Networks in their Surrounding Contexts

Theme

- **Each** individual
  - a distinctive set of personal characteristics.
- A pair of friends
  - similarities and compatibilities among two people's characteristics

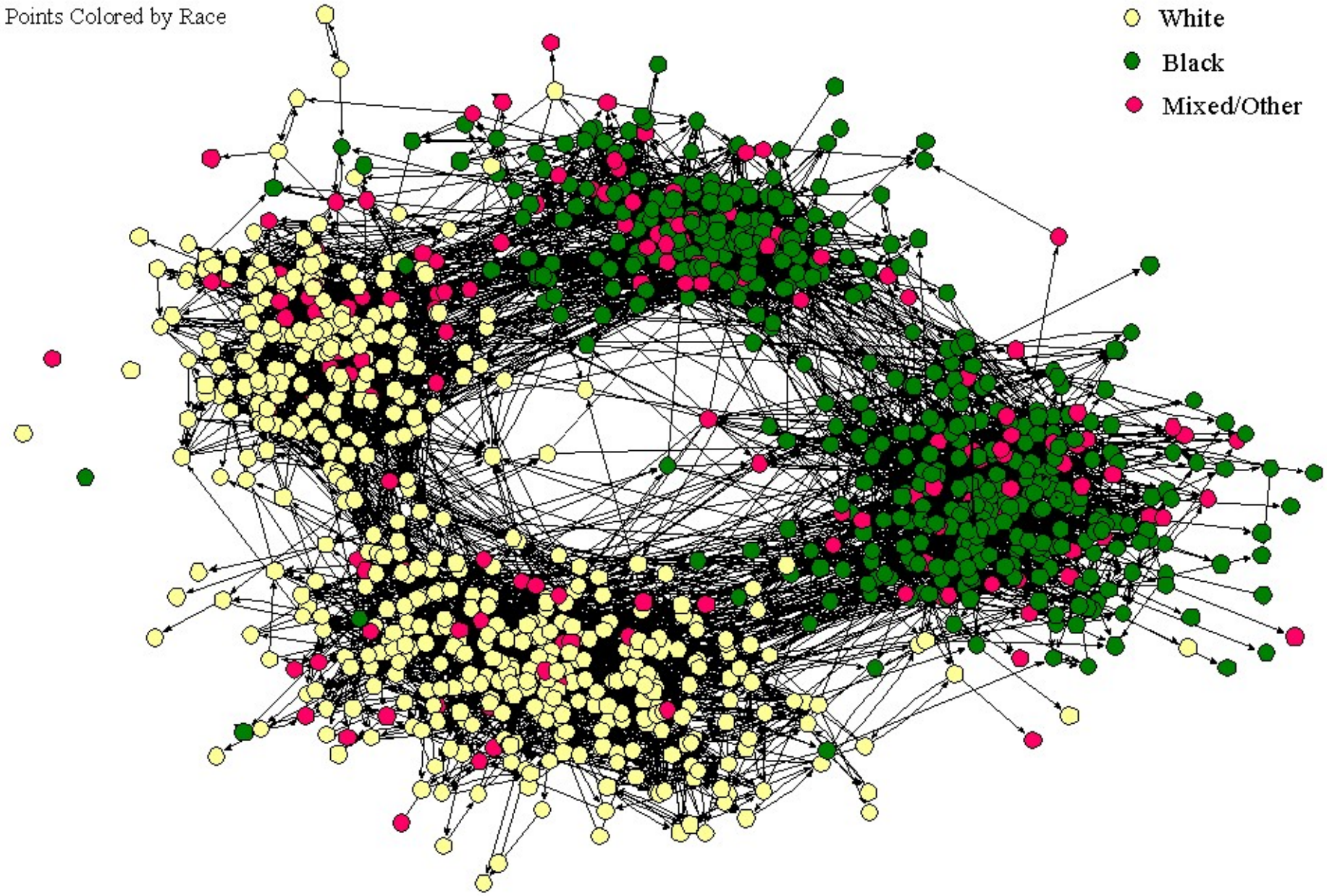
- Behaviors and activities of individuals
  - shape the formation of links within the network (surrounding contexts)

Homophily

- The tendency of individuals to associate and bond with similar others.
- Plato in Phaedrus : “similarity begets friendship”
- English Idiom : “Birds of a feather flock together”
- Chinese Idiom : 物以類聚,人以群分。 (易經·繫辭上)

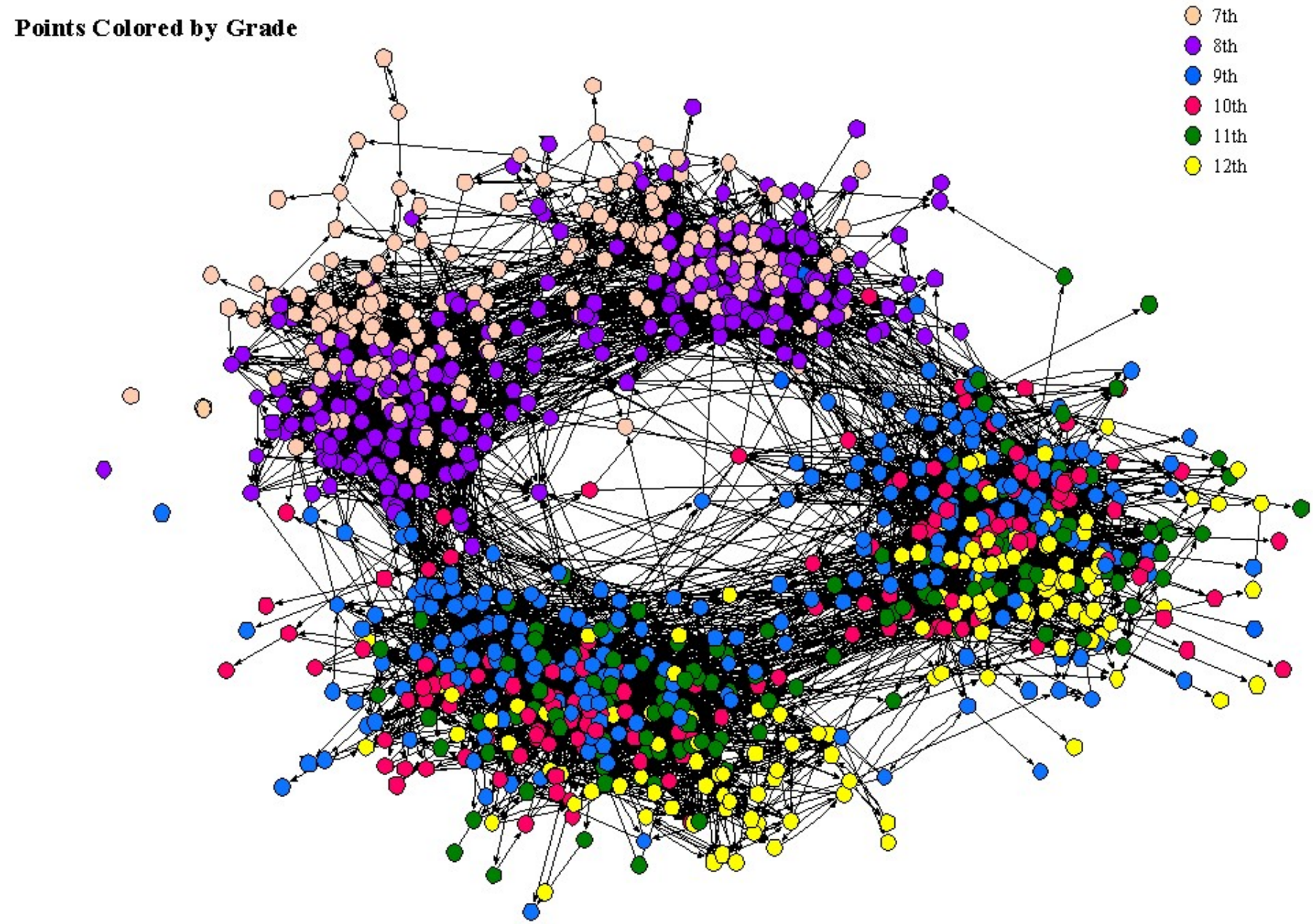
The Social Structure of "Countryside" School District

By James Moody

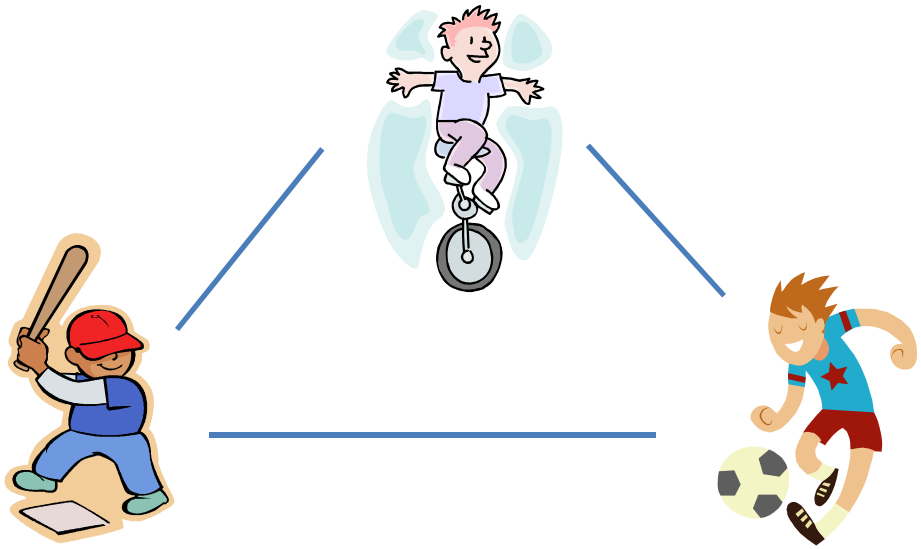


The Social Structure of "Countryside" School District

By James Moody



- Friendship that forms because two people
  - have a common friend (intrinsic network structure)
  - attend the same school or work for the same company (contextual factor)

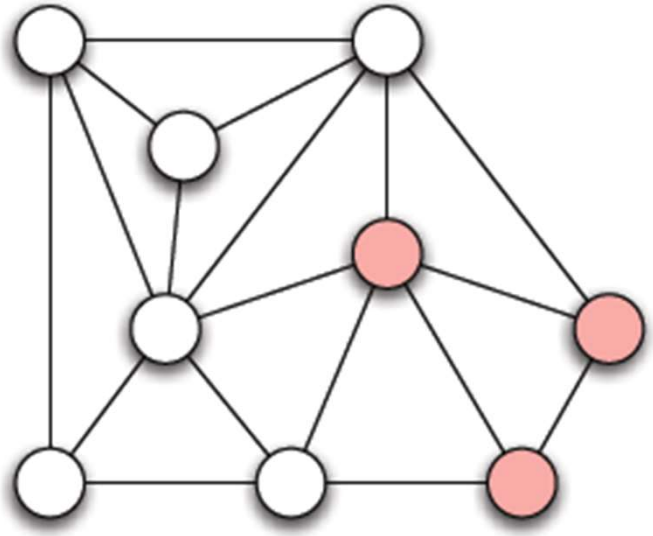


Triadic Closure : a common friend
- sources of trust
- Increase opportunity to interact

Contextual factor :
- A and B are similar in a number of dimensions
- A and C are similar …
- B and C are likely to be similar

Measuring Homophily

- Given a network of nodes with certain characteristics; e.g. gender, age or race.
- Does the network exhibit homophily ?



