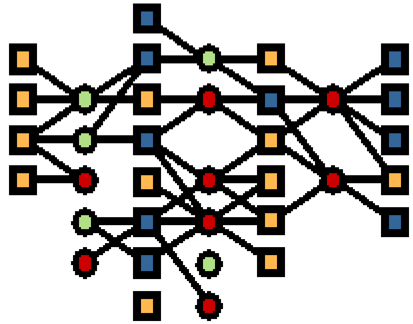


things I have wanted to say about

the analysis of 2-mode networks

but hadn't had the opportunity to. Until now.

Steve Borgatti

LINKS Center, University of Kentucky

(with help from Martin Everett & Dan Halgin)

Conference and Workshop on Two-Mode Social Network Analysis

1 October, 2009 VU Uni versity Amsterdam

Agenda

- What are 2-mode networks?
  - Establishing the domain of this conference
- Techniques
  - Conversion-based
  - Direct analyses
- Under-developed areas
  - Trajectories
  - 2-mode relational algebras
- Conclusion

Section 1

WHAT ARE 2-MODE NETWORKS?

The classic example

Names of Participants of Group I	Code Numbers and Dates of Social Events Reported in *Old City Herald*													
	(1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/16	(9) 4/8	(10) 6/10	(11) 2/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson	×	×	×	×	×	×	....	×	×	....	....	....	....	....
2. Miss Laura Mandeville	×	×	×	....	×	×	×	×	....	....	....	....	....	....
3. Miss Theresa Anderson	....	×	×	×	×	×	×	×	×	....	....	....	....	....
4. Miss Brenda Rogers	×	....	×	×	×	×	×	×	....	....	....	....	....	....
5. Miss Charlotte McDowd	....	....	×	×	×	....	×	....	....	....	....	....	....	....
6. Miss Frances Anderson	....	....	×	....	×	×	....	×	....	....	....	....	....	....
7. Miss Eleanor Nye	....	....	....	....	×	×	×	×	....	....	....	....	....	....
8. Miss Pearl Oglethorpe	....	....	....	....	....	×	....	×	×	....	....	....	....	....
9. Miss Ruth DeSand	....	....	....	....	×	....	×	×	×	....	....	....	....	....
10. Miss Verne Sanderson	....	....	....	....	....	....	×	×	×	....	....	×	....	....
11. Miss Myra Liddell	....	....	....	....	....	....	....	×	×	×	....	×	....	....
12. Miss Katherine Rogers	....	....	....	....	....	....	....	×	×	×	....	×	×	×
13. Mrs. Sylvia Avondale	....	....	....	....	....	....	×	×	×	×	....	×	×	×
14. Mrs. Nora Fayette	....	....	....	....	....	×	×	....	×	×	×	×	×	×
15. Mrs. Helen Lloyd	....	....	....	....	....	....	×	×	....	×	×	×	....	....
16. Mrs. Dorothy Murchison	....	....	....	....	....	....	....	×	×	....	....	....	....	....
17. Mrs. Olivia Carleton	....	....	....	....	....	....	....	....	×	....	×	....	....	....
18. Mrs. Flora Price	....	....	....	....	....	....	....	....	×	....	×	....	....	....

Figure 1. Davis, Gardner and Gardner (1941) *Deep South* women-by-events matrix.

More examples

- Affiliations
  - Attendance at events
  - Membership in orgs
- Items by use
  - Household items by functions
- (Binary) profile matrices
  - Species by trait
  - Componential analyses
- Correspondences/assoc
  - Authors & topics
  - Illnesses and treatments

Term	Horse	Baby	Young	Old	Male	Female	Neuter
Stallion	✓			✓	✓		
Filly	✓			✓		✓	
Gelding	✓				✓		✓
Colt	✓		✓				
Foal	✓	✓					

Modes and Ways terminology

- The ways of a matrix are its dimensions, as in rows and columns (2-way), or rows, columns and levels (3-way)
- Modes of a matrix are the distinct sets of entities pointed to by the ways

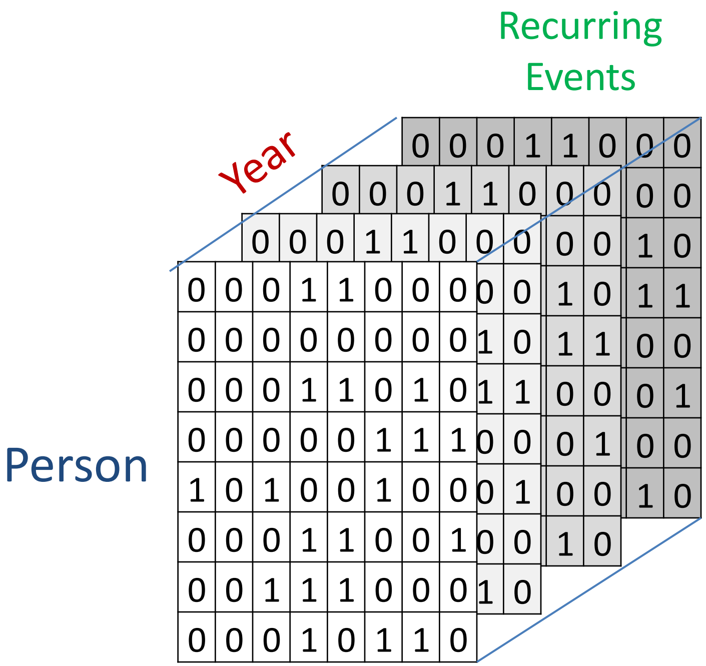
	Mary	Bill	John	Larry
Mary	0	1	0	1
Bill	1	0	0	1
John	0	1	0	0
Larry	1	0	1	0

2-way, 1-mode

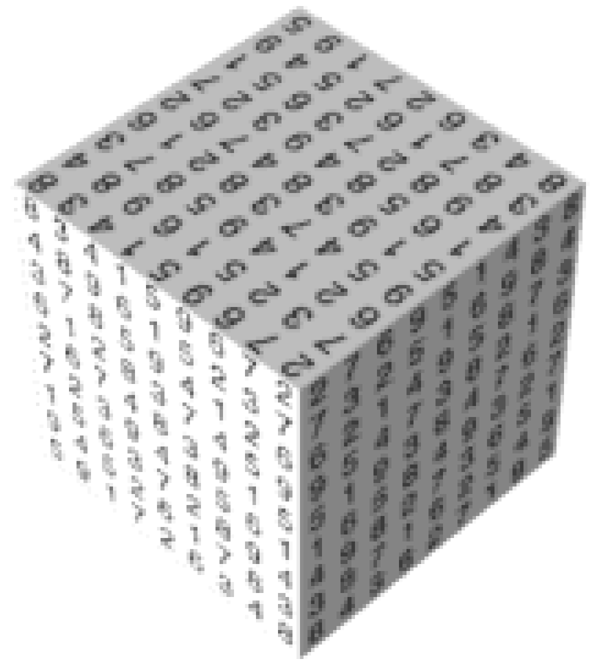
	Event 1	Event 2	Event 3	Event 4
EVELYN	1	1	1	1
LAURA	1	1	1	0
THERESA	0	1	1	1
BRENDA	1	0	1	1
CHARLO	0	0	1	1
FRANCES	0	0	1	0
ELEANOR	0	0	0	0
PEARL	0	0	0	0
RUTH	0	0	0	0
VERNE	0	0	0	0
MYRNA	0	0	0	0

2-way, 2-mode

3-way matrices

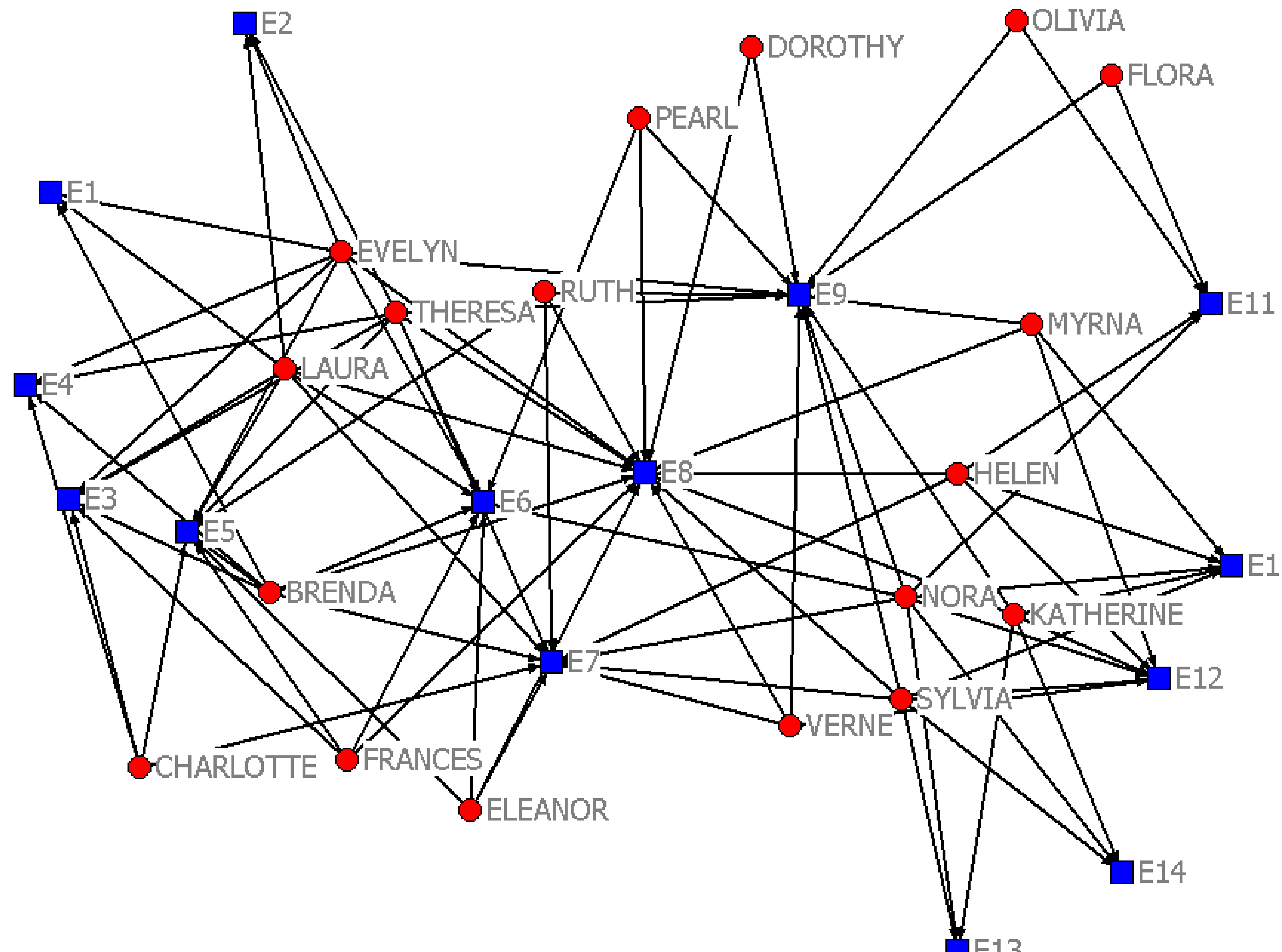


3-way, 3-mode
Longitudinal affiliations data



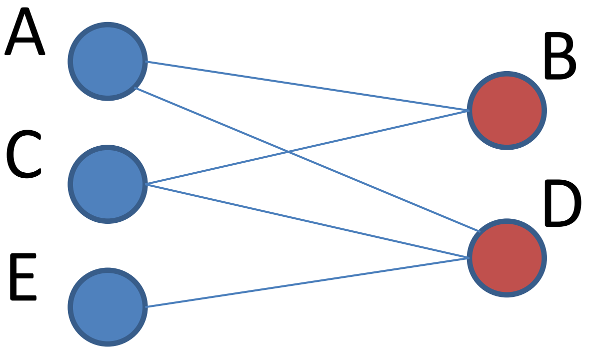
3-way, 1-mode
Krackhardt-style perceptions
by each person of relations
among all pairs of persons

Canonical visualization



Bipartite property

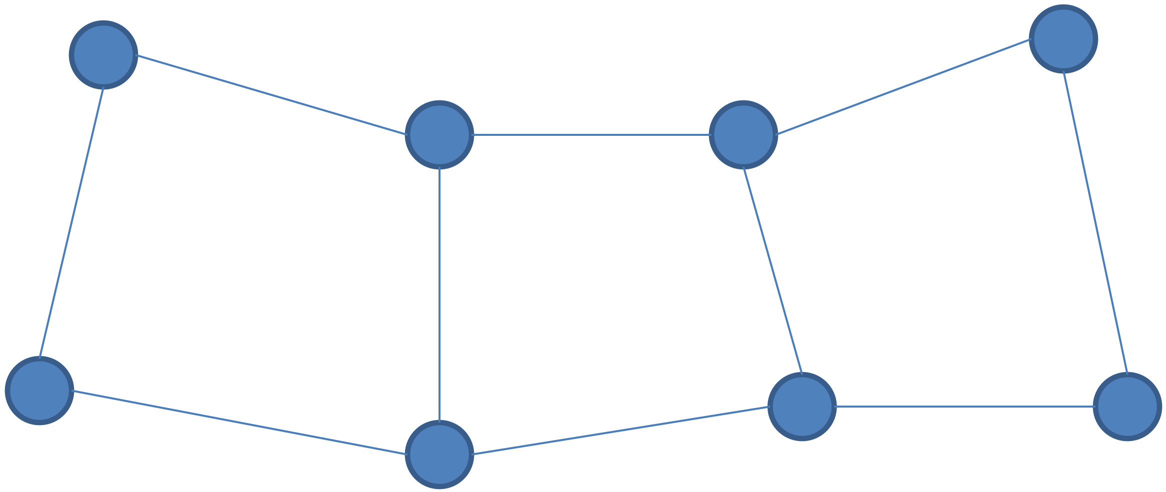
- A graph G(V,E) is bipartite if V can be partitioned into V1, V2 such that for all edges (u,v) in E, *u* belongs to V1 and *v* belongs to V2
  - No ties within V1 and V2 (independent sets), and V1+V2 = V
- 2-colorable
- No odd cycles



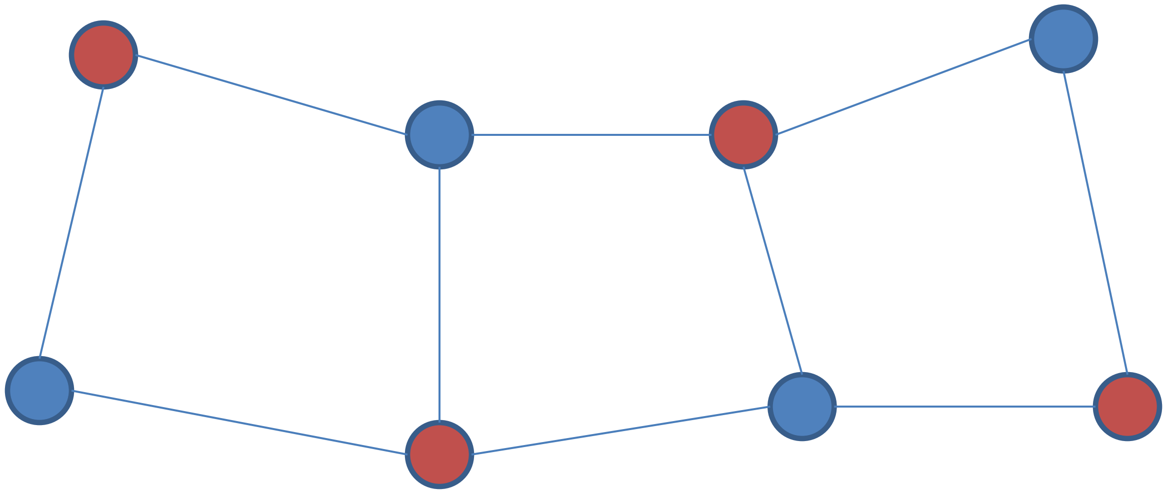
Two small points

- Bipartite and 2-mode are not interchangeable
- “2-mode network” terminology is misleading

Accidental bipartite-ism

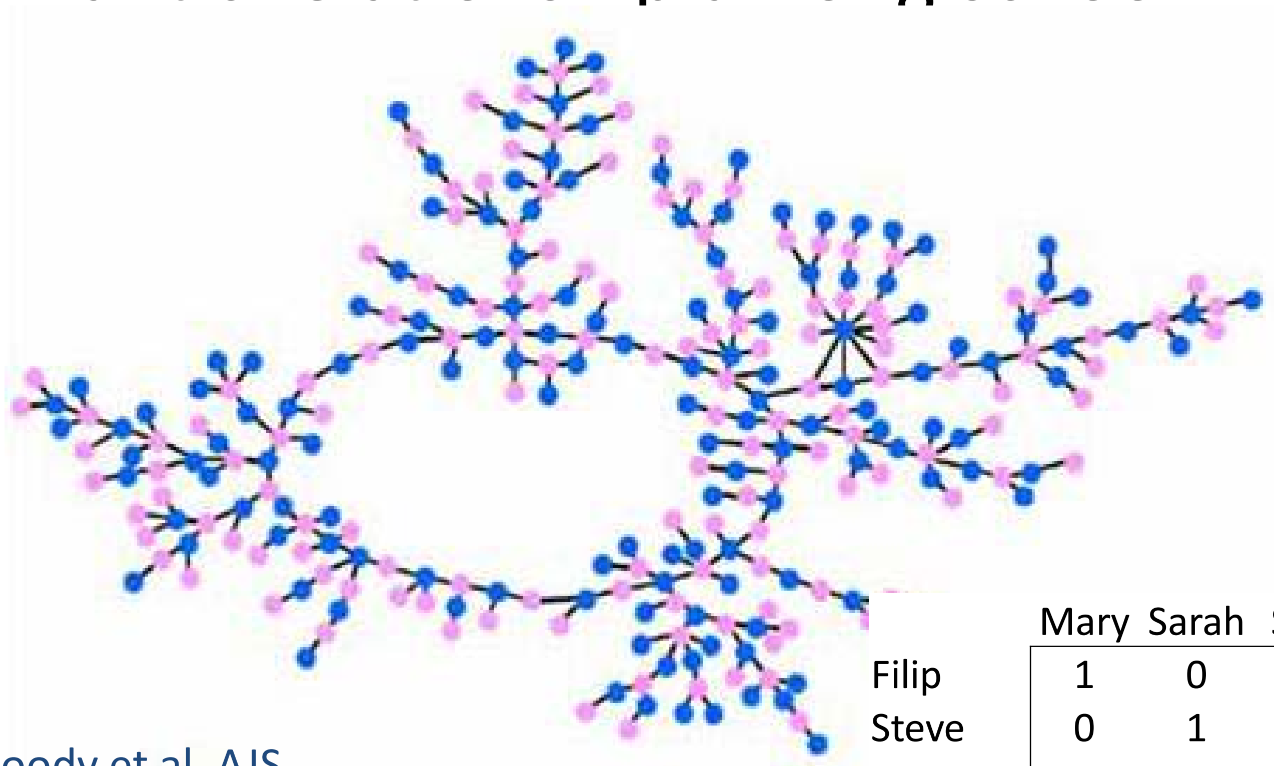


Accidental bipartite-ism



Men are from Mars, women from Venus

- Romantic relationship among school kids

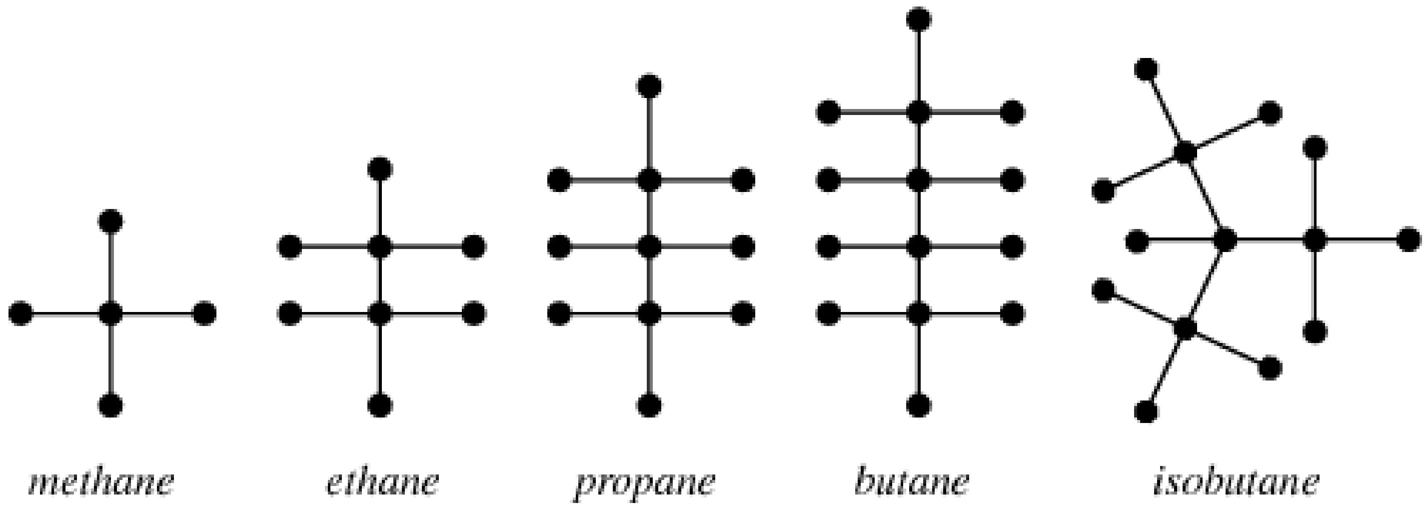


Moody et al. AJS.
Gender identified by color.

