Network Analysis:

The Hidden Structures behind the Webs We Weave

17-338 / 17-668

Affiliations and Overlapping Subgroups

Tuesday, October 7, 2025

Patrick Park

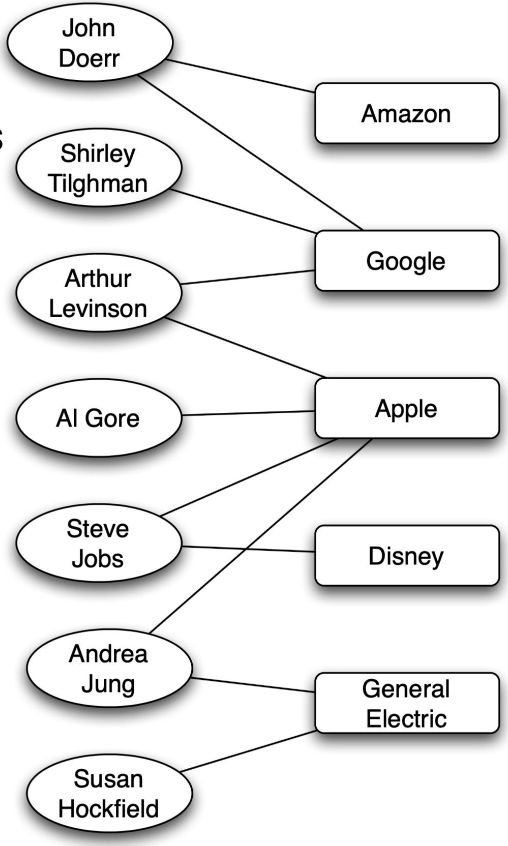
Carnegie Mellon University
School of Computer Science

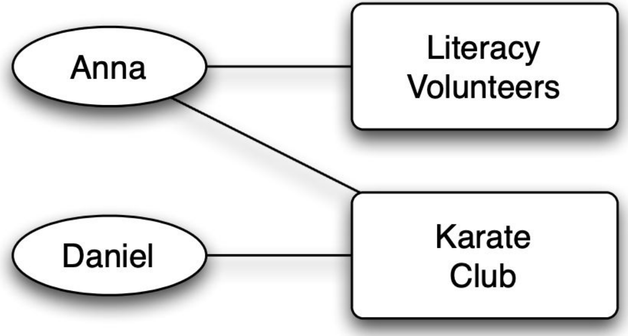
S3D
Software and Societal Systems Department

Affiliation networks are two-mode networks

We can represent the participation of a set of “actors” (people) in a set of “events” (activities) using a graph

People on corporate boards of directors

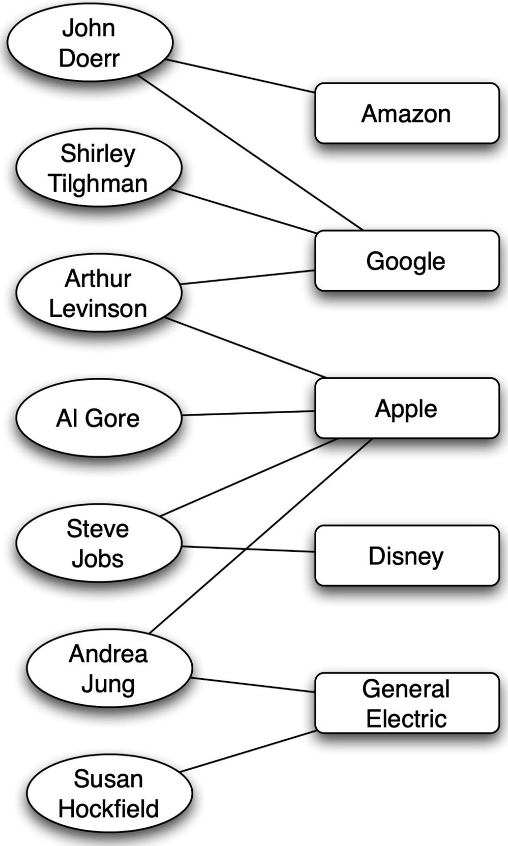


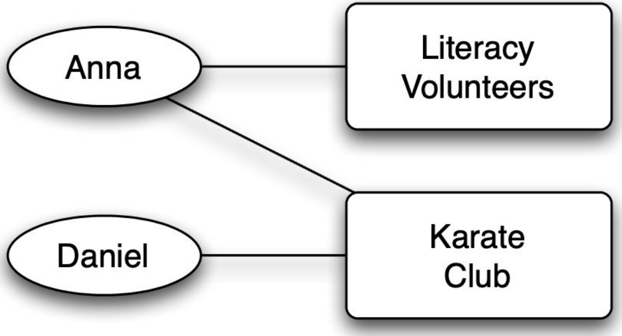


Individuals affiliated with groups or activities

Affiliation networks can reveal interesting relationships on both sides of the graph.

Events linked by authors



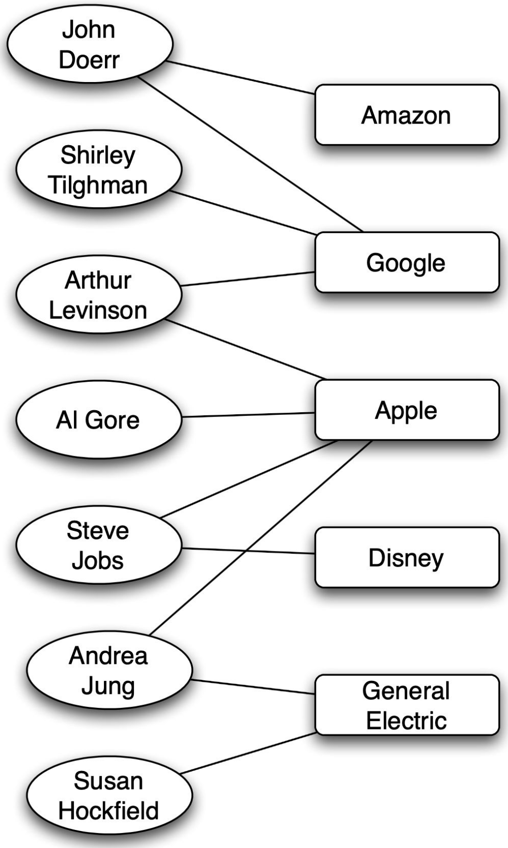


Actors linked by events

Affiliation networks can reveal interesting relationships on both sides of the graph.

Two companies are implicitly linked by having the same person sit on both company boards → possible conduits for information diffusion and influence between the two companies.

Two people are implicitly linked by serving together on a board → The board creates a "focus" of interaction for the two.



Three rationales for studying affiliation networks

Individuals' participation in events provide direct linkages between the events.