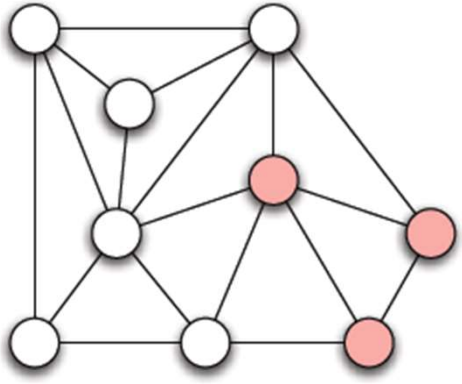
Measuring Homophily

- Gender : 6 boys and 3 girls
- Does the network exhibit homophily ?
- p = fraction of males

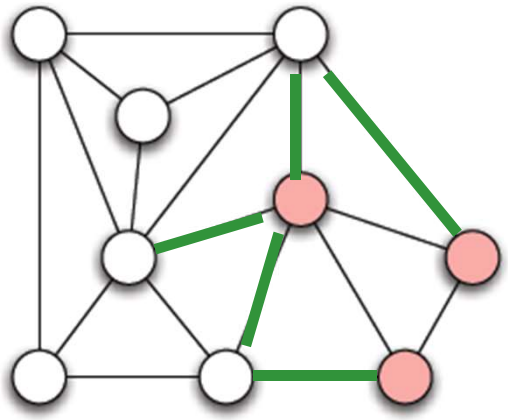
  q = fraction of females
- Prob(male-male) = p^2

  Prob(female-female) = q^2

  Prob(male-female) = 2pq



- Homophily Test: If the fraction of cross-gender edges is significantly less than 2pq, then there is evidence for homophily.



- 6 males and 3 females
- Male : p=2/3
- Female : q=1/3
- 2pq =4/9 = 8/18
- 5 out of 18 edges are cross gender
- 5/18 << 8/18
- Evidence of homophily !!

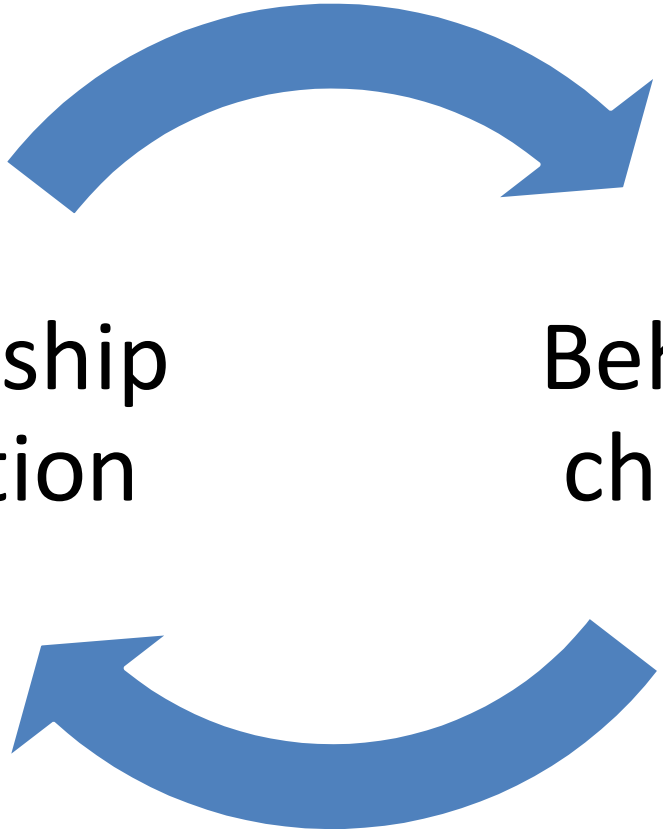
- A proper definition of “significantly less than”
- How about the case when it is “significantly more than” ? — inverse homophily
- What if the characteristic is not binary but multi-variant, such as race, native language or political orientation?
  - The edge is heterogeneous if two end nodes do not share the same characteristic.
  - Compare the no. of heterogeneous edges with the expected value.

Mechanisms Underlying Homophily

- Selection : select friends with similar characteristics
  - individual characteristics drive the formation of links
    - Live in neighborhood
  - involves immutable characteristics (determined at birth)
    - Race
  - 物以類聚,人以群分/Birds of a feather flock together”
- Socialization and social influence : modify their behaviors to bring them more closely into alignment with the behaviors of their friends.
  - reverse of selection
  - Involves mutable characteristics
  - 近朱者赤， 近墨者黑/lies down with dogs, rises with fleas

- Social influence : can produce network-wide uniformity, as new behavior spreads across the links.
- Selection : tends to drive the network towards smaller clusters of like-minded individuals (balkanization)
- Recommender systems
  - Build prediction based on others

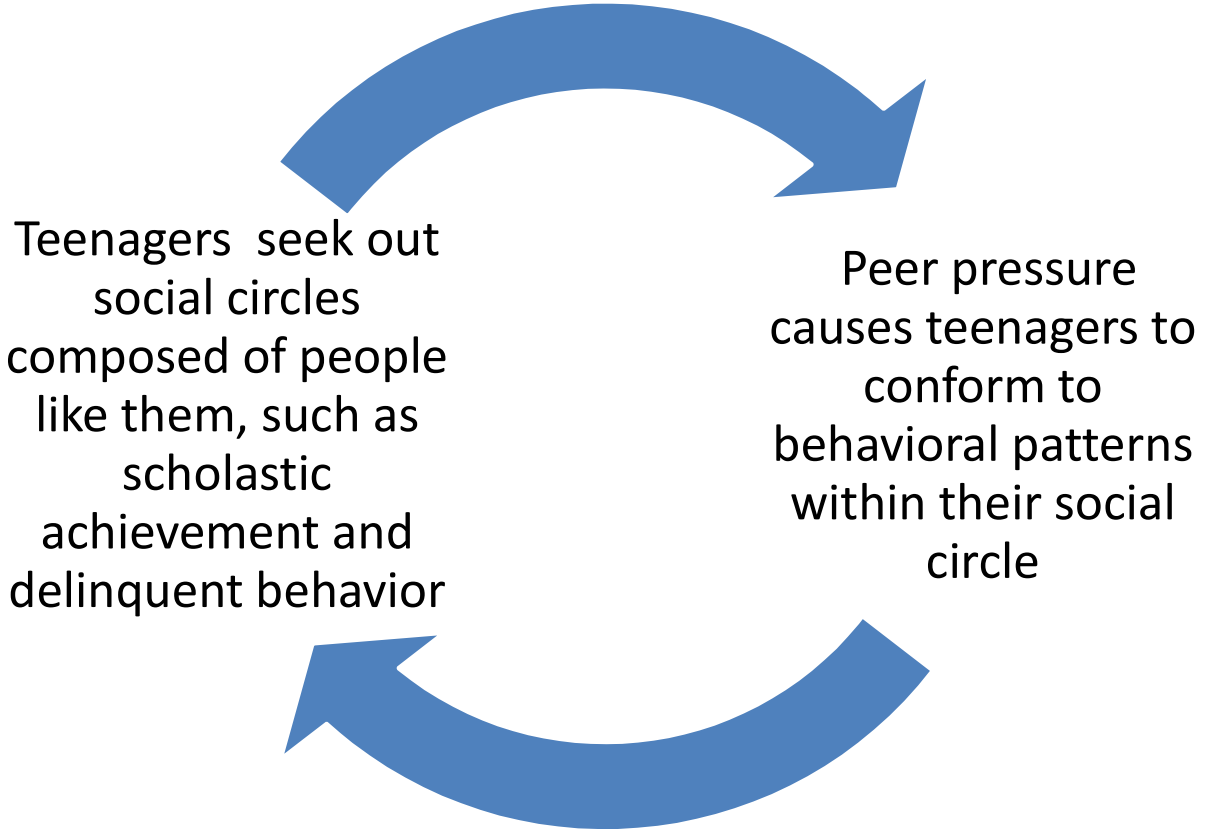
Balkanization


Friendship selection
Behavior change

The Interplay of Selection and Social Influence

- It is hard to sort out the distinct effects and relative contributions of selection and social influence as two effects interact with each other.
- Longitudinal studies - track both social connections and behaviors within a group over a period of time.
  - for identifying underlying causes
  - for reasoning about the effect of possible interventions

Teenagers


Teenagers seek out social circles composed of people like them, such as scholastic achievement and delinquent behavior
Peer pressure causes teenagers to conform to behavioral patterns within their social circle

Illicit drug use example

- Homophily across a social network — with students showing a greater likelihood to use drugs when their friends do
  - If the reason is social influence, targeting certain students and influencing them to stop using drugs would cause their friends to stop using drugs as well.
  - If the reason is almost entirely from selection, it may not reduce drug use beyond the students it directly targets.