	Mary	Sarah	Susan	Nancy	Jane
Filip	1	0	0	1	1
Steve	0	1	0	0	0
Martin	0	0	1	1	1
Tom	0	0	1	0	1
John	1	0	0	1	1
Marco	0	0	0	0	0

Trees

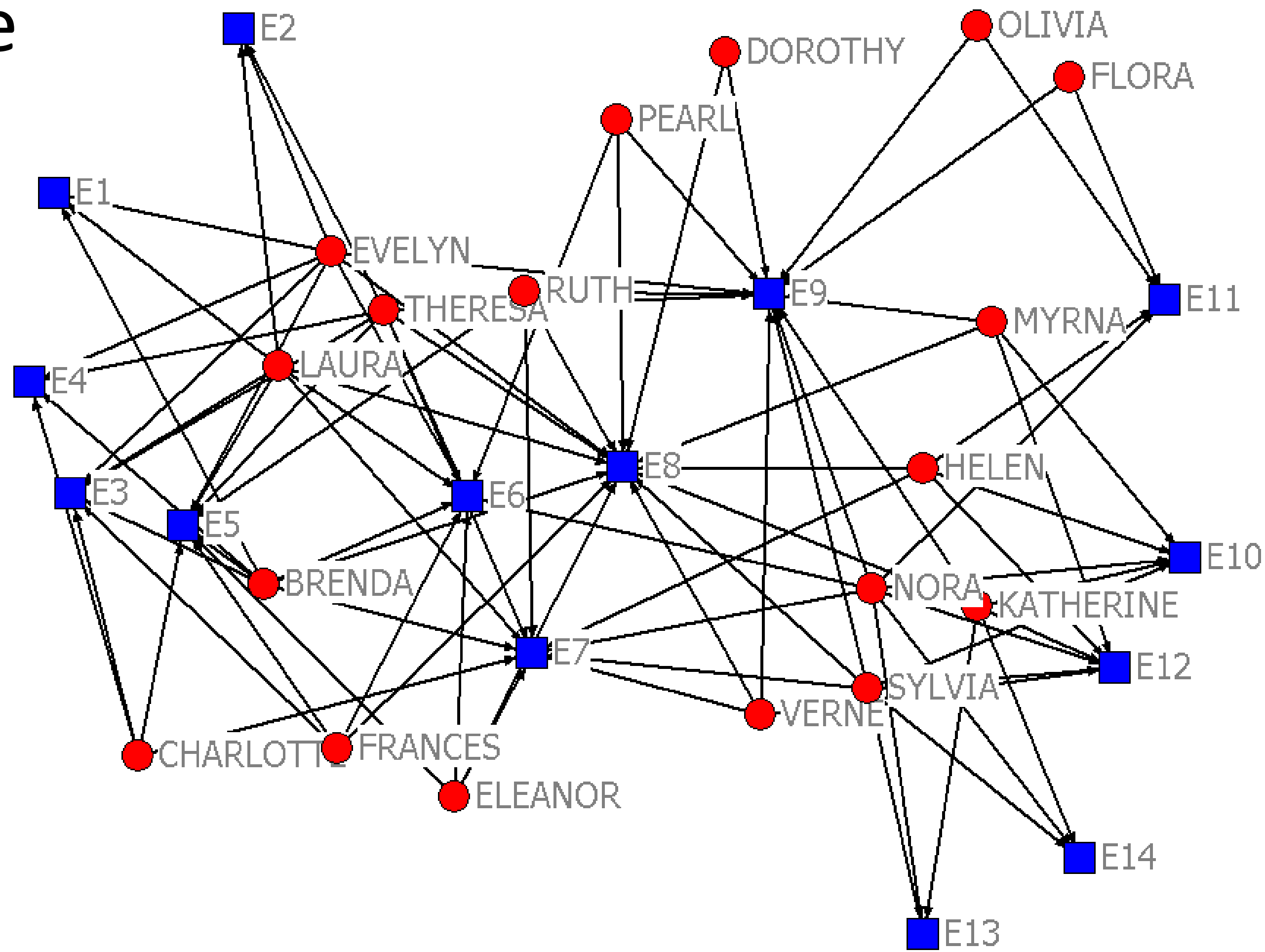
- Every tree is bipartite



- Bipartite does not imply 2-mode

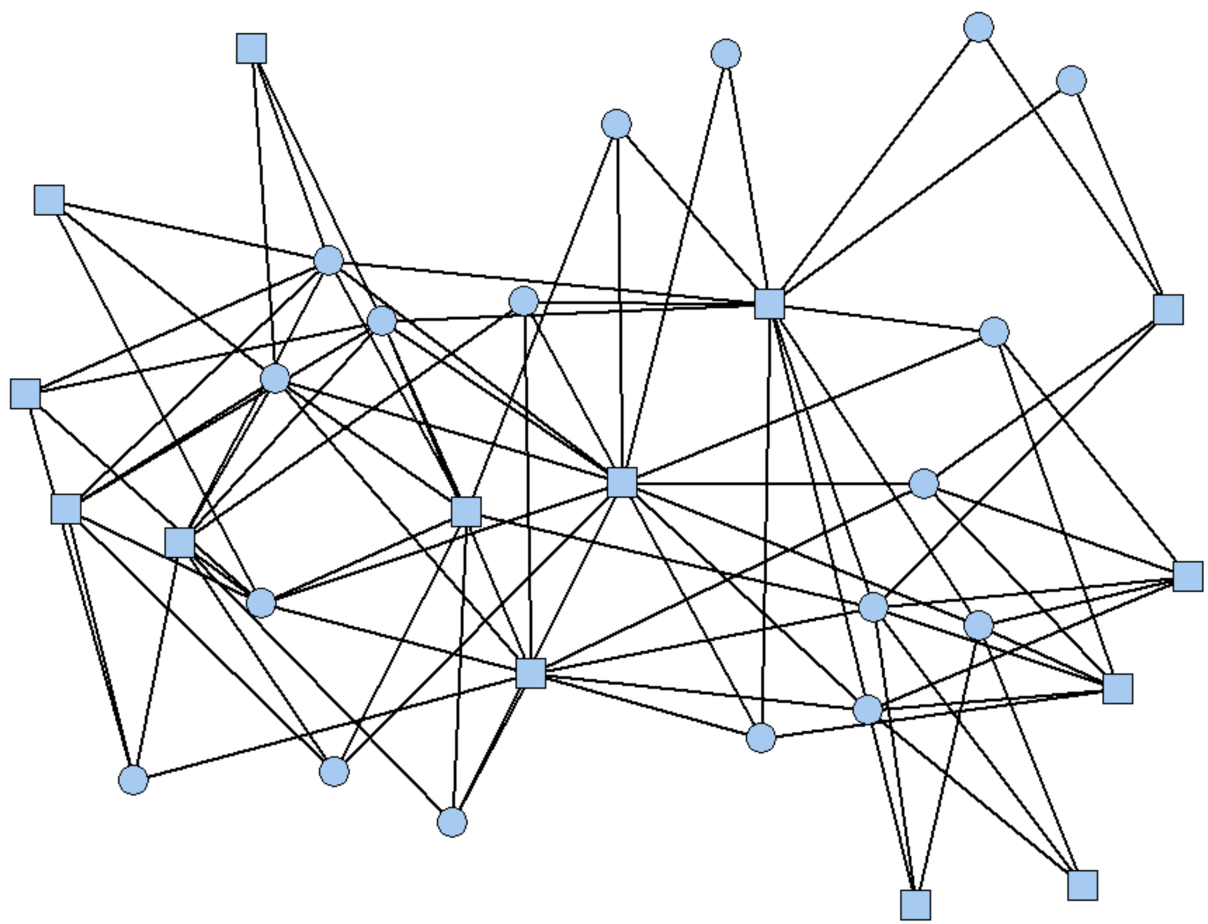
Is 2-mode terminology misleading?

1. The adjacency matrix of a “2-mode network” is 1-mode



Davis data as bipartite graph

1. The adjacency matrix of a “2-mode network” is 1-mode



Davis data without node labels

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1																			1	1	1	1	1	1		1	1					
2																			1	1	1		1	1	1	1						
3																				1	1	1	1	1	1	1	1					
4																			1		1	1	1	1	1	1						
5																					1	1	1		1							
6																					1		1	1		1						
7																							1	1	1	1						
8																								1		1	1					
9																							1		1	1	1					
10																									1	1	1			1		
11																										1	1	1		1		
12																										1	1	1		1	1	1
13																									1	1	1	1		1	1	1
14																								1	1		1	1	1	1	1	1
15																									1	1		1	1	1		
16																										1	1					
17																											1		1			
18																											1		1			
19	1	1		1																												
20	1	1	1																													
21	1	1	1	1	1	1																										
22	1		1	1	1																											
23	1	1	1	1	1	1	1		1																							
24	1	1	1	1		1	1	1						1																		
25		1	1	1	1		1		1	1			1	1	1																	
26	1	1	1	1		1	1	1	1	1	1	1	1		1	1																
27	1		1					1	1	1	1	1	1	1		1	1	1														
28											1	1	1	1	1																	
29														1	1		1	1														
30										1	1	1	1	1	1																	
31												1	1	1																		
32												1	1	1																		

Bi-adjacency matrix

Affiliation graphs

- Object of study is an affiliation graph, consisting of binary relation between 2 sets
- There are at least 2 natural matrix representations, with different modality
- 2-mode terminology confuses relation w/ its matrix representation

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN	1	1	1	1	1	1	0	1	1	0	0	0	0	0
LAURA	1	1	1	0	1	1	1	1	0	0	0	0	0	0
THERESA	0	1	1	1	1	1	1	1	1	0	0	0	0	0
BRENDA	1	0	1	1	1	1	1	1	0	0	0	0	0	0
CHARLOTTE	0	0	1	1	1	0	1	0	0	0	0	0	0	0
FRANCES	0	0	1	0	1	1	0	1	0	0	0	0	0	0
ELEANOR	0	0	0	0	1	1	1	1	0	0	0	0	0	0
PEARL	0	0	0	0	0	1	0	1	1	0	0	0	0	0
RUTH	0	0	0	0	1	0	1	1	1	0	0	0	0	0
VERNE	0	0	0	0	0	0	1	1	1	0	0	1	0	0
MYRNA	0	0	0	0	0	0	0	1	1	1	0	1	0	0
KATHERINE	0	0	0	0	0	0	0	1	1	1	0	1	1	1
SYLVIA	0	0	0	0	0	0	1	1	1	1	0	1	1	1
NORA	0	0	0	0	0	1	1	0	1	1	1	1	1	1
HELEN	0	0	0	0	0	0	1	1	0	1	1	1	0	0
DOROTHY	0	0	0	0	0	0	0	1	1	0	0	0	0	0
OLIVIA	0	0	0	0	0	0	0	0	1	0	1	0	0	0
FLORA	0	0	0	0	0	0	0	0	1	0	1	0	0	0

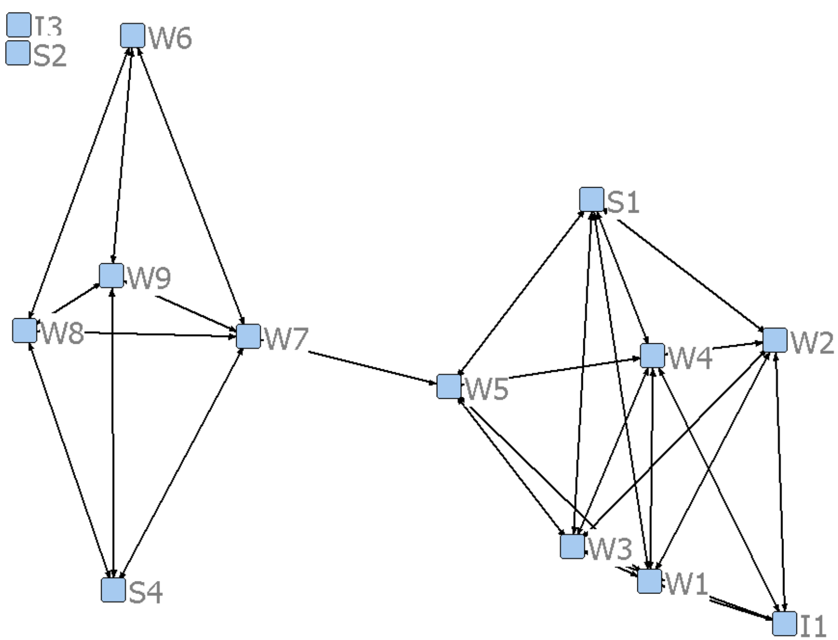
incidence matrix

	EV	LA	TH	BR	CH	FR	EL	PE	RU	VE	MY	KA	SY	NO	HE	DO	OL	FL	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN																			1	1	1	1	1	1		1	1					
LAURA																			1	1	1		1	1	1	1						
THERESA																				1	1	1	1	1	1	1	1					
BRENDA																			1		1	1	1	1	1	1						
CHARLOTTE																					1	1	1		1							
FRANCES																					1		1	1		1						
ELEANOR																							1	1	1	1						
PEARL																								1		1	1					
RUTH																							1		1	1	1					
VERNE																									1	1	1			1		
MYRNA																										1	1	1		1		
KATHERINE																										1	1	1		1	1	1
SYLVIA																									1	1	1	1		1	1	1
NORA																								1	1		1	1	1	1	1	1
HELEN																									1	1		1	1	1		
DOROTHY																										1	1					
OLIVIA																											1		1			
FLORA																											1		1			
E1	1	1		1																												
E2	1	1	1																													
E3	1	1	1	1	1	1																										
E4	1		1	1	1																											
E5	1	1	1	1	1	1	1		1																							
E6	1	1	1	1		1	1	1						1																		
E7		1	1	1	1		1		1	1			1	1	1																	
E8	1	1	1	1		1	1	1	1	1	1	1	1		1	1																
E9	1		1					1	1	1	1	1	1	1		1	1	1														
E10											1	1	1	1	1																	
E11														1	1		1	1														
E12										1	1	1	1	1	1																	
E13												1	1	1																		
E14												1	1	1																		

bi-adjacency matrix

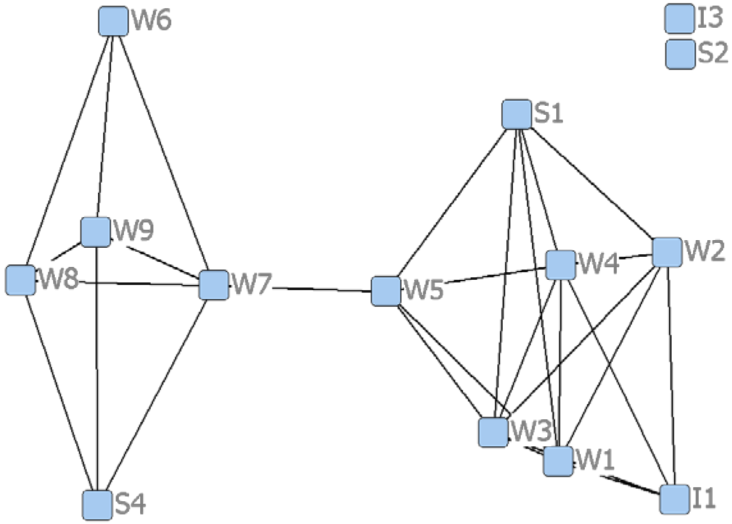
2. All graphs, including “1-mode”, have 2-mode matrix representations

- Analytical focus on nodes rather than ties



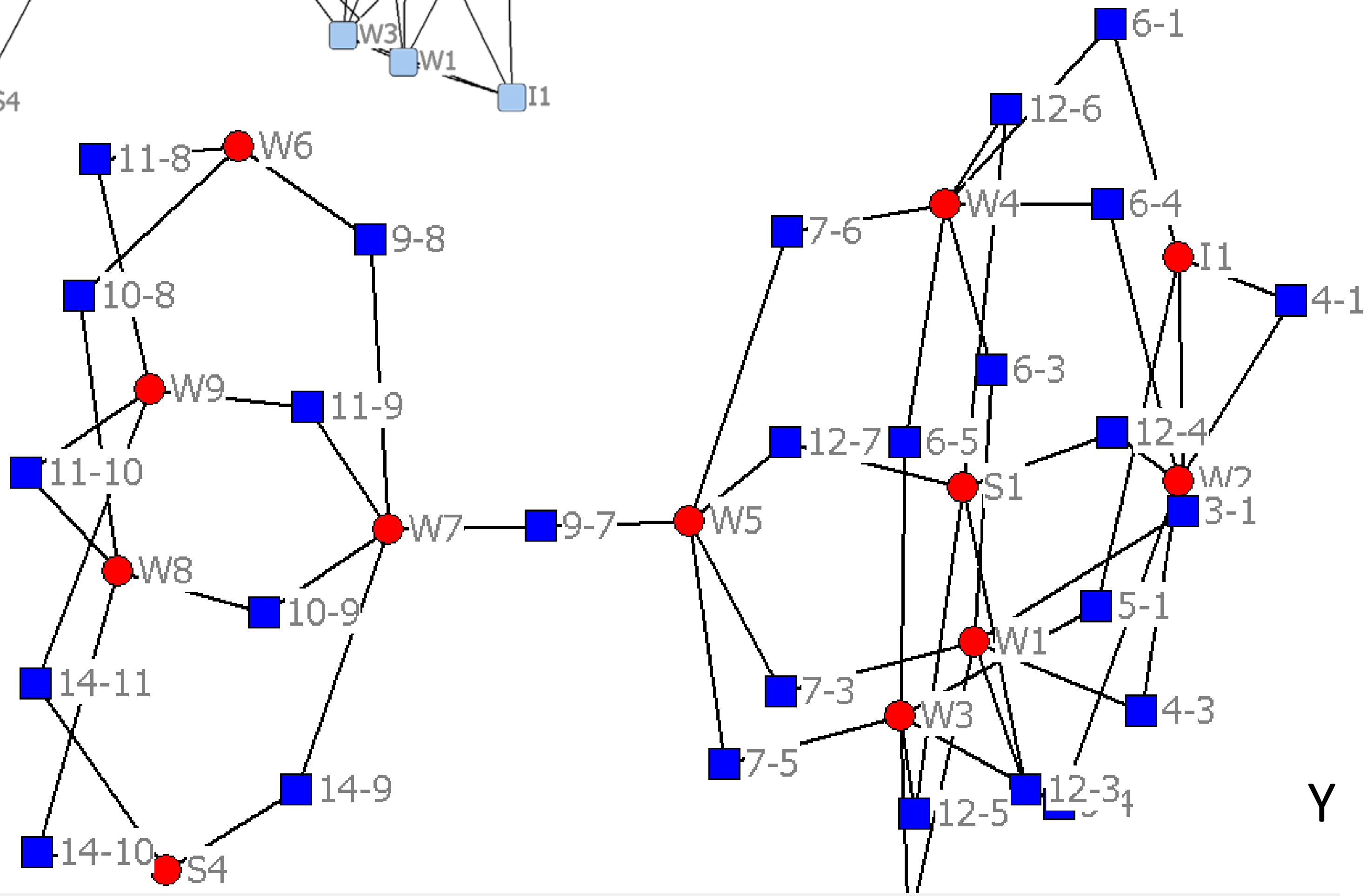
	Edges									
	3-1	4-1	4-3	5-1	5-3	5-4	6-1	6-3	6-4	6-5
I1	1	1	0	1	0	0	1	0	0	0
I3	0	0	0	0	0	0	0	0	0	0
W1	1	0	1	0	1	0	0	1	0	0
W2	0	1	1	0	0	1	0	0	1	0
W3	0	0	0	1	1	1	0	0	0	1
W4	0	0	0	0	0	0	1	1	1	1
W5	0	0	0	0	0	0	0	0	0	0
W6	0	0	0	0	0	0	0	0	0	0
W7	0	0	0	0	0	0	0	0	0	0
W8	0	0	0	0	0	0	0	0	0	0
W9	0	0	0	0	0	0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

