School of Information
University of Michigan

Information diffusion in networks

outline

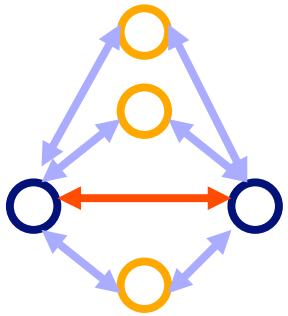
- factors influencing information diffusion
  - network structure: which nodes are connected?
  - strength of ties: how strong are the connections?
- studies in information diffusion:
  - Granovetter: the strength of weak ties
  - J-P Onnela et al: strength of intermediate ties
  - Kossinets et al: strength of backbone ties
  - Davis: board interlocks and adoption of practices
- network position and access to information
  - Burt: Structural holes and good ideas
  - Aral and van Alstyne: networks and information advantage
- networks and innovation
  - Lazer and Friedman: innovation

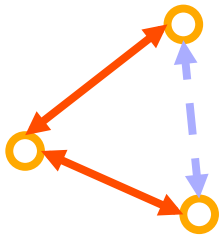
factors influencing diffusion

- 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

Strong tie defined

- A strong tie
  - frequent contact
  - affinity
  - many mutual contacts

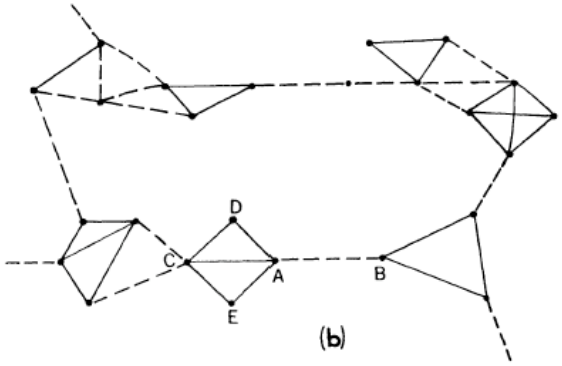




"forbidden triad":

strong ties are likely to "close"

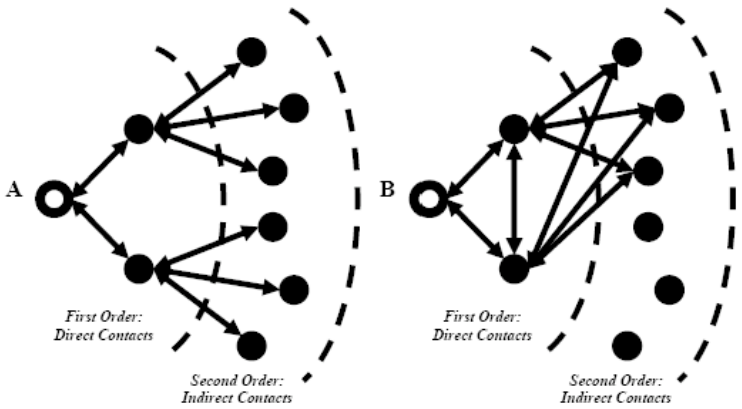
- Less likely to be a bridge (or a local bridge)



Source: Granovetter, M. (1973). "The Strength of Weak Ties", American Journal of Sociology, Vol. 78, Issue 6, May 1973, pp. 1360-1380.

school kids and 1st through 8th choices of friends

- snowball sampling:
  - will you reach more different kids by asking each kid to name their 2 best friends, or their 7th & 8th closest friend?



Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

outline

- factors influencing information diffusion
  - network structure: which nodes are connected?
  - strength of ties: how strong are the connections?
- studies in information diffusion:
  - Granovetter: the strength of weak ties
  - J-P Onnela et al: strength of intermediate ties
  - Kossinets et al: strength of backbone ties
  - Davis: board interlocks and adoption of practices
- network position and access to information
  - Burt: Structural holes and good ideas
  - Aral and van Alstyne: networks and information advantage
- networks and innovation
  - Lazer and Friedman: innovation

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”?

