CAROL	5.000	3.000	8.000	-0.250
4	PAM	6.000	3.000	9.000	-0.333
5	PAT	6.000	1.000	7.000	-0.714
6	JENNIE	5.000	0.000	5.000	-1.000
7	PAULINE	6.000	2.000	8.000	-0.500
8	ANN	7.000	1.000	8.000	-0.750
9	MICHAEL	3.000	4.000	7.000	0.143
10	BILL	2.000	1.000	3.000	-0.333
11	LEE	1.000	5.000	6.000	0.667
12	DON	2.000	3.000	5.000	0.200
13	JOHN	1.000	7.000	8.000	0.750
14	HARRY	5.000	1.000	6.000	-0.667
15	GERY	3.000	3.000	6.000	0.000
16	STEVE	3.000	4.000	7.000	0.143
17	BERT	3.000	4.000	7.000	0.143
18	RUSS	3.000	4.000	7.000	0.143

EI Degree Centrality

- We can partition degree centrality into inward and outward components
- Can we similarly partition other centrality measures the same way?

		Intern	Extern	Degree centrality
		------	------	------
1	HOLLY	4.000	5.000	9.000
2	BRAZEY	5.000	3.000	8.000
3	CAROL	5.000	3.000	8.000
4	PAM	6.000	3.000	9.000
5	PAT	6.000	1.000	7.000
6	JENNIE	5.000	0.000	5.000
7	PAULINE	6.000	2.000	8.000
8	ANN	7.000	1.000	8.000
9	MICHAEL	3.000	4.000	7.000
10	BILL	2.000	1.000	3.000
11	LEE	1.000	5.000	6.000
12	DON	2.000	3.000	5.000
13	JOHN	1.000	7.000	8.000
14	HARRY	5.000	1.000	6.000
15	GERY	3.000	3.000	6.000
16	STEVE	3.000	4.000	7.000
17	BERT	3.000	4.000	7.000
18	RUSS	3.000	4.000	7.000

EI - Eigenvector

		Internal	External	Total
		-------	-------	-------
1	HOLLY	-0.132	-0.243	-0.375
2	BRAZEY	0.000	-0.097	-0.097
3	CAROL	-0.196	0.000	-0.196
4	PAM	-0.291	0.000	-0.291
5	PAT	-0.247	0.000	-0.247
6	JENNIE	-0.176	0.000	-0.176
7	PAULINE	-0.224	-0.038	-0.262
8	ANN	-0.179	0.000	-0.179
9	MICHAEL	-0.266	-0.092	-0.357
10	BILL	-0.243	0.000	-0.243
11	LEE	-0.073	-0.024	-0.097
12	DON	-0.225	-0.092	-0.317
13	JOHN	-0.090	-0.064	-0.154
14	HARRY	-0.225	-0.092	-0.317
15	GERY	-0.206	0.000	-0.206
16	STEVE	-0.145	-0.024	-0.169
17	BERT	-0.105	-0.024	-0.128
18	RUSS	-0.161	0.000	-0.161

- $Av = \lambda v$

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

- So i's score is sum of scores of those adjacent to him
- Easily partitioned into groups