Edge incidence matrix



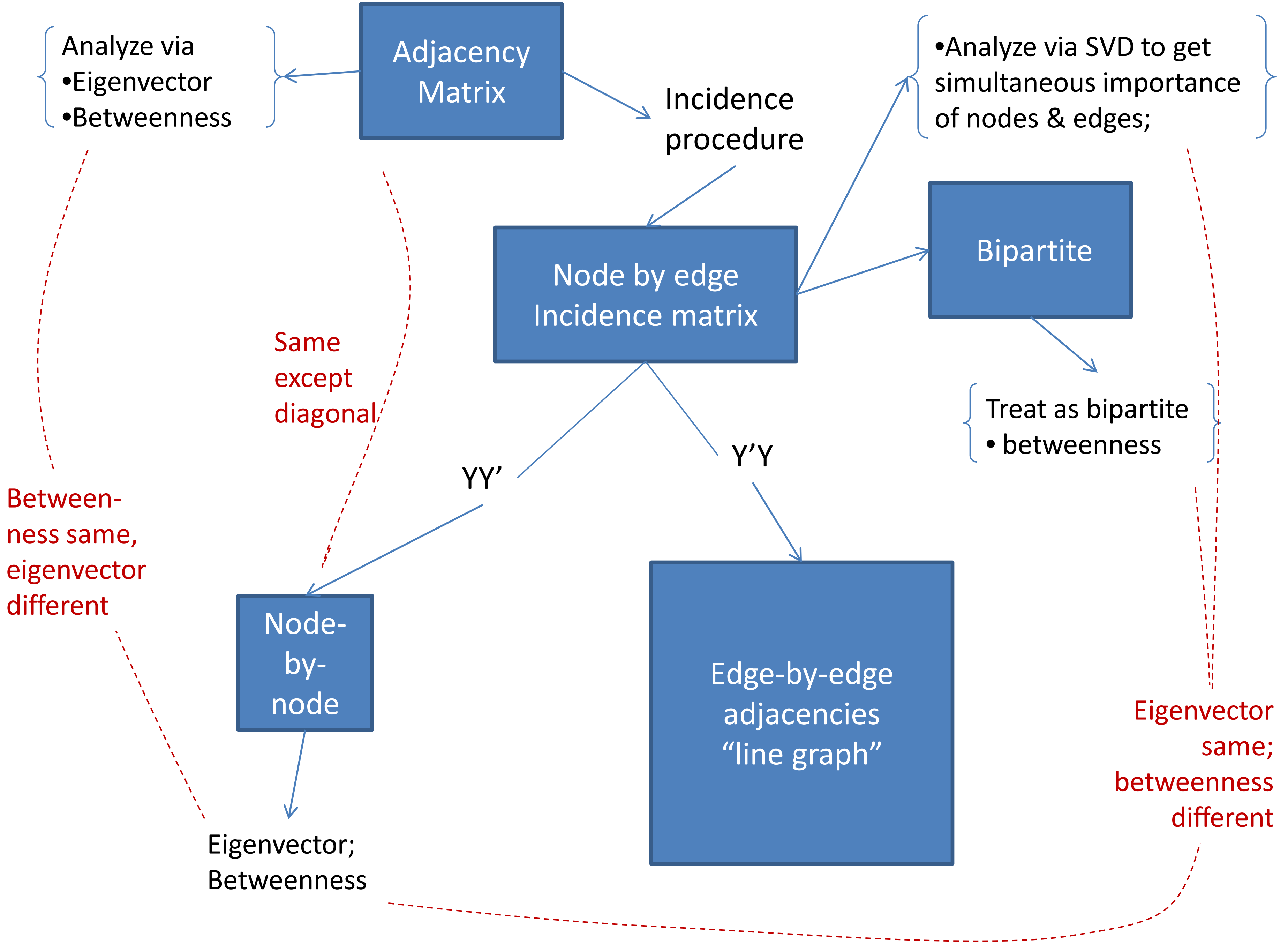
Simple expansion of the graph so that edges are also represented as nodes:
-- nodes now represent both actors and relationships on an equal footing

X



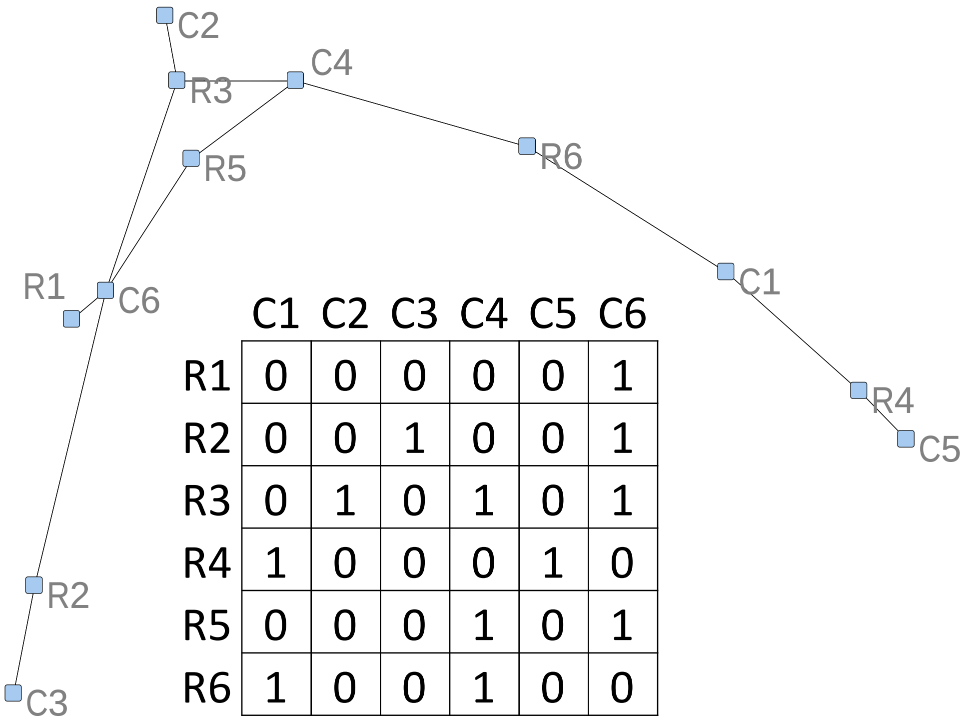
Y

Note: YY’ gives us back the adjacency matrix X, with degree down main diagonal


Analyze via
•Eigenvector
•Betweenness
Adjacency Matrix
Incidence procedure
•Analyze via SVD to get simultaneous importance of nodes & edges;
Bipartite
Node by edge Incidence matrix
Same except diagonal
Treat as bipartite
• betweenness
YY'
Y'Y
Between-ness same, eigenvector different
Node-by-node
Edge-by-edge adjacencies “line graph”
Eigenvector same; betweenness different
Eigenvector; Betweenness

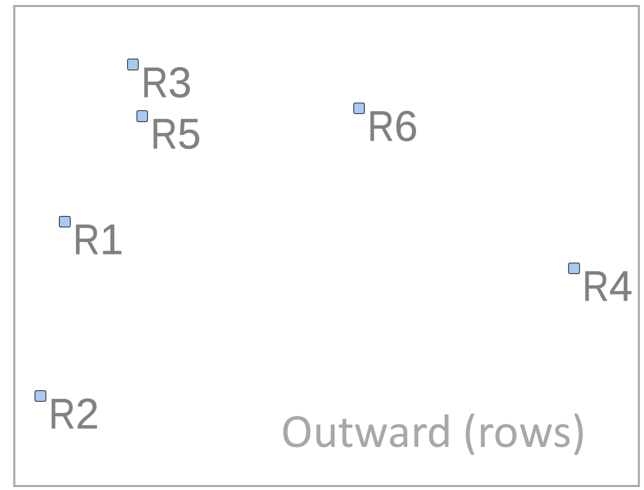
Can also treat direction as a mode

(receiver self distinguished from sender self)

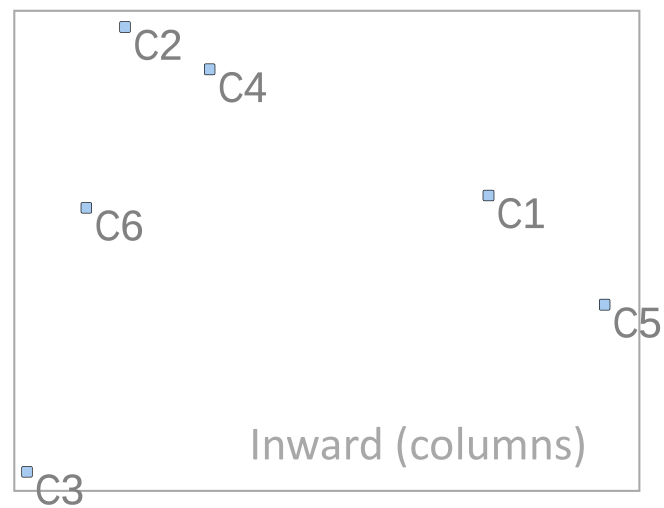


	C1	C2	C3	C4	C5	C6
R1	0	0	0	0	0	1
R2	0	0	1	0	0	1
R3	0	1	0	1	0	1
R4	1	0	0	0	1	0
R5	0	0	0	1	0	1
R6	1	0	0	1	0	0

Correspondence analysis



Outward (rows)



Inward (columns)

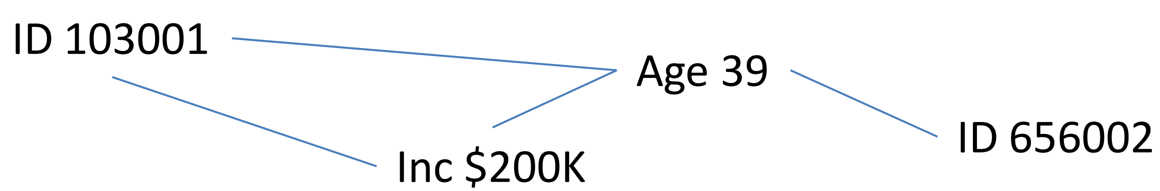
3. Not clear that we want to treat all 2-mode matrices as networks

Attribute

Person

ID	Age	Sex	Income	Gun Control	Abortion	Euthanasia
103001	39	2	$200K	5	4	5
213006	23	1	$35K	2	1	53

- Link analysis
- Actor-network theory

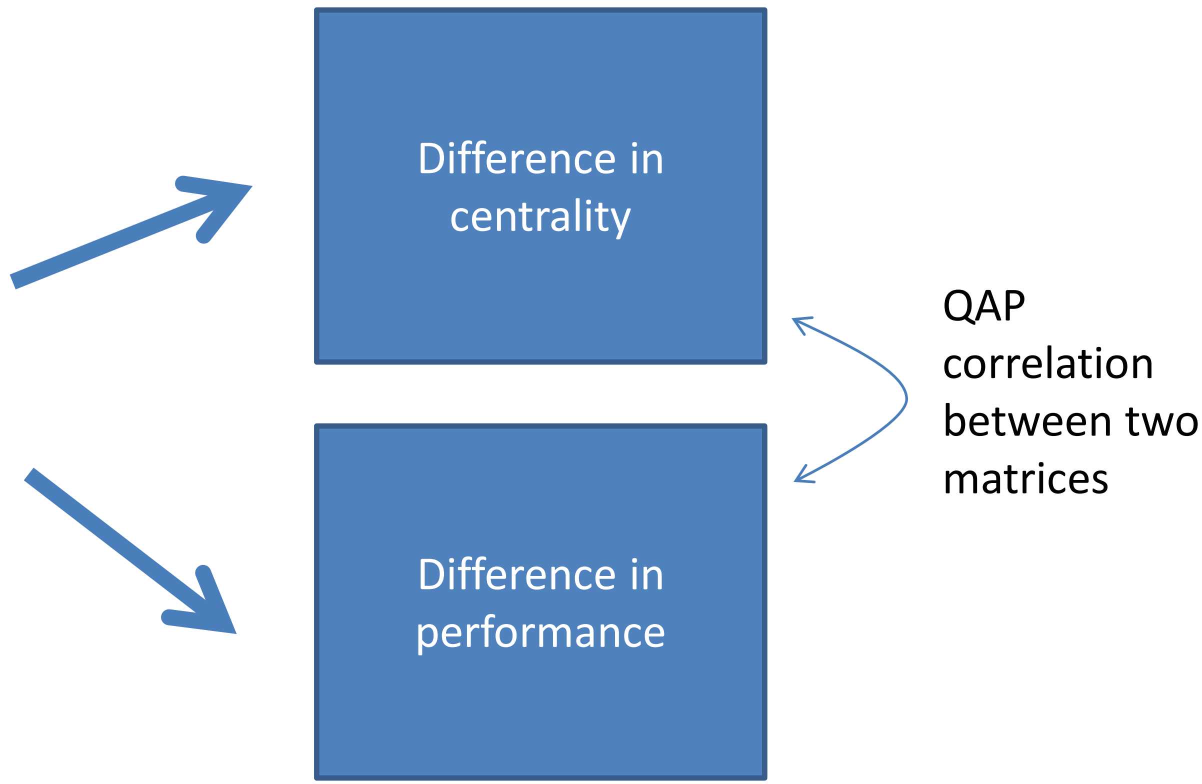


File under: Mindless Network Analysis

“Network analysis” of centrality

- How does centrality affect performance?

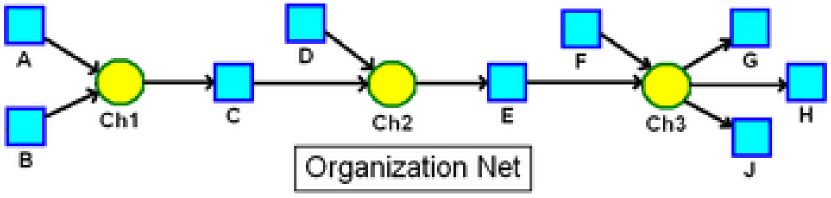
Centrality	Performance
67	3
89	4
27	2
16	1
06	1
58	3
92	5



But difficult to define …

- Standard examples
  - Affiliations
    - Attendance at events
    - Membership in orgs
  - Items by use
    - Household items by functions
  - Species by trait matrices
    - Binary trait profiles
    - Componential analyses
  - Correspondences/assoc
    - Authors and topics
    - Instructors and courses
    - Illnesses and treatments

- Binary relations defined on two sets
  - Mapping of one set to another
  - Set of unordered 2-tuples
  - But don't want to exclude valued 2-mode networks
- Directed 2-mode graphs?
  - Mapping of things or locations and events or processes (eg Petri nets)
    - Variably directed



Dangers of 2-mode analysis

- “just” a special case of standard attributes-based social science
  - Person by demographic variables
- Re-inventing actor network theory?
- Encourages use of social network techniques without social network theory
  - Network is the new pie chart – the hot way to display any and all information

Section 2

TECHNIQUES

Two approaches

Approach	Concept	Usage
Conversion*	"Convert" women by events matrix X to: - XX': Women by women - X'X: Events by events	One mode is of greater theoretical interest than the other - Events used to examine ties among the women
Direct	Treat both women and events as simply nodes in a graph with both women and events Run analyses on these "affiliation graphs" more or less as usual	Modes are of equal theoretical interest , or it is the correspondence between the modes that is of interest

- Research question: when are results different?

*Despite conversion/direct distinction, in a sense both involve "conversions" to 1-mode

Section 2A

CONVERSION APPROACH

“Converting” 2-mode to 1-mode

X
Women by events

XX’
Women by women

No. of times each pair of rows has 1 in the same column

- no. of events each pair of women attended in common

X’X
Event by Event

No. of times each pair of columns has 1 in the same row

- no. of women that co-attended each pair of events

Normalization

- No. of overlaps between women constrained or enabled by no. of events attended
  - By chance alone, women who attend many events, they will overlap more than women with few events
  - One attends fewer events than the other, so the maximum overlap is the smaller no. of events
- Bonacich (1972) adjustment of no. of overlaps for chance agreements

Measures of Similarity

- "Conversion" to 1-mode is just computing measure of similarity among rows or columns
  - XX' is $\sum_k x_{ik} x_{jk}$ which is un-normalized correlation coef

$$r_{ij} = \frac{\frac{1}{m}\sum_k x_{ik} x_{jk} - u_i u_j}{s_i s_j}$$

- In our case, X is usually binary

$$r_{ij} = \frac{ad - bc}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

		Row j 1	Row j 0	
Row i	1	a	b	a+b
	0	c	d	c+d
		a+c	b+d	m

		Pearl 1	Pearl 0	
Ruth	1	2	2	4
	0	1	9	10
		3	11	14

Measures of similarity for 1/0 data

		Row j: 1	Row j: 0	
Row i	1	a	b	a+b
	0	c	d	c+d
		a+c	b+d	m

		Pearl: 1	Pearl: 0	
Ruth	1	2	2	4
	0	1	9	10
		3	11	14

Measure	Formula
Overlaps	a
Simple matching	$\frac{a+d}{a+b+c+d}$
Jaccard	$\frac{a}{a+b+c}$
Yule's Q / G & K gamma	$\frac{ad-bc}{ad+bc}$
Dozens more …	

Normalization of cases

- If two people attend small events together, should have more implications for possible relations between them
- So when comparing rows, should probably weight certain columns more
  - Weighting events inversely by size

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN	1	1	1	1	1	1	0	1	1	0	0	0	0	0
LAURA	1	1	1	0	1	1	1	1	0	0	0	0	0	0
THERESA	0	1	1	1	1	1	1	1	1	0	0	0	0	0
BRENDA	1	0	1	1	1	1	1	1	0	0	0	0	0	0
CHARLOTTE	0	0	1	1	1	0	1	0	0	0	0	0	0	0
FRANCES	0	0	1	0	1	1	0	1	0	0	0	0	0	0
ELEANOR	0	0	0	0	1	1	1	1	0	0	0	0	0	0
PEARL	0	0	0	0	0	1	0	1	1	0	0	0	0	0
RUTH	0	0	0	0	1	0	1	1	1	0	0	0	0	0
VERNE	0	0	0	0	0	0	1	1	1	0	0	1	0	0
MYRNA	0	0	0	0	0	0	0	1	1	1	0	1	0	0
KATHERINE	0	0	0	0	0	0	0	1	1	1	0	1	1	1
SYLVIA	0	0	0	0	0	0	1	1	1	1	0	1	1	1
NORA	0	0	0	0	0	1	1	0	1	1	1	1	1	1
HELEN	0	0	0	0	0	0	1	1	0	1	1	1	0	0
DOROTHY	0	0	0	0	0	0	0	1	1	0	0	0	0	0
OLIVIA	0	0	0	0	0	0	0	0	1	0	1	0	0	0
FLORA	0	0	0	0	0	0	0	0	1	0	1	0	0	0

Large event E9

More intimate event E11

Double iterative normalization

- Deming's iterative proportional fitting algorithm
- What would data table look like if marginals were all the same?
- i.e., simultaneously control for women's gregariousness and events' size

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN	0.35	0.35	0.08	0.17	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LAURA	0.41	0.41	0.09	0.00	0.03	0.02	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
THERESA	0.00	0.52	0.12	0.26	0.04	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
BRENDA	0.52	0.00	0.12	0.26	0.04	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
CHARLOTTE	0.00	0.00	0.26	0.59	0.08	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FRANCES	0.00	0.00	0.62	0.00	0.20	0.15	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
ELEANOR	0.00	0.00	0.00	0.00	0.37	0.29	0.27	0.07	0.00	0.00	0.00	0.00	0.00	0.00
PEARL	0.00	0.00	0.00	0.00	0.00	0.73	0.00	0.18	0.09	0.00	0.00	0.00	0.00	0.00
RUTH	0.00	0.00	0.00	0.00	0.50	0.00	0.36	0.10	0.05	0.00	0.00	0.00	0.00	0.00
VERNE	0.00	0.00	0.00	0.00	0.00	0.00	0.32	0.09	0.04	0.00	0.00	0.55	0.00	0.00
MYRNA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.02	0.59	0.00	0.33	0.00	0.00
KATHERINE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.07	0.00	0.04	0.44	0.44
SYLVIA	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.07	0.00	0.04	0.43	0.43
NORA	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.07	0.00	0.04	0.42	0.42
HELEN	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.04	0.00	0.49	0.03	0.28	0.00	0.00
DOROTHY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.32	0.00	0.00	0.00	0.00	0.00
OLIVIA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.62	0.00	0.00	0.00
FLORA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.62	0.00	0.00	0.00

Analysis of converted data

- Regular sna is applicable except for
  - valuedness
  - some “additional” dependencies
- Valued graphs pose a problem for only a few techniques
  - closeness, betweenness, cliques
- Some of these handle-able by generalizing to optimal paths
  - Costs & distances:
    - optimal path has smallest sum of edge weights (see Brandes paper)
  - Probabilities:
    - optimal path has largest product of edge weights
  - Capacities:
    - optimal path has largest minimum of edge weights

Variations on Floyd's algorithm

- Recode 0s in adj matrix A to large number, n

```
for i:= 1 to n do
  for j:= 1 to n do if A(j,i) < n then
    for k:= 1 to n do
      if A(j,i)+A(I,k) < A(j,k))
        then A(j,k) = A(j,i) + A(i,k)
```

Outer: MIN
Inner: SUM

- Distances or costs (min of sum): no change
- Probabilities (max of product):