- Event – event

Contact among individuals who participate in the same social events increases the likelihood of tie formation.

- Individual – individual

The interaction between individuals and events forms a social system that can be studied as a whole.

- Individual – event

Affiliation networks can be represented as a matrix

The affiliation matrix A = {a_{ij}}:

- rows are individuals
- columns as events.

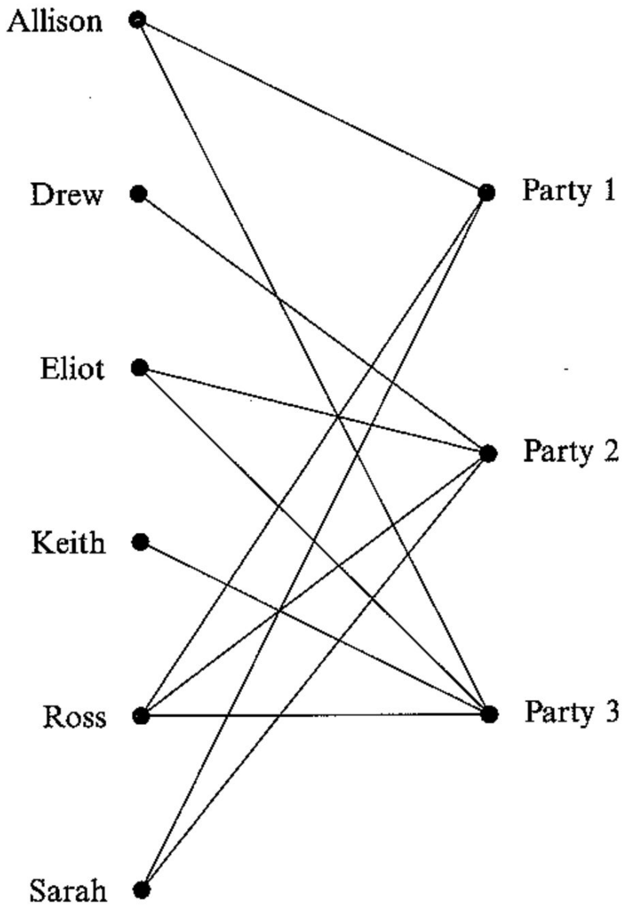
a_{ij} = 1 if individual *i* is affiliated with event *j*

Actor	Event Party 1	Event Party 2	Event Party 3
Allison	1	0	1
Drew	0	1	0
Eliot	0	1	1
Keith	0	0	1
Ross	1	1	1
Sarah	1	1	0

(Wasserman & Faust)

… or as a bipartite graph

Note the context-dependent meaning of “degree.”



(Wasserman & Faust)

The affiliation network can also be represented as a sociomatrix

	Allison	Drew	Eliot	Keith	Ross	Sarah	Party 1	Party 2	Party 3
Allison	-	0	0	0	0	0	1	0	1
Drew	0	-	0	0	0	0	0	1	0
Eliot	0	0	-	0	0	0	0	1	1
Keith	0	0	0	-	0	0	0	0	1
Ross	0	0	0	0	-	0	1	1	1
Sarah	0	0	0	0	0	-	1	1	0
Party 1	1	0	0	0	1	1	-	0	0
Party 2	0	1	1	0	1	1	0	-	0
Party 3	1	0	1	1	1	0	0	0	-

(Wasserman & Faust)

The bipartite graph can also be represented as a sociomatrix

A:

	Allison	Drew	Eliot	Keith	Ross	Sarah	Party 1	Party 2	Party 3
Allison	-	0	0	0	0	0	1	0	1
Drew	0	-	0	0	0	0	0	1	0
Eliot	0	0	-	0	0	0	0	1	1
Keith	0	0	0	-	0	0	0	0	1
Ross	0	0	0	0	-	0	1	1	1
Sarah	0	0	0	0	0	-	1	1	0
Party 1	1	0	0	0	1	1	-	0	0
Party 2	0	1	1	0	1	1	0	-	0
Party 3	1	0	1	1	1	0	0	0	-

(Wasserman & Faust)

We can summarize the co-membership frequencies

The product of A and A' (transpose) $\mathbf{X}^{\mathcal{N}} = \mathbf{AA}'$.

	Party 1	Party 2	Party 3
Allison	1	0	1
Drew	0	1	0
Eliot	0	1	1
Keith	0	0	1
Ross	1	1	1
Sarah	1	1	0

X

	Allison	Drew	Eliot	Keith	Ross	Sarah
Party 1	1	0	0	0	1	1
Party 2	0	1	1	0	1	1
Party 3	1	0	1	1	1	0

=

	Allison	Drew	Eliot	Keith	Ross	Sarah
Allison	**2**	0	1	1	2	1
Drew	0	**1**	1	0	1	1
Eliot	1	1	**2**	1	2	1
Keith	1	0	1	**1**	1	0
Ross	2	1	2	1	**3**	2
Sarah	1	1	1	0	2	**2**

Similarly, we can summarize event overlap frequencies

The product of A' and A

$$\mathbf{X}^{\mathscr{M}} = \mathbf{A}'\mathbf{A}$$

	Allison	Drew	Eliot	Keith	Ross	Sarah
Party 1	1	0	0	0	1	1
Party 2	0	1	1	0	1	1
Party 3	1	0	1	1	1	0

X

	Party 1	Party 2	Party 3
Allison	1	0	1
Drew	0	1	0
Eliot	0	1	1
Keith	0	0	1
Ross	1	1	1
Sarah	1	1	0

=

	Party 1	Party 2	Party 3
Party 1	3	2	2
Party 2	2	4	2
Party 3	2	2	4

A take on “community” in this context

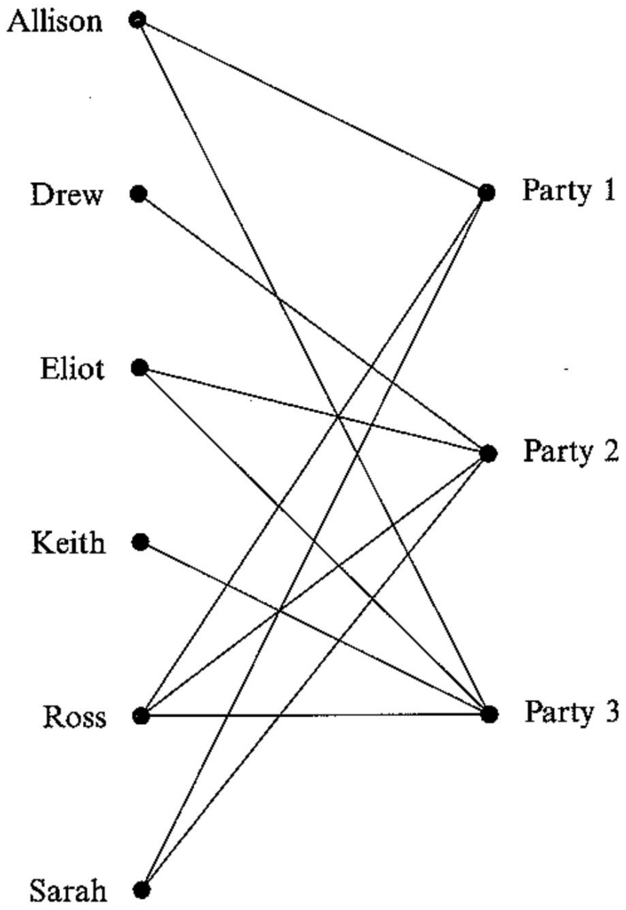
Community as clique at level c

For the co-membership relation for actors:

- Subgraph in which all *pairs* of actors share memberships in at least c events

For the overlap relation for events:

- Subgraph in which all *pairs* of events share at least c members



(Wasserman & Faust)

Affiliation Network Studies

Examples of Two-Mode Networks

Political polarization and the structure of the board interlock network

- Longer geodesic, less cohesion

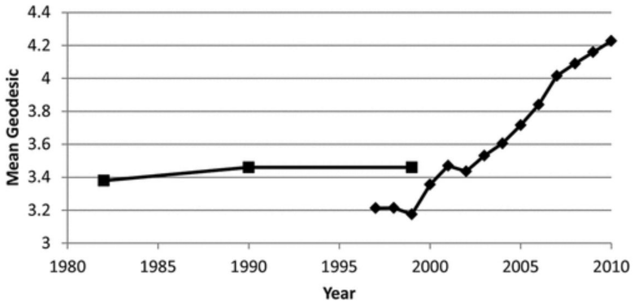


Fig. 1. Mean geodesic in main component of board interlock networks, 1982–2010. Data for 1982–99 are from Davis et al. (2003); 1997–2010, this study; study population differs across the sources.

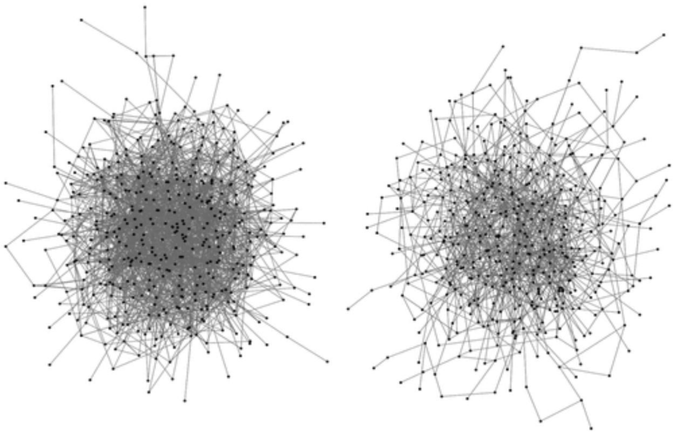


Fig. 2. S&P 500 interlock network main component, 1996 (*left*) and 2010 (*right*)

Chu and Davis 2016

Examples of Two-Mode Networks

Political polarization and the structure of the board interlock network

- Longer geodesic, less cohesion

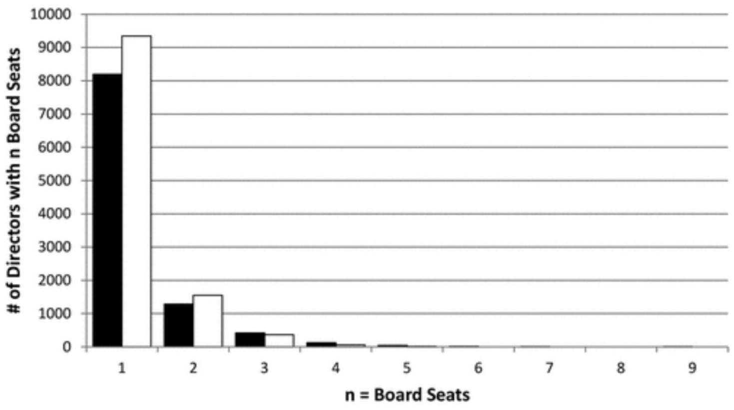


Fig. 3. Distribution of directors by number of S&P 1500 board seats, 2000 (*black bars*) and 2010 (*white bars*).

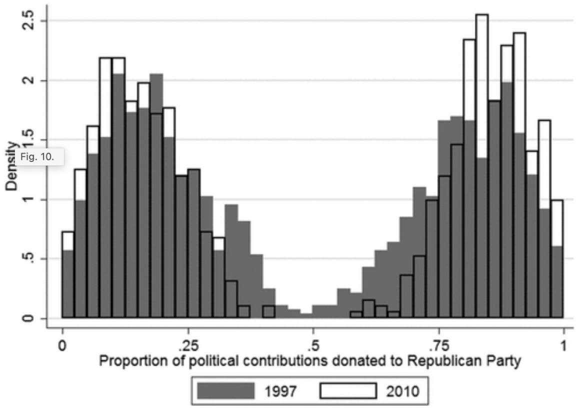


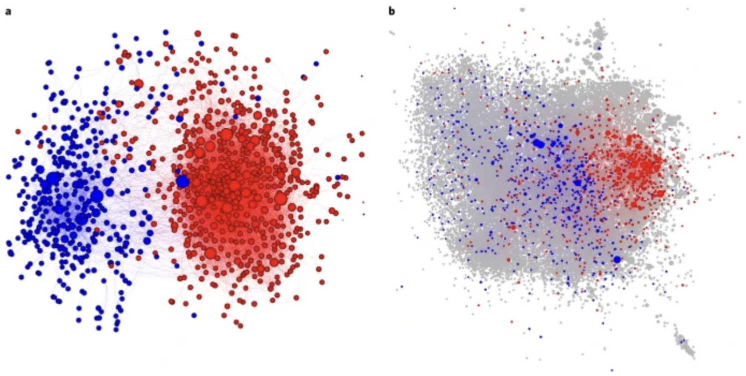
Fig. 10. Simulated distribution of number of executives by proportion of political contributions allocated to the Republican Party.

