Contagion

Complex Networks, Course 295A, Spring, 2008

Prof. Peter Dodds

Department of Mathematics & Statistics
University of Vermont

Outline

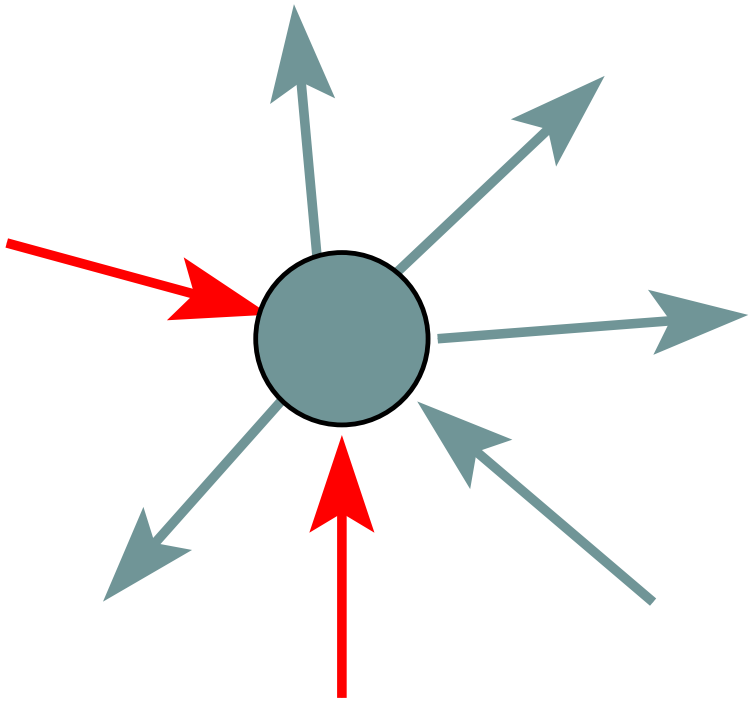
Contagion models

Some large questions concerning network contagion:

1. For a given spreading mechanism on a given network, what's the probability that there will be global spreading?
2. If spreading does take off, how far will it go?
3. How do the details of the network affect the outcome?
4. How do the details of the spreading mechanism affect the outcome?

- Next up: We'll look at some fundamental kinds of spreading on generalized random networks.

Spreading mechanisms



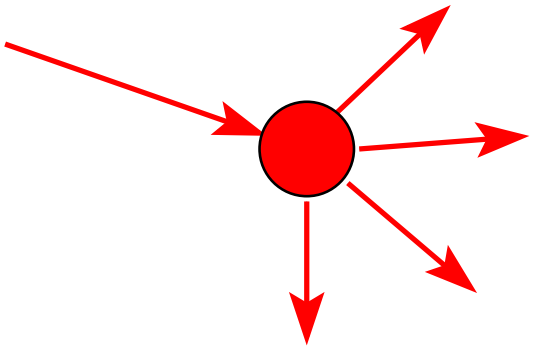
uninfected
infected

- General spreading mechanism:
  State of node *i* depends on history of *i* and *i*'s neighbors' states.
- Doses of entity may be stochastic and history-dependent.
- May have multiple, interacting entities spreading at once.

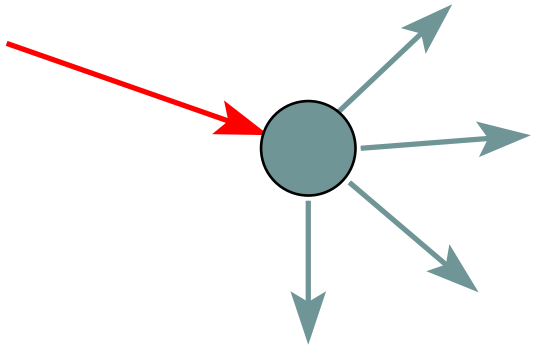
Spreading on Random Networks

- For random networks, we know local structure is pure branching.
- Successful spreading is ∴ contingent on single edges infecting nodes.

Success



Failure:



- Focus on binary case with edges and nodes either infected or not.

Contagion condition

- We need to find:
  r = the average # of infected edges that one random infected edge brings about.
- Define β_k as the probability that a node of degree k is infected by a single infected edge.
-

$$r = \sum_{k=0}^{\infty} \underbrace{\frac{kP_k}{\langle k \rangle}}_{\substack{\text{prob. of} \\ \text{connecting to} \\ \text{a degree } k \text{ node}}} \cdot \underbrace{\beta_k}_{\substack{\text{Prob. of} \\ \text{infection}}} \cdot \underbrace{(k-1)}_{\substack{\text{\# outgoing} \\ \text{infected} \\ \text{edges}}}$$

$$+ \sum_{k=0}^{\infty} \overbrace{\frac{kP_k}{\langle k \rangle}} \cdot \underbrace{(1-\beta_k)}_{\substack{\text{Prob. of} \\ \text{no infection}}} \cdot \underbrace{0}_{\substack{\text{\# outgoing} \\ \text{infected} \\ \text{edges}}}$$

Contagion condition

- Our contagion condition is then:

$$\boxed{r = \sum_{k=0}^{\infty} \frac{(k-1)kP_k}{\langle k \rangle} \beta_k > 1.}$$

- Case 1: If $\beta_k = 1$ then

$$r = \frac{\langle k(k-1) \rangle}{\langle k \rangle} > 1.$$

- Good: This is just our giant component condition again.

Contagion condition

- Case 2: If $\beta_k = \beta < 1$ then

$$r = \beta \frac{\langle k(k-1) \rangle}{\langle k \rangle} > 1.$$

- A fraction (1-β) edges do not transmit the infection.
- Analogous phase transition to giant component case but critical value of $\langle k \rangle$ is increased.
- Aka bond percolation.
- Resulting degree distribution P'_k:

$$P'_k = \beta^k \sum_{i=k}^{\infty} \binom{i}{k} (1-\beta)^{i-k} P_i.$$

- We can show $F_{P'}(x) = F_P(\beta x + 1 - \beta)$.

Contagion condition

- Cases 3, 4, 5, ...: Now allow β_k to depend on k
- Asymmetry: Transmission along an edge depends on node's degree at other end.
- Possibility: β_k increases with k... unlikely.
- Possibility: β_k is not monotonic in k... unlikely.
- Possibility: β_k decreases with k... hmmm.
- $\beta_k \searrow$ is a plausible representation of a simple kind of social contagion.
- The story:
  More well connected people are harder to influence.

Contagion condition

- Example: $\beta_k = 1/k$.
- $$r = \sum_{k=1}^{\infty} \frac{(k-1)kP_k}{\langle k \rangle} \beta_k = \sum_{k=1}^{\infty} \frac{(k-1)kP_k}{\langle k \rangle k}$$
$$= \sum_{k=1}^{\infty} \frac{(k-1)P_k}{\langle k \rangle} = \frac{\langle k \rangle - 1}{\langle k \rangle} = 1 - \frac{1}{\langle k \rangle}$$
- Since r is always less than 1, no spreading can occur for this mechanism.
- Decay of β_k is too fast.
- Result is independent of degree distribution.

Contagion condition

- Example: $\beta_k = H(\frac{1}{k} - \phi)$
  where $0 < \phi \leq 1$ is a threshold and H is the Heaviside function.
- Infection only occurs for nodes with low degree.
- Call these nodes vulnerables:
  they flip when only one of their friends flips.
-

$$r = \sum_{k=1}^{\infty} \frac{(k-1)kP_k}{\langle k \rangle} \beta_k = \sum_{k=1}^{\infty} \frac{(k-1)kP_k}{\langle k \rangle} H(\frac{1}{k} - \phi)$$

$$= \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} \frac{(k-1)kP_k}{\langle k \rangle} \quad \text{where } \lfloor \cdot \rfloor \text{ means floor.}$$

Contagion condition

- The contagion condition:

$$r = \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} \frac{(k-1)kP_k}{\langle k \rangle} > 1.$$

- As $\phi \to 1$, all nodes become resilient and $r \to 0$.
- As $\phi \to 0$, all nodes become vulnerable and the contagion condition matches up with the giant component condition.
- Key: If we fix ϕ and then vary $\langle k \rangle$, we may see two phase transitions.
- Added to our standard giant component transition, we will see a cut off in spreading as nodes become more connected.

Thresholds

- What if we now allow thresholds to vary?
- We need to backtrack a little...