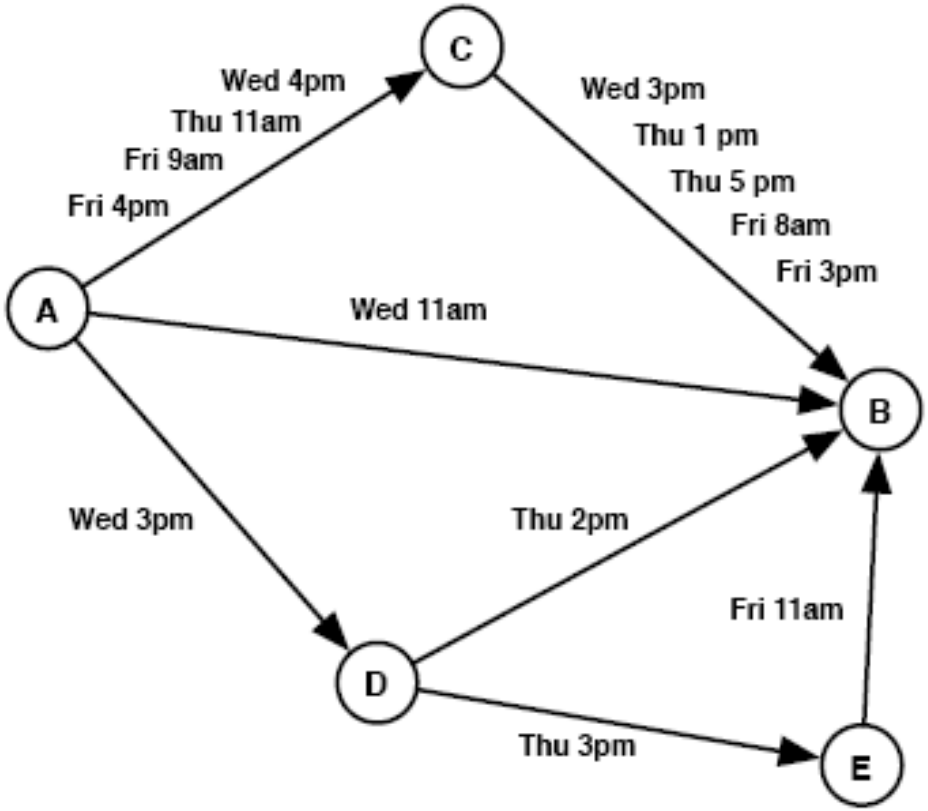
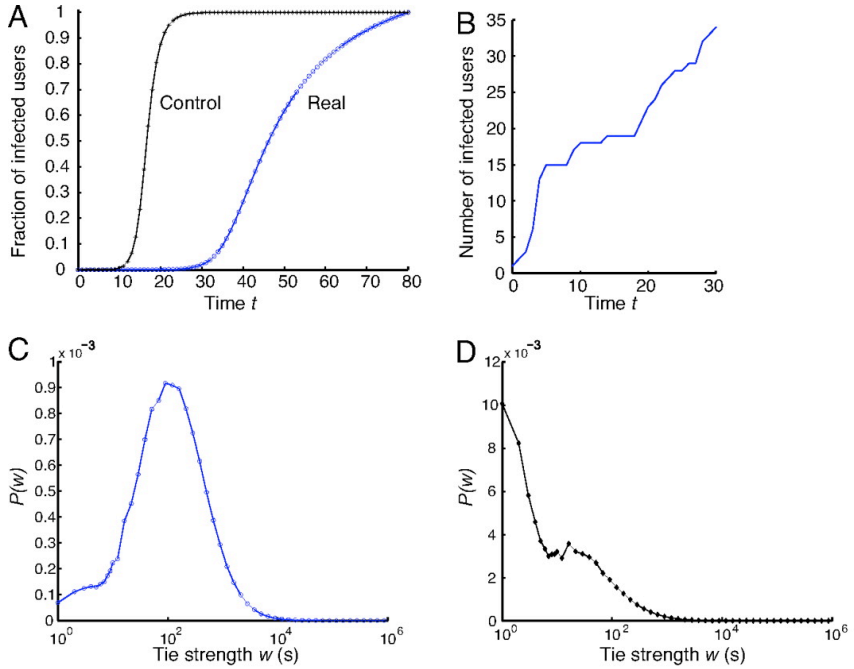


image source: Kossinets et al. “The structure of information pathways in a social communication network”, KDD 2008

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

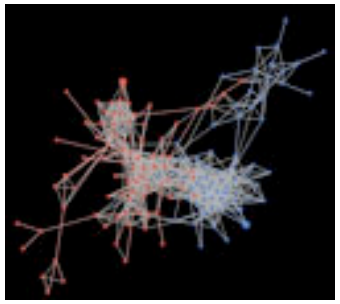
Localized strong ties slow infection spread.



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

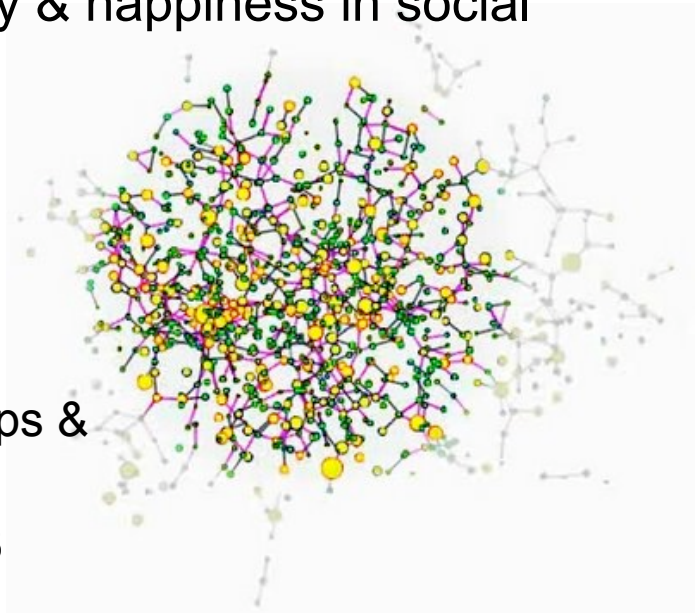
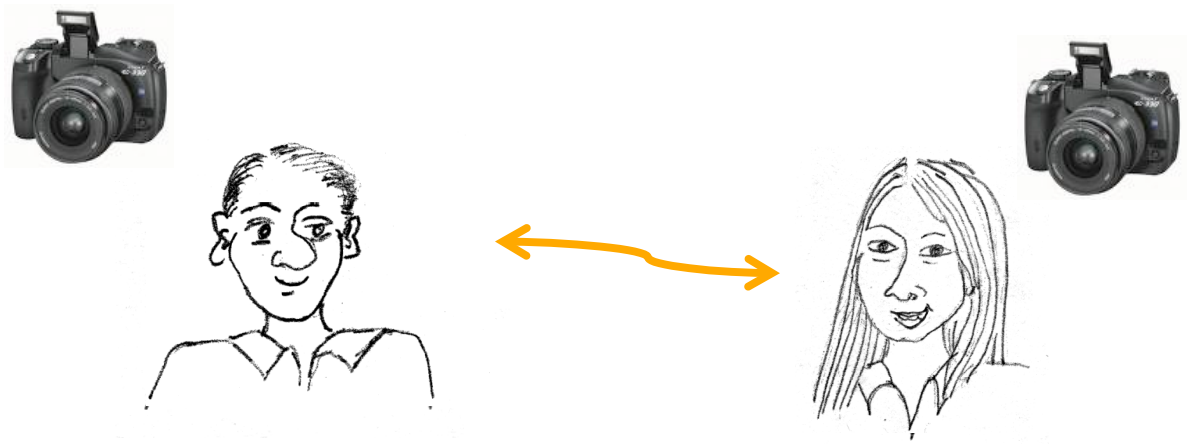