$$v_{ig} = \lambda^{-1} \sum_{j \in g} a_{ij} v_j$$

Closeness Centrality

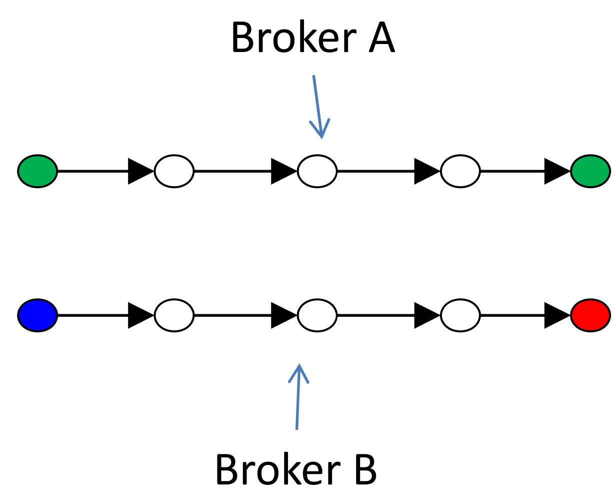
		Internal	External	Total
1	HOLLY	14	24	38
2	BRAZEY	34	25	59
3	CAROL	14	34	48
4	PAM	12	30	42
5	PAT	13	30	43
6	JENNIE	15	40	55
7	PAULINE	12	28	40
8	ANN	14	34	48
9	MICHAEL	16	20	36
10	BILL	22	28	50
11	LEE	24	35	59
12	DON	22	21	43
13	JOHN	20	18	38
14	HARRY	22	21	43
15	GERY	14	22	36
16	STEVE	17	28	45
17	BERT	22	29	51
18	RUSS	17	23	40

- Separately sum distances to group insiders and outsiders
- Call this EI Closeness

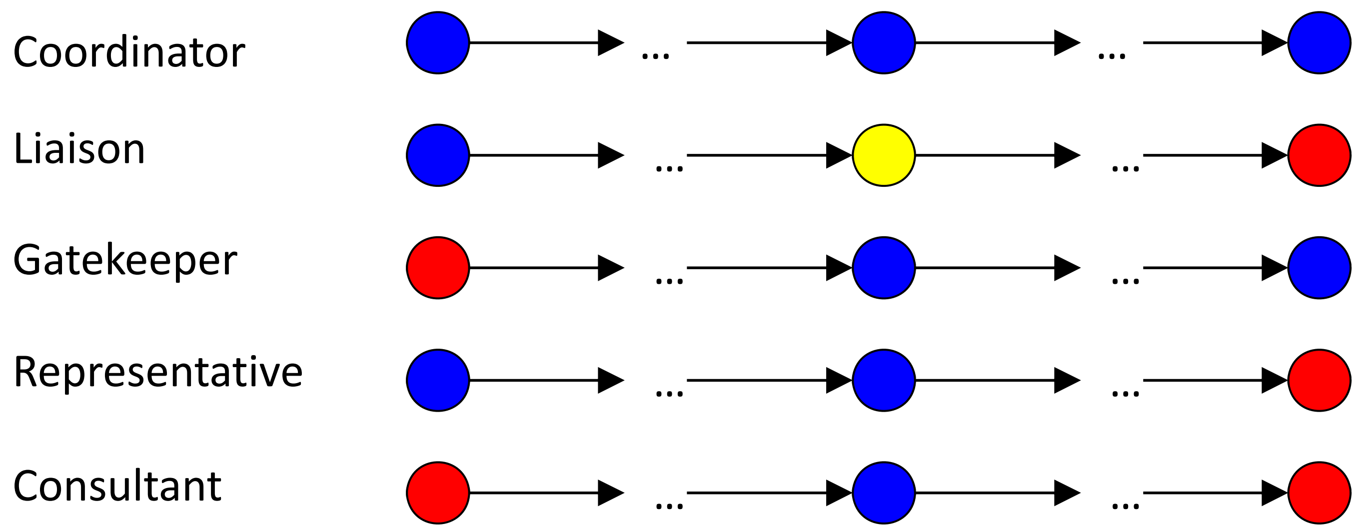
EI Betweenness

		Same	Different	Total
1	HOLLY	1.86	53.04	54.91
2	BRAZEY	0.00	0.00	0.00
3	CAROL	0.50	0.00	0.50
4	PAM	8.39	24.47	32.86
5	PAT	5.39	16.47	21.86
6	JENNIE	1.50	0.00	1.50
7	PAULINE	11.14	55.34	66.47
8	ANN	1.21	3.45	4.66
9	MICHAEL	39.36	27.52	66.88
10	BILL	0.00	0.00	0.00
11	LEE	0.00	0.00	0.00
12	DON	0.00	4.57	4.57
13	JOHN	10.64	54.96	65.59
14	HARRY	0.00	4.57	4.57
15	GERY	48.58	30.11	78.69
16	STEVE	27.79	23.79	51.58
17	BERT	5.21	5.21	10.42
18	RUSS	16.42	26.51	42.93