Social Contagion

Social Contagion

Examples abound

- being polite/rude
- strikes
- innovation
- residential segregation
- ipods
- obesity
- Harry Potter
- voting
- gossip
- Rubik's cube 
- religious beliefs
- leaving lectures

SIR and SIRS contagion possible

- Classes of behavior versus specific behavior: dieting

Social Contagion

We need to understand influence

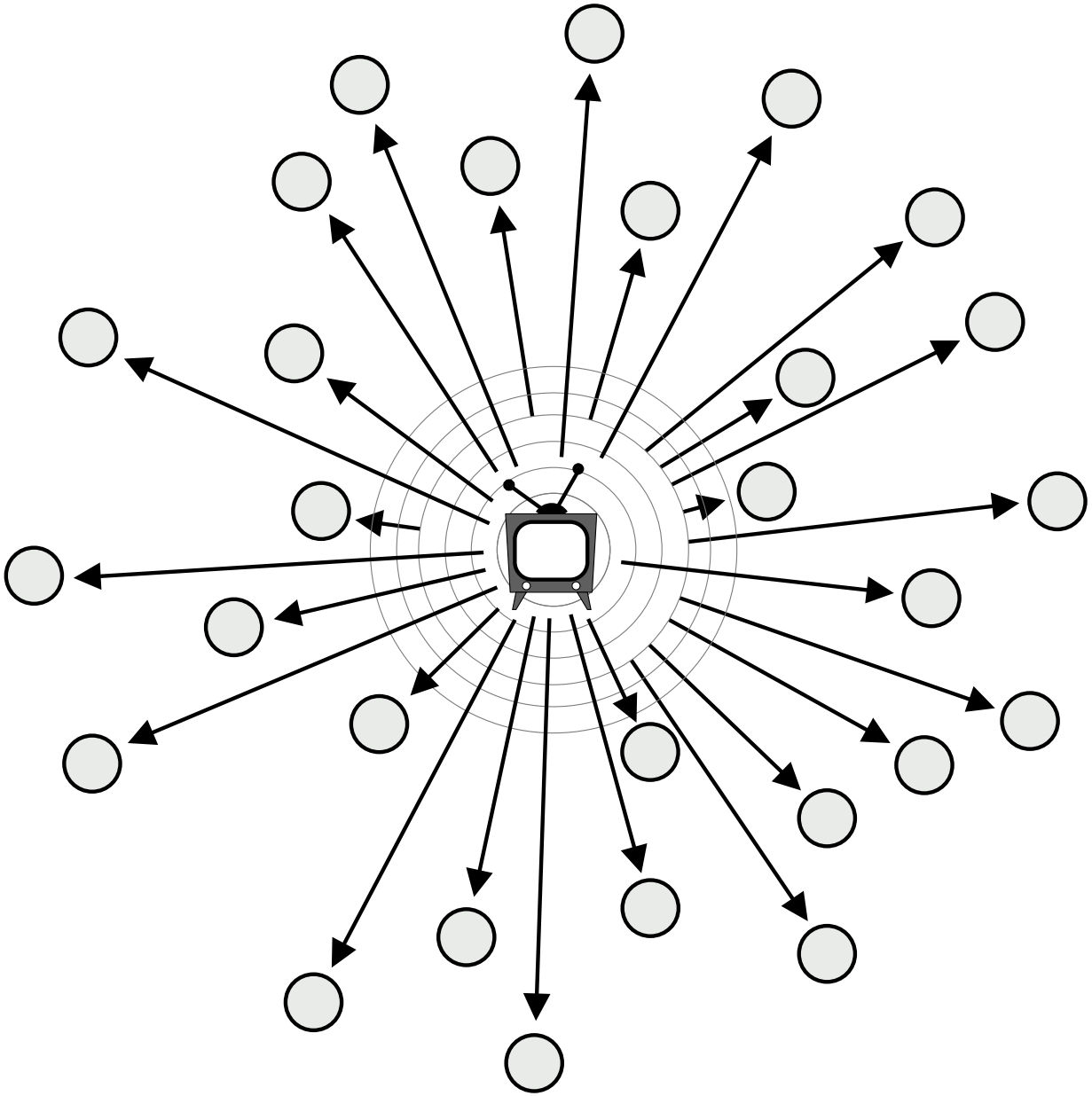
- Who influences whom? Very hard to measure...
- What kinds of influence response functions are there?
- Are some individuals super influencers? Highly popularized by Gladwell [5] as ‘connectors’
- The infectious idea of opinion leaders (Katz and Lazarsfeld) [8]

One perspective

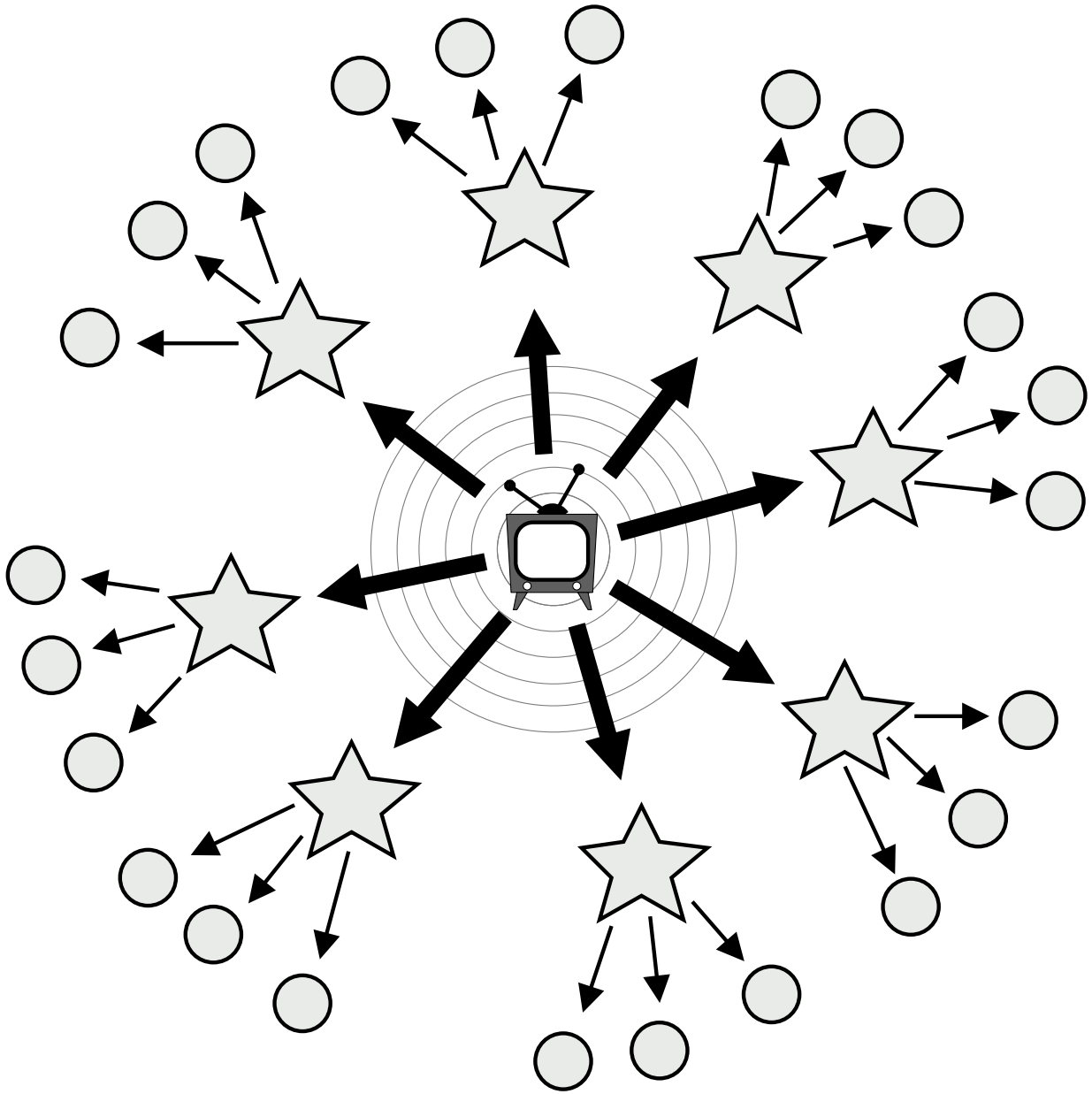
"In historical events great men—so-called—are but labels serving to give a name to the event, and like labels they have the least possible connection with the event itself. Every action of theirs, that seems to them an act of their own free will, is in an historical sense not free at all, but in bondage to the whole course of previous history, and predestined from all eternity."

—Leo Tolstoy, War and Peace.