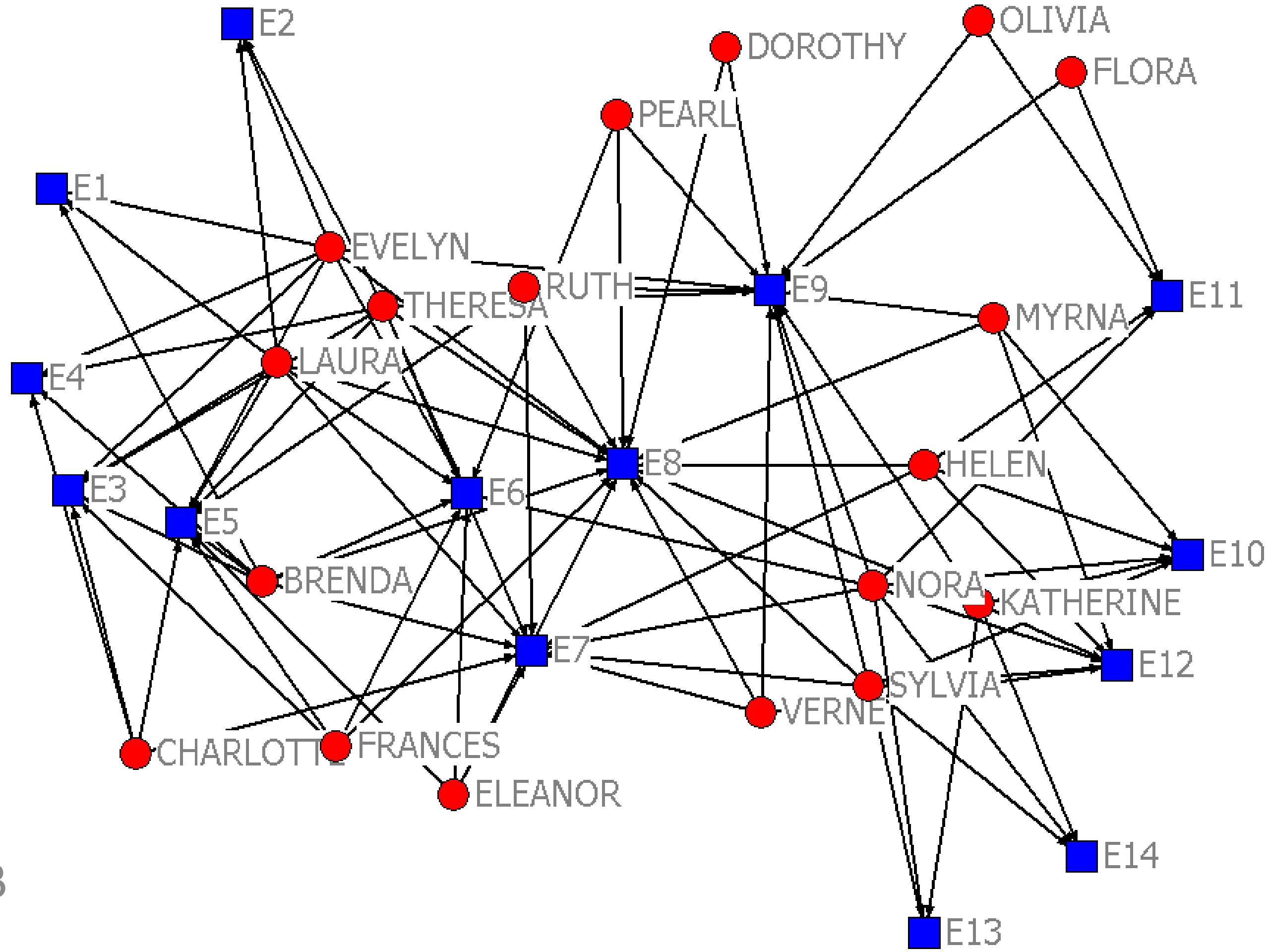
```
for i:= 1 to n do
  for j:= 1 to n do if A(j,i) < n then
    for k:= 1 to n do
      if A(j,i) * A(I,k) > A(j,k))
        then A(j,k) = A(j,i) * A(i,k)
```

Outer: MAX
Inner: PROD

- Capacities (max of min):

```
      if A(j,i) * A(I,k) > A(j,k))
        then A(j,k) = Min(A(j,i),A(i,k))
```

Outer: MAX
Inner: MIN



Section 2B

DIRECT ANALYSIS OF AFFILIATION GRAPHS

Matrix representations for direct analysis

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN	1	1	1	1	1	1	0	1	1	0	0	0	0	0
LAURA	1	1	1	0	1	1	1	1	0	0	0	0	0	0
THERESA	0	1	1	1	1	1	1	1	1	0	0	0	0	0
BRENDA	1	0	1	1	1	1	1	1	0	0	0	0	0	0
CHARLOTTE	0	0	1	1	1	0	1	0	0	0	0	0	0	0
FRANCES	0	0	1	0	1	1	0	1	0	0	0	0	0	0
ELEANOR	0	0	0	0	1	1	1	1	0	0	0	0	0	0
PEARL	0	0	0	0	0	1	0	1	1	0	0	0	0	0
RUTH	0	0	0	0	1	0	1	1	1	0	0	0	0	0
VERNE	0	0	0	0	0	0	1	1	1	0	0	1	0	0
MYRNA	0	0	0	0	0	0	0	1	1	1	0	1	0	0
KATHERINE	0	0	0	0	0	0	0	1	1	1	0	1	1	1
SYLVIA	0	0	0	0	0	0	1	1	1	1	0	1	1	1
NORA	0	0	0	0	0	1	1	0	1	1	1	1	1	1
HELEN	0	0	0	0	0	0	1	1	0	1	1	1	0	0
DOROTHY	0	0	0	0	0	0	0	1	1	0	0	0	0	0
OLIVIA	0	0	0	0	0	0	0	0	1	0	1	0	0	0
FLORA	0	0	0	0	0	0	0	0	1	0	1	0	0	0

	EV	LA	TH	BR	CH	FR	EL	PE	RU	VE	MY	KA	SY	NO	HE	DO	OL	FL	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN																			1	1	1	1	1	1		1	1					
LAURA																			1	1	1		1	1	1	1						
THERESA																				1	1	1	1	1	1	1	1					
BRENDA																			1		1	1	1	1	1	1						
CHARLOTTE																					1	1	1		1							
FRANCES																					1		1	1		1						
ELEANOR																							1	1	1	1						
PEARL																								1		1	1					
RUTH																							1		1	1	1					
VERNE																									1	1	1			1		
MYRNA																										1	1	1		1		
KATHERINE																										1	1	1		1	1	1
SYLVIA																									1	1	1	1		1	1	1
NORA																								1	1		1	1	1	1	1	1
HELEN																									1	1		1	1	1		
DOROTHY																										1	1					
OLIVIA																											1		1			
FLORA																											1		1			
E1	1	1		1																												
E2	1	1	1																													
E3	1	1	1	1	1	1																										
E4	1		1	1	1																											
E5	1	1	1	1	1	1	1		1																							
E6	1	1	1	1		1	1	1						1																		
E7		1	1	1	1		1		1	1			1	1	1																	
E8	1	1	1	1		1	1	1	1	1	1	1	1		1	1																
E9	1		1					1	1	1	1	1	1	1		1	1	1														
E10											1	1	1	1	1																	
E11														1	1		1	1														
E12										1	1	1	1	1	1																	
E13												1	1	1																		
E14												1	1	1																		

Incidence matrix B

bi-Adjacency matrix $A = \begin{pmatrix} O & B \\ B^T & O \end{pmatrix}$,

Approaches for network measures

recycle

- Do std SNA on bipartite graph, then normalize measures for bipartite
- Do std SNA on bipartite graph, but don't normalize
  - just compare with expected values from appropriate 2-mode baseline models
- Do nothing

de novo

- Construct new formulae specifically designed for affiliation graphs
- Generalize/redefine/invent new constructs

Centrality

Normalization Strategy

- Degree $$d_i^* = \frac{d_i}{n_2} \quad d_j^* = \frac{d_j}{n_1}$$

  For
  $i \in V_1 \quad j \in V_2$

  - General strategy: divide scores by different constant depending on node set it belongs to

- Closeness $$c_i^* = \frac{n_2 + 2(n_1 - 1)}{c_i} \qquad c_j^* = \frac{n_1 + 2(n_2 - 1)}{c_j}$$

Betweenness

Normalization Strategy

$$b_k = \tfrac{1}{2}\sum_{i\neq k}^{n}\sum_{j\neq k,i}^{n}\frac{g_{ikj}}{g_{ij}} \qquad b_i^* = \frac{b_i}{b^{V_1\max}} \qquad b_j^* = \frac{b_j}{b^{V_2\max}}$$

For $i \in V_1 \quad j \in V_2$

$$b^{V_1\max} = \tfrac{1}{2}[n_2{}^2(s+1)^2 + n_2(s+1)(2t-s-1) - t(2s-t+3)]$$

$$s = (n_1-1)\ div\ n_2 \qquad t = (n_1-1)\bmod n_2$$

$$b^{V_2\max} = \tfrac{1}{2}[n_1{}^2(p+1)^2 + n_1(p+1)(2r-p-1) - r(2p-r+3)]$$

$$p = (n_2-1)\ div\ n_1 \qquad r = (n_1-1)\bmod n_2$$

Approaches for network measures

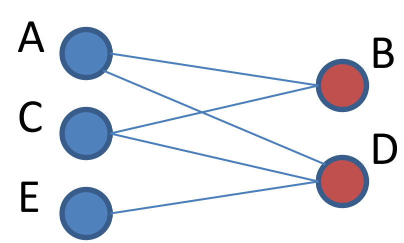
recycle

- Do std SNA on bipartite graph, then normalize measures for bipartite
  - Case of density
- Do std SNA on bipartite graph, but don't normalize
  - just compare with expected values from appropriate 2-mode baseline models
- Do nothing
  - Case of avg distance

==

de novo

- Construct new formulae designed for affiliations
  - incidence matrices
- Generalize/redefine/invent new constructs
  - Case of transitivity
    - Transitive quads?



The case of density

- Network density

  T is no. of ties, n is no. of nodes

$$d = \frac{2T}{n(n-1)}$$

- Max density in a 2-mode net

  n1, n2 are size of vertex sets

$$\frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)}$$

- So normalize density like this:

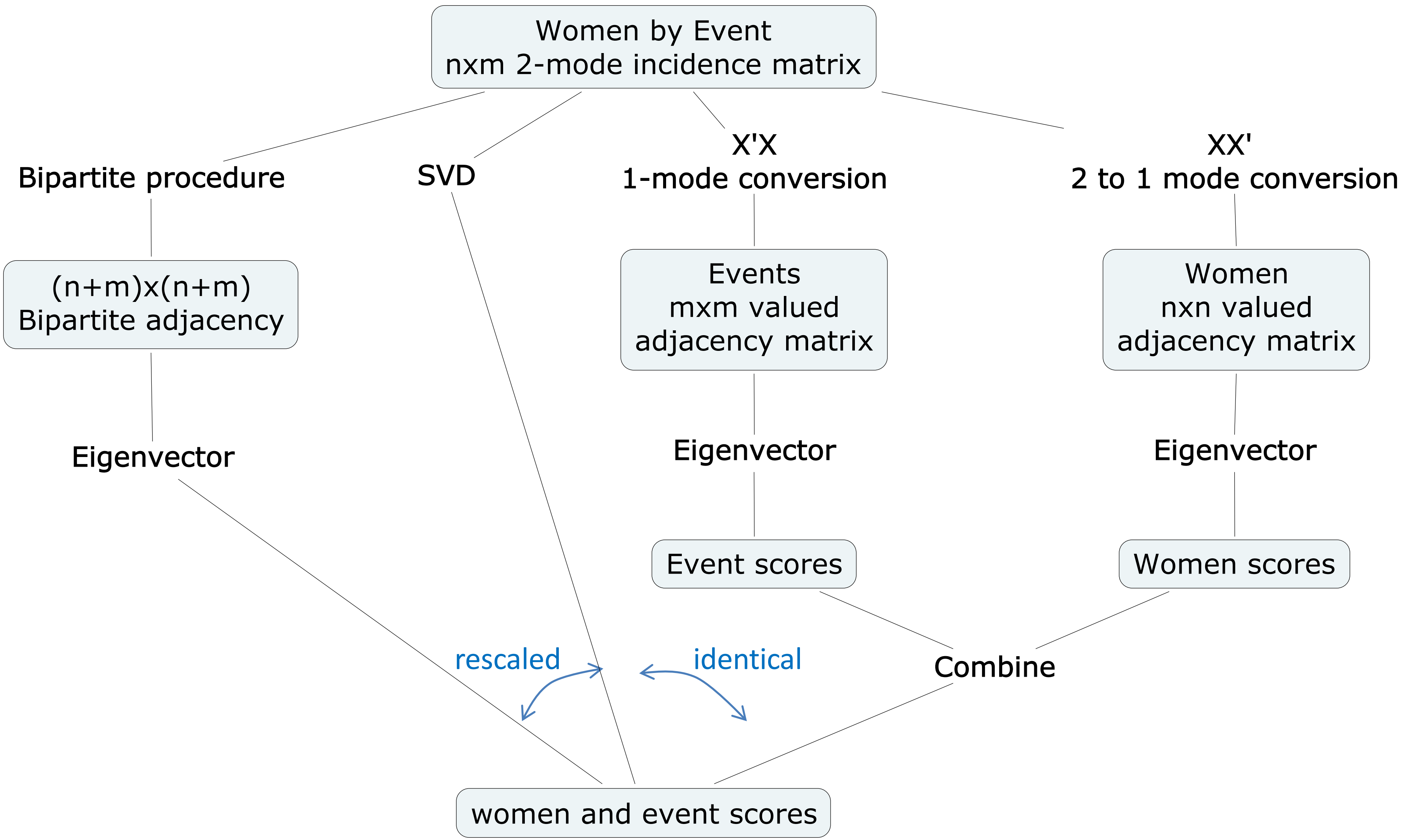
$$d^{2-\mathrm{mode}} = \frac{d}{\frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}}$$

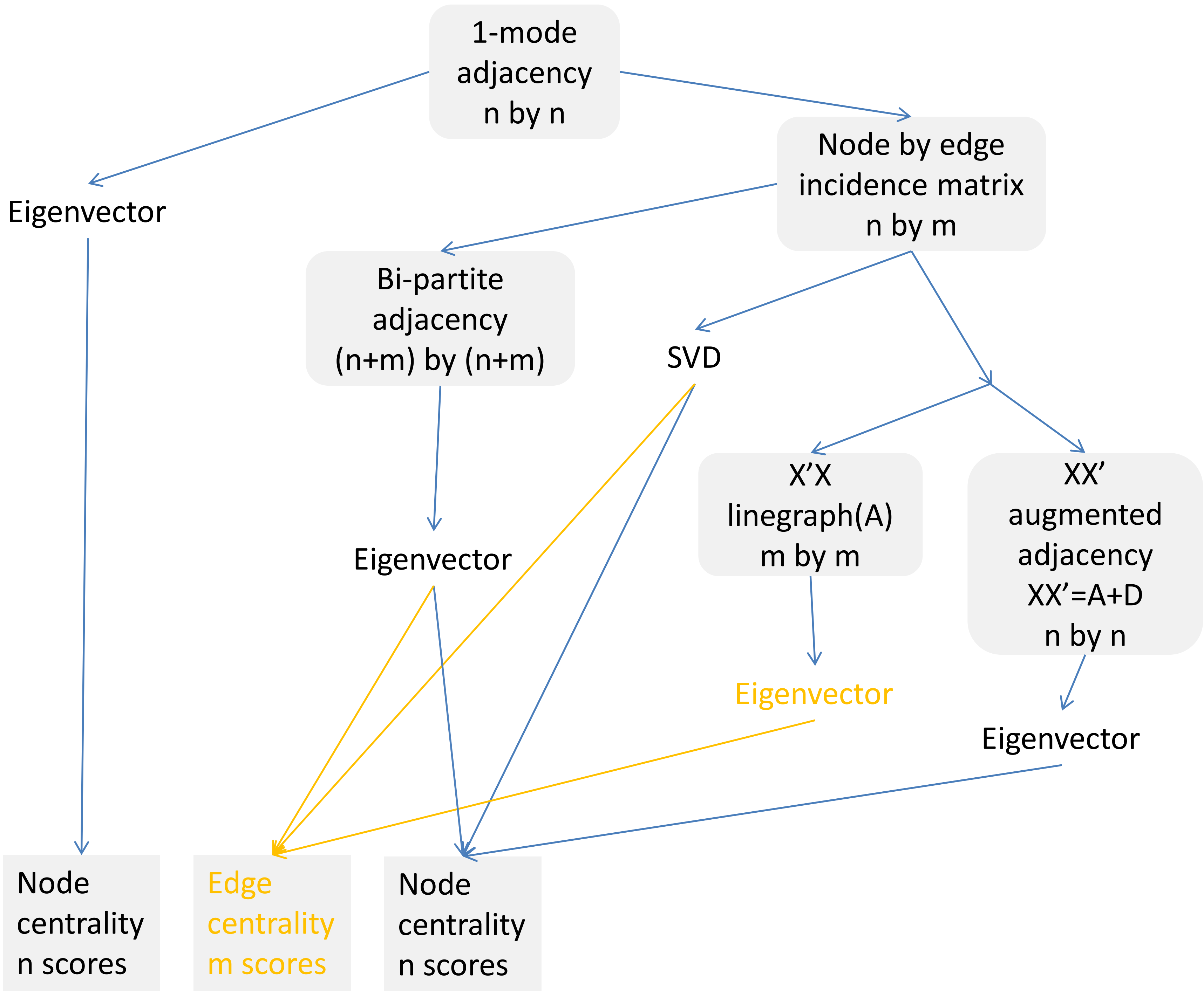
- Or …

$$d^{2-\mathrm{mode}} = \frac{d}{\frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}} = \frac{\frac{2T}{n(n-1)}}{\frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)}} = \frac{2T}{n_1 n_2}$$

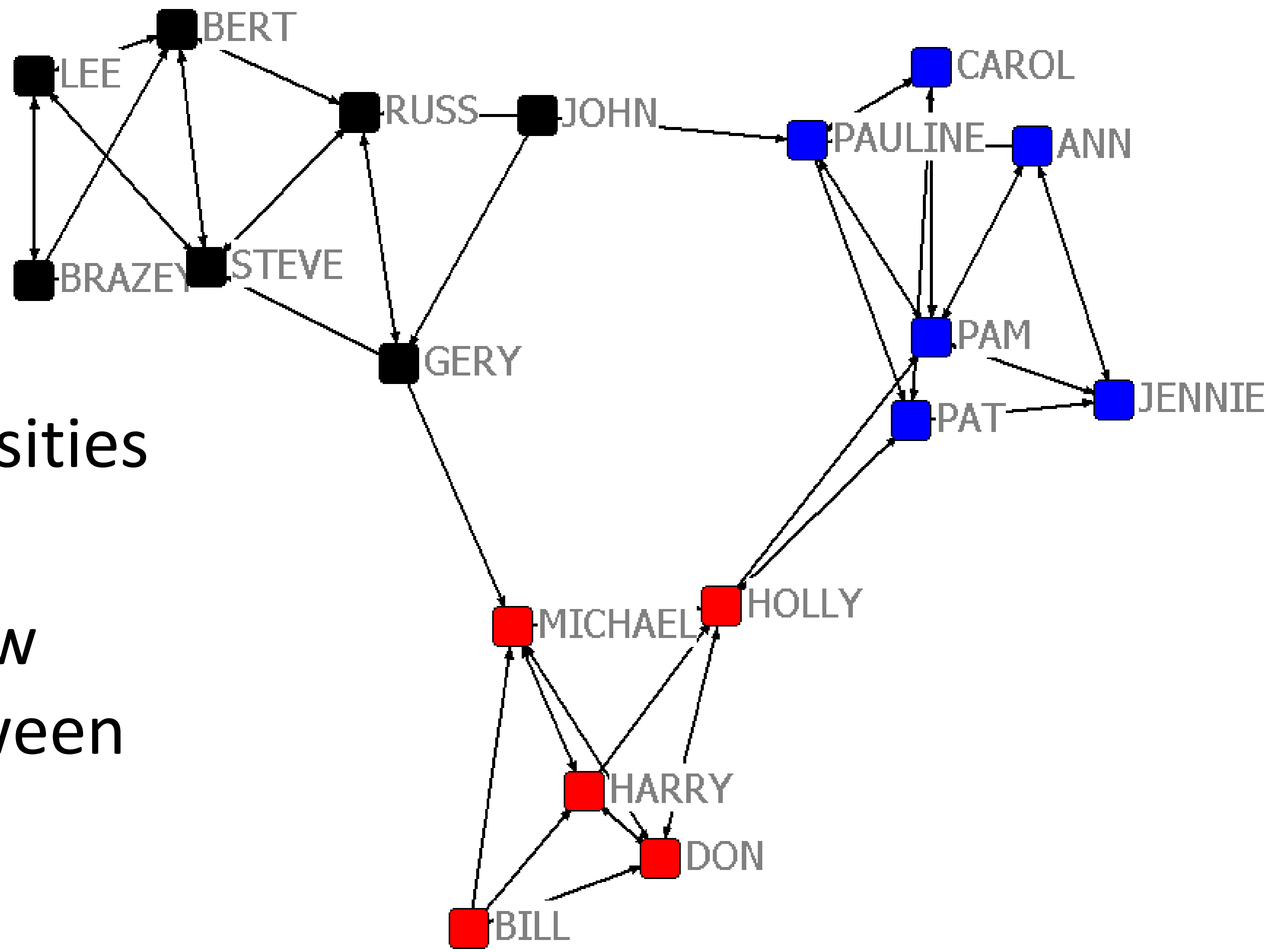
  - Bipartite completeness

Eigenvector Centrality/Core-periphery




1-mode adjacency n by n
Node by edge incidence matrix n by m
Eigenvector
Bi-partite adjacency (n+m) by (n+m)
SVD
X’X linegraph(A) m by m
XX’ augmented adjacency XX’=A+D n by n
Eigenvector
Eigenvector
Eigenvector
Node centrality n scores
Edge centrality m scores
Node centrality n scores