Obesity status tracking example

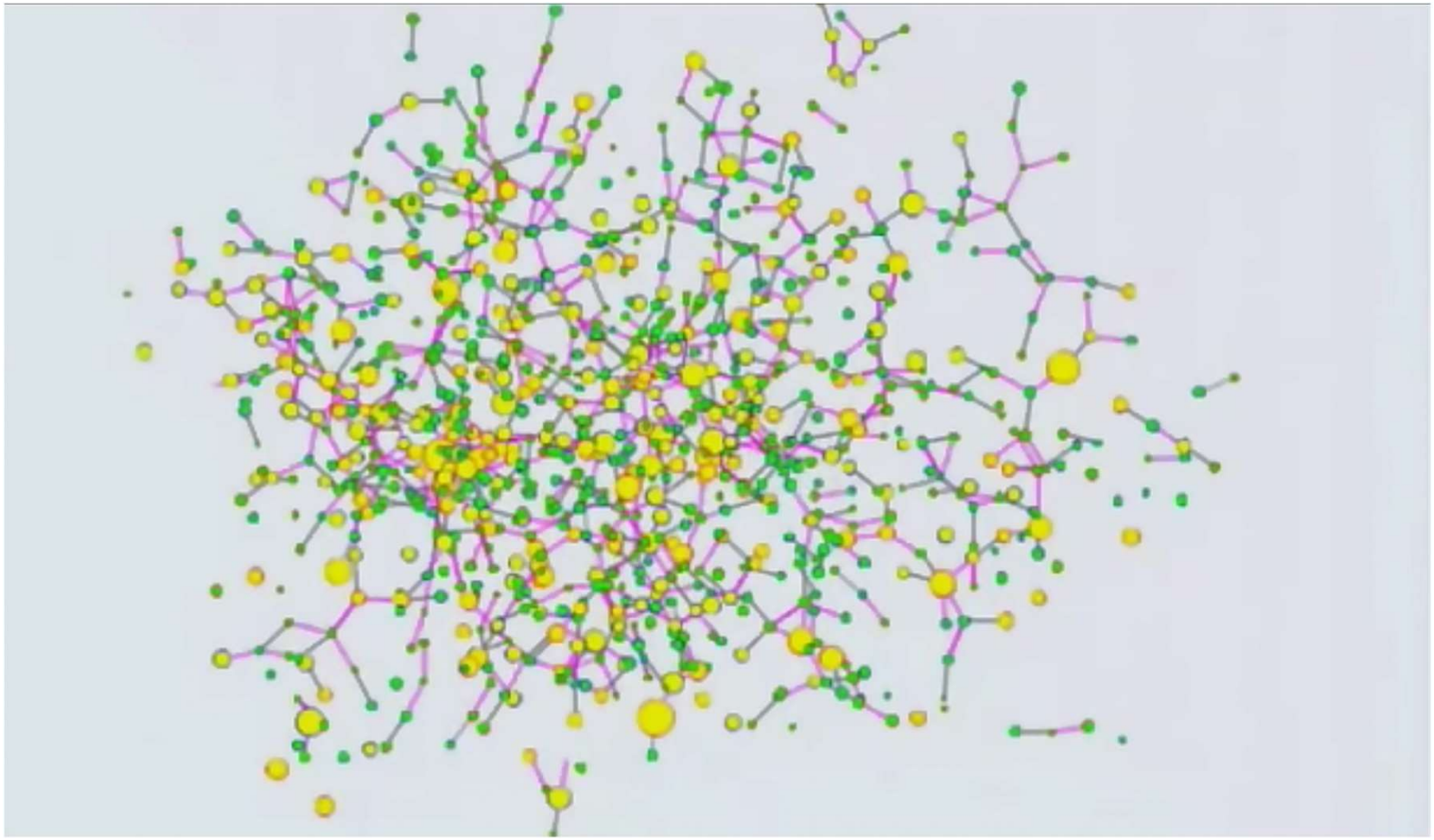
Friendship selection based on obesity status

Change in living styles and hence obesity status

Obesity Status Tracking

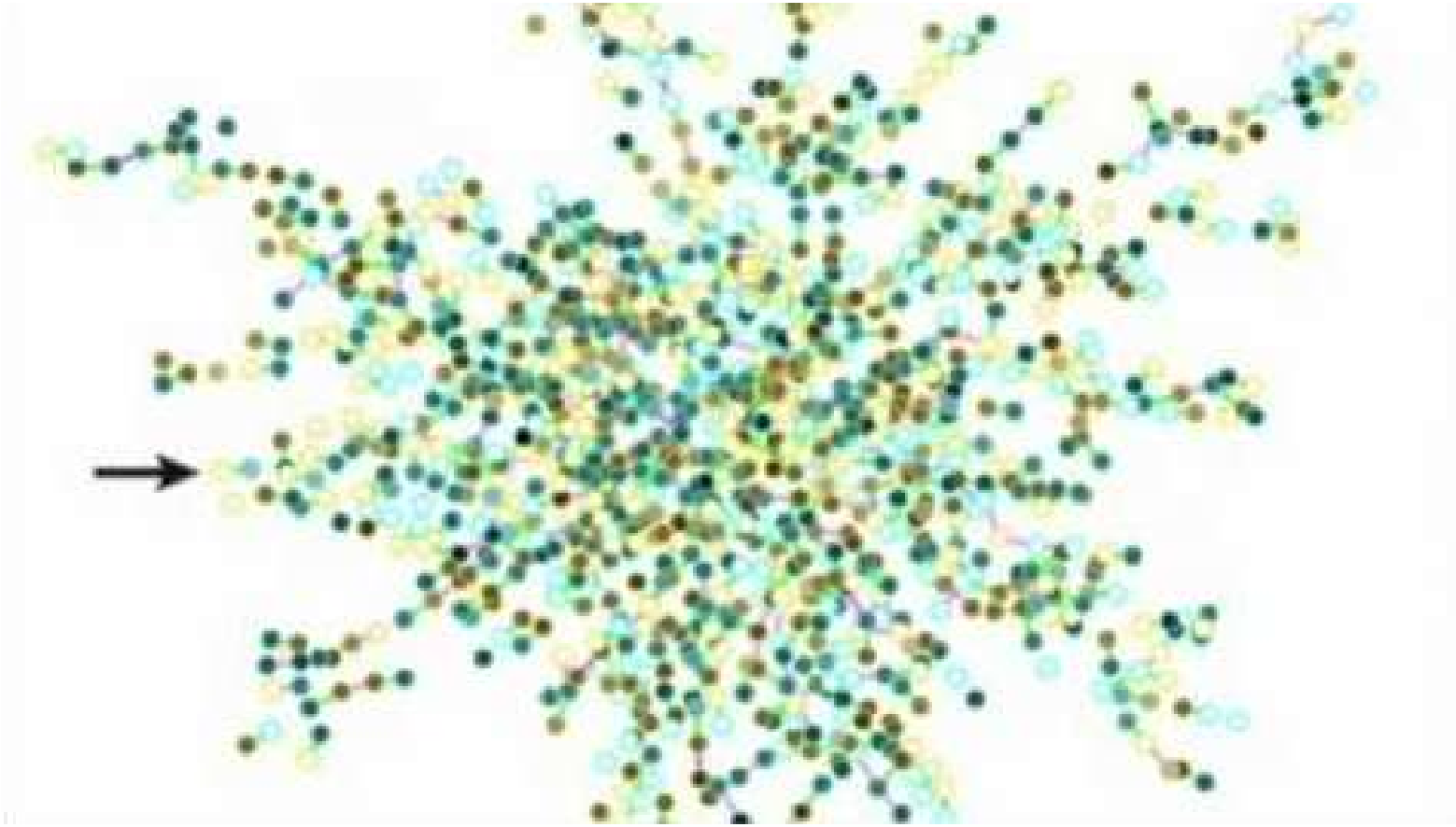
- Christakis and Fowler studies the longitudinal data covering roughly 12,000 people over a 32-year period.
- Homophily : obese and non-obese people clustered in the network.
- One of the reasons
  - selection effects
  - confounding effects of homophily according to other characteristics
  - changes in the obesity status of the friends was exerting a behavioral influence that affects ones future obesity status
    - Contagion effect : spread through the underlying social network

The Spread of Obesity in Social Networks



N.A. Christakis and J.H. Fowler, "The Spread of Obesity in a Large Social Network Over 32 Years," New England Journal of Medicine 357(4): 370-379 (July 2007)

The Spread of Smoking in Social Networks



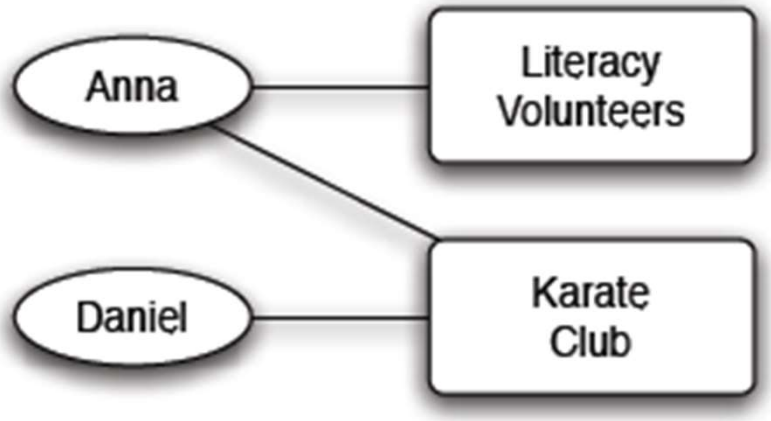
N.A. Christakis and J.H. Fowler, "The Collective Dynamics of Smoking in a Large Social Network," New England Journal of Medicine, 358(21): 2249-2258 (May 2008)

Homophily

- How the underlying mechanisms of homophily will affect the further evolution of the network ?
- How these mechanisms interact with possible outside attempts to influence the behavior of people in the network?