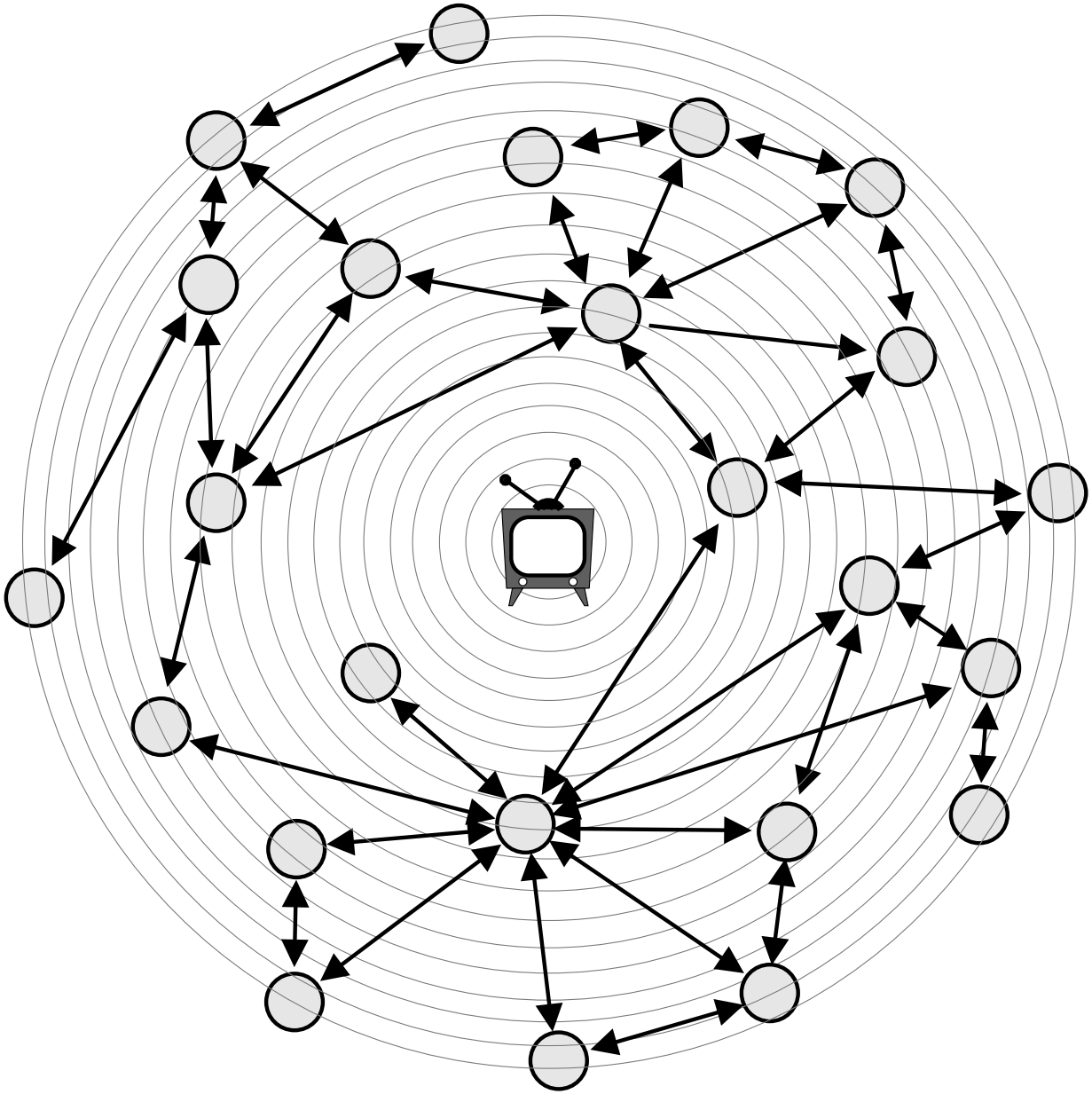
The hypodermic model of influence



The two step model of influence [8]



The general model of influence



Social Contagion

Why do things spread?

- Because of system level properties?
- Or properties of special individuals?
- Is the match that lights the fire important?
- Yes. But only because we are narrative-making machines...
- We like to think things happened for reasons...
- System/group properties harder to understand
- Always good to examine what is said before and after the fact...

Social Contagion

Some important models

- Tipping models—Schelling (1971) [9, 10, 11]
  - Simulation on checker boards.
  - Idea of thresholds.
- Threshold models—Granovetter (1978) [7]
- Herding models—Bikhchandani et al. (1992) [1, 2]
  - Social learning theory, Informational cascades,...

Social contagion models

Thresholds

- Basic idea: individuals adopt a behavior when a certain fraction of others have adopted
- 'Others' may be everyone in a population, an individual's close friends, any reference group.
- Response can be probabilistic or deterministic.
- Individual thresholds can vary
- Assumption: order of others' adoption does not matter... (unrealistic).
- Assumption: level of influence per person is uniform (unrealistic).

Social Contagion

Some possible origins of thresholds:

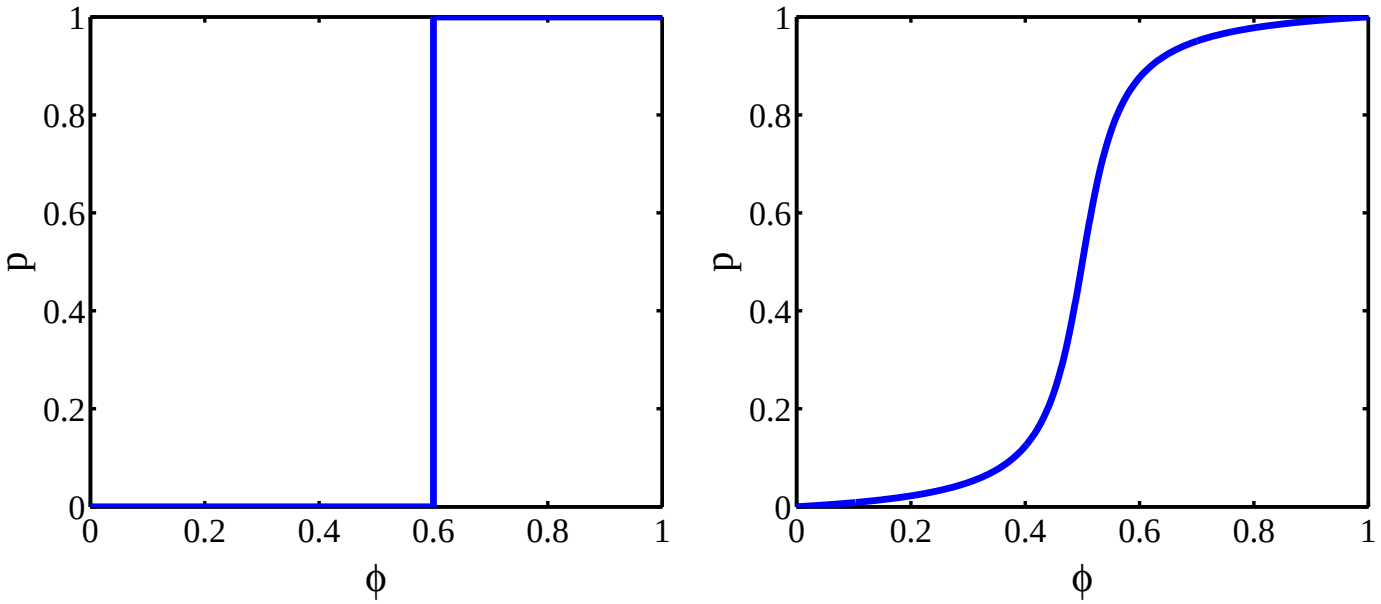
- Desire to coordinate, to conform.
- Lack of information: impute the worth of a good or behavior based on degree of adoption (social proof)
- Economics: Network effects or network externalities
- Externalities = Effects on others not directly involved in a transaction
- Examples: telephones, fax machine, Facebook, operating systems
- An individual's utility increases with the adoption level among peers and the population in general

Social Contagion

Granovetter's Threshold model—definitions

- γ = threshold of an individual.
- $f(\gamma)$ = distribution of thresholds in a population.
- $F(\gamma)$ = cumulative distribution = $\int_{\gamma'=0}^{\gamma} f(\gamma')\mathrm{d}\gamma'$
- ϕ_t = fraction of people 'rioting' at time step t.