image source: Christakis & Fowler, 'The Spread of Obesity in a Large Social Network over 32 years', NEJM 357(4):370-379, 2007

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

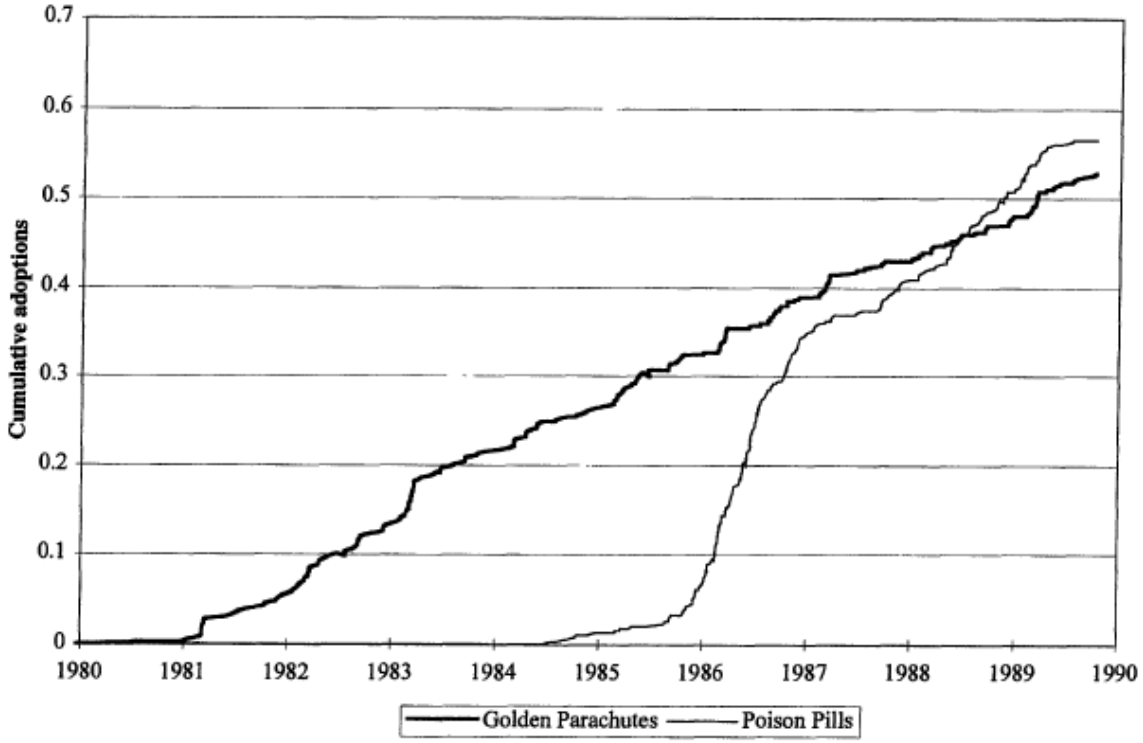


FIG. 1—Diffusion of poison pills and golden parachutes among 1986 Fortune 500 firms, 1980–89.

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?