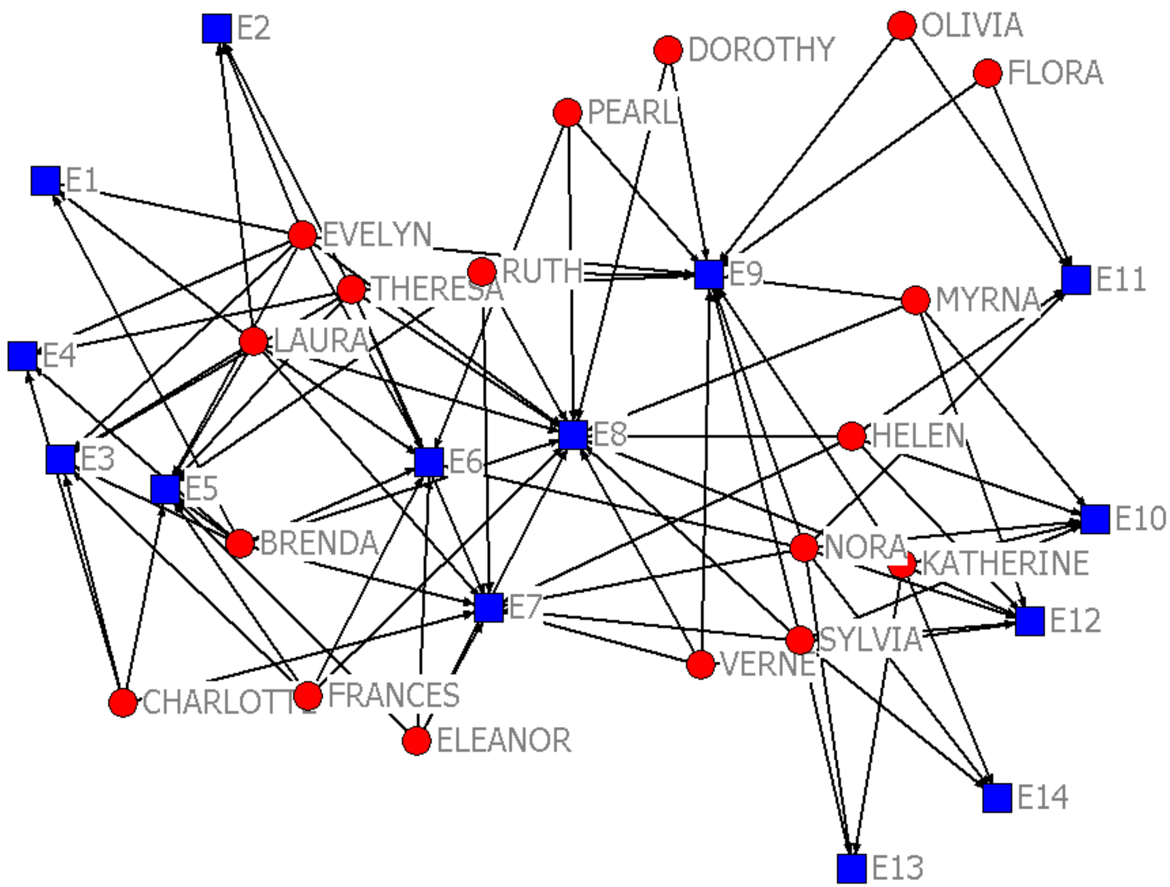
Cohesive subgroups



- Seek high densities within groups
- And, often, low densities between groups

Subgroups in affiliation graphs

- Affiliations graphs are locally sparse
  - Density of every ego-net is zero
  - Transitivity zero
- No cliques as generally defined



graph theoretic approaches

- Regions
  - Components, k-cores work reasonably well
- Cliques
  - No cliques, as normally defined
- Relaxations of clique
  - Too numerous to use directly
  - 2-stage analyses work well
  - N-cliques of special interest

K-Plexes

K	No.
2	394
3	5553
4	37633

Bi-cliques

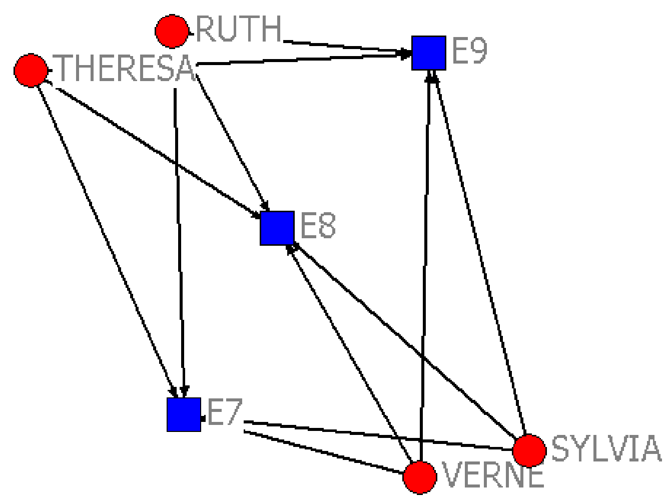
- 2-cliques of affiliation graphs are a bit less numerous
  - can yield pleasing results
- 2-cliques are bipartite complete
  - all possible ties present
  - captures clique notion better than clique does

- Bi-clique is a 2-clique of an affiliations graph, a maximally complete subgraph of bipartite graph
  - Typically require min no. of nodes from each node set
  - e.g., at least one of each type

Bi-clique analysis of Davis data

68 2-cliques found
Includes 5 that are mono-modal

2: THERESA E2 E3 E4 E5 E6 E7 E8 E9
4: THERESA RUTH VERNE SYLVIA E7 E8 E9
5: VERNE SYLVIA E7 E8 E9 E12
6: SYLVIA E7 E8 E9 E10 E12 E13 E14
7: THERESA RUTH E5 E7 E8 E9
8: EVELYN THERESA PEARL RUTH
VERNE MYRNA KATHERINE
SYLVIA DOROTHY E8 E9
9: EVELYN THERESA PEARL E6 E8 E9
10: VERNE MYRNA KATHERINE
SYLVIA E8 E9 E12
11: MYRNA KATHERINE SYLVIA
E8 E9 E10 E12
12: KATHERINE SYLVIA E8 E9 E10 E12 E13 E14
13: EVELYN THERESA E2 E3 E4 E5 E6 E8 E9
…