Social Sciences—Threshold models



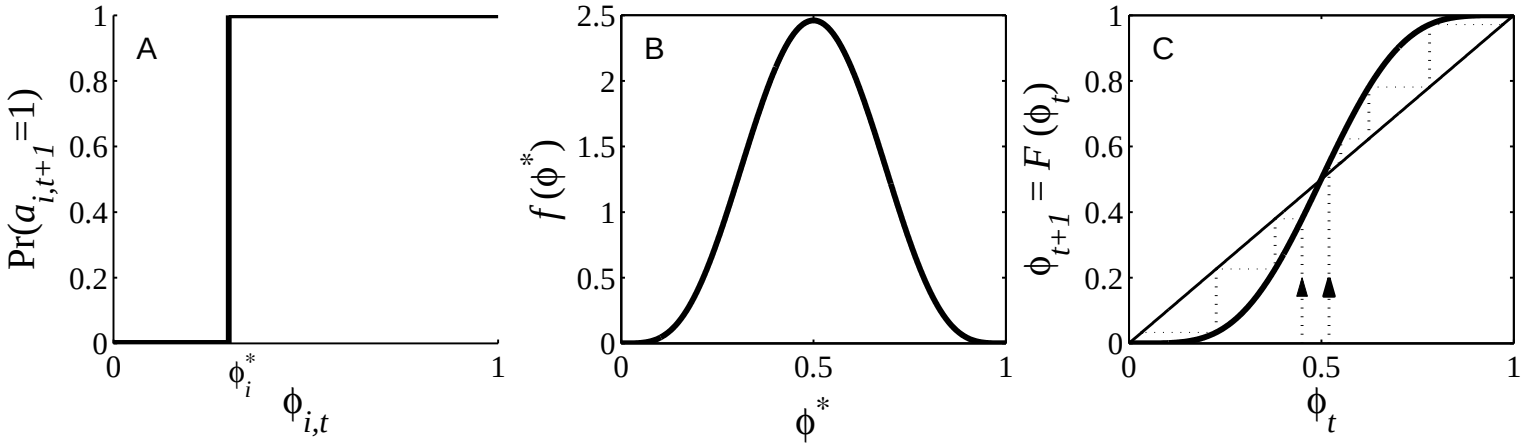
- Example threshold influence response functions: deterministic and stochastic
- ϕ = fraction of contacts 'on' (e.g., rioting)
- Two states: S and I.

Social Sciences: Threshold models

- At time $t+1$, fraction rioting = fraction with $\gamma \leq \phi_t$.
- $$\phi_{t+1} = \int_0^{\phi_t} f(\gamma)\mathrm{d}\gamma = F(\gamma)|_0^{\phi_t} = F(\phi_t)$$
- $\Rightarrow$ Iterative maps of the unit interval $[0, 1]$.

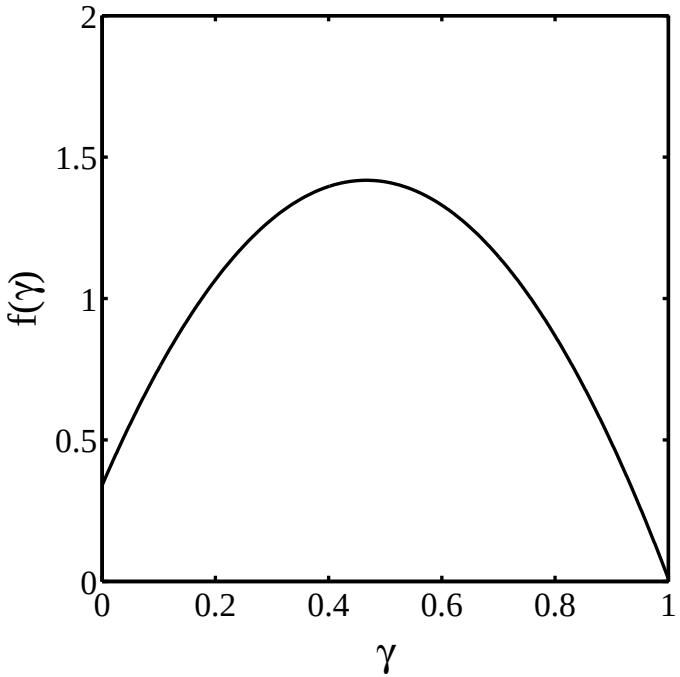
Social Sciences—Threshold models

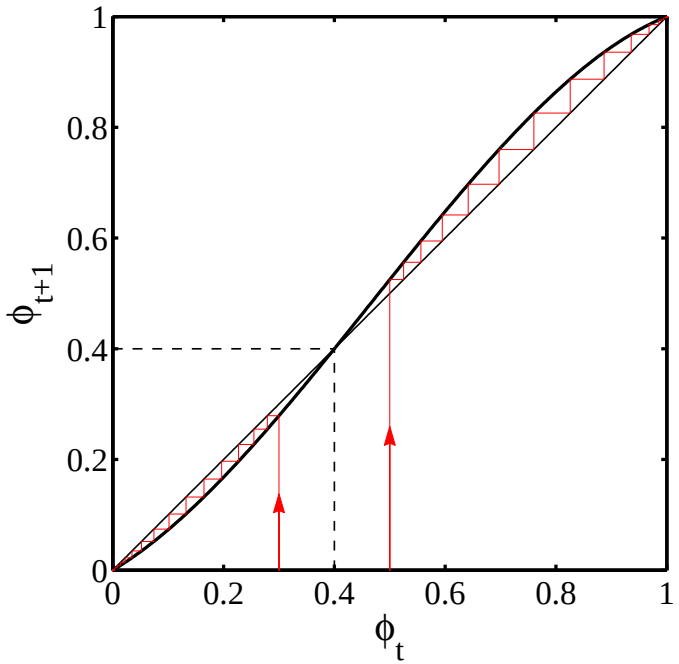
Action based on perceived behavior of others.



- Two states: S and I.
- ϕ = fraction of contacts 'on' (e.g., rioting)
- Discrete time, synchronous update (strong assumption!)
- This is a Critical mass model

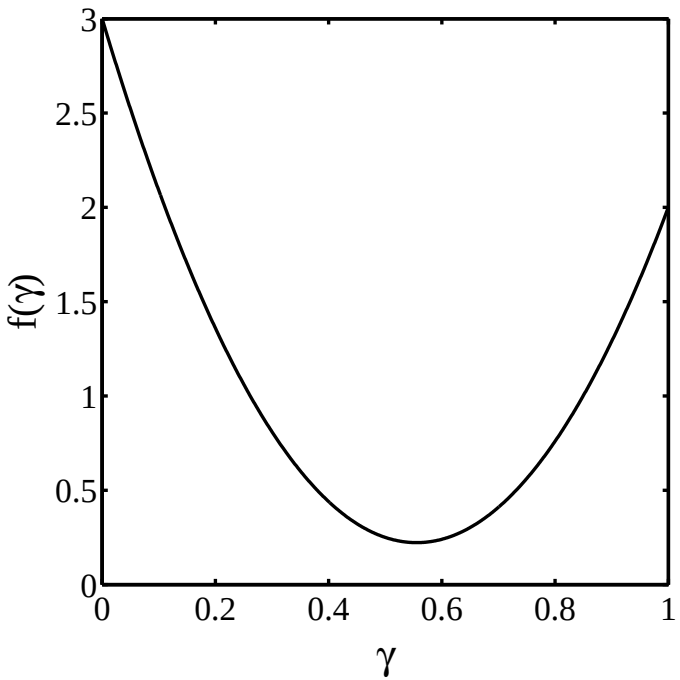
Social Sciences: Threshold models

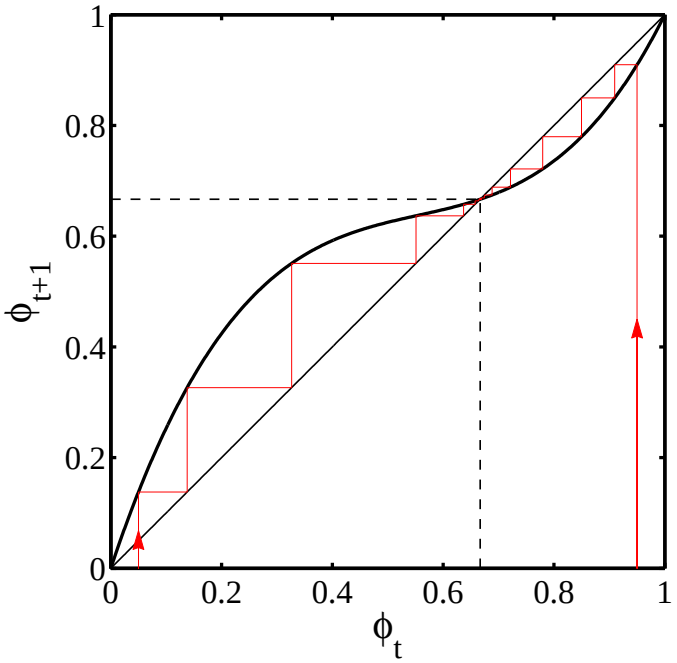




- Critical mass model

Social Sciences: Threshold models





- Example of single stable state model

Social Sciences—Threshold models

Implications for collective action theory:

1. Collective uniformity $\not\Rightarrow$ individual uniformity
2. Small individual changes $\Rightarrow$ large global changes

Threshold model on a network

Many years after Granovetter and Soong's work:

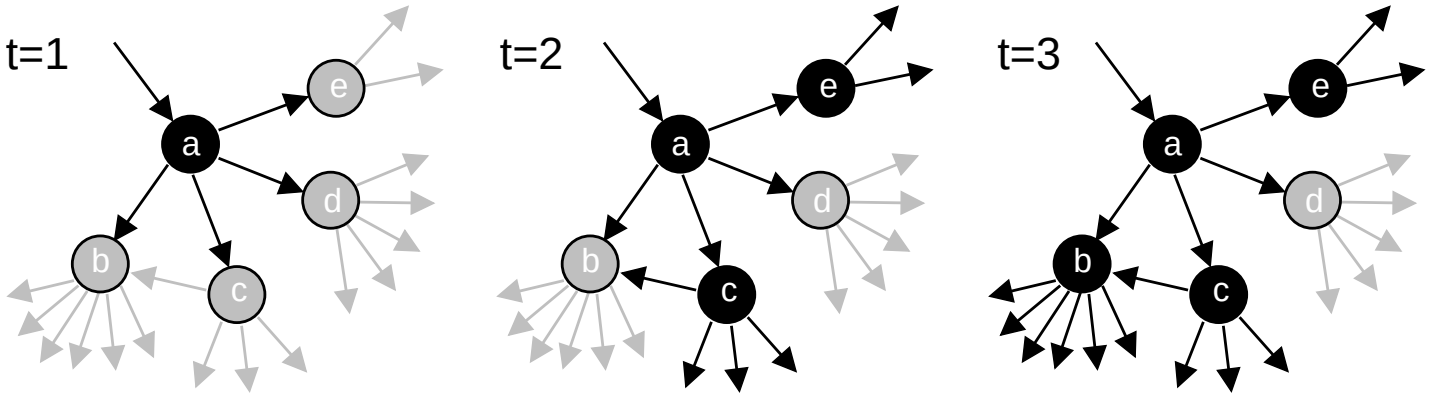
"A simple model of global cascades on random networks"
D. J. Watts. Proc. Natl. Acad. Sci., 2002 [13]

- Mean field model $\longrightarrow$ network model
- Individuals now have a limited view of the world

Threshold model on a network

- Interactions between individuals now represented by a network
- Network is sparse
- Individual i has k_i contacts
- Influence on each link is reciprocal and of unit weight
- Each individual i has a fixed threshold ϕ_i
- Individuals repeatedly poll contacts on network
- Synchronous, discrete time updating
- Individual i becomes active when fraction of active contacts $a_i \geq \phi_i k_i$

Threshold model on a network



- All nodes have threshold $\phi = 0.2$.

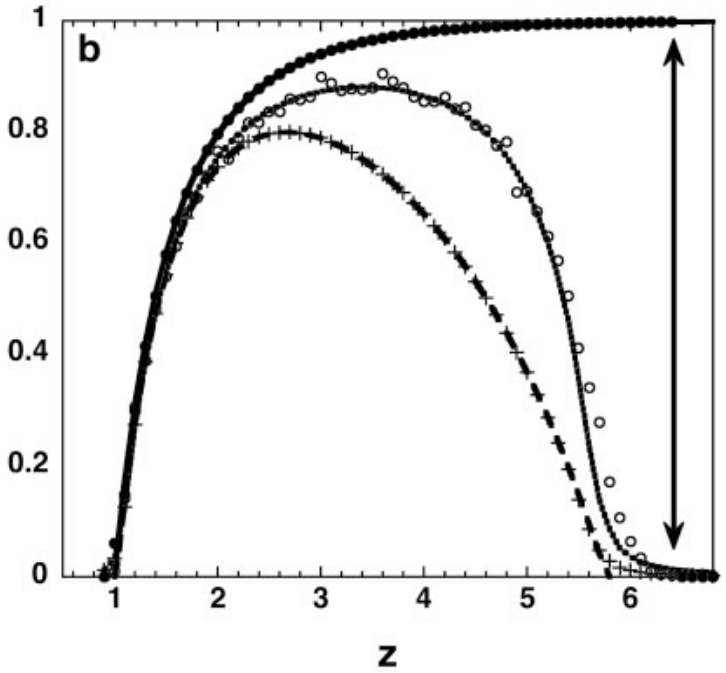
The most gullible

Vulnerables:

- Recall definition: individuals who can be activated by just one contact being active are vulnerables.
- The vulnerability condition for node i: $1/k_i \geq \phi_i$.
- Means # contacts $k_i \leq \lfloor 1/\phi_i \rfloor$.
- Key: For global cascades on random networks, must have a *global component of vulnerables* [13]
- For a uniform threshold ϕ, our contagion condition tells us when such a component exists:

$$r = \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} \frac{(k-1)kP_k}{\langle k \rangle} > 1.$$

Cascades on random networks

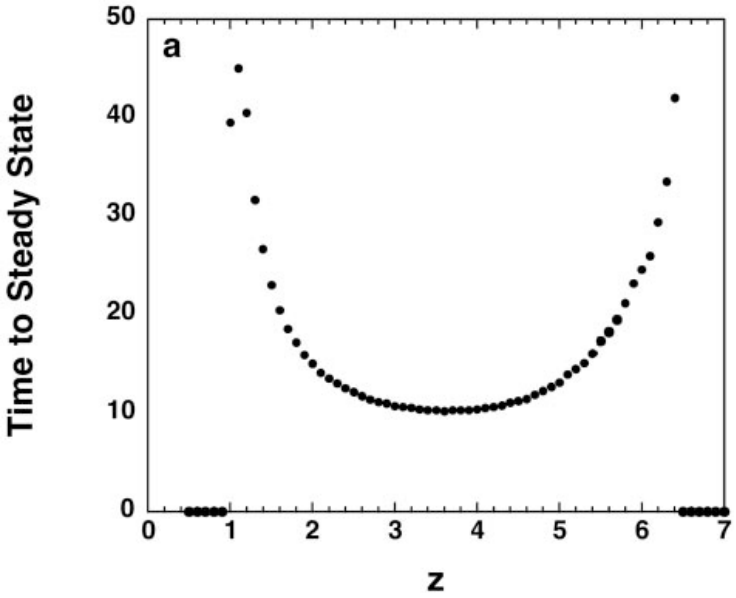


- Top curve: final fraction infected if successful.
- Middle curve: chance of starting a global spreading event (cascade).
- Bottom curve: fractional size of vulnerable subcomponent. [13]

(n.b., $z = \langle k \rangle$)

- Cascades occur only if size of vulnerable subcomponent > 0.
- System is robust-yet-fragile just below upper boundary [3, 4, 12]
- 'Ignorance' facilitates spreading.

Cascades on random networks

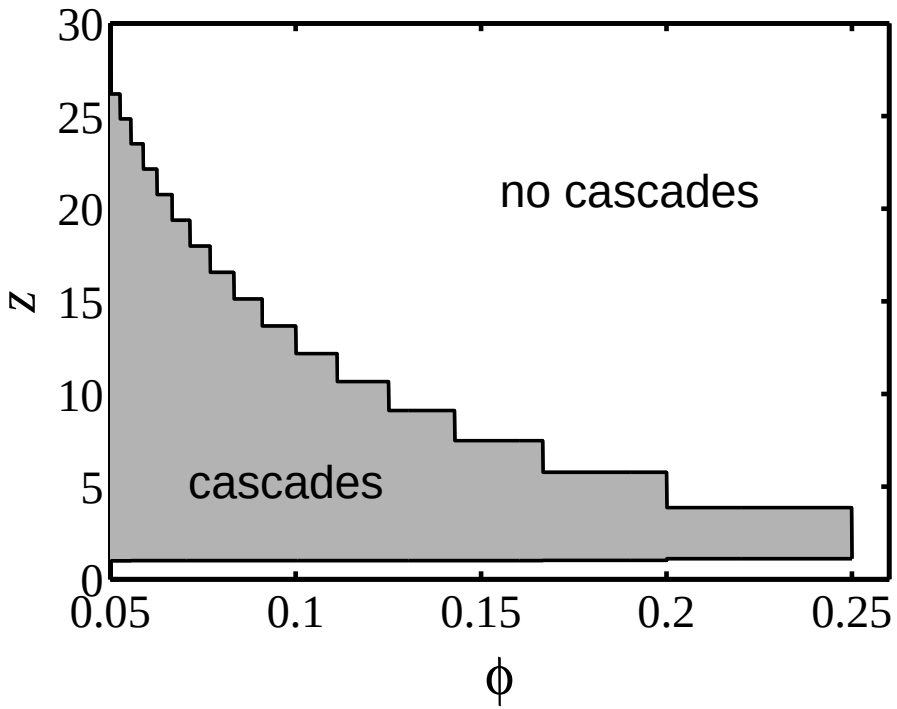


- Time taken for cascade to spread through network. [13]
- Two phase transitions.