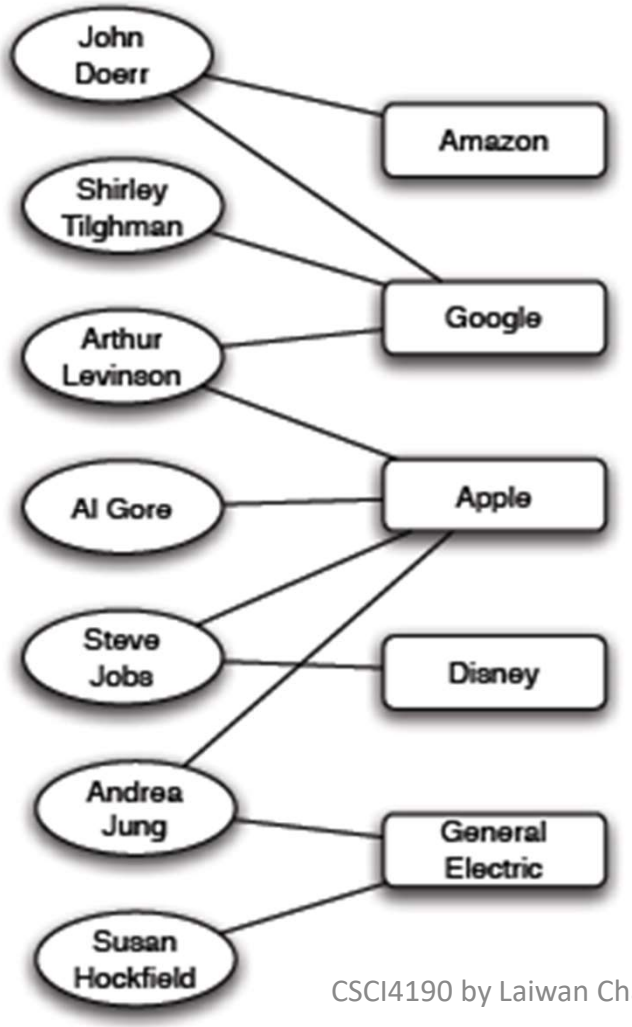
Affiliation Networks

- Include the contextual factors (the set of activities) into the network
- A node
  - a person, or
  - a focus ("focal point" of social interaction)
- An edge
  - Connecting a person to a focus if participation

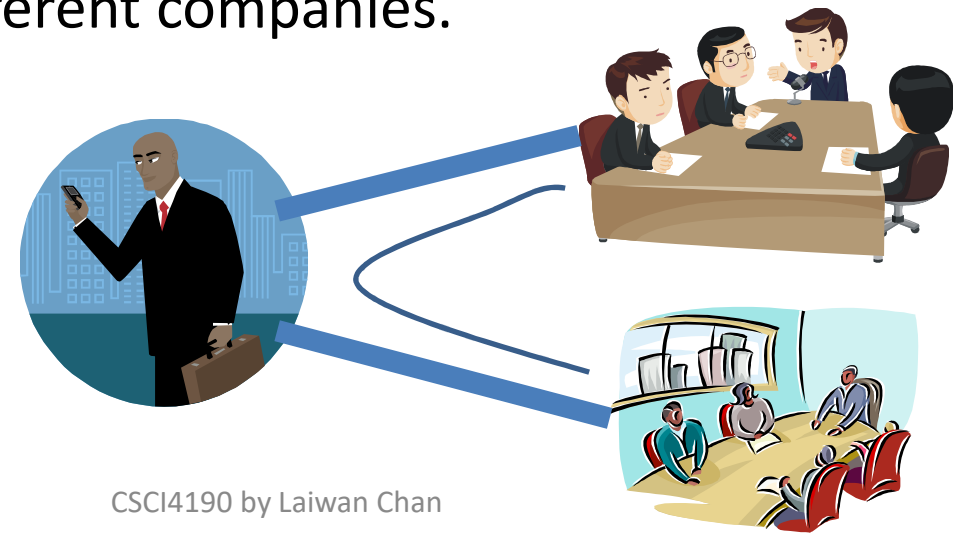


- Affiliation network is expressed as a bipartite graph.
- A graph is bipartite if its nodes can be divided into two sets in such a way that every edge connects a node in one set to a node in the other set. (i.e. no edges joining a pair of nodes that belong to the same set; all edges go between the two sets.)

Memberships on corporate boards of directors



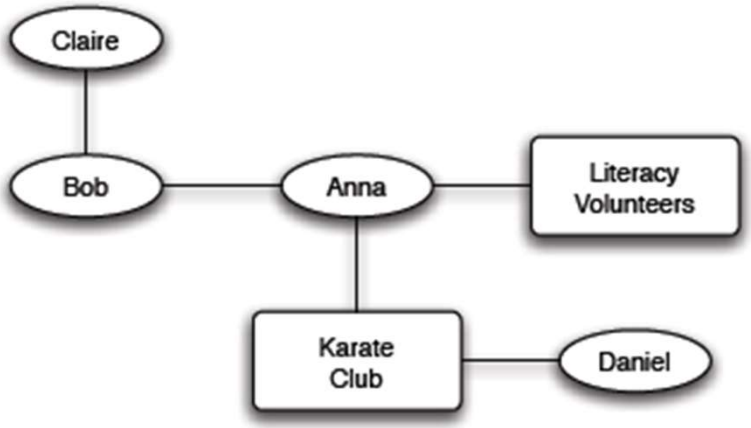
- The boards of directors reveal relationship on both sides of the graph
  - Two companies are implicitly linked by having the same person sit on their boards;
    - possible conduits for information and influence to flow between different companies.



- Two people are implicitly linked by serving together on a board,
  - particular patterns of social interaction among some of the most powerful members of society.

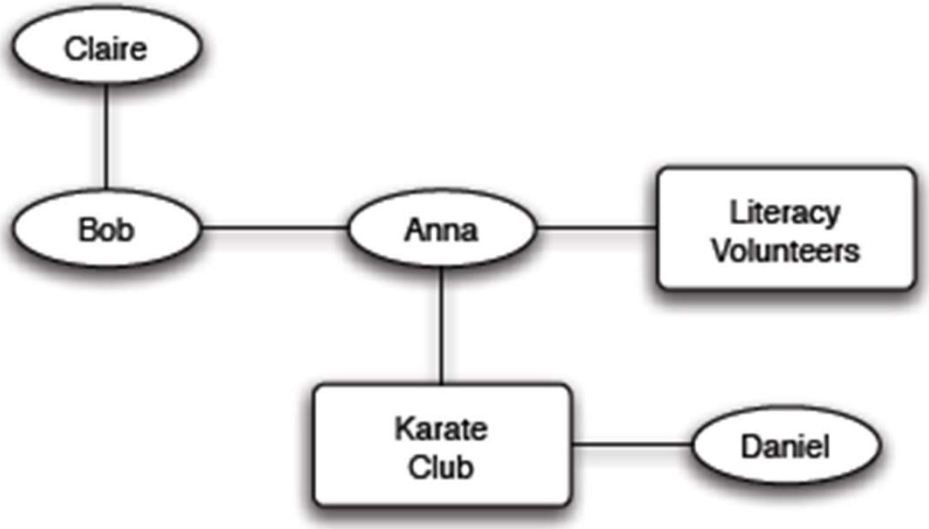
Co-evolution

- Two people sharing the same focus → an opportunity to become friends.
- Two people are friends → influence each other's choice of foci.



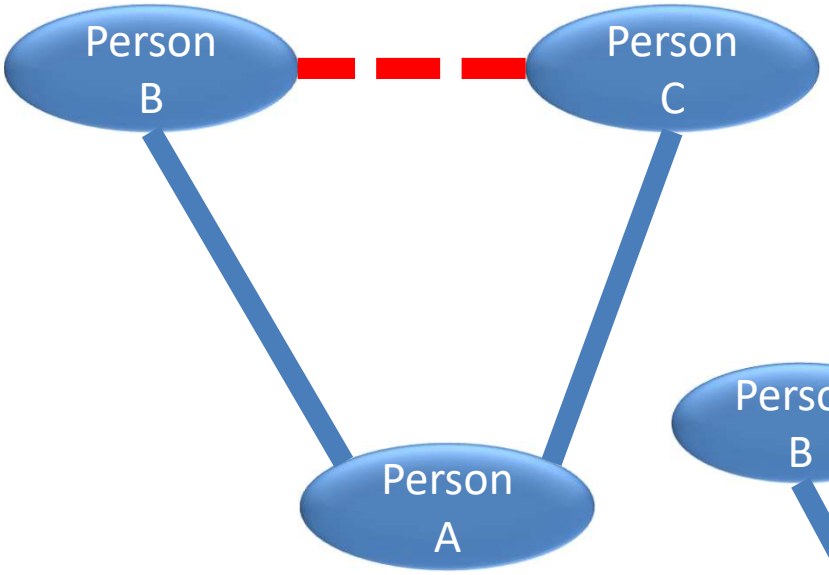
Social-affiliation Network

- An extension of an affiliation network.
- Node
  - People
  - Foci
- Edge
  - Between 2 people
  - Between a person and a focus