2-clique No. 4

STEP 2:
Hierarchical clustering of
the clique overlap matrix

Hierarchical clustering of bi-clique co-membership

```
                                                         K
                 H                                       A
           D     A     F E         T                     T
           O O   R     R L     E   H B                   H         S
         P R L F L     A E     V L E R             H M E       V Y
         E O I L O     N A R   E A R E             E Y R   N   E L
         A T V O T     C N U   L U E N           E L R I E O   R V E E E
         R H I R T E E E O T E Y R S D E E E E E 1 E N N 1 R E N I 1 1 1
         L Y A A E 2 1 S R H 4 N A A A 7 3 6 5 8 1 N A E 0 A 9 E A 2 3 4
------   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
22.000   . . . . . . . . . . . . . XXX . . . XXX . . . . . . . . . . . .
20.000   . . . . . . . . . . . . XXXXX . . XXXXX . . . . . . . . . . . .
18.000   . . . . . . . . . . . . XXXXX . . XXXXX . . . . . . . XXX . . .
17.000   . . . . . . . . . . . . XXXXX . XXXXXXX . . . . . . . XXX . . .
15.250   . . . . . . . . . . . XXXXXXX . XXXXXXX . . . . . . . XXX . . .
15.000   . . . . . . . . . . . XXXXXXX XXXXXXXXX . . . . . . . XXX . . .
14.667   . . . . . . . . . . . XXXXXXX XXXXXXXXX . . . . . . . XXXXX . .
14.033   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . . . . . . XXXXX . .
13.000   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . . . XXXXX . .
12.000   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . . XXXXXXX . .
11.747   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . XXXXXXXXX . .
 9.878   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX XXXXXXXXXXX . .
 9.000   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . . XXX XXXXXXXXXXX . .
 8.975   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . . XXXXXXXXXXXXXXX . .
 8.288   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 7.069   . . . . . . . . XXX XXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 6.245   . . . . . . . . XXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 6.000   . . . . . . . . XXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 5.143   . . . . . . . XXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 5.000   XXX . . . . . XXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 4.675   XXX . . . . XXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 4.464   XXX . . . XXXXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 4.000   XXX XXX . XXXXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 3.703   XXX XXX . XXXXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXXXXXX
 3.126   XXX XXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXXXXXX
 2.188   XXX XXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXX
 2.000   XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXX
 1.788   XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 0.860   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

Hierarchical clustering of bi-clique co-membership

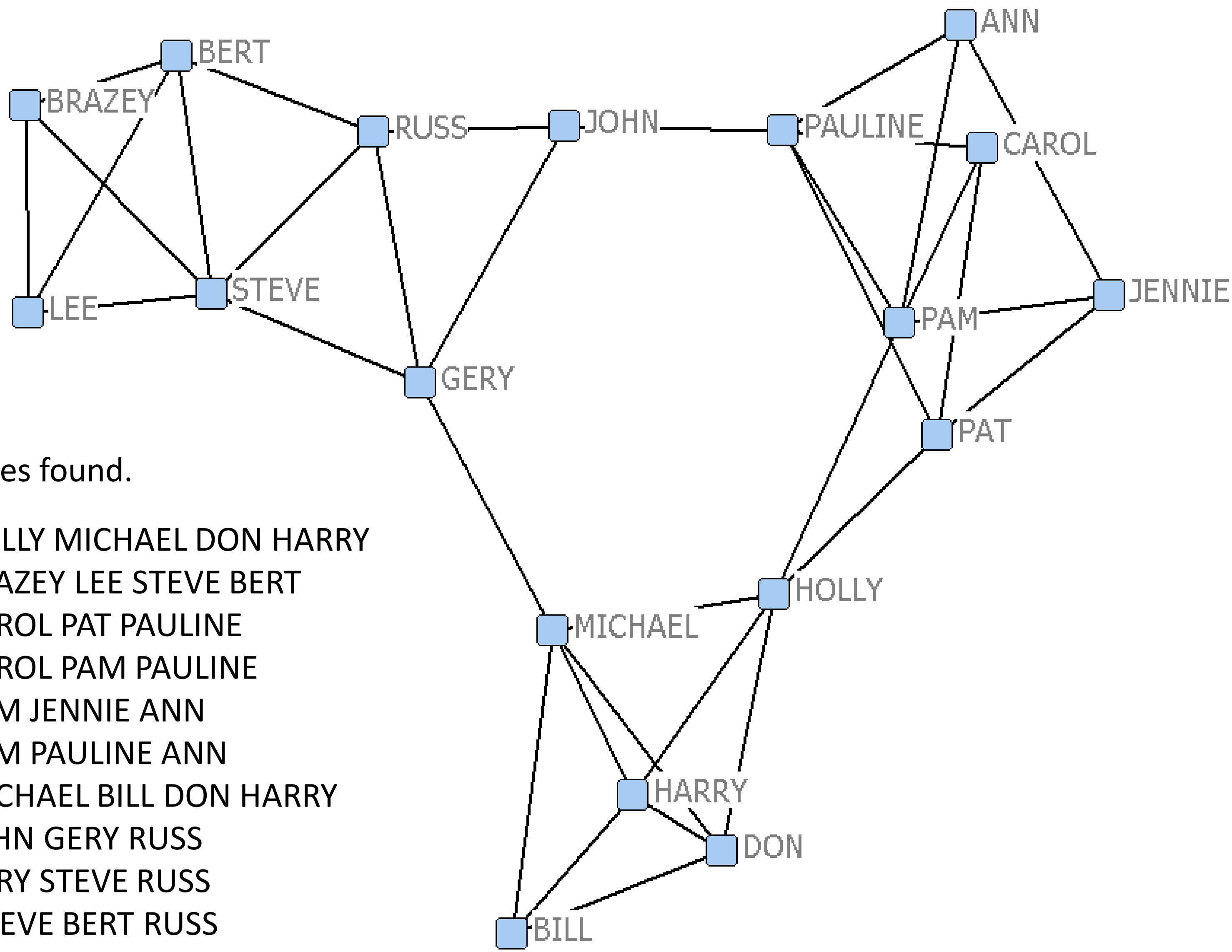
```
                                                       K
                 H                                     A
           D     A     F E         T                   T
           O O   R     R L     E   H B                 H         S
         P R L F L     A E     V L E R             H M E       V Y
         E O I L O     N A R   E A R E             E Y R   N   E L
         A T V O T     C N U   L U E N           E L R I E O   R V E E E
         R H I R T E E E O T E Y R S D E E E E E 1 E N N 1 R E N I 1 1 1
         L Y A A E 2 1 S R H 4 N A A A 7 3 6 5 8 1 N A E 0 A 9 E A 2 3 4
------   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
22.000   . . . . . . . . . . . . . XXX . . . XXX . . . . . . . . . . . .
20.000   . . . . . . . . . . . . XXXXX . . XXXXX . . . . . . . . . . . .
18.000   . . . . . . . . . . . . XXXXX . . XXXXX . . . . . . . XXX . . .
17.000   . . . . . . . . . . . . XXXXX . XXXXXXX . . . . . . . XXX . . .
15.250   . . . . . . . . . . . XXXXXXX . XXXXXXX . . . . . . . XXX . . .
15.000   . . . . . . . . . . . XXXXXXX XXXXXXXXX . . . . . . . XXX . . .
14.667   . . . . . . . . . . . XXXXXXX XXXXXXXXX . . . . . . . XXXXX . .
14.033   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . . . . . . XXXXX . .
13.000   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . . . XXXXX . .
12.000   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . . XXXXXXX . .
11.747   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX . XXXXXXXXX . .
 9.878   . . . . . . . . . . . XXXXXXXXXXXXXXXXX . . XXX XXXXXXXXXXX . .
 9.000   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . . XXX XXXXXXXXXXX . .
 8.975   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . . XXXXXXXXXXXXXXX . .
 8.288   . . . . . . . . XXX . XXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 7.069   . . . . . . . . XXX XXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 6.245   . . . . . . . . XXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX . .
 6.000   . . . . . . . . XXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 5.143   . . . . . . . XXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 5.000   XXX . . . . . XXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 4.675   XXX . . . . XXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
 4.464   XXX . . . XXXXXXXXXXXXXXXXXXXXXXXXXXXXX . XXXXXXXXXXXXXXXXX XXX
```