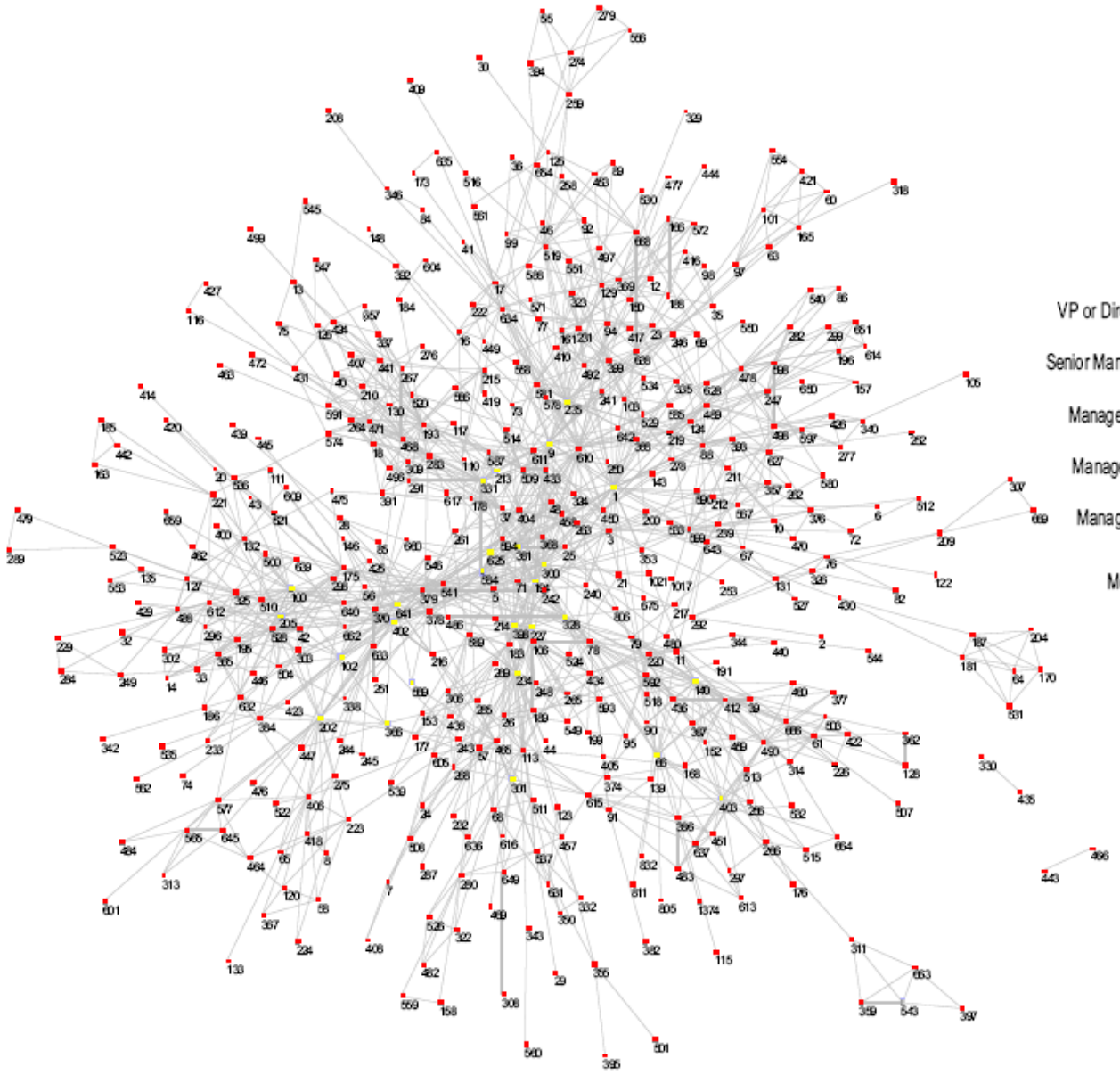
outline

- factors influencing information diffusion
  - network structure: which nodes are connected?
  - strength of ties: how strong are the connections?
- studies in information diffusion:
  - Granovetter: the strength of weak ties
  - J-P Onnela et al: strength of intermediate ties
  - Kossinets et al: strength of backbone ties
  - Davis: board interlocks and adoption of practices
- network position and access to information
  - Burt: Structural holes and good ideas
  - Aral and van Alstyne: networks and information advantage
- networks and innovation
  - Lazer and Friedman: innovation

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)



	Percent Social Isolates	Mean Network Size	Mean Network Constraint	Mean Number Cited as Discussion Partners	Mean Network Constraint Cited Discussn. Partners	Mean Path Distance (min-max) for the 476 connected managers in graph
VP or Director (25)	0%	12.6	29.8	4.9	70.2	3.3 (2.7-4.2)
Senior Manager (41)	5%	8.5	37.3	3.8	78.1	3.7 (2.9-6.4)
Manager III (121)	11%	6.4	50.2	3.7	77.9	4.0 (3.0-6.4)
Manager II (199)	27%	4.1	65.0	2.8	83.1	4.3 (2.8-6.4)
Manager I (287)	44%	3.4	73.6	2.4	83.4	4.6 (3.4-7.4)
Mean (673)	29%	5.0	60.5	2.9	81.0	4.2 (2.7-7.4)

Figure 2. Supply-Chain Discussion Network

(excludes 193 social isolates)

Source: Structural Holes and Good Ideas; R. Burt, American Journal of Sociology, 2004