(n.b., $z = \langle k \rangle$)

- Largest vulnerable component = critical mass.
- Now have endogenous mechanism for spreading from an individual to the critical mass and then beyond.

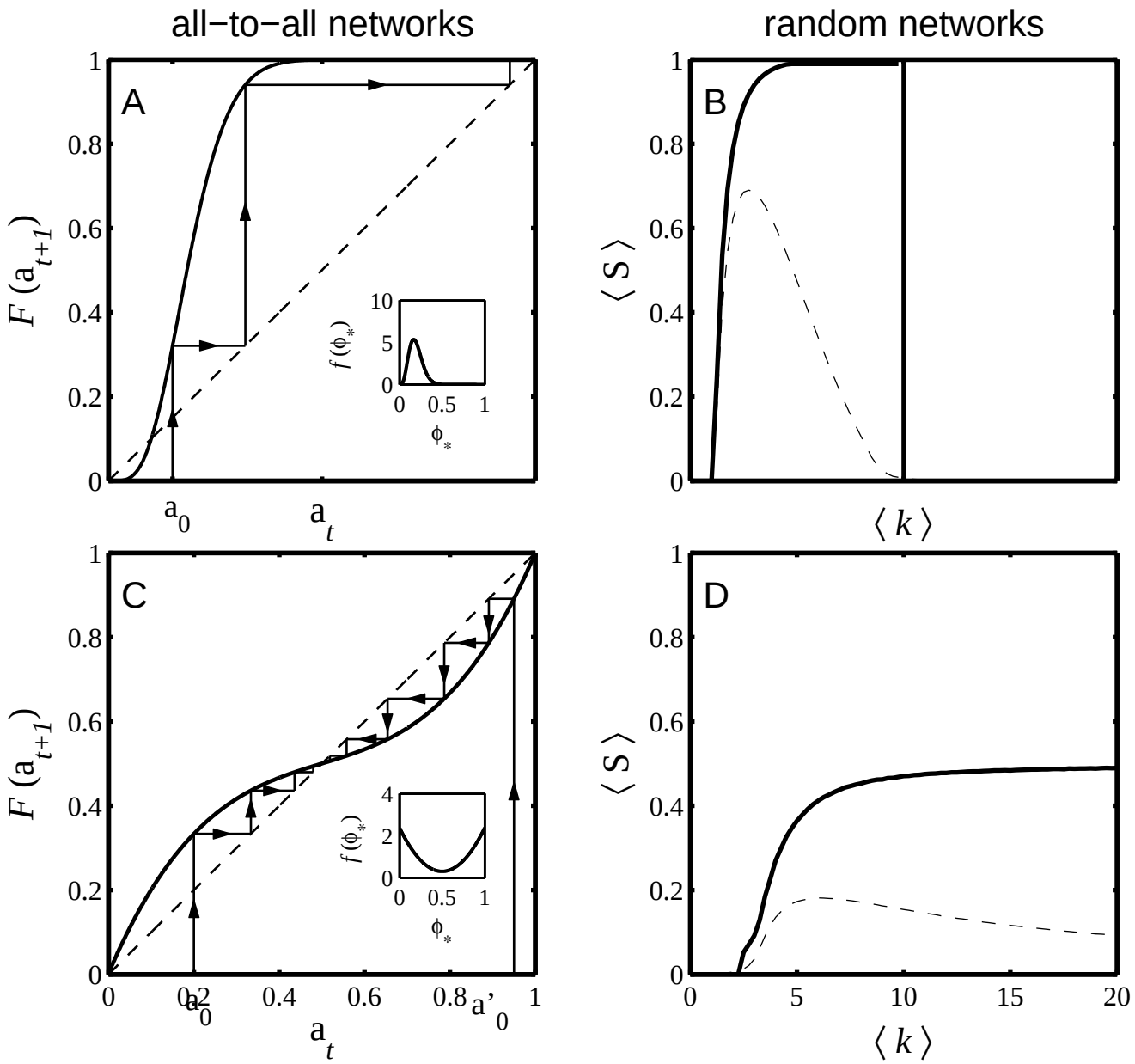
Cascade window for random networks



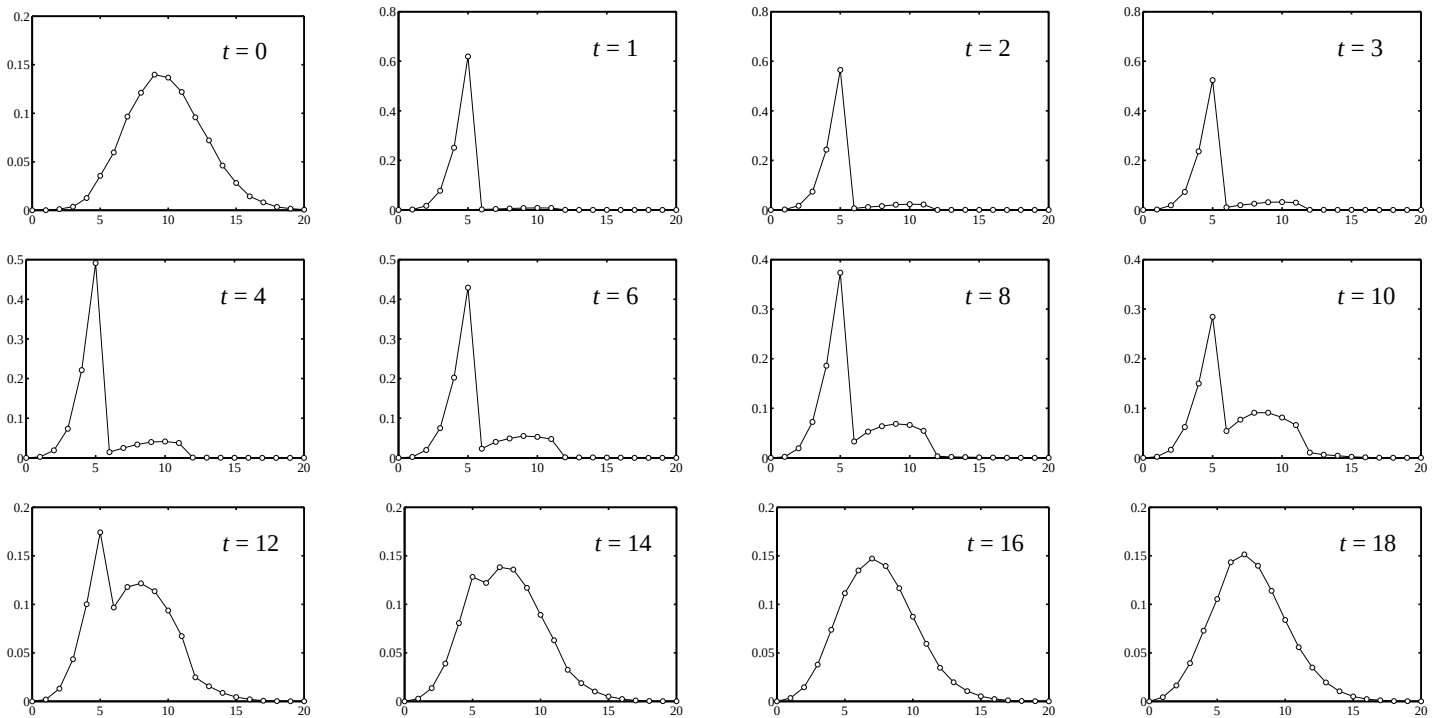
(n.b., $z = \langle k \rangle$)

- Outline of cascade window for random networks.