2-mode secondary analysis of clique structures

- Start with any graph (e.g., 1-mode)
- Run cliques (or something else)
- Construct valued node-by-clique membership matrix
  - x_{ij} = no. of ties that node *i* has to clique *j*
- Visualize as 2-mode network
- Find subgroups, etc.

File in: 2-mode methods applied to 1-mode data

Example: campnet dataset



10 cliques found.

1: HOLLY MICHAEL DON HARRY
2: BRAZEY LEE STEVE BERT
3: CAROL PAT PAULINE
4: CAROL PAM PAULINE
5: PAM JENNIE ANN
6: PAM PAULINE ANN
7: MICHAEL BILL DON HARRY
8: JOHN GERY RUSS
9: GERY STEVE RUSS
10: STEVE BERT RUSS

Xij = no. of ties that node i has with clique j

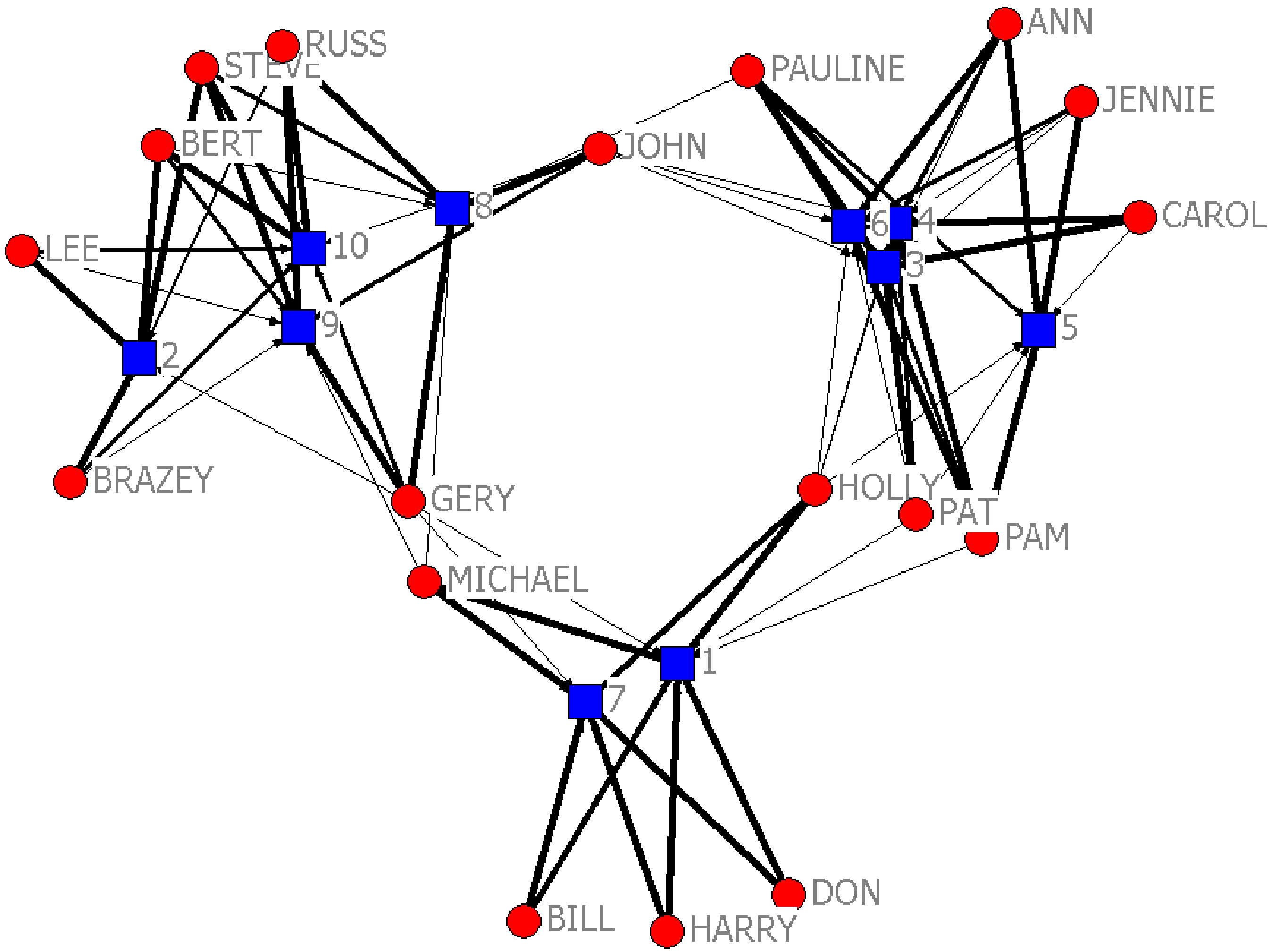
Valued clique membership matrix

	C2	C10	C9	C8	C7	C1	C3	C4	C6	C5
LEE	1.0	0.7	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BRAZEY	1.0	0.7	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BERT	1.0	1.0	0.7	0.3	0.0	0.0	0.0	0.0	0.0	0.0
STEVE	1.0	1.0	1.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0
RUSS	0.5	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
GERY	0.3	0.7	1.0	1.0	0.3	0.3	0.0	0.0	0.0	0.0
JOHN	0.0	0.3	0.7	1.0	0.0	0.0	0.3	0.3	0.3	0.0
MICHAEL	0.0	0.0	0.3	0.3	1.0	1.0	0.0	0.0	0.0	0.0
BILL	0.0	0.0	0.0	0.0	1.0	0.8	0.0	0.0	0.0	0.0
HARRY	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
DON	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
HOLLY	0.0	0.0	0.0	0.0	0.8	1.0	0.3	0.3	0.3	0.3
PAULINE	0.0	0.0	0.0	0.3	0.0	0.0	1.0	1.0	1.0	0.7
PAT	0.0	0.0	0.0	0.0	0.0	0.3	1.0	0.7	0.3	0.3
PAM	0.0	0.0	0.0	0.0	0.0	0.3	0.7	1.0	1.0	1.0
CAROL	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.7	0.3
ANN	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.7	1.0	1.0
JENNIE	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3	0.7	1.0

Xij = 1.0 if node i is member of clique j

Rows & cols sorted by Optimal Scaling scores

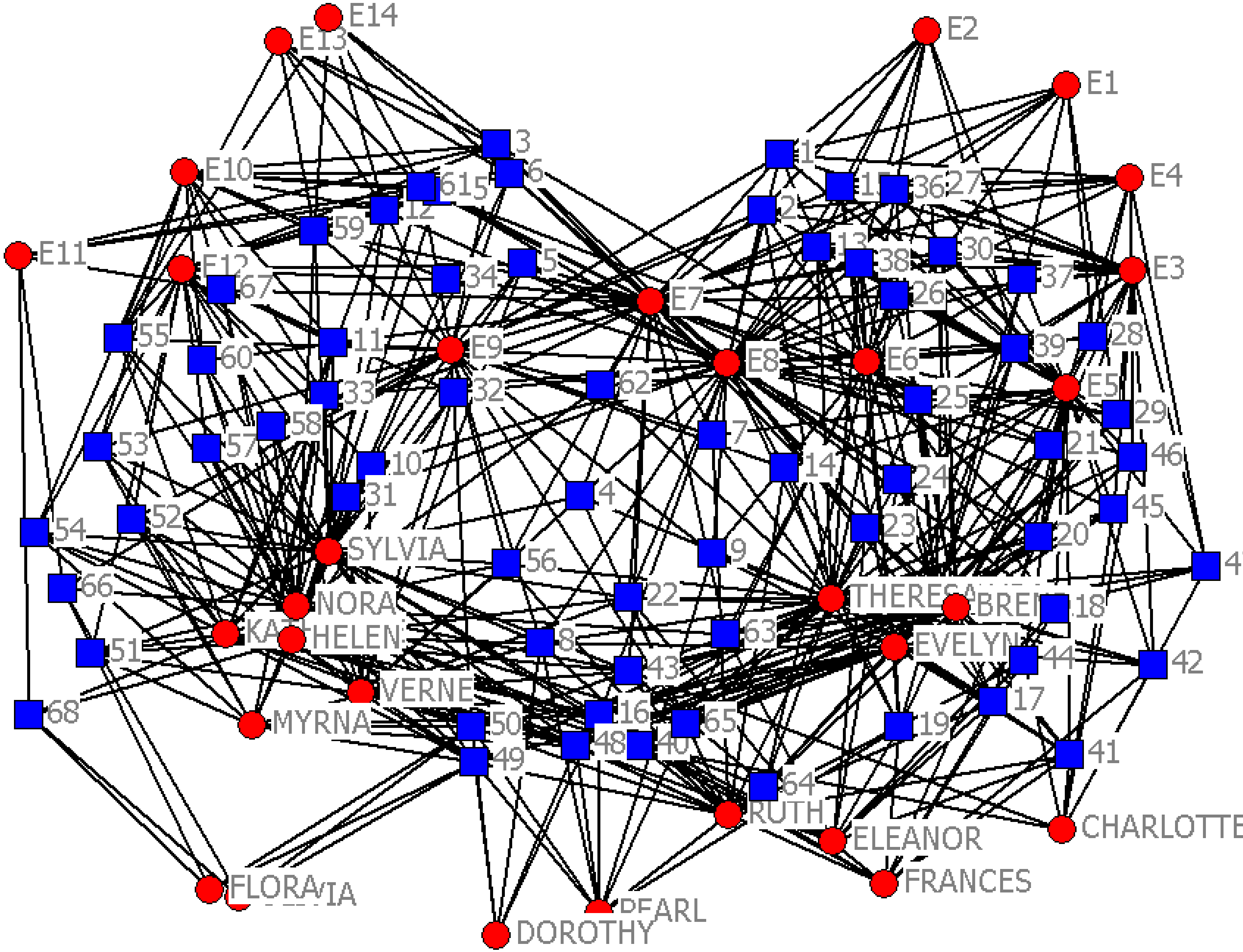
2-mode visualization of valued clique membership



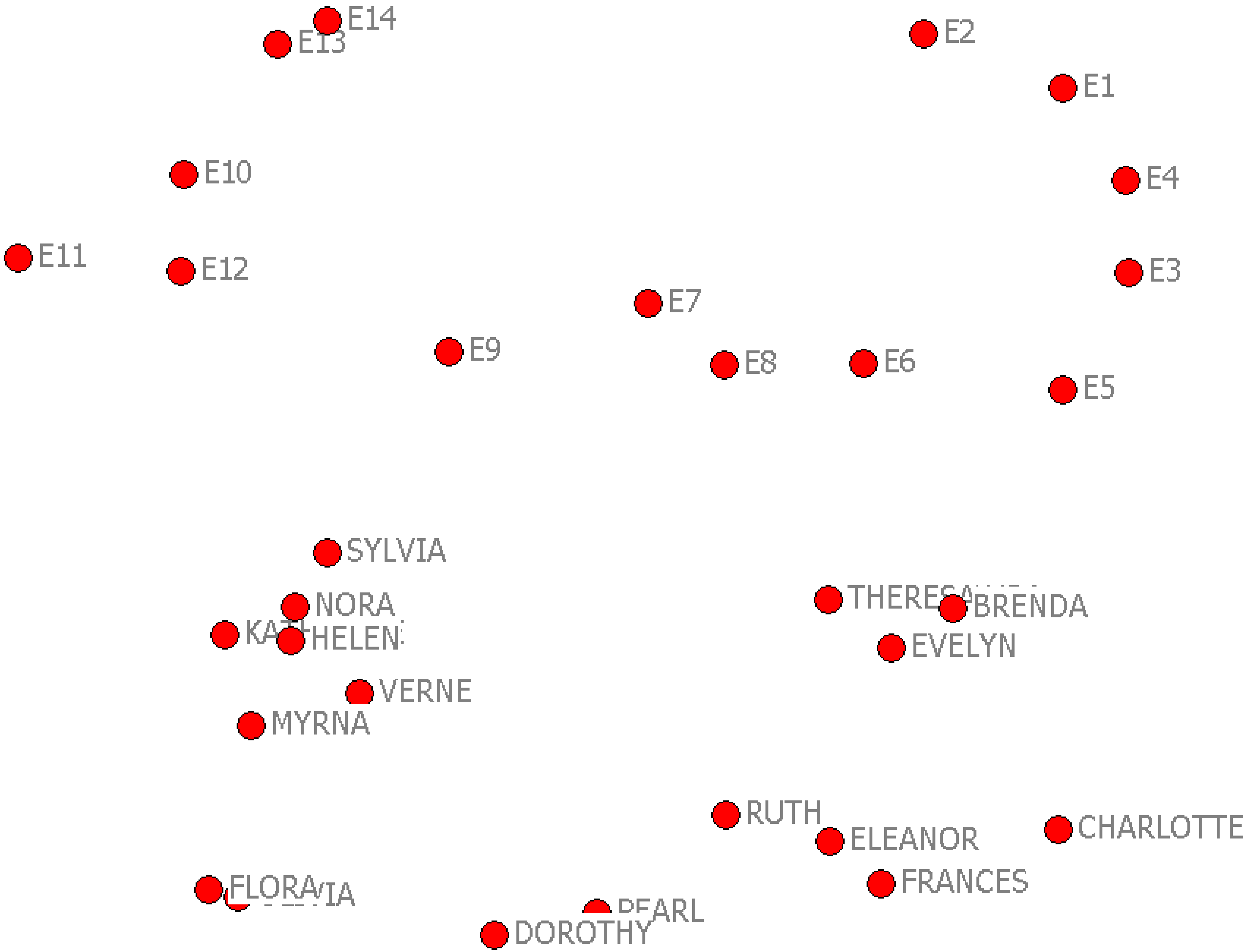
Valued clique participation analysis of bi-clique analysis of Davis dataset

- Run bi-cliques on 2-mode Davis dataset
  - Equivalently, run 2-cliques on 32x32 bi-partite adjacency matrix
  - 68 bi-cliques are found*
- Construct 32x68 valued node by bi-clique matrix X, where
  - x_{ij} = no. of ties node *i* has with clique *j*

* 63 if we remove bi-cliques that are mono-modal



E14
E13
E2
E1
E10
E4
E11
E12
E3
E7
E9
E8
E6
E5
SYLVIA
NORA
HELEN
VERNE
MYRNA
EVELYN
RUTH
ELEANOR
CHARLOTTE
FRANCES
FLORA
PEARL
DOROTHY


E14
E13
E2
E1
E10
E4
E11
E12
E3
E7
E9
E8
E6
E5
SYLVIA
NORA
THERESA
BRENDA
HELEN
EVELYN
VERNE
MYRNA
RUTH
ELEANOR
CHARLOTTE
FRANCES
FLORA
PEARL
DOROTHY

Structural blockmodeling of affiliations

		Corporate Board				
		1	2	3	4	5
	1	1	1	0	0	0
	2	1	1	0	0	0
Person	3	1	1	1	1	1
	4	1	1	1	1	1
	5	0	0	1	1	1
	6	0	0	1	1	1
	7	0	0	1	1	1

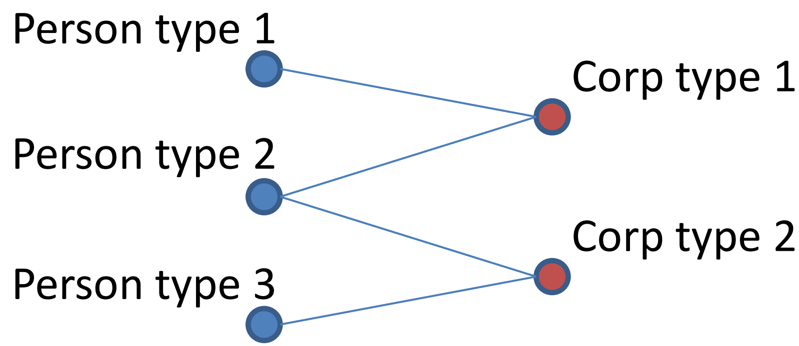


Fig. 4. Structural blockmodel of a 2-way, 2-mode matrix.

- 3 kinds of people, 2 kinds of corporations
- Characteristic pattern of 1-blocks and 0-blocks

Borgatti, S. P., & Everett, M. G. 1992. Regular blockmodels of multiway, multimode matrices. Social Networks, 14: 91-120.

Definition & computation

- A blocking is structural if for all p(i)=p(k) and p(j)=p(m), $x_{ij} = x_{km}$
  - If cell (i,j) and cell (k,m) belong to same matrix block, they must have same value
- Optimal scaling
  - SVD of X gives same score to identical rows/columns
- Combinatorial optimization
  - Find partitions P1, P2 so as to minimize sum of variances within matrix blocks

	1	2	3	4	5
1	1	1	0	0	0
2	1	1	0	0	0
3	1	1	1	1	1
4	1	1	1	1	1
5	0	0	1	1	1
6	0	0	1	1	1
7	0	0	1	1	1

Row Scores

1	p1	-1.863
2	p2	-1.863
3	p3	-0.149
4	p4	-0.149
5	p5	0.994
6	p6	0.994
7	p7	0.994

Column Scores

1	c1	-1.369
2	c2	-1.369
3	c3	0.730
4	c4	0.730
5	c5	0.730

Regular blockmodeling of affiliations

			Restaurant				
			A		B		
			1	2	3	4	5
Person	X	1	1	1	1		1
		2	1	1		1	
		3	1			1	1
		4	1		1		1
		5		1		1	
	Y	6			1		1
		7				1	

Fig. 9. Regular blockmodel of matrix in Figure 8.

	A	B
X	1	1
Y		1

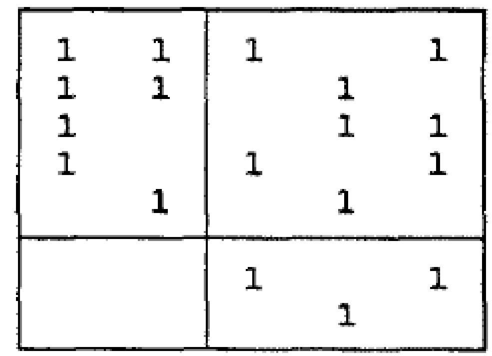
Fig. 10. Image matrix

- 2 kinds of people (e.g., rich,poor), 2 kinds of restaurants
- Characteristic pattern of regular 1-blocks and 0-blocks
  - Must be a 1 in every row and every column of matrix block, or all zeros

Borgatti, S. P., & Everett, M. G. 1992. Regular blockmodels of multiway, multimode matrices. Social Networks, 14: 91-120.

Definition & computation

- A regular blocking is one in which if any cell in a matrix block contains a given value, then every row and column in that block must contain an instance of that value



- Combinatorial optimization
  - Find partitions P1, P2 so as to minimize errors
    - In a 1-block, lack of a 1 in any row or column is an error
    - In a 0-block, presence of a 1 is an error

Side note: regular blockmodels of valued multi-mode matrices

		Restaurant				
		A		B		
		1	2	3	4	5
Person X	1	1	1	1	2	1
	2	1	1	2	1	2
Person Y	3	1	2	2	1	1
	4	1	2	1	2	1
	5	2	1	2	1	2
Person Z	6	1	1	2	2	2
	7	1	1	2	2	2

1 = Lunch
2 = Dinner

- In an adj matrix, if we recode nulls to 2s and apply this definition of regular equiv, the resulting blockmodel will hold for the complement of the graph

Section 3

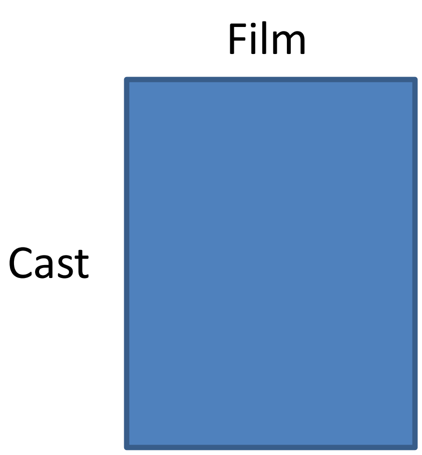
FUTURE DIRECTIONS, APPLICATIONS ETC

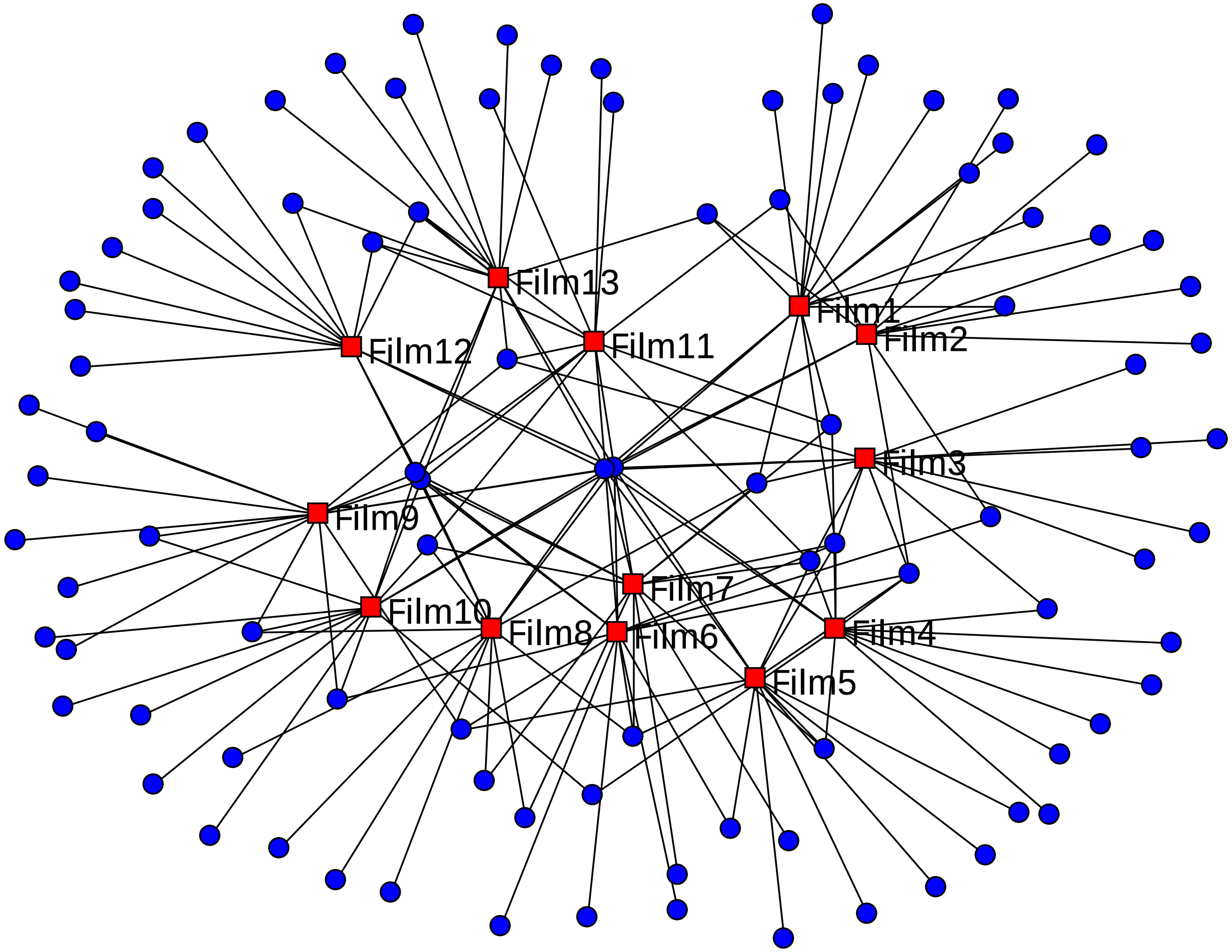
Under-developed areas

- Trajectories
  - e.g., careers
    - Movement of actors through a resource or achievement space
- Relational algebras

Social trajectories