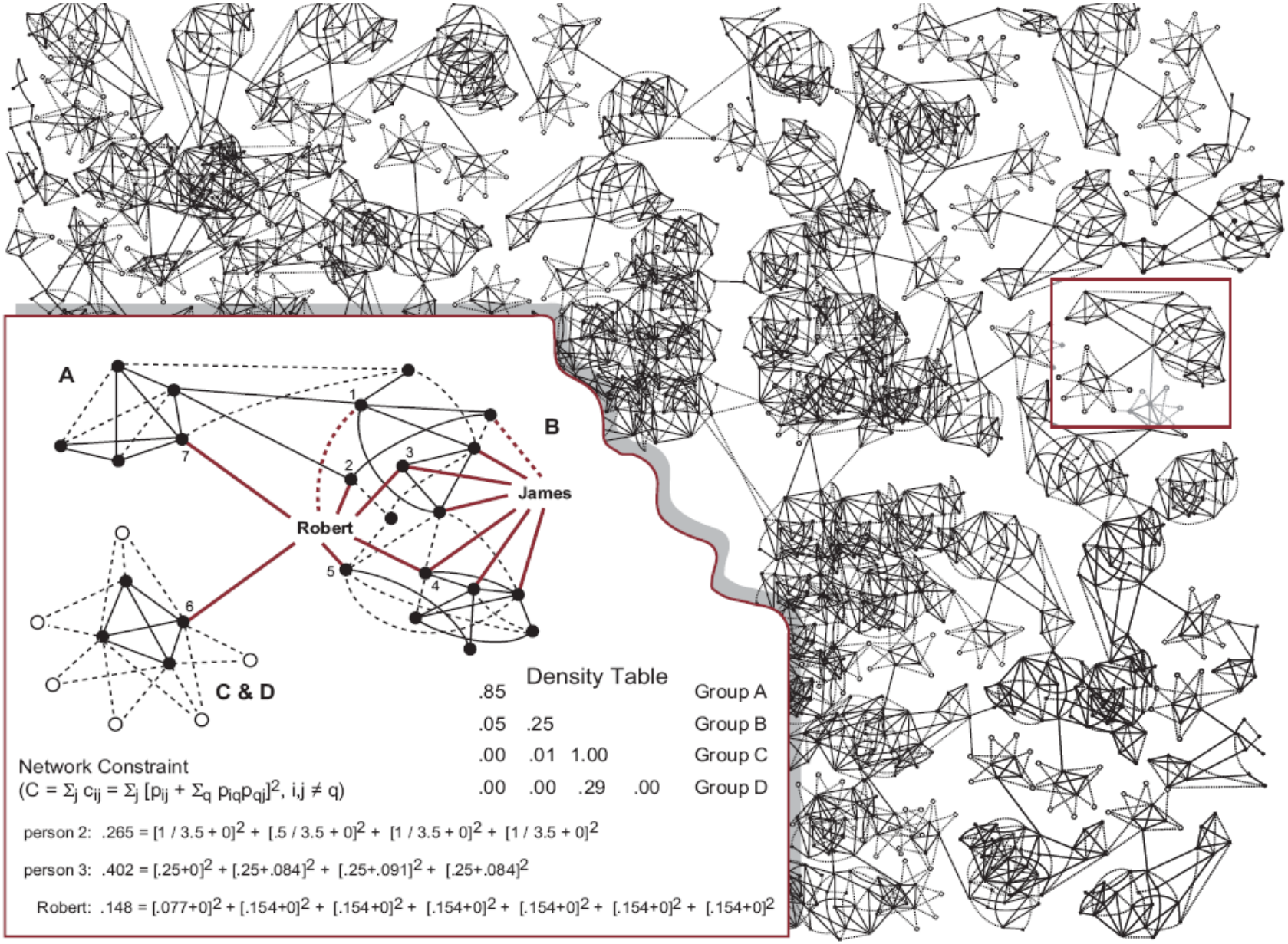


Figure 1. The Small World of Markets and Organizations

Source: Structural Holes and Good Ideas; R. Burt, American Journal of Sociology, 2004

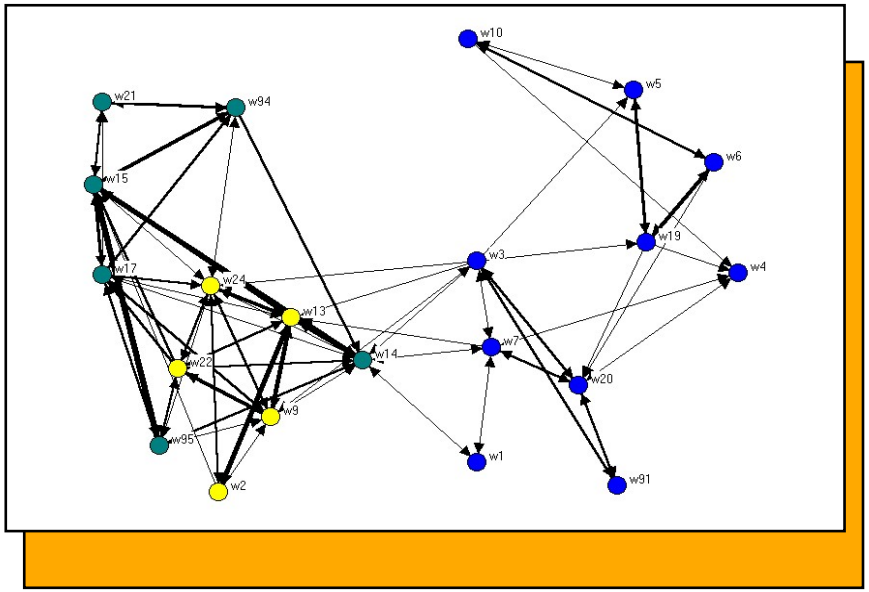
results

- people whose networks bridge structural holes have
  - higher compensation
  - positive performance evaluations
  - more promotions
  - more good ideas