Examples of Two-Mode Networks

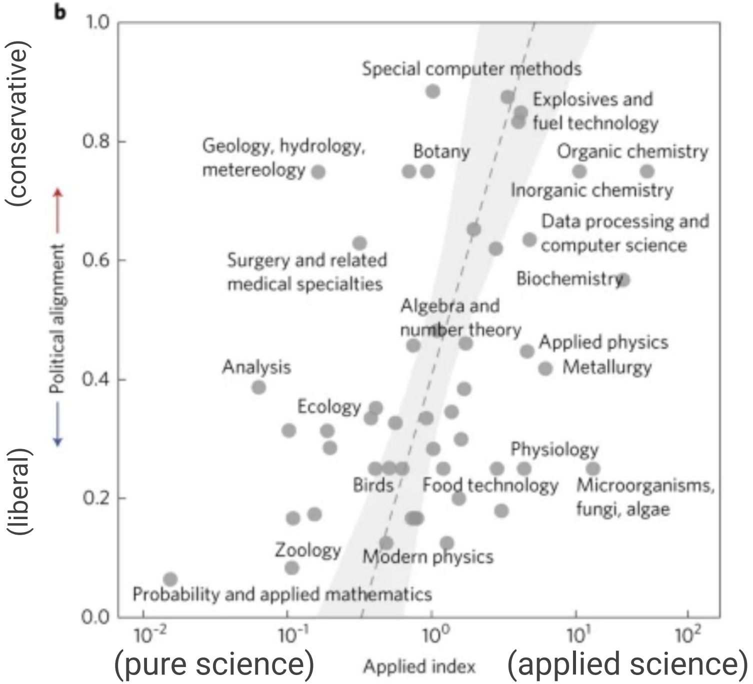
Is science politicized?: Partisan difference in the consumption of science

- Amazon book co-purchase data

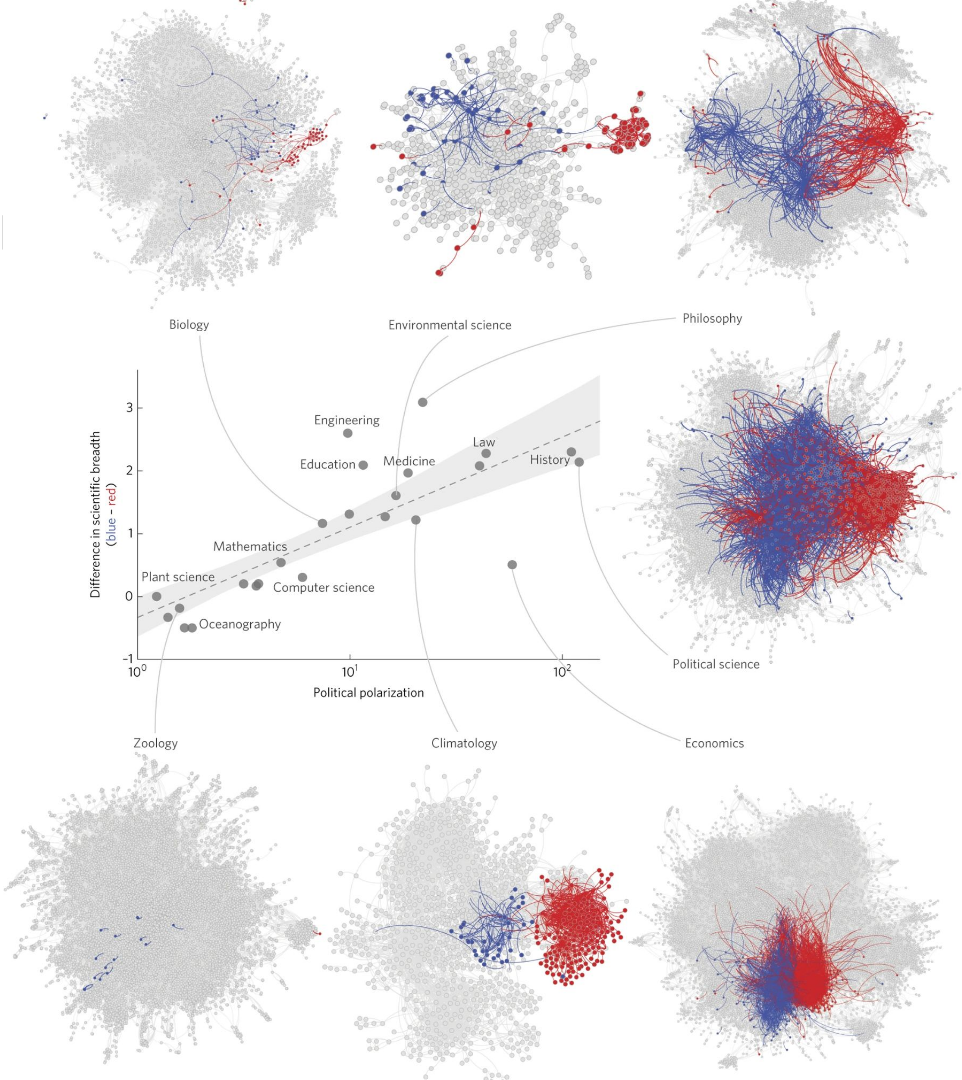
Figure 1: Visualization of the co-purchase network among liberal, conservative and scientific books.



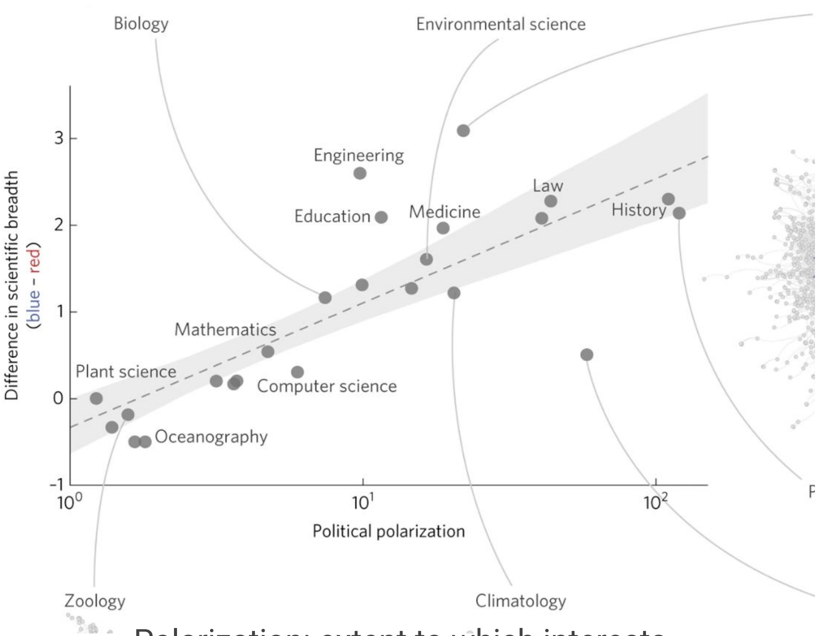
a, Links between 583 liberal (blue) and 673 conservative (red) books. **b**, Links between these books and science (grey) books. As shown in **a**, 97.2% of red books linked to other reds and 93.7% of blue books



Shi et al. 2017



Scientific breadth: Number of science books in a discipline connected to red (blue) books divided by number of red books connected to science discipline



