

Strength of weak ties

Long-range links are often casual acquaintances,

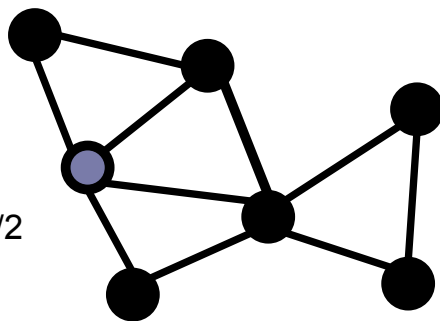
... but are very important for search and other network phenomena

Explaining triadic closure

1. **Opportunity.** If you spend a lot of time with your best friend and your girlfriend, there is an increased chance they will meet.
2. **Incentive.** If your best friend hates your girlfriend, it stresses both relationships.
3. **Homophily.** If you have things in common with both your best friend and your girlfriend, they have things in common too.

Definition: The **clustering coefficient** of a node v is the fraction of pairs of v 's friends that are connected to each other by edges.

Clustering Coefficient = $1/2$



The higher the clustering coefficient of a node, the more strongly triadic closure is acting on it

factors influencing diffusion

- network structure (unweighted)
 - density
 - degree distribution
 - clustering
 - connected components
 - community structure
- strength of ties (weighted)
 - frequency of communication
 - strength of influence
- spreading agent
 - attractiveness and specificity of information

how does strength of a tie influence diffusion?

- M. S. Granovetter: *The Strength of Weak Ties*, AJS, 1973:
- finding a job through a contact that one saw
 - frequently (2+ times/week) 16.7%
 - occasionally (more than once a year but < 2x week) 55.6%
 - rarely 27.8%
- but... length of path is short
 - contact directly works for/is the employer
 - or is connected directly to employer

strength of tie: frequency of communication

- Kossinets, Watts, Kleinberg, KDD 2008:
 - which paths yield the most up to date info?
 - how many of the edges form the “backbone”?

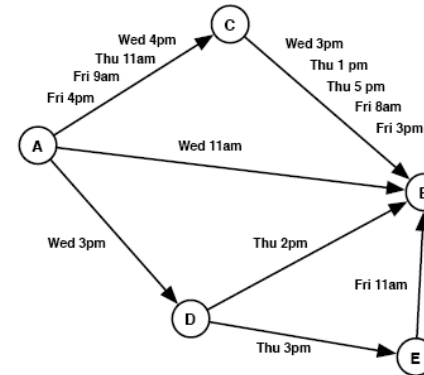
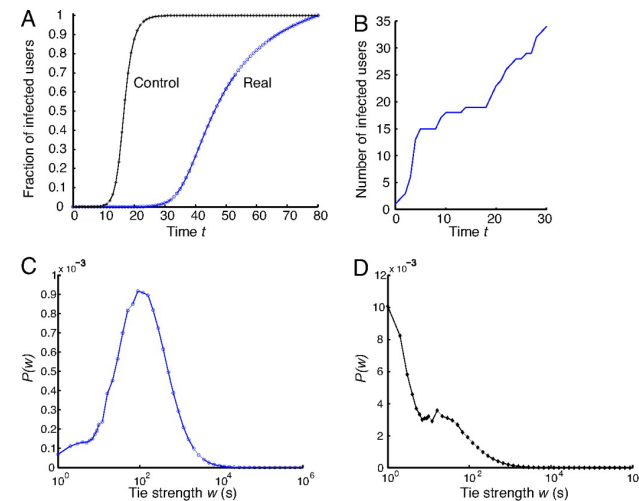


image source: Kossinets et al. “The structure of information pathways in a social communication network”, KDD 2008

the strength of intermediate ties

- strong ties
 - frequent communication, but ties are redundant due to high clustering
- weak ties
 - reach far across network, but communication is infrequent...
- Onnela J. et.al. PNAS 2007;104:7332-7336
 - use nation-wide cellphone call records and simulate diffusion using actual call timing
 - in simulation, individuals are most likely to obtain novel information through ties of intermediate strength

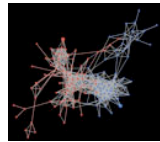
Localized strong ties slow infection spread.



source: Onnela J. et.al. PNAS 2007;104:7332-7336

how can information diffusion be different from simple contagion (e.g. a virus)?