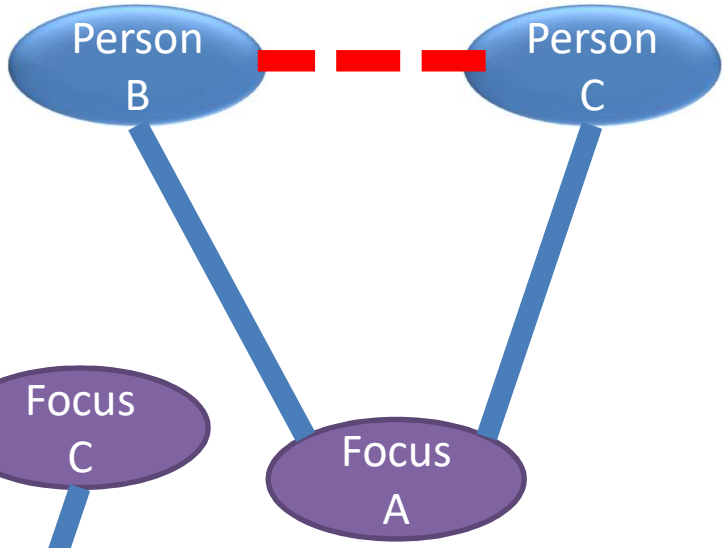


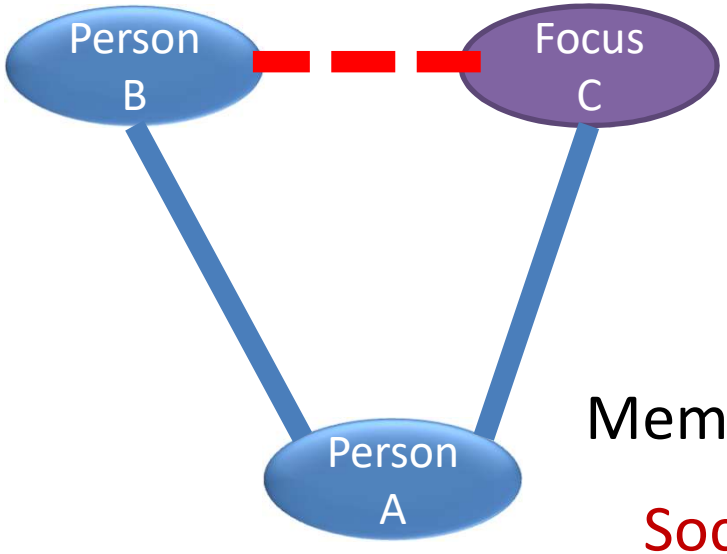
3 Closure Processes selection

Triadic closure



Focal closure

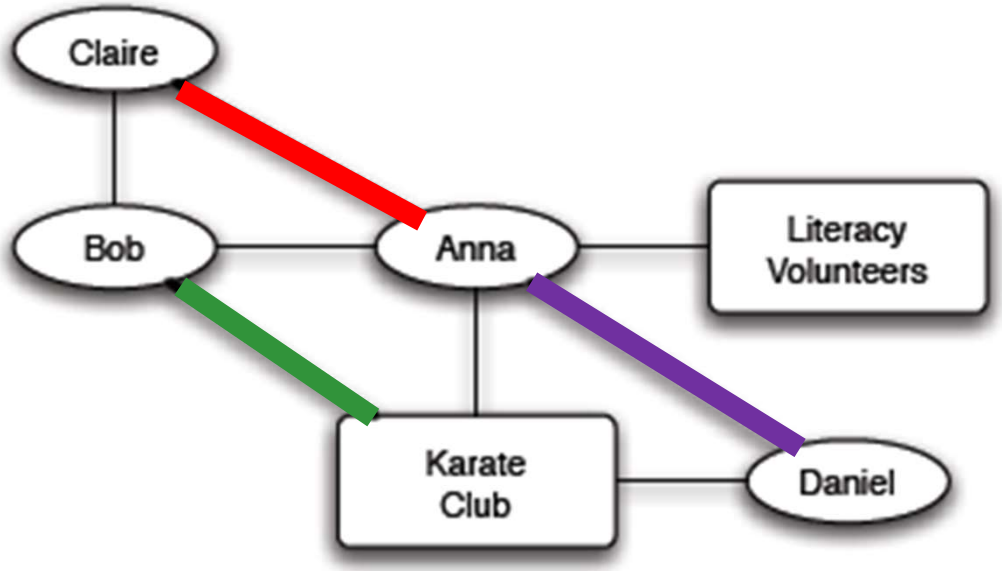




Membership closure

Social influence

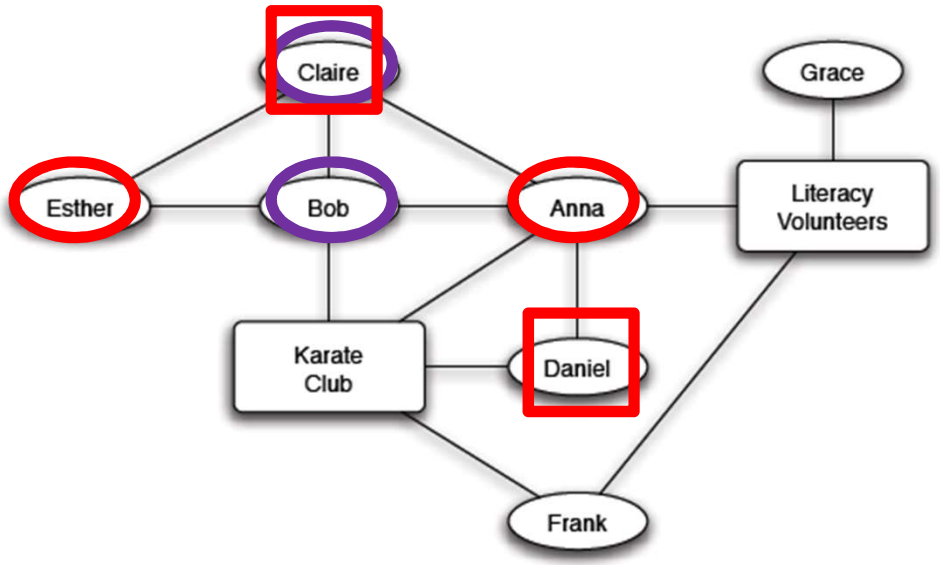
- Bob introduces Anna to Claire (Triadic Closure)
- Karate introduces Anna to Daniel (Focal Closure)
- Anna introduces Bob to Karate (Membership Closure)



Tracking Link Formation in On-Line Data

- How much more likely is a link to form between two people in a social network if they already have a friend in common? (Triadic Closure)
- How much more likely is an edge to form between two people if they have multiple friends in common?

- Anna and Esther have two friends in common.
- Claire and Daniel only have one friend in common.
- How much more likely is the formation of a link in the first of these two cases? (more sources of opportunity and trust for the interaction)



Tracking Link Formation

- Take two snapshots of the network at different times .
- First snapshot
  - For each k, identify all pairs of nodes who have exactly k friends in common, but who are not directly connected by an edge.
- Second snapshot
  - T (k) = the fraction of these pairs that have formed an edge.

E-mail communication

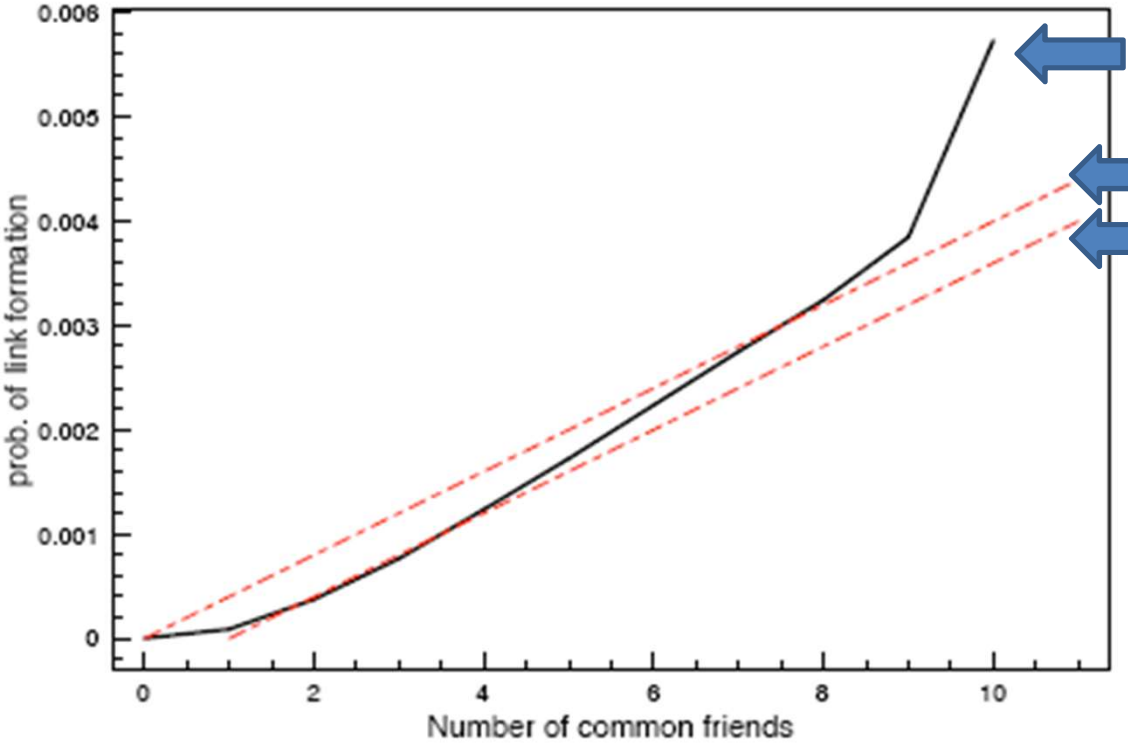
- By Kossinets and Watts
- Full history of e-mail communication among roughly 22,000 undergraduate and graduate students over a one-year period at a large U.S. university
- A link between two people if they had exchanged e-mail in each direction in the past 60 days.
- 1 day between each consecutive snapshot
- Average over multiple snapshots to get T(k)

Tracking Link Formation

- An independent probability *p* of forming a link between 2 people for each common friend that they have.
- 2 people have k friends in common => the probability they fail to form a link is $(1-p)^k$
- 2 people have k friends in common => probability that they form a link is $1-(1-p)^k$

$$T_{baseline}(k) = 1 - (1-p)^k$$

Triadic Closure : T(k) verses no. of common friends



T(k)

$1 - (1-p)^k$

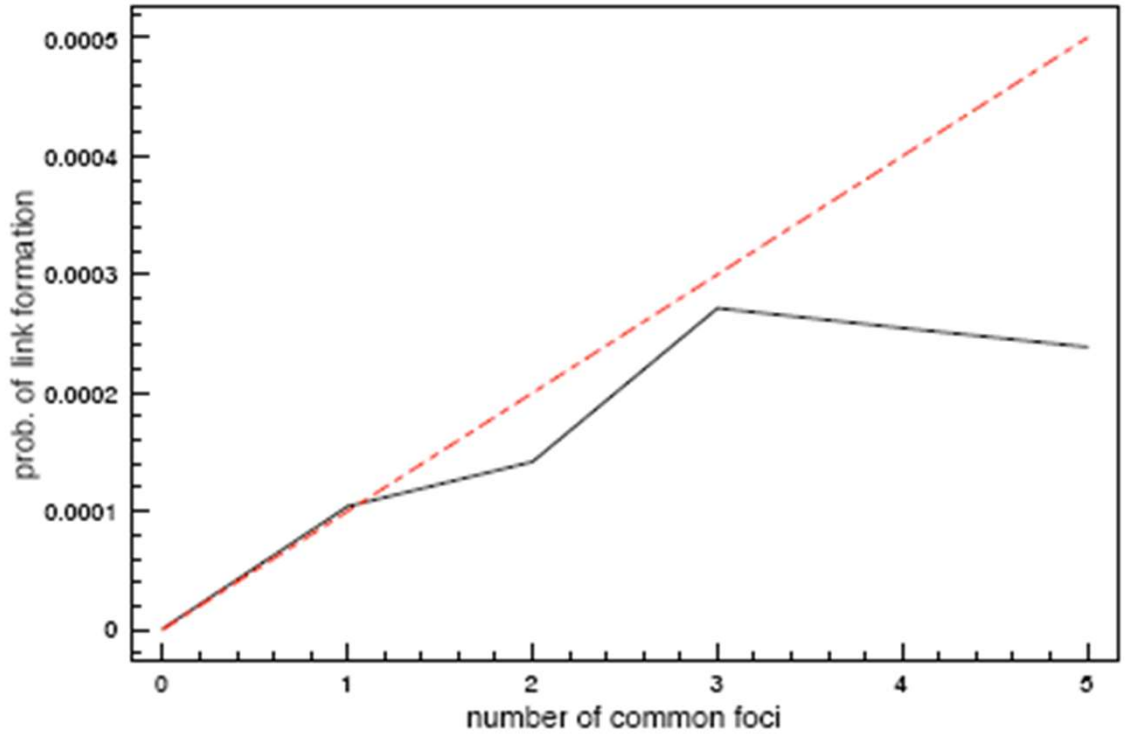
$1 - (1-p)^{k-1}$

When we have more friends : assumption of independent effects from common friends is too simple

- Focal closure : find the probability that two people form a link as a function of the number of foci that are jointly affiliated with
- Membership closure : find the probability that a person becomes involved with a particular focus as a function of the number of friends who are already involved in it ?