All-to-all versus random networks

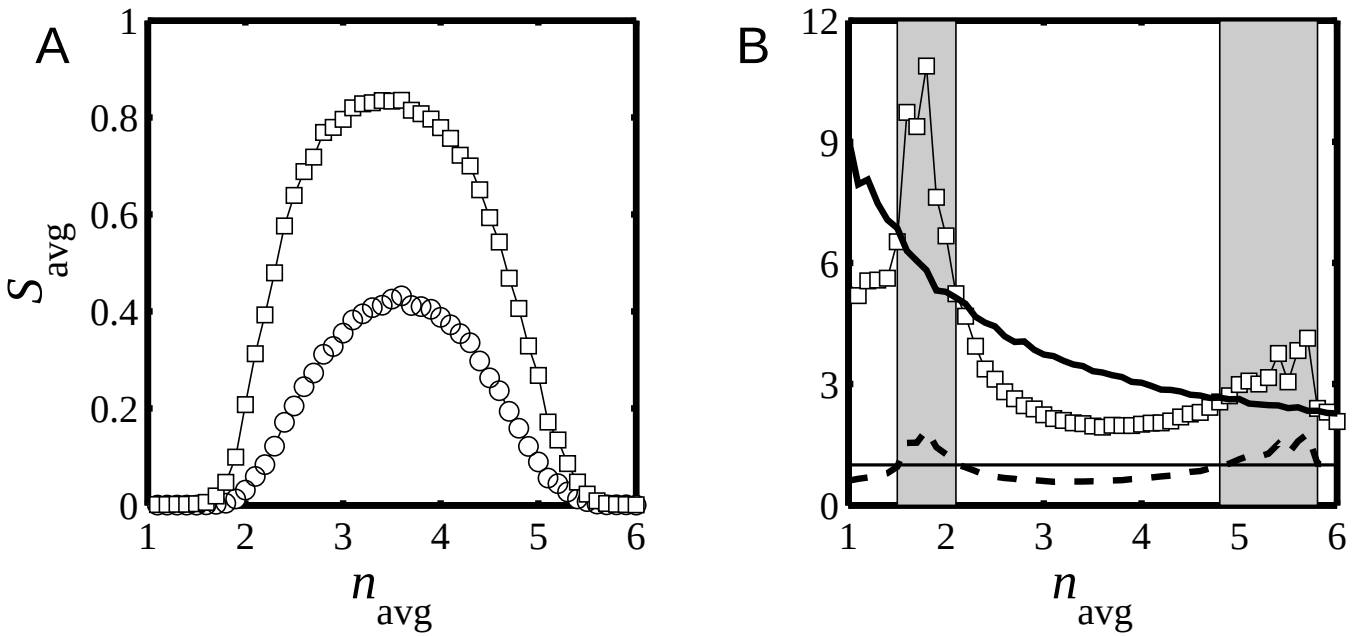


Early adopters—degree distributions



$P_{k,t}$ versus k

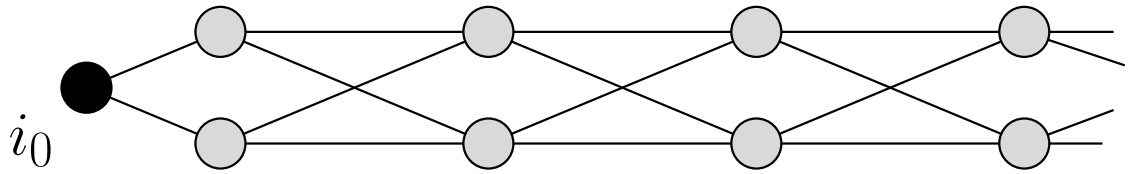
The multiplier effect



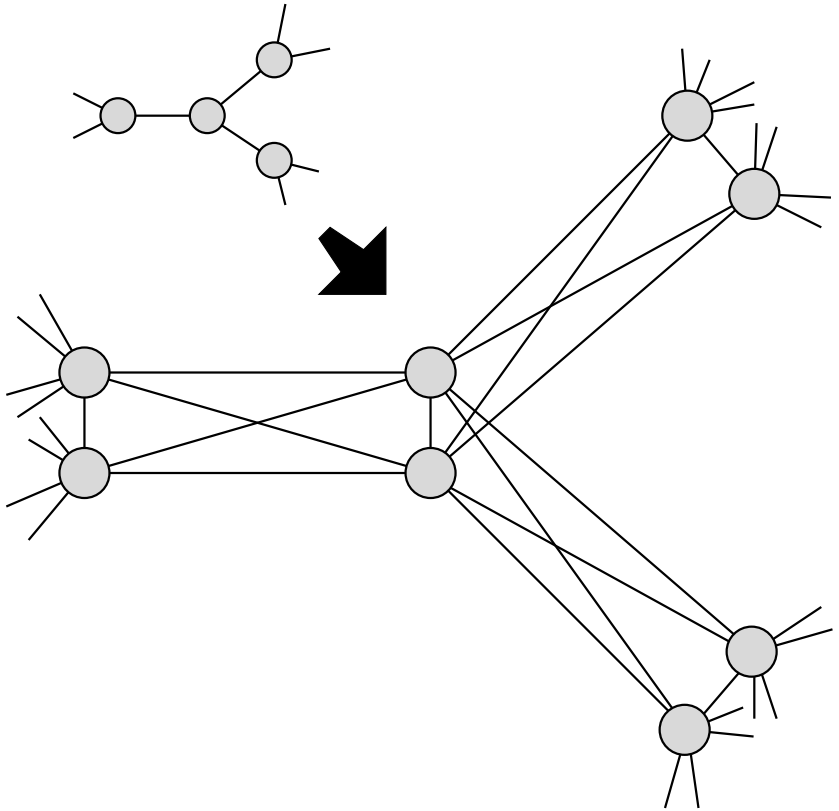
Gamma distributed degrees (skewed)

Special subnetworks can act as triggers

A



B



- $\phi = 1/3$ for all nodes

Threshold contagion on random networks

- Three pieces (among many) to describe analytically:
  1. The fractional size of the largest subcomponent of vulnerable nodes.
  2. The chance of starting a global spreading event (or cascade)
  3. The final size of any succesful spread.

Threshold contagion on random networks

- First goal: Find the largest component of vulnerable nodes.
- Recall that for finding the giant component's size, we had to solve:

$$F_\pi(x) = xF_P\left(F_\rho(x)\right) \text{ and } F_\rho(x) = xF_R\left(F_\rho(x)\right)$$

- We'll find a similar result for the subset of nodes that are vulnerable.
- This is a node-based percolation problem.
- For a general threshold distribution $f(\phi)$, a degree k node is vulnerable with probability

$$\beta_k = \int_0^{1/k} f(\phi)\mathrm{d}\phi\,.$$

Threshold contagion on random networks

- Everything now revolves around the modified generating function:

$$F_P^{(v)}(x) = \sum_{k=0}^{\infty} \beta_k P_k x^k.$$

- Generating function for friends-of-friends distribution is related in same way as before:

$$F_R^{(v)}(x) = \frac{F'^{(v)}_P(x)}{F'^{(v)}_P(1)}.$$

Threshold contagion on random networks

- Functional relations for component size g.f.'s are almost the same...

$$F_\pi^{(v)}(x) = \underbrace{1 - F_P^{(v)}(1)}_{\substack{\text{central node} \\ \text{is not} \\ \text{vulnerable}}} + xF_P^{(v)}\left(F_\rho^{(v)}(x)\right)$$

$$F_\rho^{(v)}(x) = \underbrace{1 - F_R^{(v)}(1)}_{\substack{\text{first node} \\ \text{is not} \\ \text{vulnerable}}} + xF_R^{(v)}\left(F_\rho^{(v)}(x)\right)$$

- Can now solve as before to find $S_1^{(v)} = 1 - F_\pi^{(v)}(1)$.

Threshold contagion on random networks

- Second goal: Find probability of triggering largest vulnerable component.
- Assumption is first node is randomly chosen.
- Same set up as for vulnerable component except now we don't care if the initial node is vulnerable or not:

$$F_{\pi}^{(v)}(x) = xF_P\left(F_{\rho}^{(v)}(x)\right)$$

$$F_{\rho}^{(v)}(x) = 1 - F_R^{(v)}(1) + xF_R^{(v)}\left(F_{\rho}^{(v)}(x)\right)$$

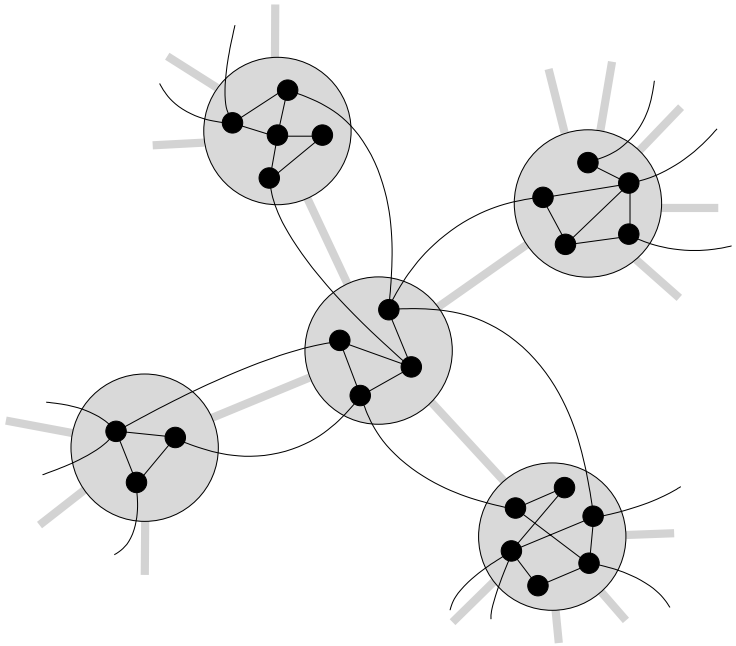
Threshold contagion on random networks

- Third goal: Find expected fractional size of spread.
- Not easy even for uniform threshold problem.
- Difficulty is in figuring out if and when nodes that need ≥ 2 hits switch on.
- See recent progress by Gleeson and Cahalane [6] for variable seed size on random networks.