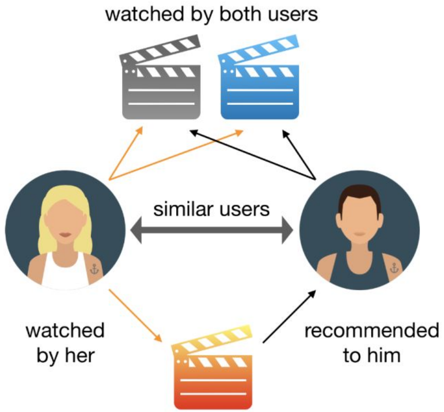
Polarization: extent to which interests from conservatives and liberals in a given topic diverge

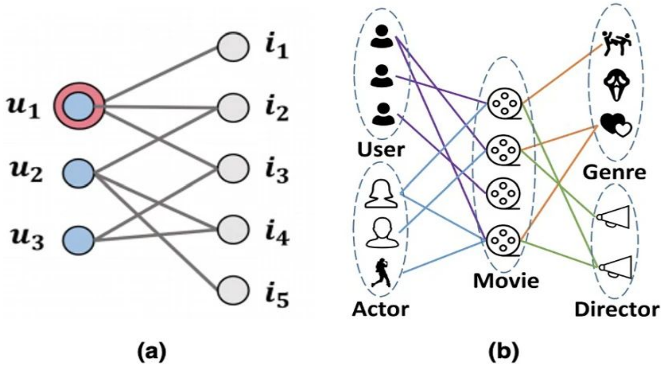
Shi et al. 2017

Affiliation networks and Prediction

Affiliation networks are also widely used to predict and recommend products to online users

- based on similarities in people's connections to artifacts (affiliations)





Higher-Order Interactions

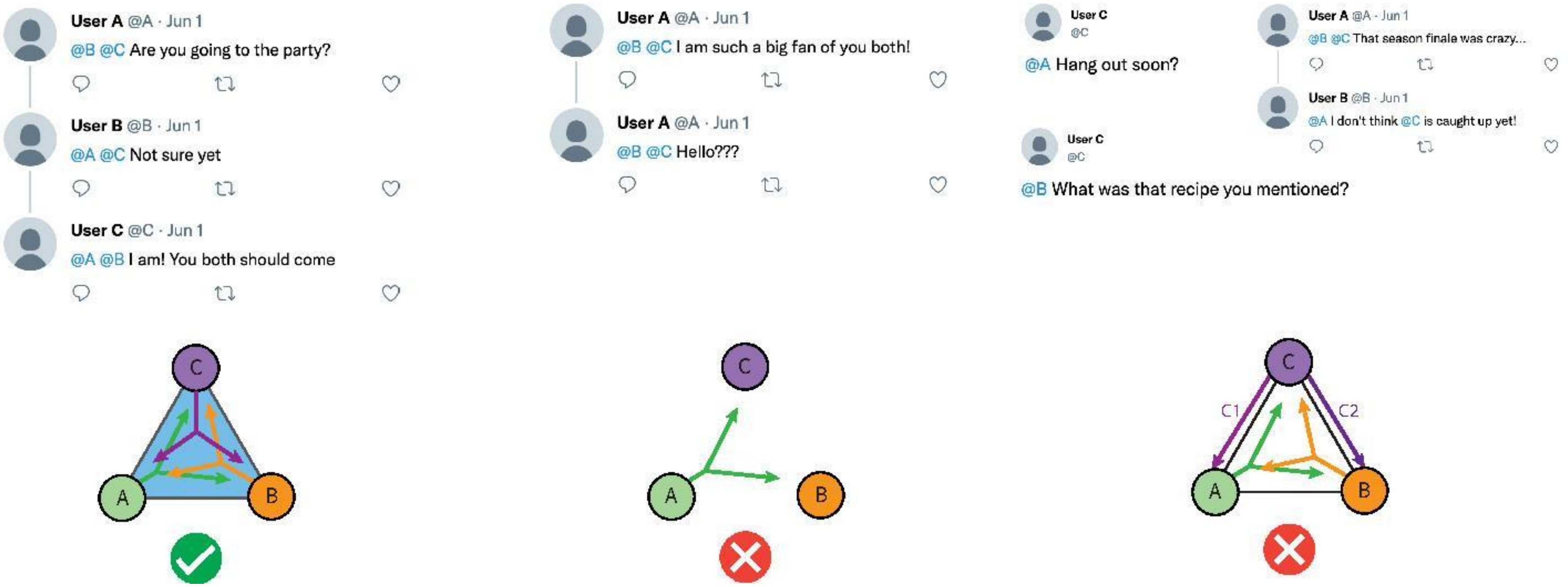
Networks consist of 1:1 interactions (dyad)

Affiliation networks assume complete connections within a group

However, interactions in group contexts cannot be reduced to the sum of 1:1 dyadic interactions

Higher-Order Interactions

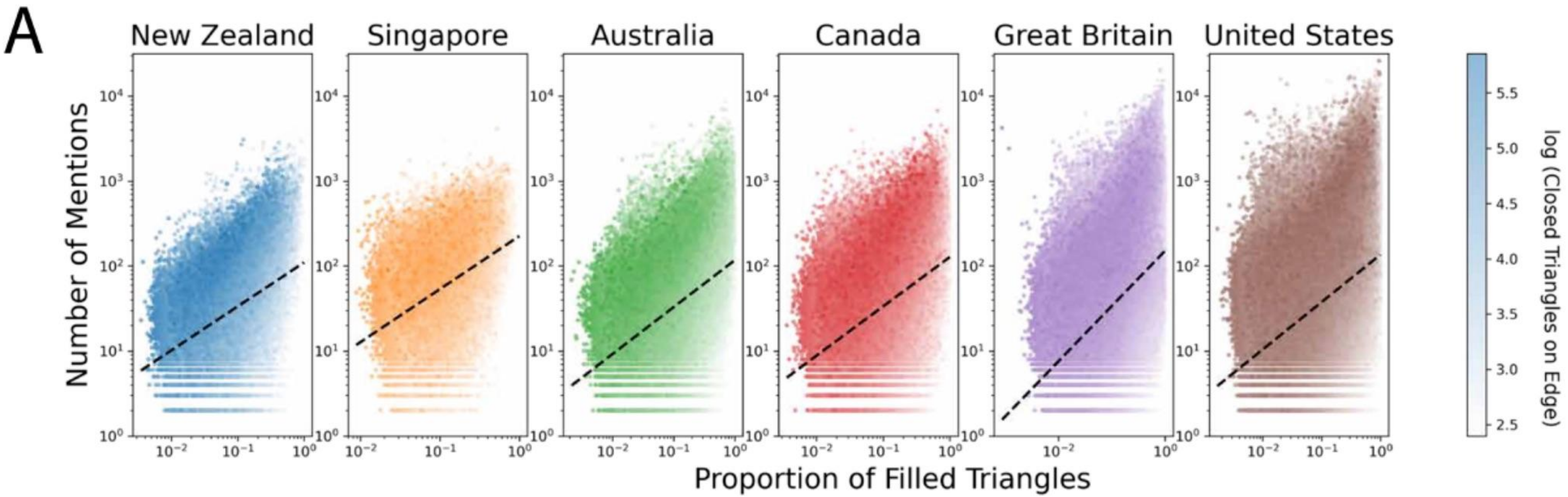
Constructing higher-order interactions among Twitter users