Focal closure
T(k) verses no. of common foci

- Supplemented with class schedules
- Focus — taking same class



$1 - (1 - p)^k$

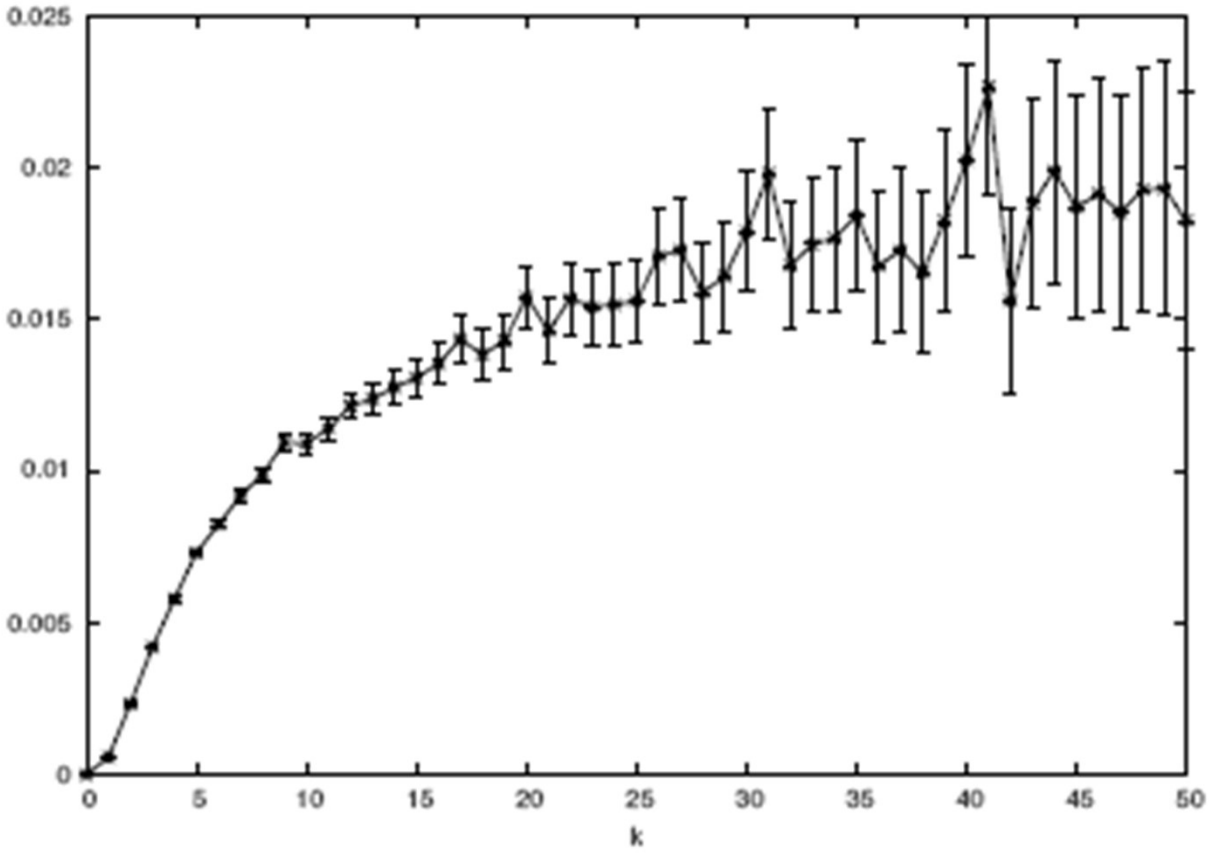
T(k)

Diminishing returns effect when joining attending many classes

Membership closure

LiveJournal : a blogging site

- foci correspond to membership in user-defined communities

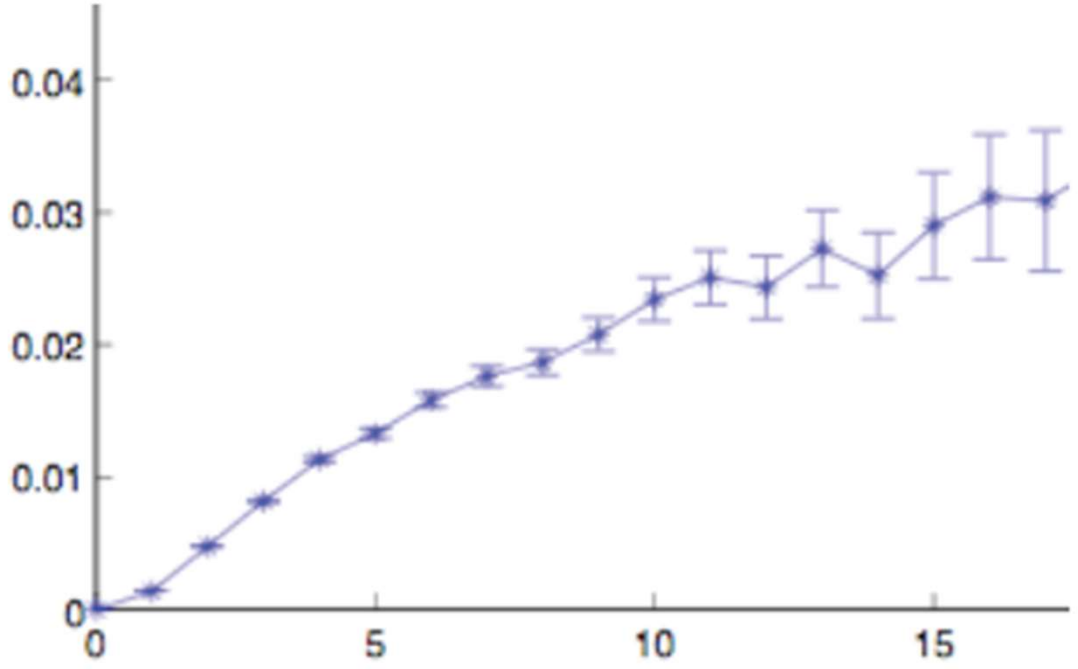


Probability of joining a community when k friends are already members

Wikipedia editors

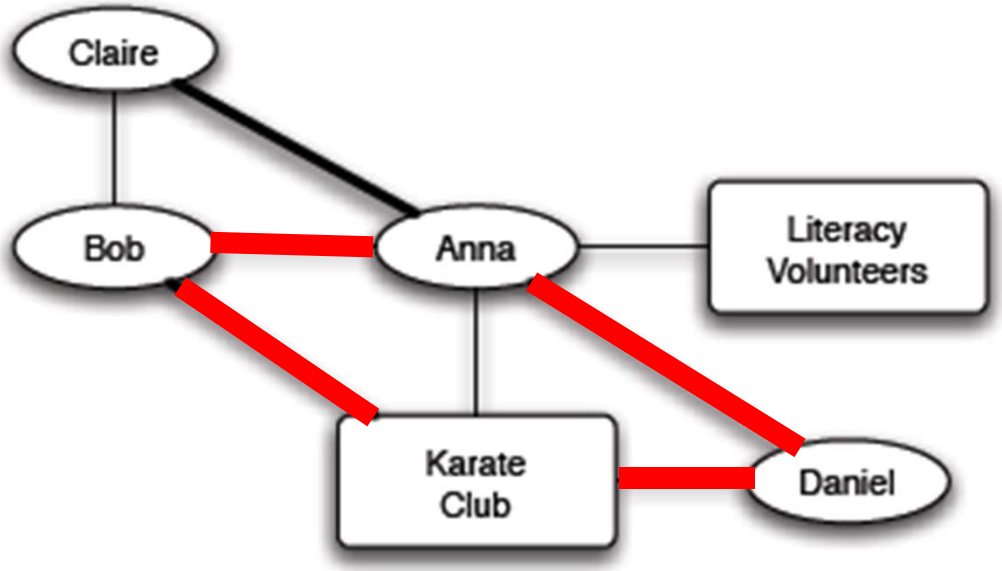
- Two activities : editing articles and editing the discussion page.
- Link between editors if they have communicated using “user talk page”.
- Each Wikipedia article defines a focus (editor associated with the articles he/she edited).

Prob of editing a Wikipedia articles as a function of the number of friends who have already done so



Multiple effects

- Bob and Daniel become friends because of their common friend Anna (Triadic Closure)
- Both Bob and Daniel go to Karate Club (Focal Closure)