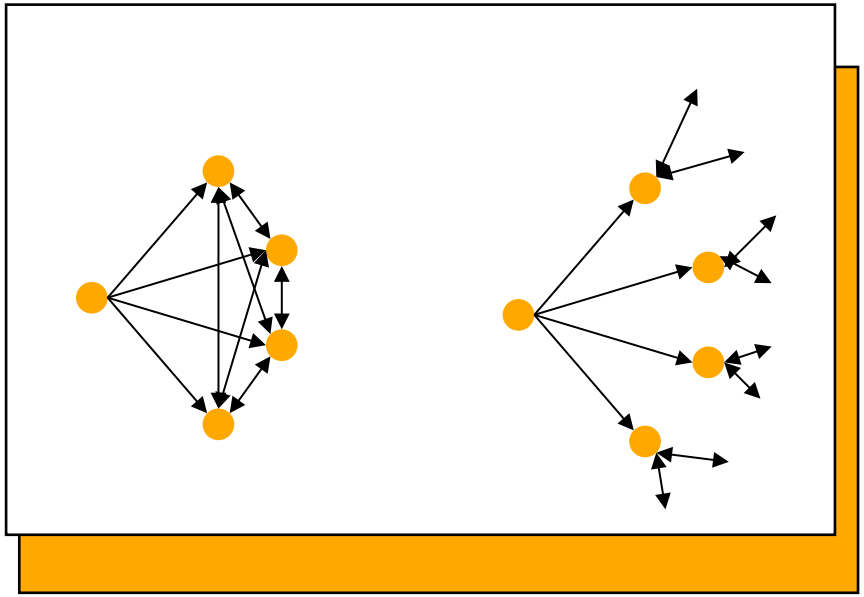
- these brokers are
  - more likely to express ideas
  - less likely to have their ideas dismissed by judges
  - more likely to have their ideas evaluated as valuable

networks & information advantage

Betweenness



Constrained vs. Unconstrained



Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital (formerly titled 'Network Structure & Information Advantage'), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

slide: Marshall van Alstyne

Aral & Alstyne: Study of a head hunter firm

- Three firms initially
- Unusually measurable inputs and outputs
  - 1300 projects over 5 yrs and
  - 125,000 email messages over 10 months (avg 20% of time!)
  - Metrics

    (i) Revenues per person and per project,
    (ii) number of completed projects,
    (iii) duration of projects,
    (iv) number of simultaneous projects,
    (v) compensation per person
- Main firm 71 people in executive search (+2 firms partial data)
  - 27 Partners, 29 Consultants, 13 Research, 2 IT staff
- Four Data Sets per firm
  - 52 Question Survey (86% response rate)
  - E-Mail
  - Accounting
  - 15 Semi-structured interviews

Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital (formerly titled 'Network Structure & Information Advantage'), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

slide: Marshall van Alstyne

Email structure matters

New Contract Revenue — **Coefficients[a]**

	Unstandardized Coefficients			
	B	Std. Error	Adj. R^2	Sig. F Δ
(Base Model)			0.40	
Best structural pred.	12604.0***	4454.0	0.52	.006
Ave. E-Mail Size	-10.7**	4.9	0.56	.042
Colleagues' Ave. Response Time	-198947.0	168968.0	0.56	.248

a. Dependent Variable: **Bookings02**
b. Base Model: YRS_EXP, PARTDUM, %_CEO_SRCH, SECTOR(dummies), %_SOLO.

Contract Execution Revenue — **Coefficients[a]**

	Unstandardized Coefficients			
	B	Std. Error	Adj. R^2	Sig. F Δ
(Base Model)			0.19	
Best structural pred.	1544.0**	639.0	0.30	.021
Ave. E-Mail Size	-9.3*	4.7	0.34	.095
Colleagues' Ave. Response Time	-368924.0**	157789.0	0.42	.026

a. Dependent Variable: **Billings02**
b. N=39. *** p<.01, ** p<.05, * p<.1

Sending *shorter* e-mail helps get contracts and finish them.

Faster response from colleagues helps finish them.

Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital (formerly titled 'Network Structure & Information Advantage'), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

slide: Marshall van Alstyne

H5: Recruiters with larger personal rolodexes generate no more or less output

	Revenue $	$ for completed searches	Completed searches	Multitasking	Duration	Duration controlling for multitasking
Size of rolodex (Q50)	-10.2 (60.3)	-22.9 (32.6)	0.000 (0.001)	0.000 (0.001)	-0.013 (0.021)	-0.013 (0.016)

*** p < 0.10, ** p < 0.05, *** p < 0.01, Standard err in paren.**

Instead, a larger private rolodex is associated with:

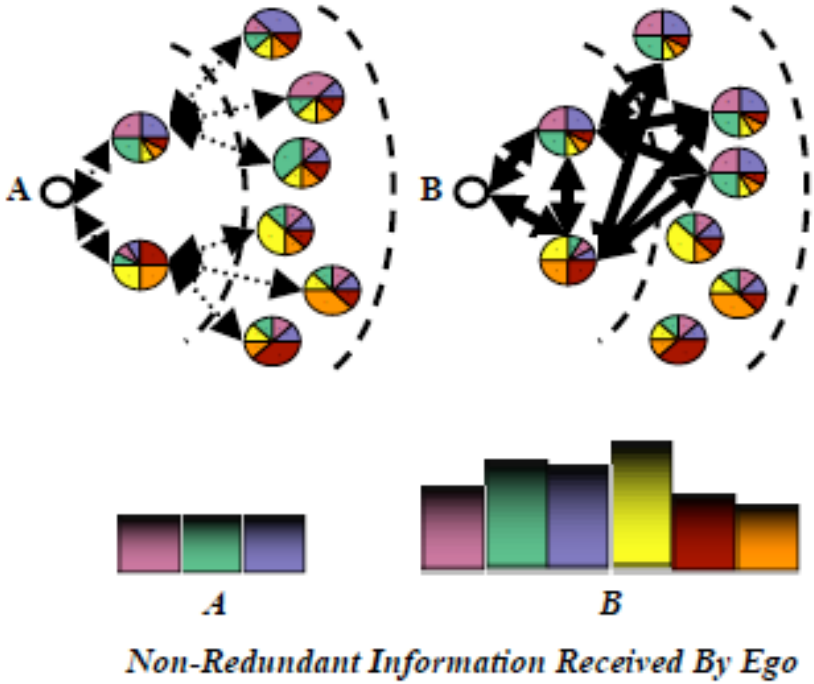
- Less information sharing
- Less DB proficiency
- Lower % of e-mail read
- Less learning from others
- Less perceived credit for ideas given to colleagues
- More dissembling on the phone

Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital (formerly titled 'Network Structure & Information Advantage'), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

slide: Marshall van Alstyne

diverse networks drive performance by providing access to novel information

- network structure (having high degree) correlates with receiving novel information sooner (as deduced from hashed versions of their email)
- getting information sooner correlates with $$ brought in
  - controlling for # of years worked
  - job level
  - ….



Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

Network Structure Matters

New Contract Revenue — **Coefficients[a]**

	Unstandardized Coefficients			
	B	Std. Error	Adj. R^2	Sig. F Δ
(Base Model)			0.40	
Size Struct. Holes	13770***	4647	0.52	.006
Betweenness	1297*	773	0.47	.040

a. Dependent Variable: **Bookings02**
b. Base Model: YRS_EXP, PARTDUM, %_CEO_SRCH, SECTOR(dummies), %_SOLO.

Contract Execution Revenue — **Coefficients[a]**

	Unstandardized Coefficients			
	B	Std. Error	Adj. R^2	Sig. F Δ
(Base Model)			0.19	
Size Struct. Holes	7890*	4656	0.24	.100
Betweenness	1696**	697	0.30	.021

a. Dependent Variable: **Billings02**
b. N=39. *** p<.01, ** p<.05, * p<.1

Bridging diverse communities is significant.

Being in the thick of information flows is significant.

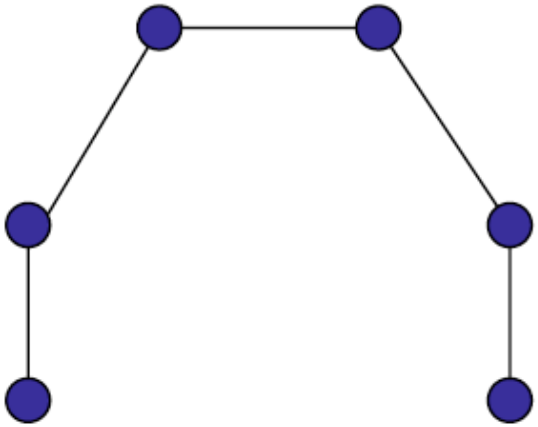
Source: M. van Alstyne, S. Aral. Networks, Information & Social Capital (formerly titled 'Network Structure & Information Advantage'), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958158

slide: Marshall van Alstyne

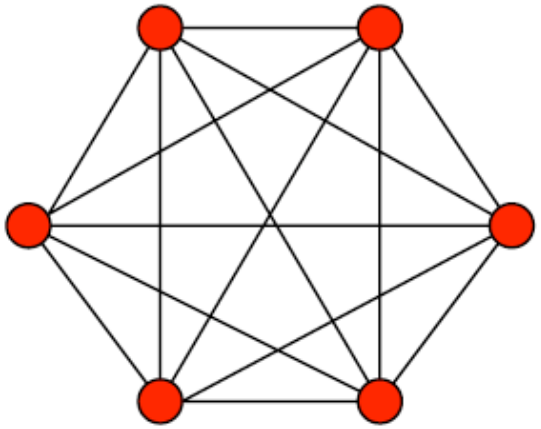
outline

- factors influencing information diffusion
  - network structure: which nodes are connected?
  - strength of ties: how strong are the connections?
- studies in information diffusion:
  - Granovetter: the strength of weak ties
  - J-P Onnela et al: strength of intermediate ties
  - Kossinets et al: strength of backbone ties
  - Davis: board interlocks and adoption of practices
- network position and access to information
  - Burt: Structural holes and good ideas
  - Aral and van Alstyne: networks and information advantage
- networks and innovation
  - Lazer and Friedman: innovation

networks and innovation: is more information diffusion always better?