Sarker and Park, working paper

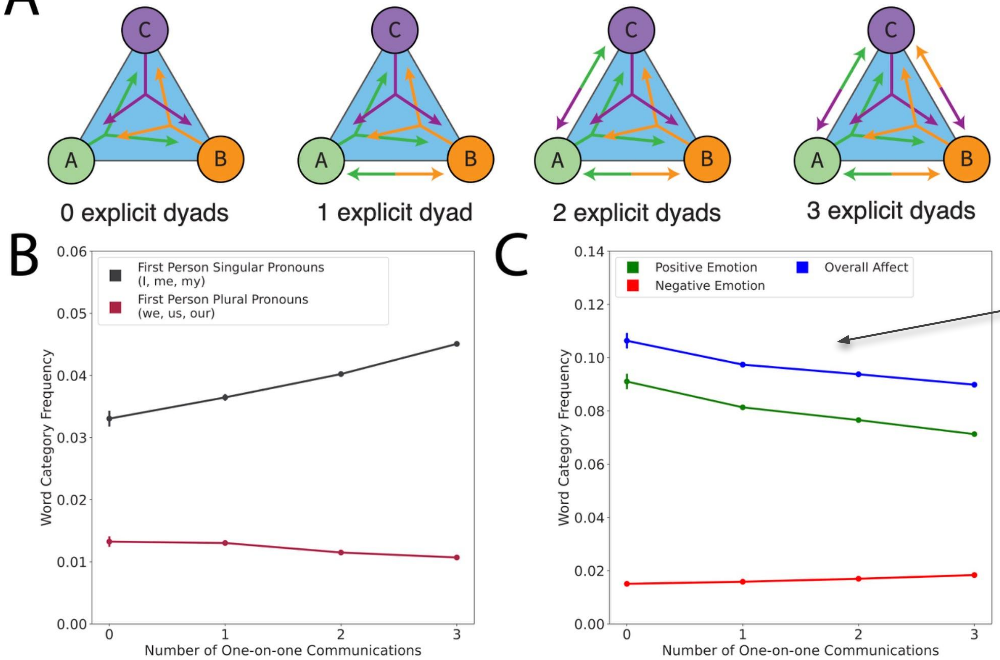
Higher-Order Interactions

Edges involved in higher-order interactions (filled triangles) are three times stronger



Sarker and Park, working paper

Higher-Order Interactions



Filled triangles with fewer 1:1 interactions express more positive emotions in their tweets

Sarker and Park, working paper

Affiliation Network is a Proxy

Affiliations are proxies of contexts of social interaction

However, they do not necessarily reflect the participants' perceptions or actual interactions in that context

Back to triadic closure

The projection of two-mode networks creates a number of issues

1. Tie formation

Each tie in a prototypical one-mode network is assumed to be created separately, e.g., a standard phone call creates a communication tie from one person to another.

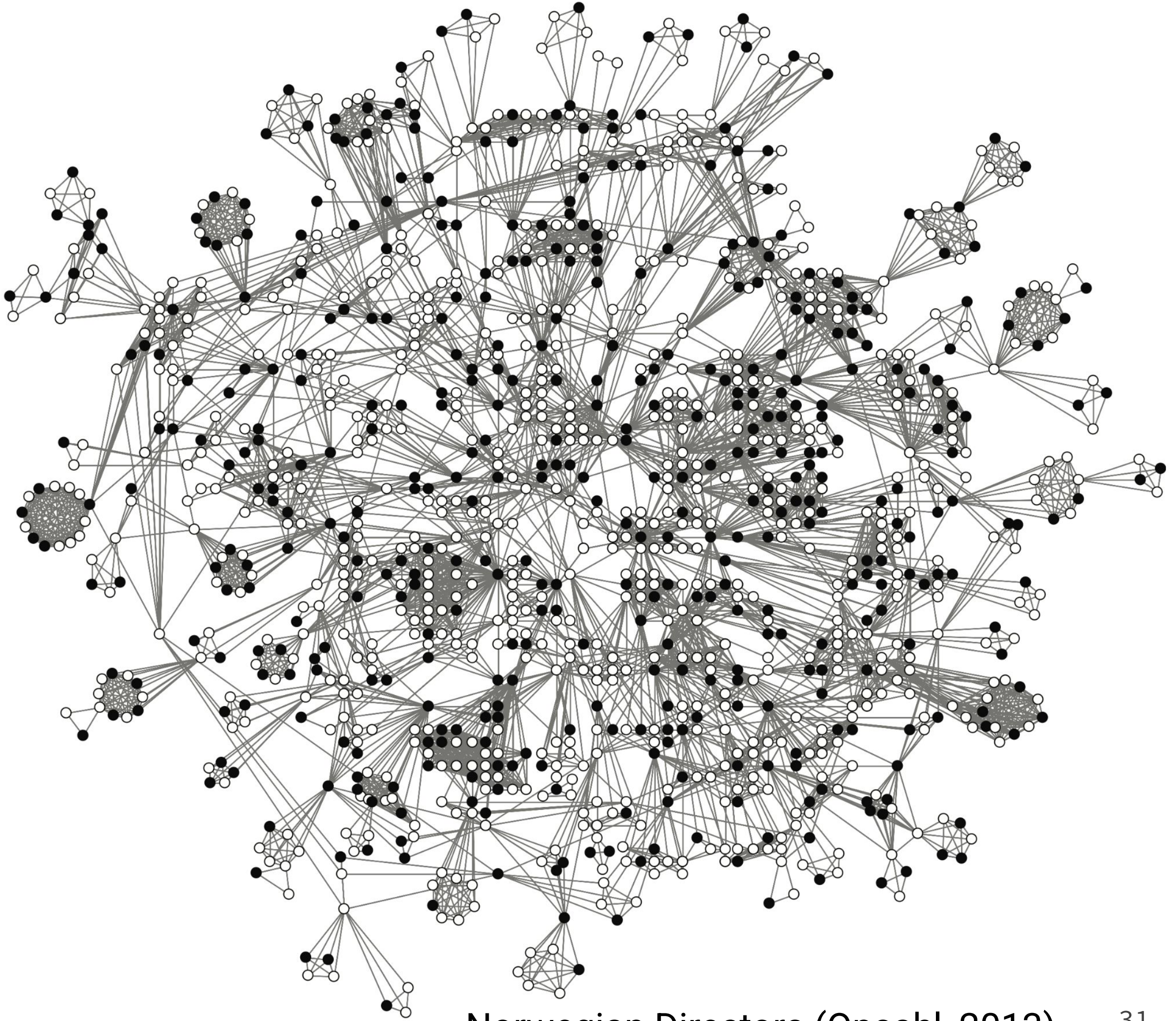
This is not the case in projected two-mode networks, e.g., a director forms ties with all the other Directors on a board when they joins that board.