Quantifying the Interplay Between Selection and Social Influence

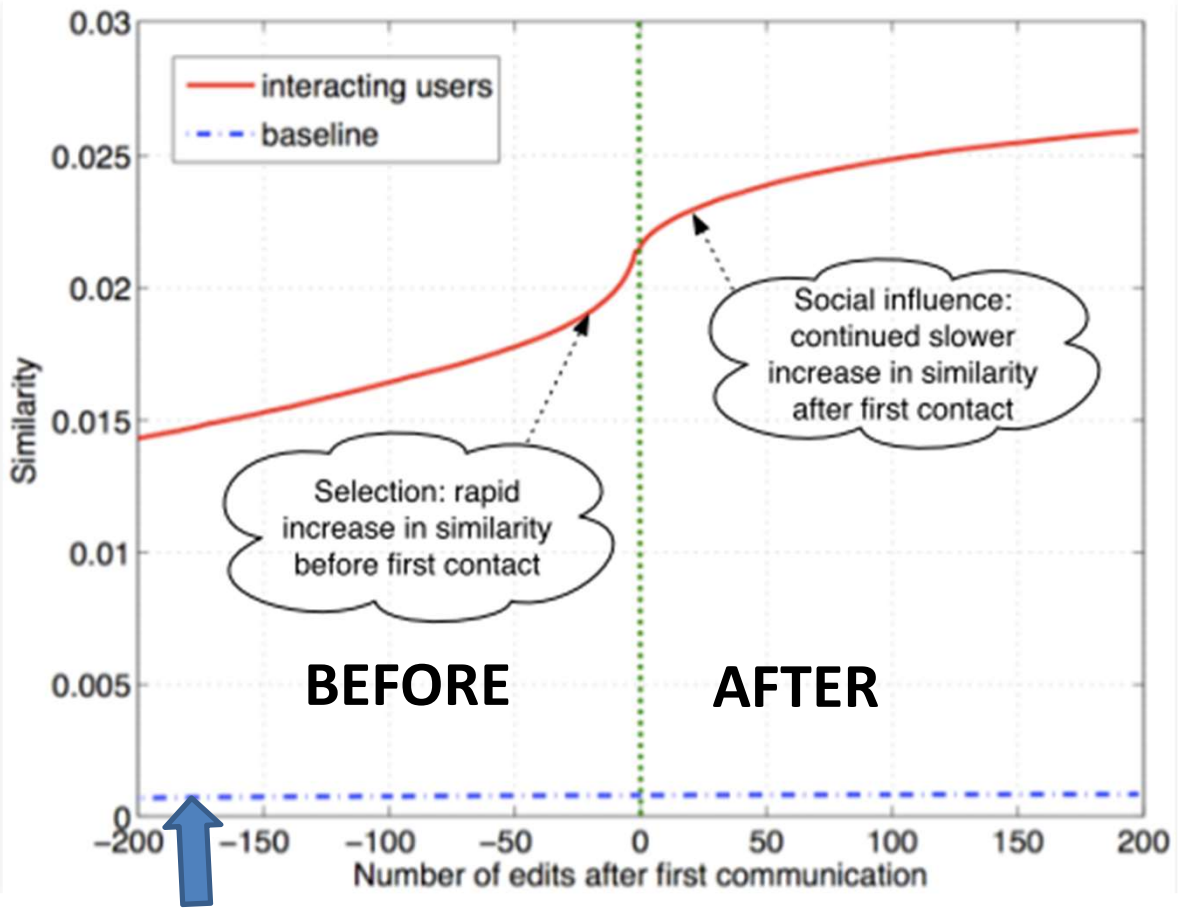
- How selection and social influence work together to produce homophily?
  - How do similarities in behavior between two Wikipedia editors relate to their pattern of social interaction over time?
- Similarity between 2 Wikipedia editors A, B

$$= \frac{number\ of\ articles\ edited\ by\ both\ A\ and\ B}{number\ of\ articles\ edited\ by\ at\ least\ one\ of\ A\ or\ B}$$

- similar to the definition of neighborhood overlap
- Is homophily arising because
  - editors are connecting (talk) with those edited the same articles (selection), or
  - editors are led to the articles of those they talk to (social influence)?

Quantifying the Interplay Between Selection and Social Influence.

- Average similarity of two editors relative to the time they began communication
- Similarity increases BEFORE and AFTER
  - Show both selection and social inference
  - Rapid rise just before they start talking (the role of selection)



Non-interacting editors

Homophily - Why Social Media Has Made You More Polarized

- http://www.science20.com/news_articles/homophily_why_social_media_has_made_you_more_polarized-128637

Homophily - Why Social Media Has Made You More Polarized

By News Staff | January 29th 2014 10:42 AM | Print | E-mail | Track Comments

RSS Share / Save ... Tweet Like

SCIENTIFIC BLOGGING

News Staff

Search This Blog

Everyone claims to care about diversity, individualism and tolerance. Very few people (R.I.P. Pete Seeger) really do. Instead, they want their beliefs affirmed and they want to demonize the opposition at every turn

The remoteness and anonymity of social media makes aggressive and cultural political posturing easy - that is why people who think the majority of their friends have differing opinions than their own engage less on Facebook. Politically active tend to stick in their own circles, ignore those on the other side and become more polarized.

A new paper suggests that politics remain the great divider - your grandparents told you never to talk politics or religion at a party and that holds true for the Internet also, people just ignore the advice these days.

Academics aren't likely to have much to offer to Facebook in how to be successful but they do anyway, saying that displaying shared interests between friends during prickly conversations could help diffuse possible arguments and alleviate tension. The paper also notes that increasing exposure and engagement to weak ties could make people more resilient in the face of political disagreement.

Currently, Facebook's news feed instead highlights people they interact with the most. It's trivial to figure out where people are politically so the authors suggest sprinking in a few opposing ideas. Though other research has shown that people are quick to 'un-friend' anyone with opposing opinions anyway (who is quickest to do so? Sorry, liberals, this time you are the least tolerant.)

"People are mainly friends with those who share similar values and interests. They tend to interact with them the most, a phenomenon called homophily," said Catherine Grevet, the Georgia Tech Ph.D. student who led the study. "But that means they rarely interact with the few friends with differing opinions. As a result, they aren't exposed to opposing viewpoints. Designing social media toward nudging users to strengthen relationships with weak ties with different viewpoints could have beneficial consequences for the platform, users and society."

The paper was based on surveys of more than 100 politically active Facebook users in the spring of 2013 amid debates about budgets cuts, gay marriage and gun control regulations. The majority of participants were liberal, female and under the age of 40, mirroring the traditional Facebook user. More than 70 percent said they don't talk about politics with their friends with different opinions. When they saw something they didn't agree with, 60 percent said they ignored it and didn't comment. When they did, sometimes it made the person question the relationship and disassociate and from the friend.

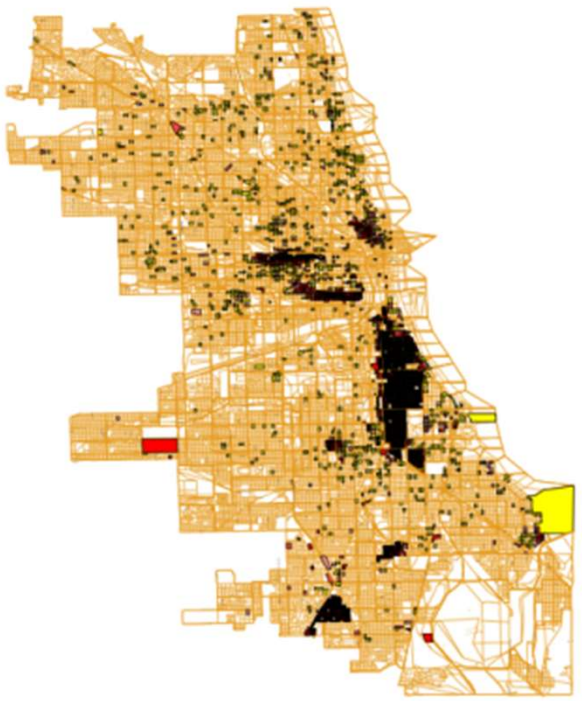
"Even though people could simply unfriend someone with different opinions, and there were certainly those who did that, there were many relationships that were able to be maintained," said Grevet. "Through a combination of behaviors on Facebook like hiding, tuning out, logging off or avoiding certain conversations, people negotiated around those differences to stay connected."

Spatial patterns of segregation

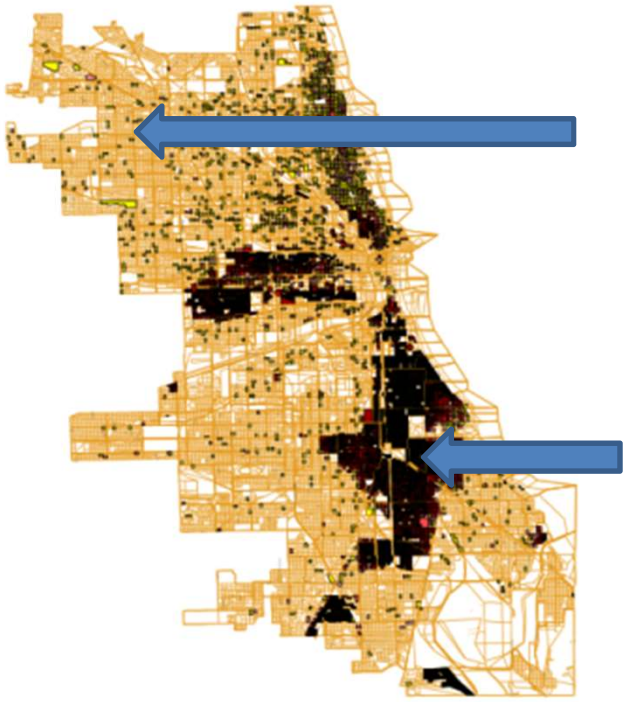
- Homophily : the formation of ethnically and racially homogeneous neighborhoods in cities
  - a process with a dynamic aspect
- What mechanisms?

Spatial patterns of segregation

Homophily : the formation of ethnically and racially homogeneous neighborhoods in cities



(a) *Chicago, 1940*



(b) *Chicago, 1960*

African-Americans < 25%

African-Americans > 75%

The Schelling Model

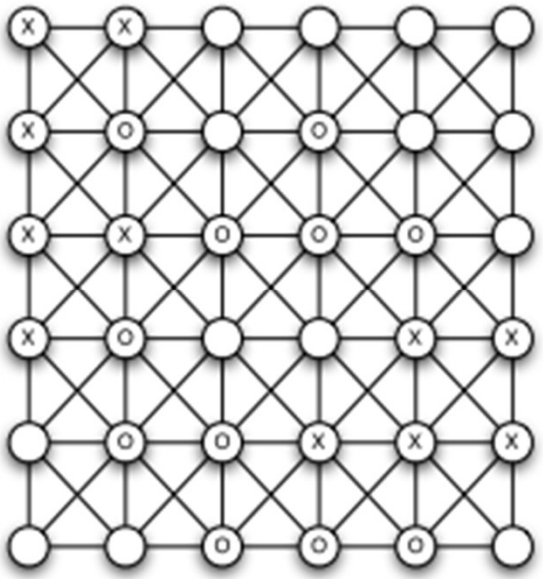
- How global patterns of spatial segregation can arise from the effect of homophily operating at a local level (Thomas Schelling)
  - an intentionally simplified mechanism
  - works even when no one individual explicitly wants a segregated outcome