- simple contagion:
 - infected individual infects neighbors with information at some rate
- threshold contagion:
 - individuals must hear information (or observe behavior) from a number or fraction of friends before adopting
- in lab: complex contagion (Centola & Macy, AJS, 2007)
 - how do you pick individuals to “infect” such that your opinion prevails
 - try it out in NetLogo:
 - <http://projects.si.umich.edu/netlearn/NetLogo4/DiffusionCompetition.html>



diffusion of innovation

- surveys:
 - farmers adopting new varieties of hybrid corn by observing what their neighbors were planting (Ryan and Gross, 1943)
 - doctors prescribing new medication (Coleman et al. 1957) (see lab to play with data set)
 - Christakis and Fowler (spread of obesity & happiness in social networks) 2008
- online behavioral data:
 - Lerman (spread of Flickr photos & Digg stories) 2007
 - Backstrom et al. (joining LiveJournal groups & CS conferences) 2006
 - + others e.g. Anagnostopoulos et al. 2008

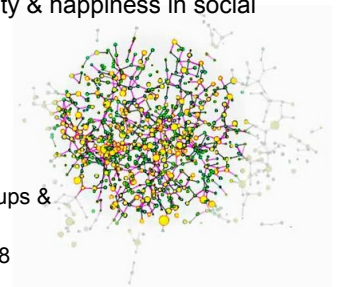
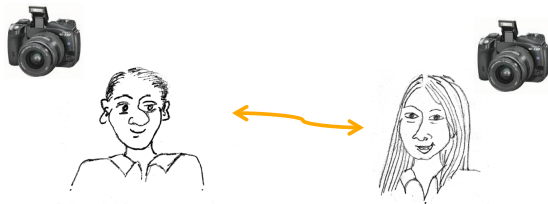


image source: Christakis & Fowler, 'The Spread of Obesity in a Large Social Network over 32 years', NEJM 357(4):370-379, 2007

Open question: how do we tell influence from correlation?



- approaches:
 - time resolved data: if adoption time is shuffled, does it yield the same patterns?
 - if edges are directed: does reversing the edge direction yield less predictive power?

Example from reading: adopting new practices

- Davis, corporate governance in the 1980s

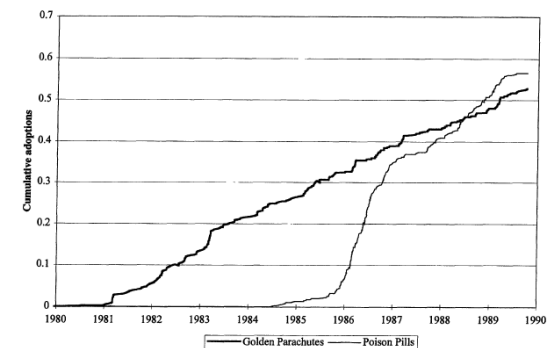


FIG. 1—Diffusion of poison pills and golden parachutes among 1986 Fortune 500 firms, 1980–89.

Source: Corporate Elite Networks and Governance Changes in the 1980s; Gerald F. Davis, Henrich R. AJS Volume 103 Number 1 (July 1997): 1– 37.

differences

- poison pills
 - diffused through interlocks
 - geography had little to do with it
 - more likely to be influenced by tie to firm doing something similar & having similar centrality
- golden parachutes
 - did not diffuse through interlocks
 - geography was a significant factor
 - more likely to follow “central” firms
- why did one diffuse through the “network” while the other did not?

Burt: structural holes and good ideas

- Managers asked to come up with an idea to improve the supply chain
- Then asked:
 - whom did you discuss the idea with?
 - whom do you discuss supply-chain issues with in general
 - do those contacts discuss ideas with one another?

- 673 managers (455 (68%) completed the survey)
- ~ 4000 relationships (edges)