→ How to use random networks to detect a baseline level? How to compare observed measures with those found in corresponding random networks?

2. Connectedness

A projected two-mode network tends to have more and larger fully-connected cliques than typical one-mode networks.

→ Impacts network measures based on triangles, e.g., clustering coefficients.



Norwegian Directors (Opsahl, 2013)

Clustering coefficients for one-mode networks

Global clustering coefficient: fraction of triplets or 2-paths (i.e., three nodes connected by two ties) that are closed by the presence of a tie between the first and the third node.

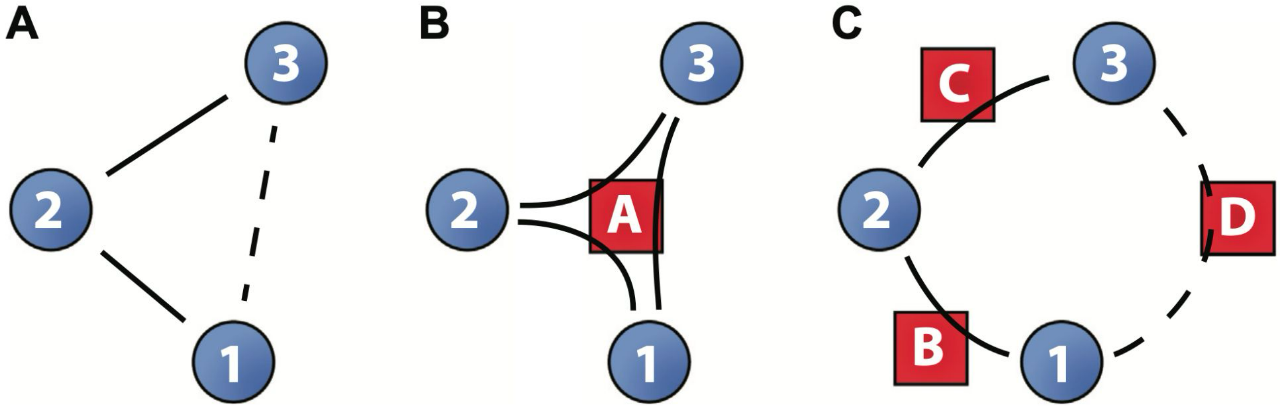
$$C = \frac{3 \times triangles}{triplets} = \frac{closed\ triplets}{triplets} = \frac{\tau_{\Delta}}{\tau}$$

Local clustering coefficient: the fraction of ties among a node's contacts over the possible number of ties between them.

$$C(i) = \frac{number\ of\ actual\ ties\ among\ node\ i's\ contacts}{number\ of\ possible\ ties\ among\ node\ i's\ contacts} = \frac{\tau_{i,\Delta}}{\tau}$$

A triangle in a projected two-mode network can be formed by two possible configurations

(Opsahl, 2013)



B: Three scholars coauthor an article, *A*. This article creates three-way links among the authors (triadic closure).

C: Scholars 2 and 3 coauthor *C* and scholars 1 and 2 coauthor *B*. 1 and 3 coauthored article *D*. Triadic closure occurs through three double-authored collaborations

The clustering coefficient in one-mode random networks greatly underestimates the baseline-level of clustering in projected two-mode networks

Opsahl randomized the two-mode structure of a scientific collaboration network while maintaining the degree distributions (i.e., randomly assigning the ties in the two-mode network while holding constant each author's number of co-authored papers, and each paper's number of authors) before projecting it onto a one-mode network and calculating the global clustering coefficient.

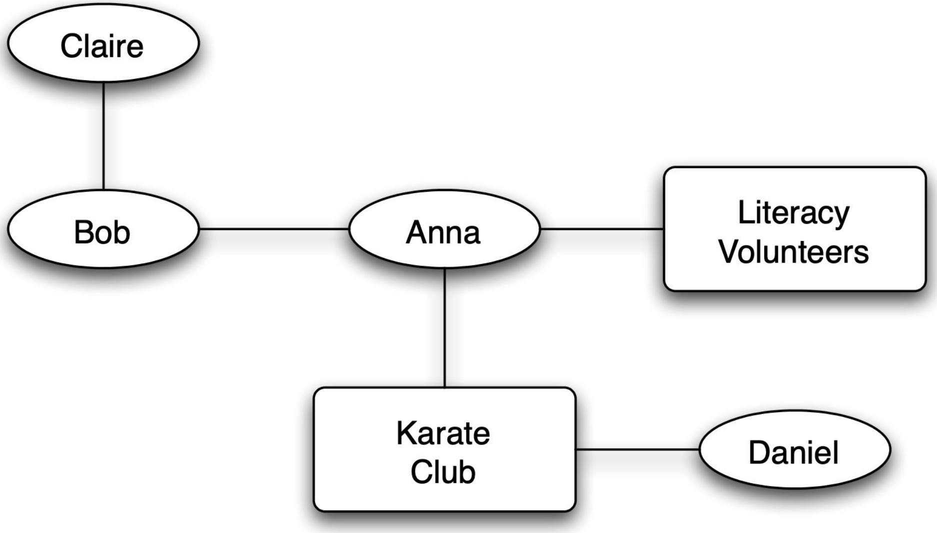
Across 1000 projected random two-mode networks, the average global clustering coefficient was 0.1236, which is over 350 times larger than the coefficient in corresponding one-mode classical random networks.