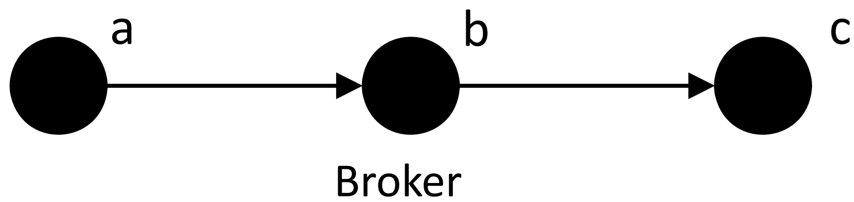
- A node can be along the shortest path between nodes that belong to the same group as each other (broker A), or to different groups (broker B)



Extending betweenness ala Gould & Fernandez brokerage

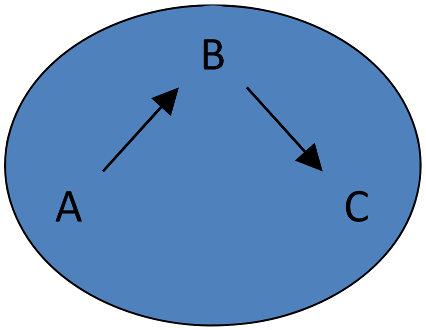


Brokerage Roles

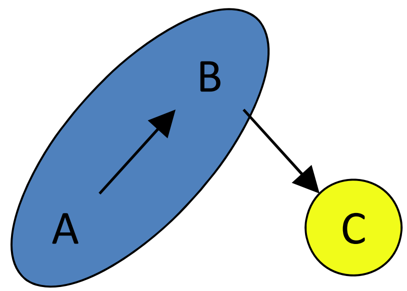


- Gould & Fernandez
- Broker is middle node of directed triad
- What if nodes belong to different organizations?

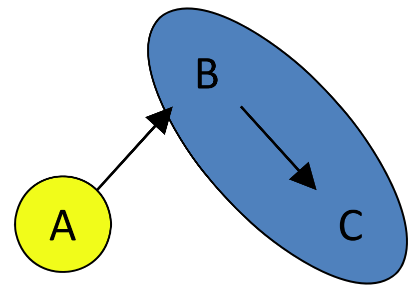
Gould & Fernandez
Brokerage Roles



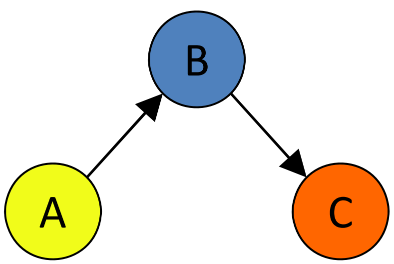
Coordinator



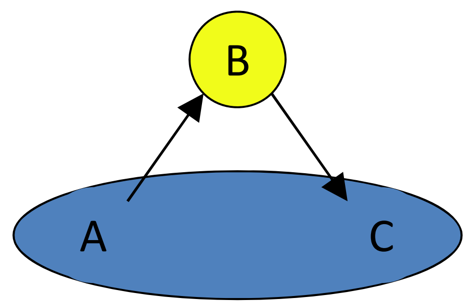
Representative



Gatekeeper



Liaison



Consultant

Example

	Coord	Gate	Rep	Cons	Liais	Total
JB	3	17	1	0	3	24
TB	0	5	0	4	5	14
MC	1	0	0	0	0	1
CC	0	0	0	0	5	5
BD	1	0	40	0	0	41
TD	5	5	45	8	25	88
PD	0	0	0	0	0	0
JF	0	0	0	0	0	0
KG	7	22	9	0	15	53
SM	0	1	0	0	0	1
BS	1	0	0	0	0	1
AS	0	0	0	0	0	0
JT	0	0	0	0	0	0
PW	0	30	0	0	0	30
CW	0	6	0	3	5	14
TW	0	0	0	0	0	0
Total	18	86	95	15	58	272

Role Profiles