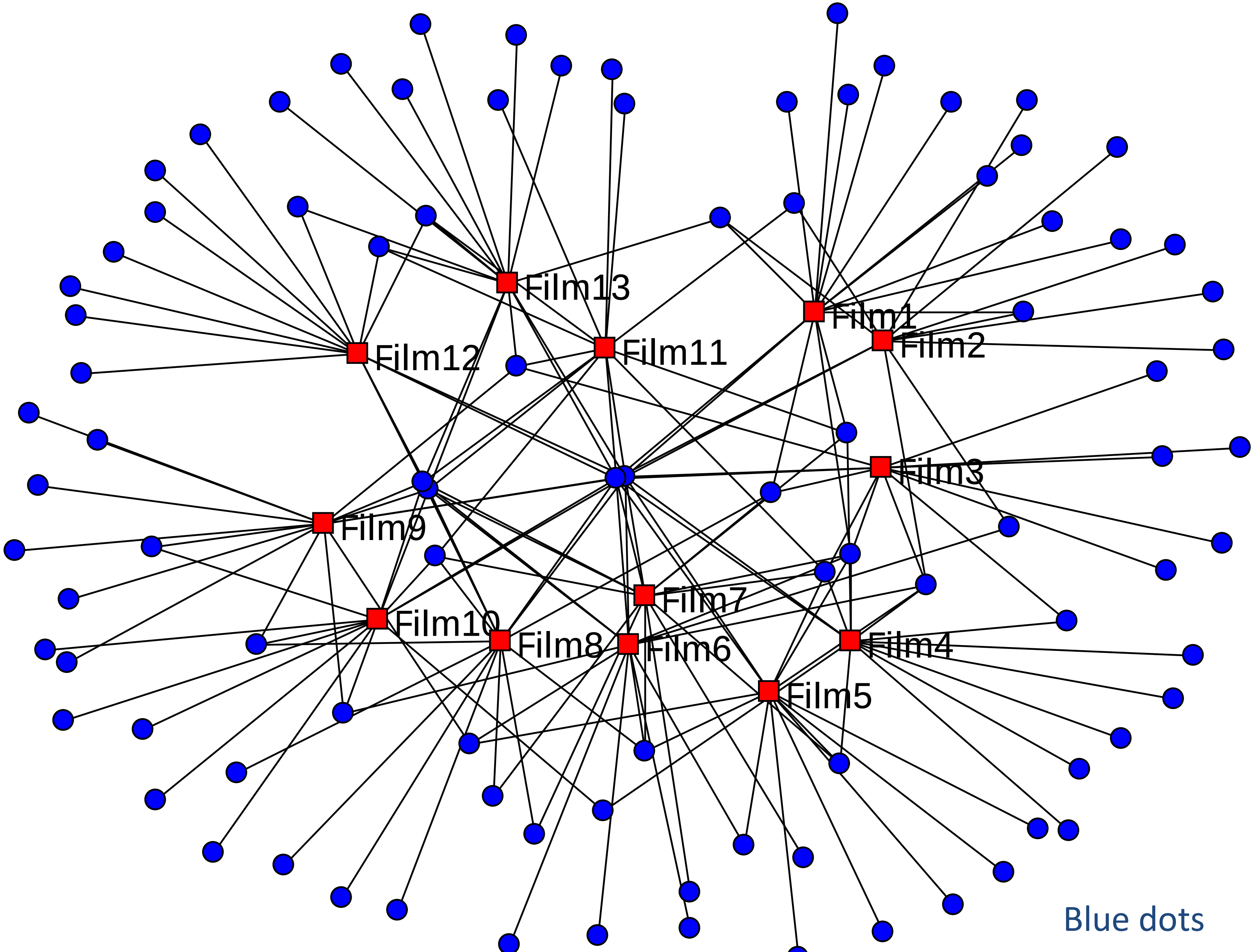
- First 13 films of Pedro Almodóvar
- Actor by film incidence matrix
- Films identified by chronological order





Blue dots are actors

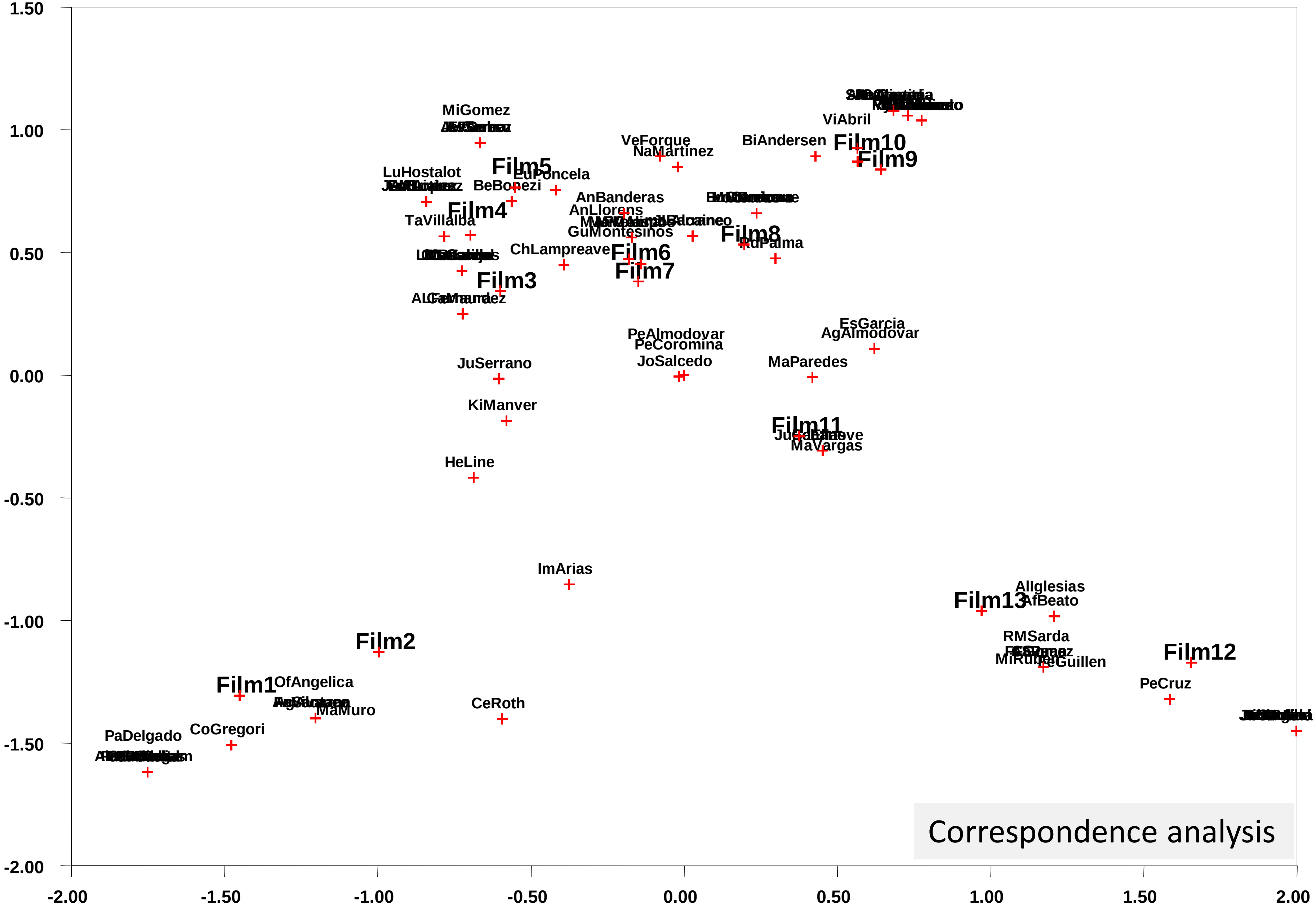
- First 13 films of Pedro Almodovar
- Actor by film incidence matrix
- Films identified by chronological order



Blue dots are actors

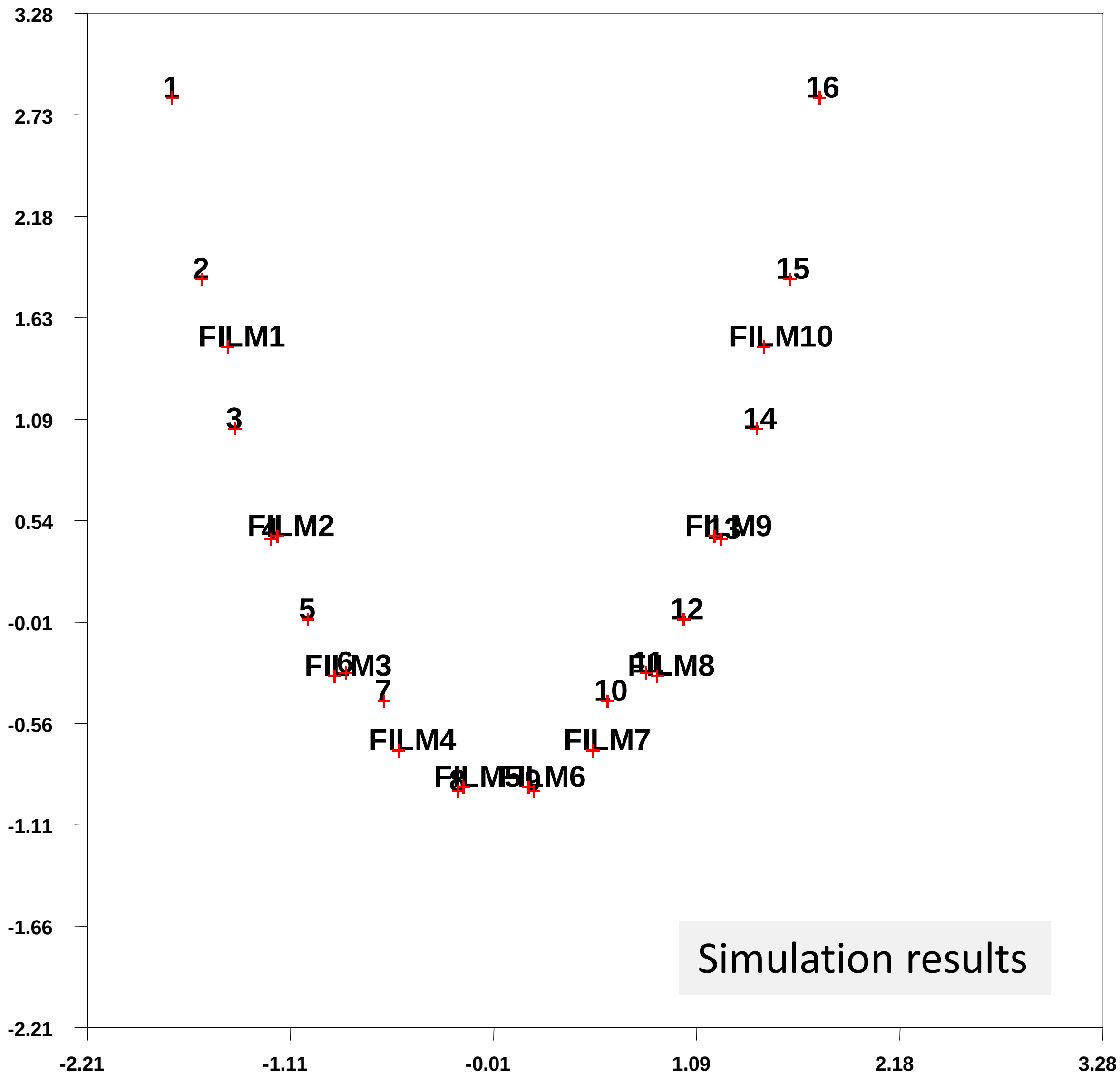
Films adjacent in time have overlapping sets of actors

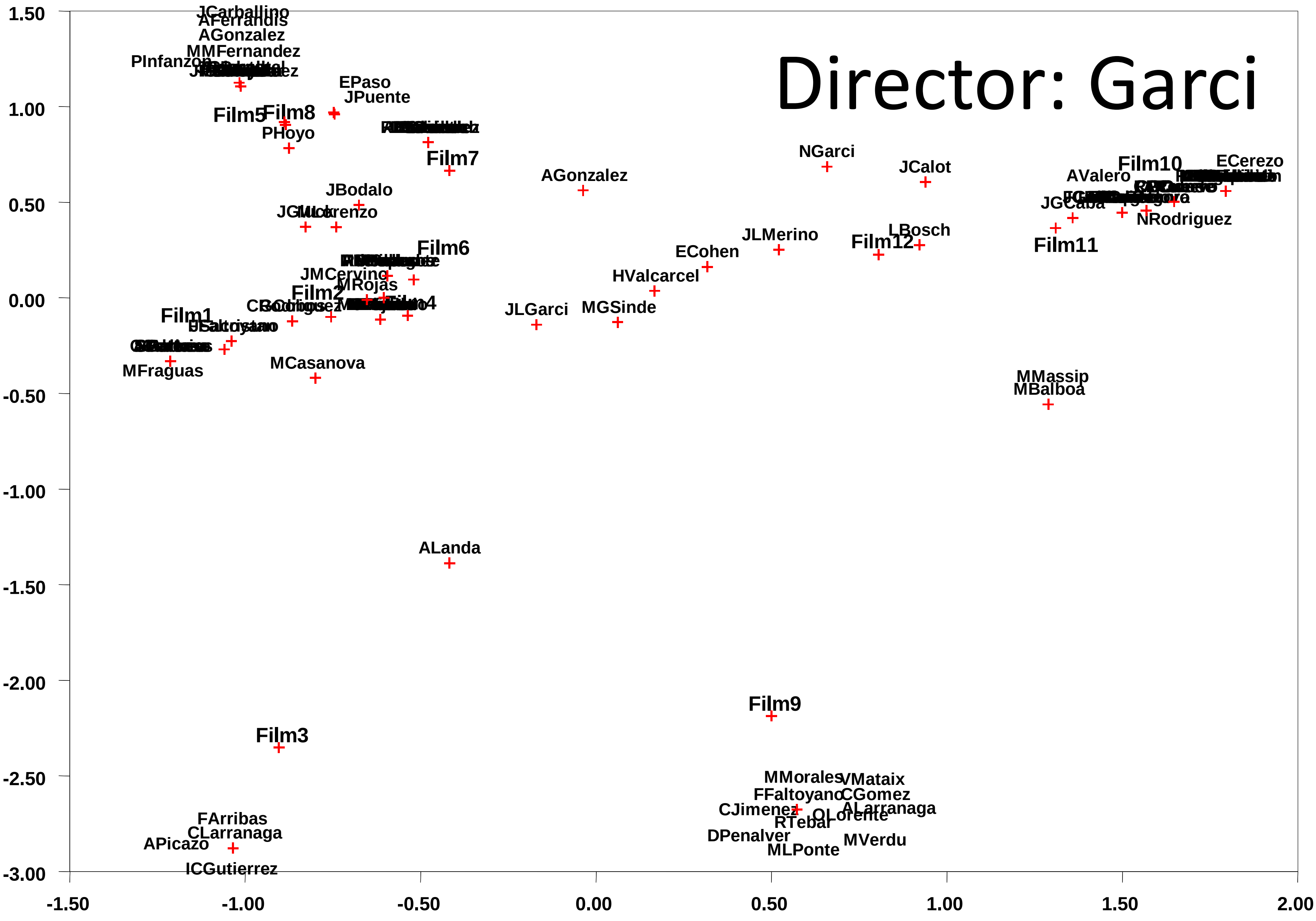
Director: Almodovar



Idealized pattern

Each successive film carries over a number of actors from previous film, and adds new ones



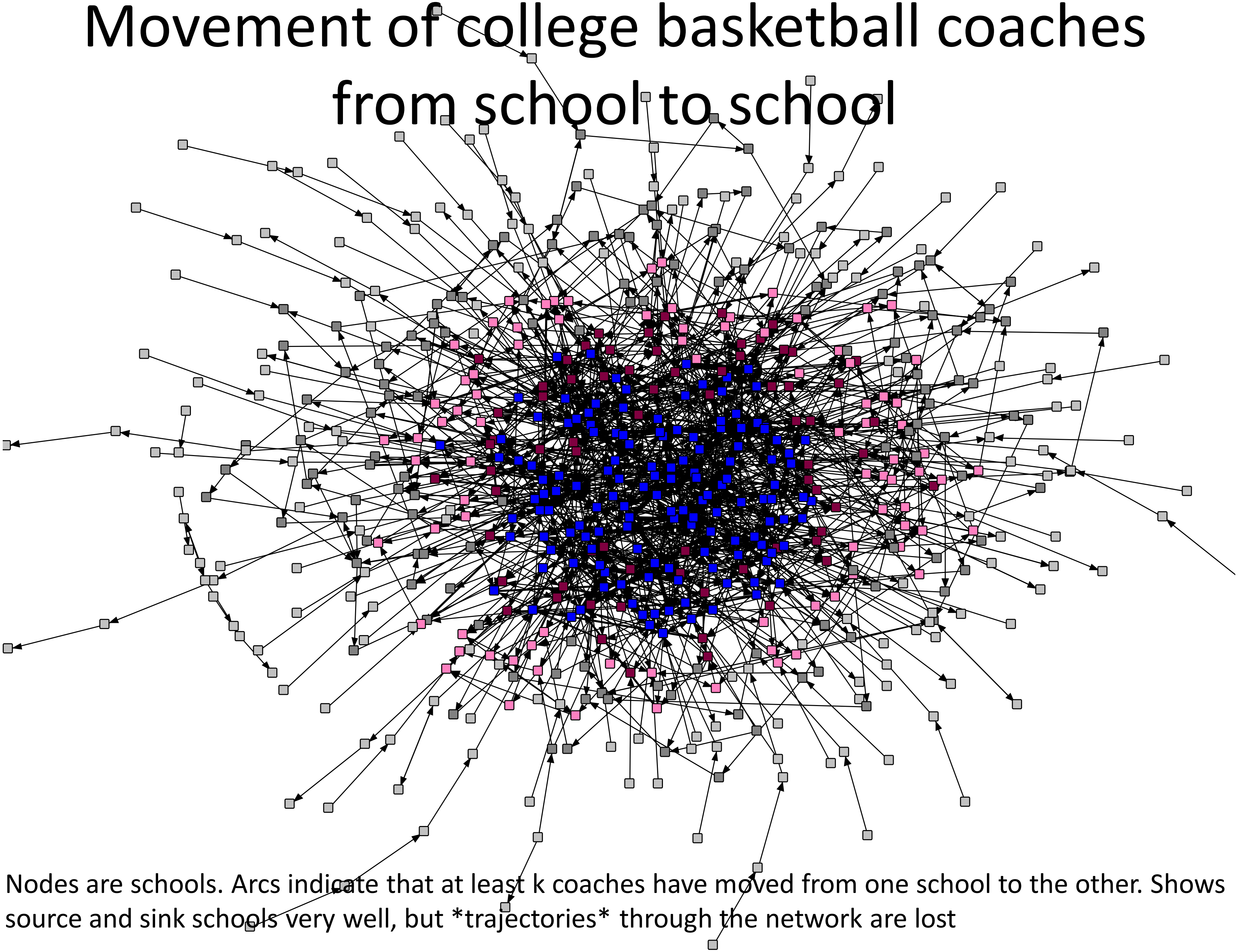


Director: Garci
Film1
Film2
Film3
Film4
Film5
Film6
Film7
Film8
Film9
Film10
Film11
Film12
JCarballino
AFerrandis
AGonzalez
MMFernandez
PInfanzon
EPaso
JPuente
PHoyo
JBodalo
MCasanova
MFraguas
JMCervino
MRojas
AGonzalez
NGarci
JCalot
AValero
ECerezo
JGCaba
NRodriguez
LBosch
JLMerino
ECohen
HValcarcel
JLGarci
MGSinde
MMassip
MBalboa
ALanda
MMorales
VMataix
FFaltoyano
CGomez
CJimenez
ALarranaga
OLorente
RTebar
DPenalver
MVerdu
MLPonte
FArribas
CLarranaga
APicazo
ICGutierrez
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
-3.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00

Different kind of trajectories

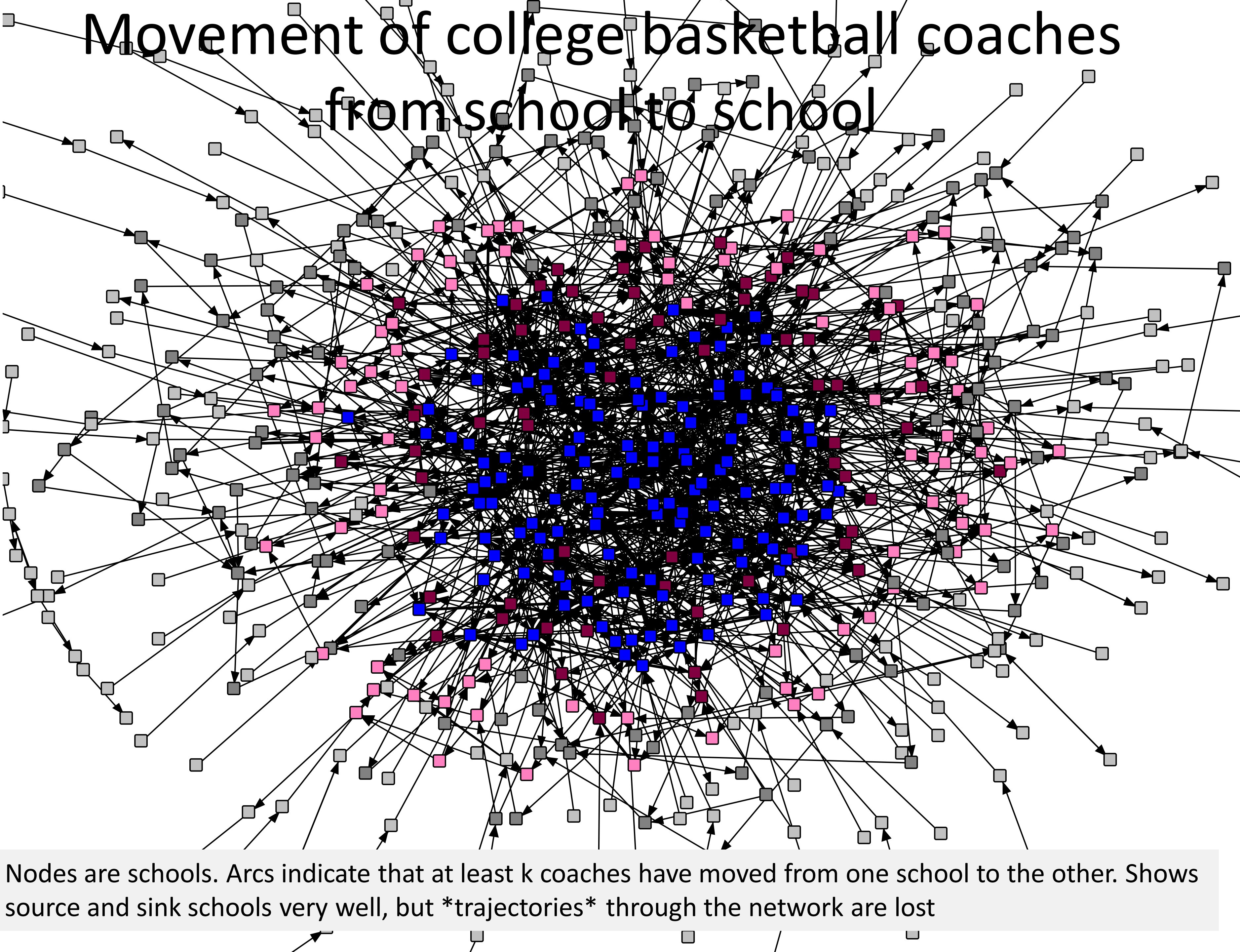
- Halgin's (2009) study of careers of college basketball coaches
- Raw data is coach by school by year matrix X
  - x_{ijk} = 1 if coach *i* was coaching for school j in year k
    x_{ijk} = 0 otherwise
- Conversion approach: reduction to 1-mode
  - School by school transition matrix T
  - t_{ij} = number of coaches that have moved from school *i* to school *j*
    - Can make row-stochastic, etc.

Movement of college basketball coaches from school to school



Nodes are schools. Arcs indicate that at least k coaches have moved from one school to the other. Shows source and sink schools very well, but *trajectories* through the network are lost

Movement of college basketball coaches from school to school

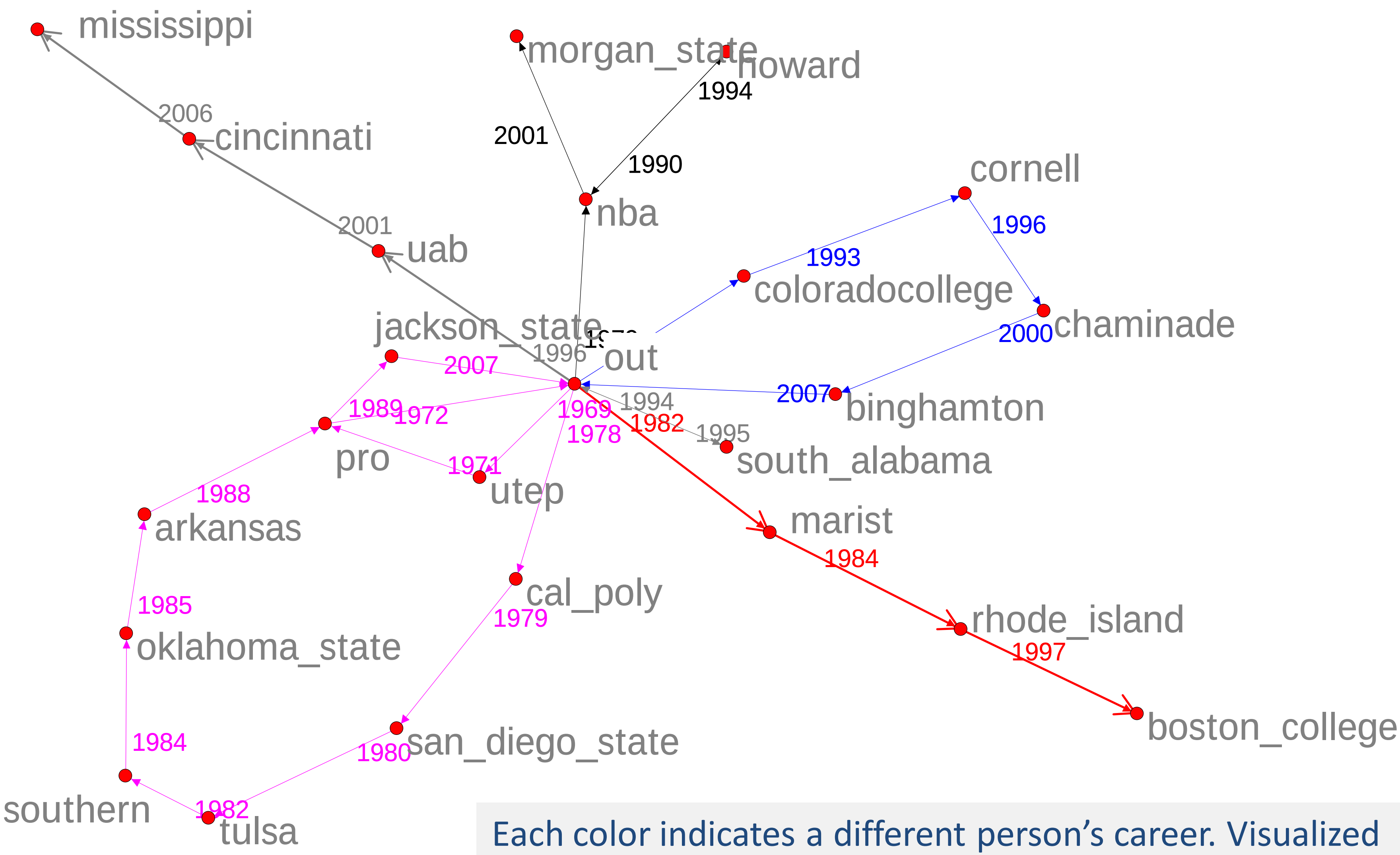


Nodes are schools. Arcs indicate that at least k coaches have moved from one school to the other. Shows source and sink schools very well, but *trajectories* through the network are lost

Retaining the paths

- Construct directed multi-relational valued graph
  - Nodes are schools
  - Relations are individual coaches
    - An arc from school A to school B in relation X indicates that coach X moved from A to B
  - Arc weights represent the year of the move
- Trajectories are paths (or trails, or walks) through an unknown network of possibilities

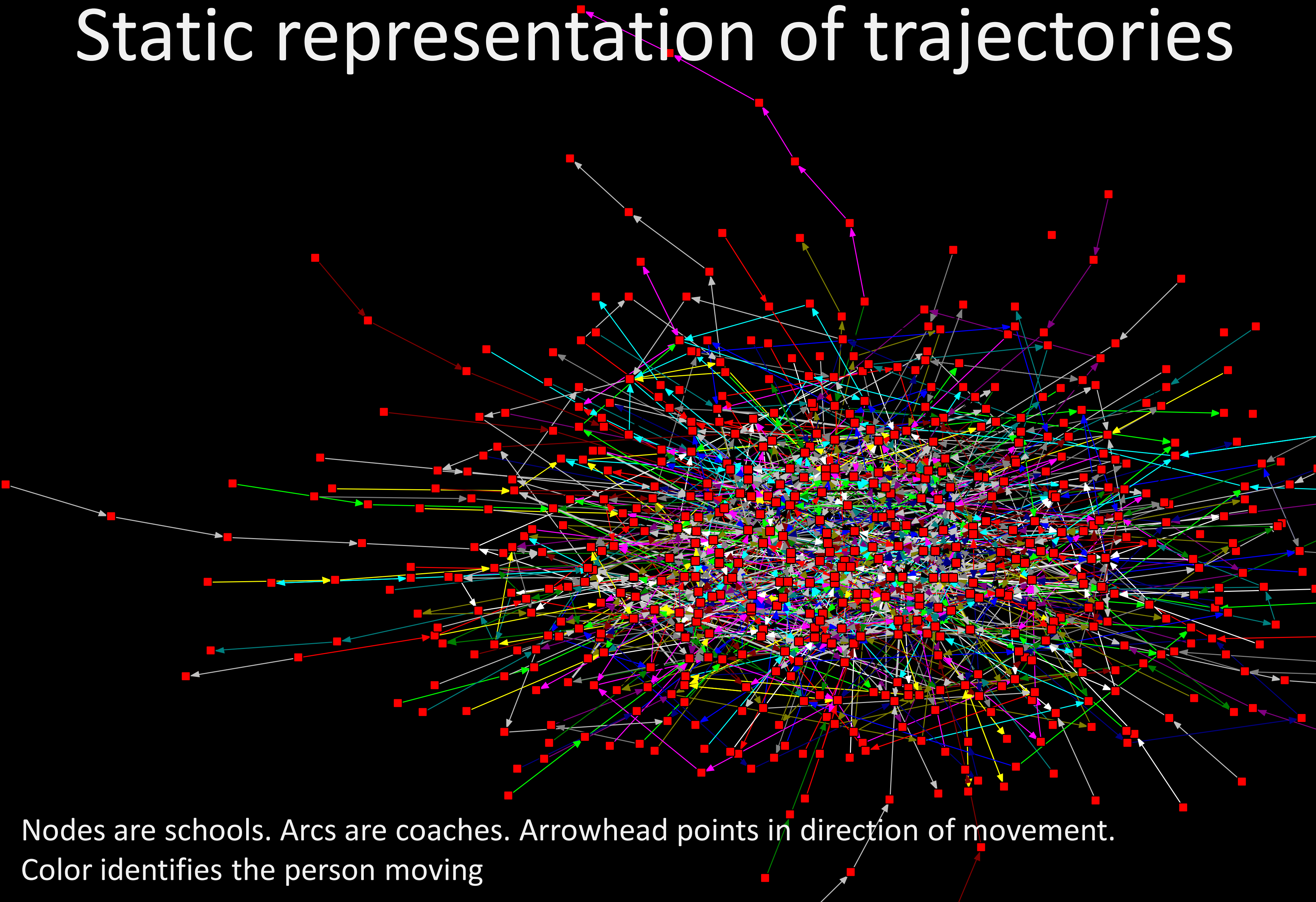
Retaining the paths*



Each color indicates a different person's career. Visualized in NetDraw as a valued graph multi-relational graph. Each "relation" is a coach. Values are years.

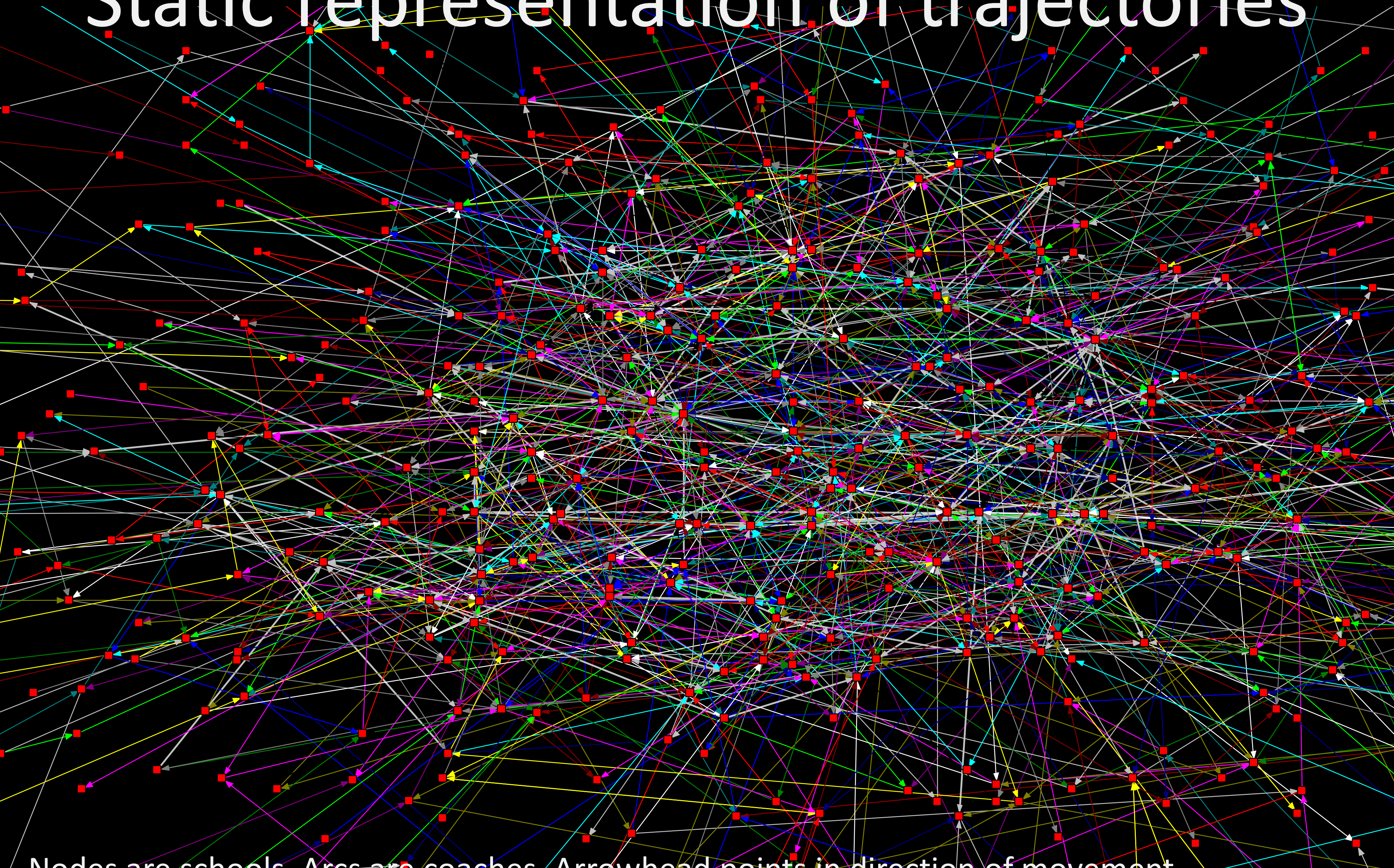
*Well, most of them

Static representation of trajectories



Nodes are schools. Arcs are coaches. Arrowhead points in direction of movement.
Color identifies the person moving

Static representation of trajectories

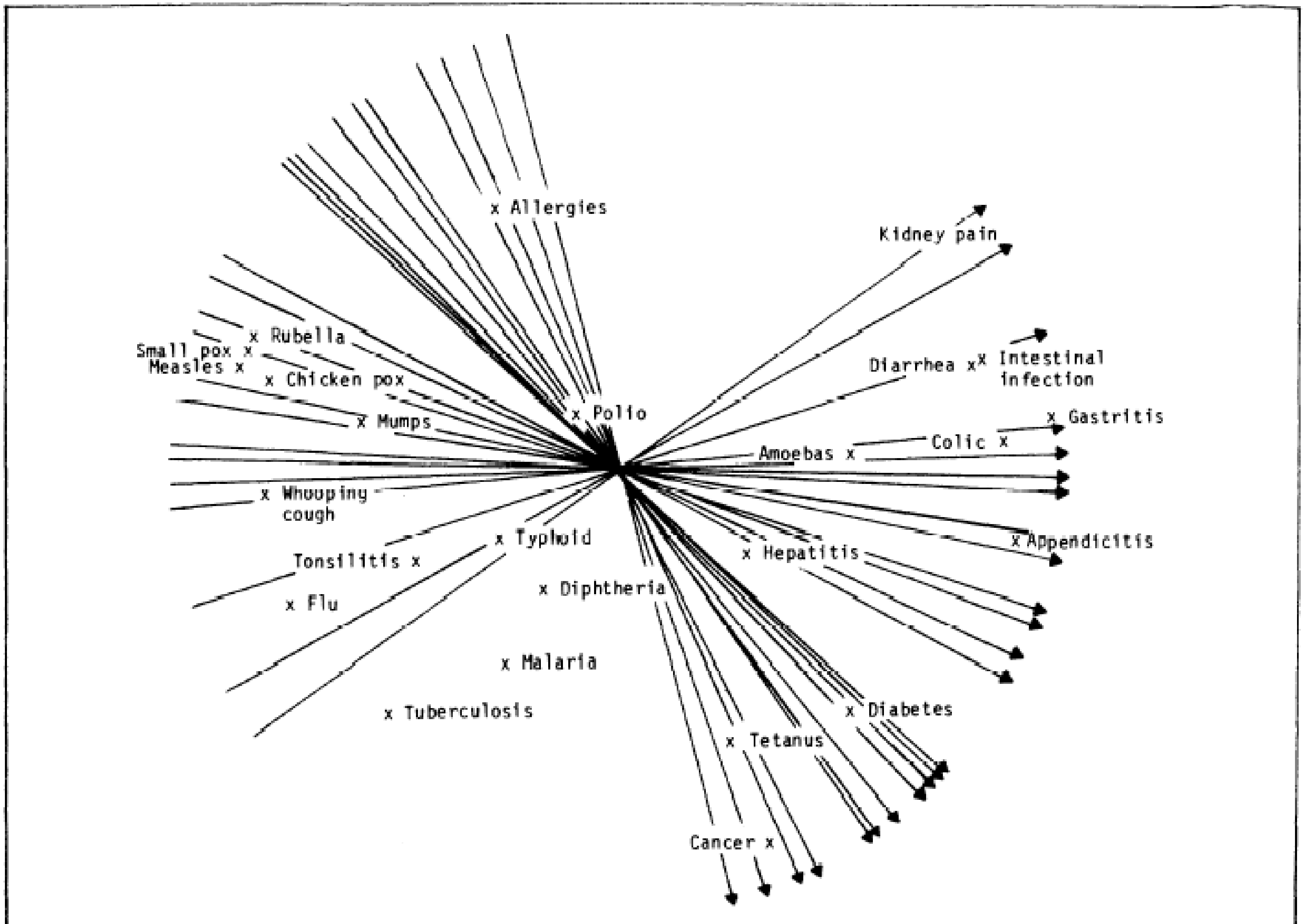


Nodes are schools. Arcs are coaches. Arrowhead points in direction of movement

trajectories

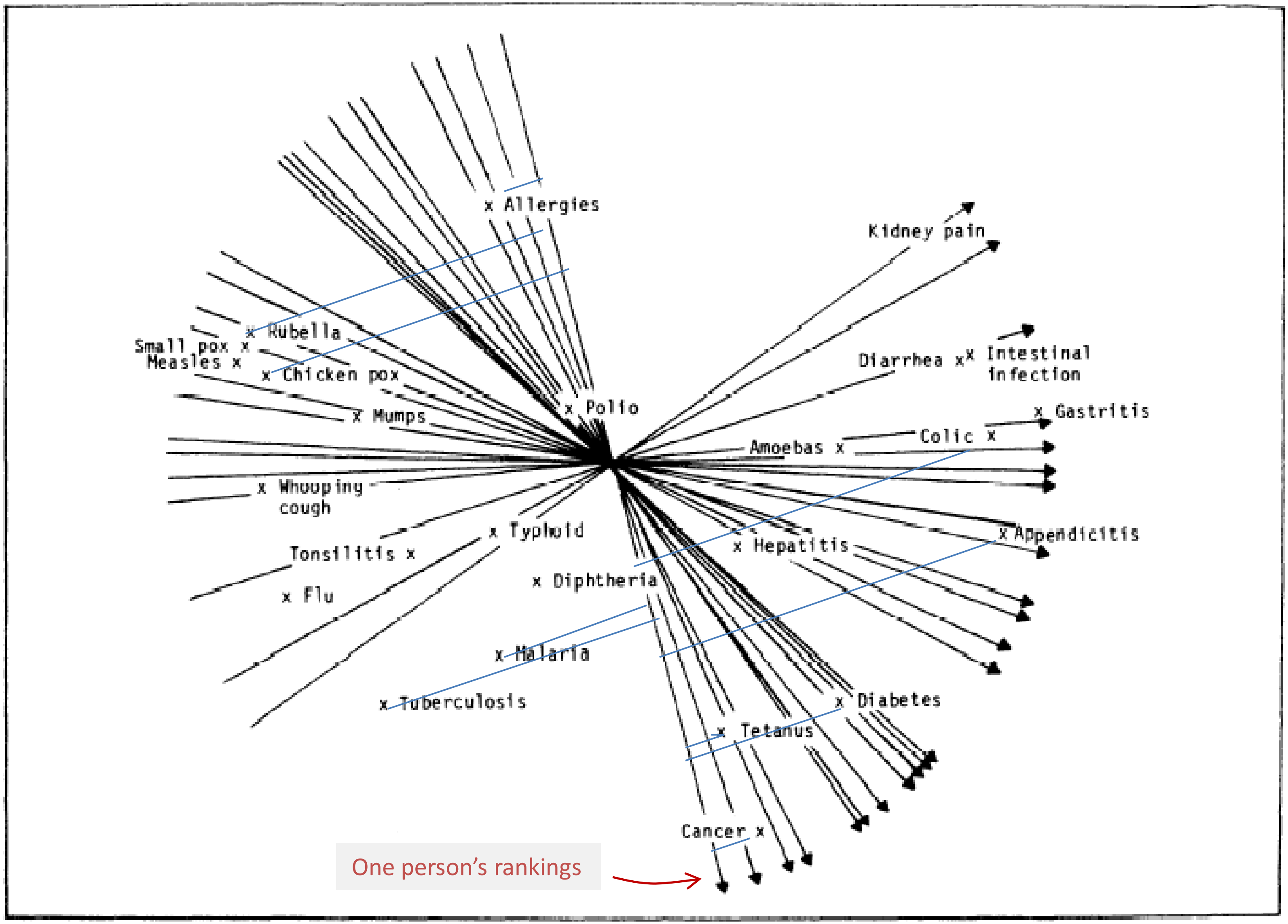
- How to characterize trajectories
  - Paths, trails, walks; Length
  - Node attributes (e.g., each move is to higher status node)
  - Typicality
- Visualization
  - If node distance proportional to overall transition probabilities, long arcs represent non-typical moves
  - Clustering of nodes and trajectories to maximize simplicity
  - Vector models. Position nodes so that each trajectories through the space can be represented by a vector

Age of the Infirm (Guatemala)

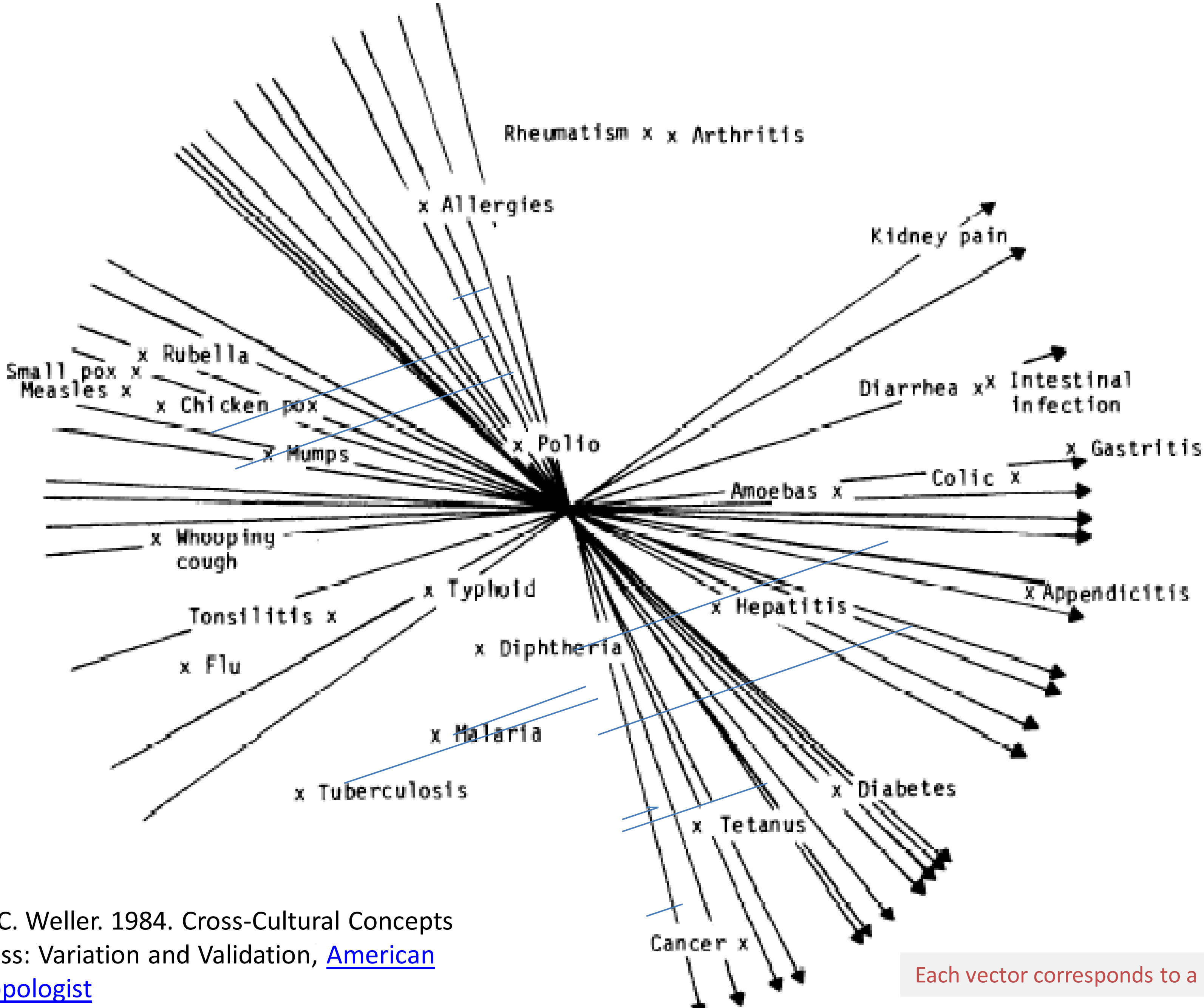


Susan C. Weller. 1984. Cross-Cultural Concepts of Illness: Variation and Validation, American Anthropologist

Age of the Infirm (Guatemala)



Susan C. Weller. 1984. Cross-Cultural Concepts of Illness: Variation and Validation, [American Anthropologist]



Susan C. Weller. 1984. Cross-Cultural Concepts of Illness: Variation and Validation, American Anthropologist

Each vector corresponds to a respondent

Relational Algebras

- E.g., bibliometric data
  - Each article has authors, topical keywords, year,
    journal, etc.
  - View as k-mode, or interlinked 2-mode matrices
    - Author by article (A)
    - Keyword by article (K)
    - Journal by article, etc. (J)
- Some 2-word compositions
  - AA': co-authorship
  - AK': author by keyword
  - KK': keyword co-occurrence
  - AJ': author by journal
  - KJ': keyword appears in journal

- Higher compositions
  - (AK')(AK')' = AK'KA' = similarity of authors across keywords
    - Does AK'KA at T1 predict AA' (co-authorship) at T2?
    - i.e., are people who work on same topics more likely to co-author papers?
  - AK'KJ' whether an author has published on a topic that has appeared in a given journal
    - Does AK'KJ' predict AJ'?
    - i.e., people publish in journals that are relevant to their research

Conclusion

- Comments on the terminology & boundaries of 2-mode domain
  - Affiliation graphs
  - A few caveats about 2-mode
- Overview of techniques currently in use
  - Normalization still rare in conversion approach
  - Many opportunities in the direct approach
- Some 2-mode application areas
  - Trajectories
  - Relational algebras

Better get to our lunch before someone else does …