Types of Closures in Affiliation Networks

Let’s slightly extend the notion of closure

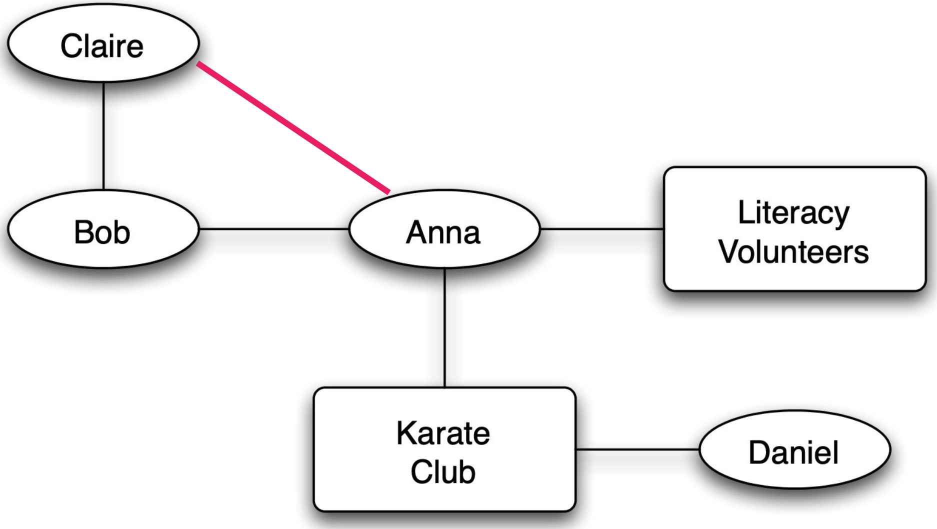
Two two kinds of edges:

- Social (e.g., friendship)
- Affiliation (e.g., participation in activity)



Different mechanisms for link formation can now all be viewed as types of closure processes.

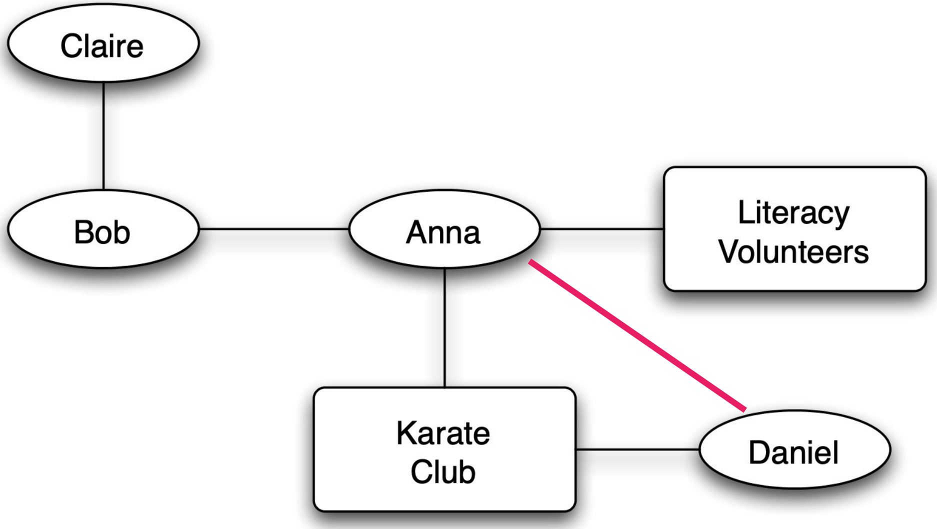
“Triadic closure”: friend of a frend



Different mechanisms for link formation can now all be viewed as types of closure processes.

“Focal closure”

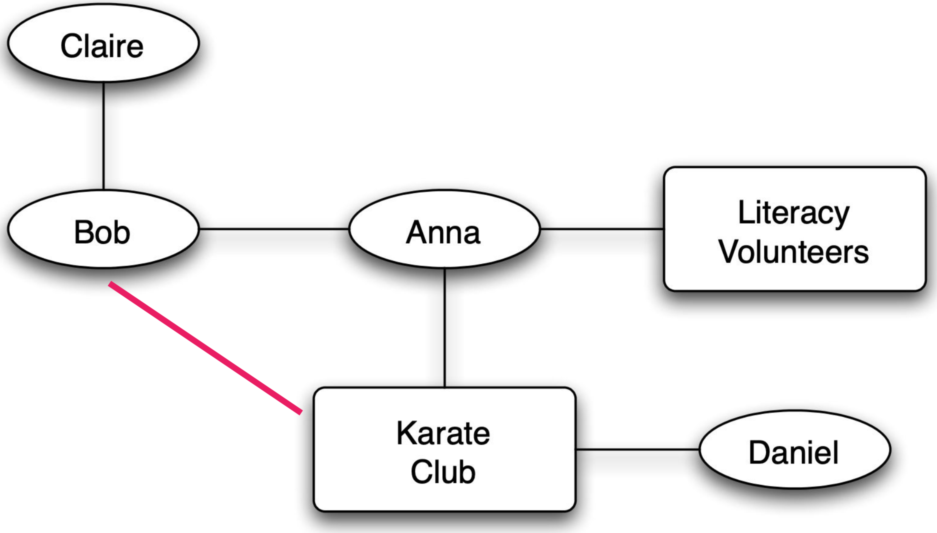
Two people form a link when they have a focus in common (self-selection mechanism).



Different mechanisms for link formation can now all be viewed as types of closure processes.

"Membership closure"

Bob takes part in a focus that his friend Anna is already involved in (social influence mechanism).

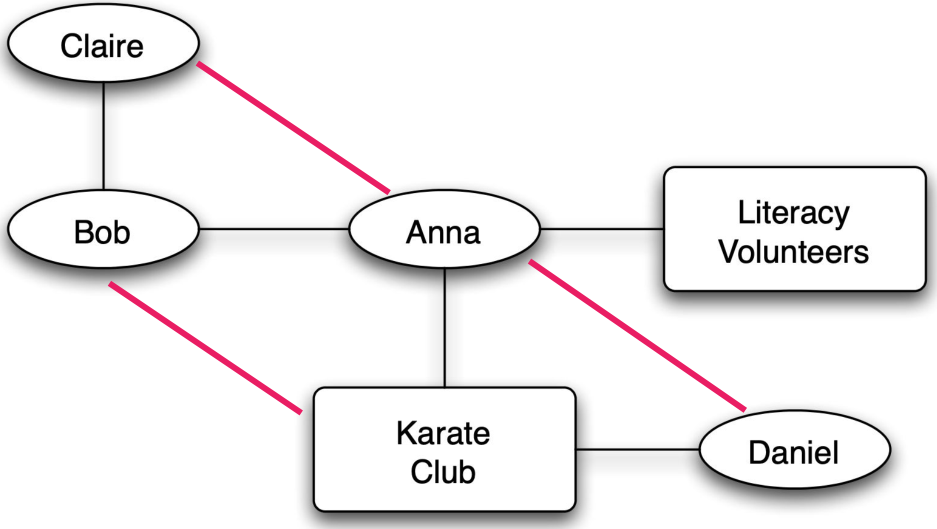


Different mechanisms for link formation can now all be viewed as types of closure processes.

(i) Bob introduces Anna to Claire.

(ii) Karate introduces Anna to Daniel.

(iii) Anna introduces Bob to Karate.



Summary

Affiliation networks can reveal interesting relationships on both sides of the bipartite graph.

We need to rethink many of our one-mode measures.