The Schelling Model

- Population of individuals called agents
- Each agent is of type X or type O
  - The (immutable) characteristic (race, ethnicity, country of origin, or native language) for homophily
  - Agents reside in cells of a grid (simple model of a 2-D city map)
  - Some cells contain agents while others are unpopulated
  - Each cell has 8 neighbors, except those on the boundary

x	x				
x	o		o		
x	x	o	o	o	
x	o			x	x
	o	o	x	x	x
		o	o	o	

grid



graph

The Schelling Model

- Each agent wants to have at least *t* agents of its own type as neighbors
  - t=3 in the example
- Unsatisfied agents
  - fewer than *t* neighbors of the same type as itself
  - move to a new cell

X1	X2				
X3	O1		O2		
X4	X5	O3	O4	O5	
X6	O6			X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

X1	X2				
X3	O1		O2		
X4	X5	O3	O4	O5	
X6	O6			X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

	X2				
X3	O1		O2		
X4	X5	O3	O4	O5	
X6	O6		X1	X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

X3	O1		O2		
X4	X5	O3	O4	O5	
X6	O6	X2	X1	X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

X3		O1	O2		
X4	X5	O3	O4	O5	
X6	O6	X2	X1	X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

X3		O1	O2		
X4	X5	O3	O4	O5	
X6	O6	X2	X1	X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

X3		O1	O2		
X4	X5	O3	O4		
X6	O6	X2	X1	X7	X8
	O7	O8	X9	X10	X11
	O5	O9	O10	O11	

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
	O7	O8	X9	X10	X11
	O5	O9	O10	O11	

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
O11	O7	O8	X9	X10	X11
	O5	O9	O10		

X1	X2				
X3	O1		O2		
X4	X5	O3	O4	O5	
X6	O6			X7	X8
	O7	O8	X9	X10	X11
		O9	O10	O11	

initial

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
O11	O7	O8	X9	X10	X11
	O5	O9	O10		

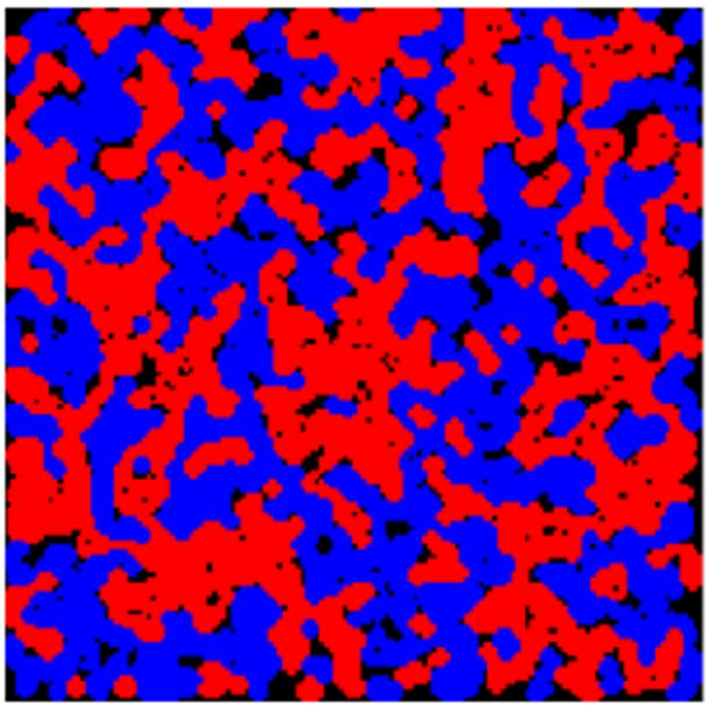
After 1 round of movement

The Dynamics of Movement

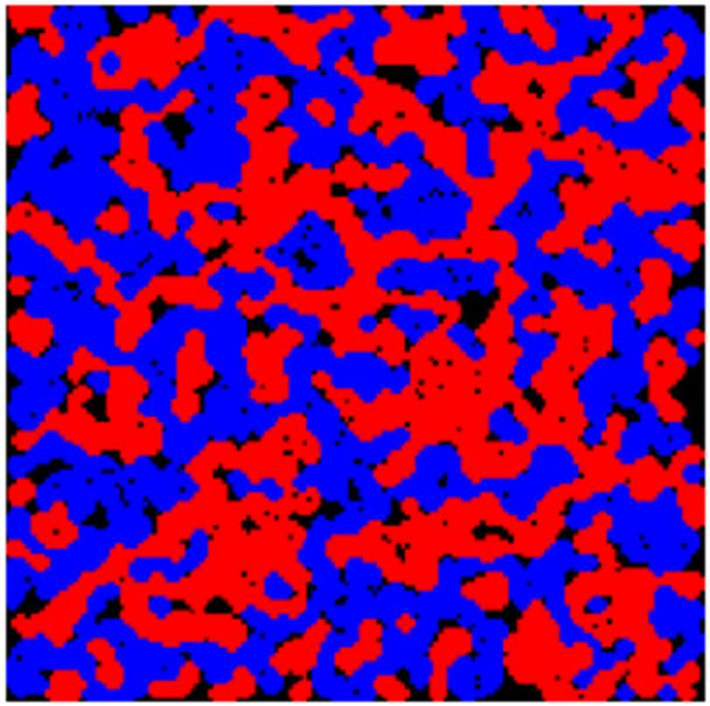
- Agents move in a sequence of rounds
  - In each round, the unsatisfied agents in some order will move to an unoccupied cell where it will be satisfied.
  - These new locations may cause different agents to be unsatisfied, and this leads to a new round of movement.
  - If no new cells make the agent satisfied, stay or move to a random cell.

Larger examples

- Agent characteristic : Red and Blue
- Empty : Black
- t =3, grid size = 150 x 150, 10,000 agents, 2500 empty cells, ~50 rounds



(a) *A simulation with threshold 3.*



(b) *Another simulation with threshold 3.*

Interpretations of the Model

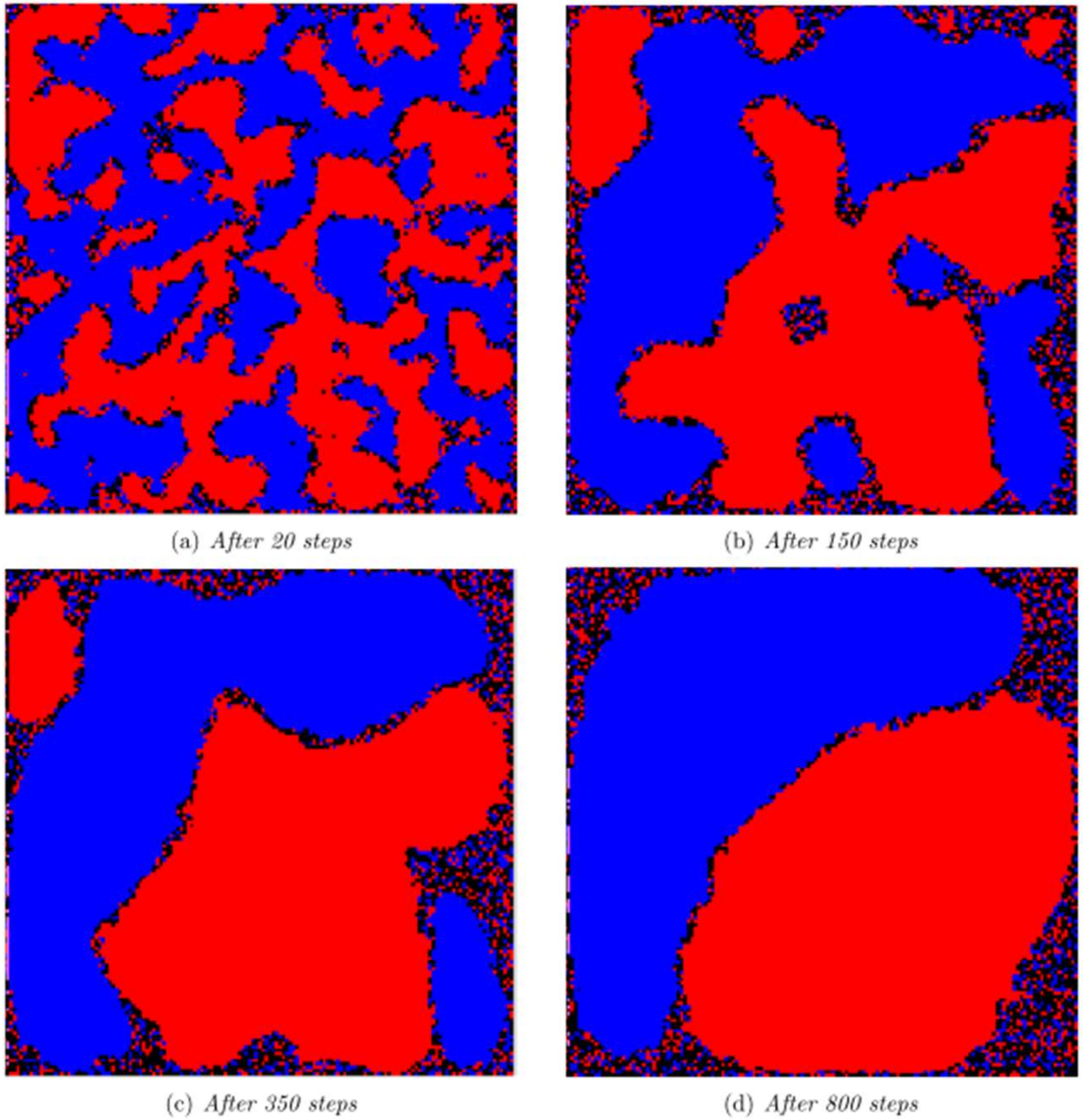
- When t=3, requirement is not harsh,
  - Minority : 3 of the same type and 5 others
- When t=4,
  - equal no. of both types.
- In the long run, integrated regions→ segregated regions

X	X	O	O	X	X
X	X	O	O	X	X
O	O	X	X	O	O
O	O	X	X	O	O
X	X	O	O	X	X
X	X	O	O	X	X

Interpretations of the Model

- More typically, agents form larger clusters
  - agents become unsatisfied and attach to larger clusters (where higher probability to be satisfied)
- The overall effect
  - local preferences of individual agents have produced a global pattern that none of them necessarily intended

t =4,
grid size = 150 x 150,
10,000 agents



(a) *After 20 steps*

(b) *After 150 steps*

(c) *After 350 steps*

(d) *After 800 steps*

Schelling model and Homophily

The Schelling model is an example that, as homophily draws people together along immutable characteristics (race or ethnicity), it creates a natural tendency for mutable characteristics (decision about where to live) to change in accordance with the network structure