linear network



fully connected network

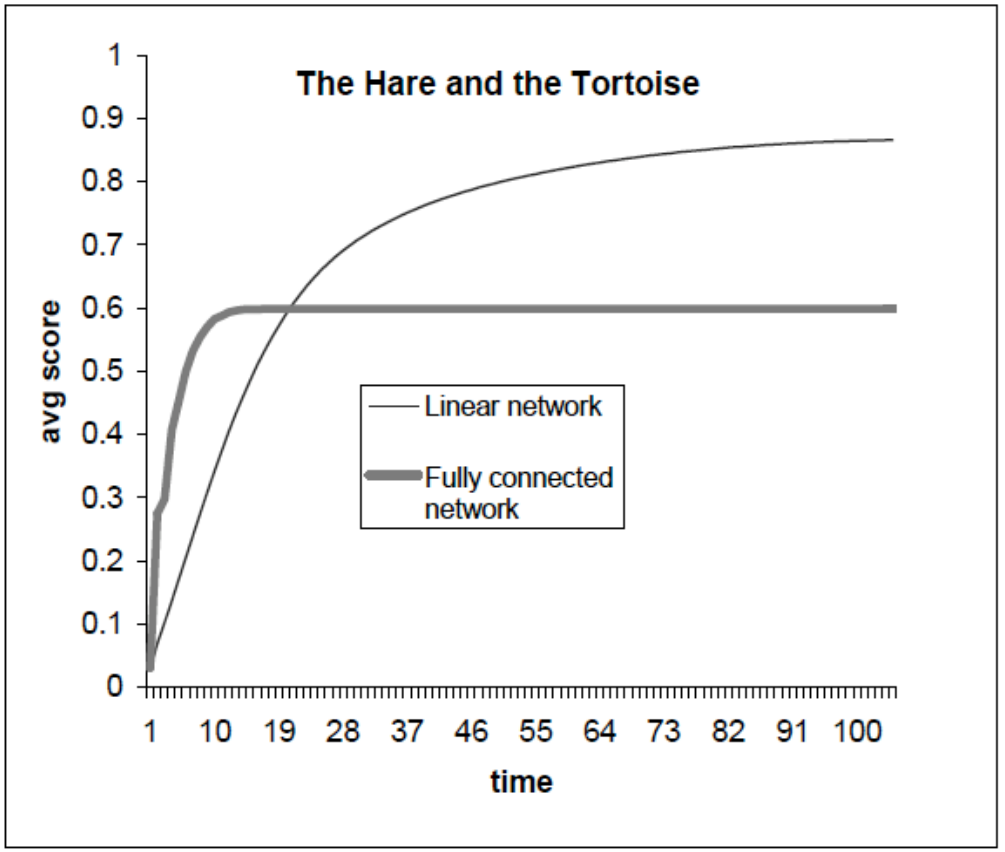
- Nodes can innovate on their own (slowly) or adopt their neighbor's solution
- Best solutions propagate through the network

Tortoise, Hare: David Eppstein. http://commons.wikimedia.org/wiki/File:Tortoise_and_hare_algorithm.svg

source: Lazer, David and Friedman, Allan,The Parable of the Hare and the Tortoise: Small Worlds, Diversity, and System Performance: http://ssrn.com/abstract=832627

networks and innovation

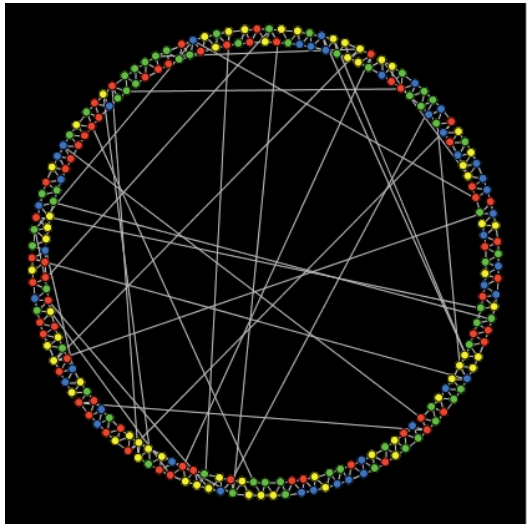
- fully connected network converges more quickly on a solution, but if there are lots of local maxima in the solution space, it may get stuck without finding optimum.
- linear network (fewer edges) arrives at better solution eventually because individuals innovate longer



source: Lazer, David and Friedman, Allan,The Parable of the Hare and the Tortoise: Small Worlds, Diversity, and System Performance: http://ssrn.com/abstract=832627

lab: networks and coordination

- Kearns et al. Science 313 (5788), pp. 824 – 827, 2006:
  - network structure affects convergence in coordination games, e.g. graph coloring
  - try it out in NetLogo:
    - http://projects.si.umich.edu/netlearn/NetLogo4/GraphColoring.html



to sum up

- network structure influences information diffusion
- strength of tie matters
- diffusion can be simple (person to person) or complex (individuals having thresholds)
- people in special network positions (the brokers) have an advantage in receiving novel info & coming up with “novel” ideas
- in some scenarios, information diffusion may hinder innovation