Extensions

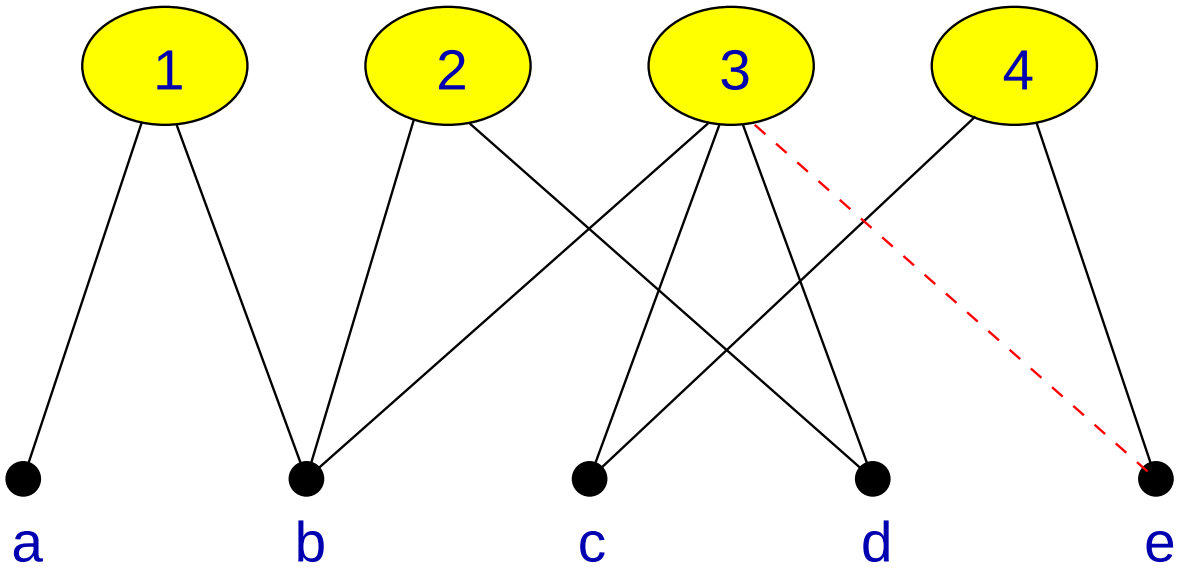
- Assumption of sparse interactions is good
- Degree distribution is (generally) key to a network's function
- Still, random networks don't represent all networks
- Major element missing: group structure

Group structure—Ramified random networks



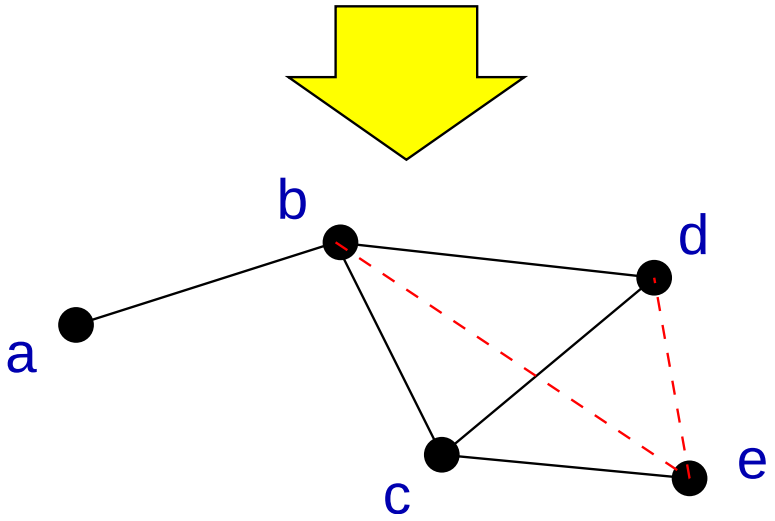
p = intergroup connection probability
q = intragroup connection probability.

Bipartite networks



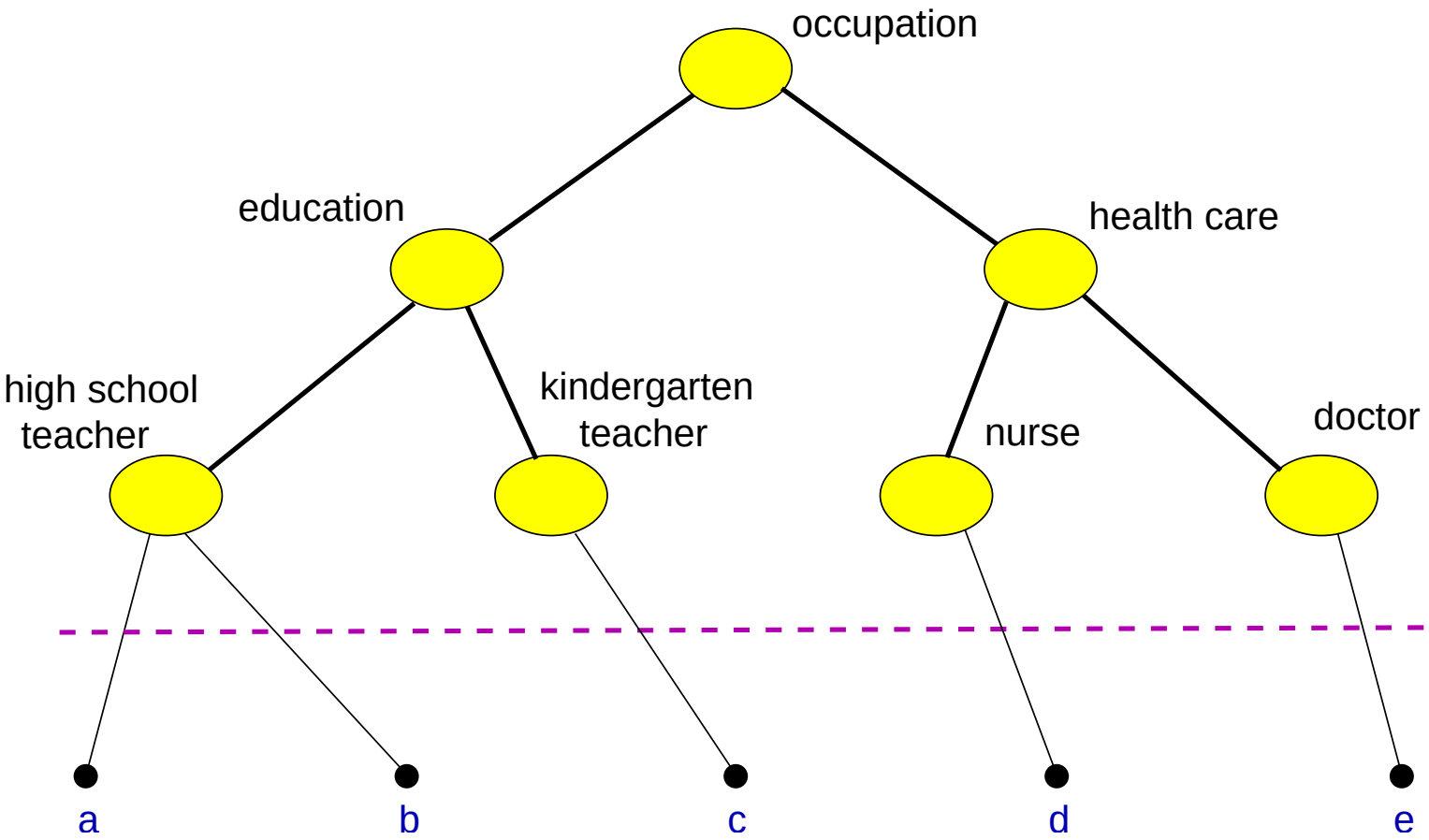
contexts

individuals

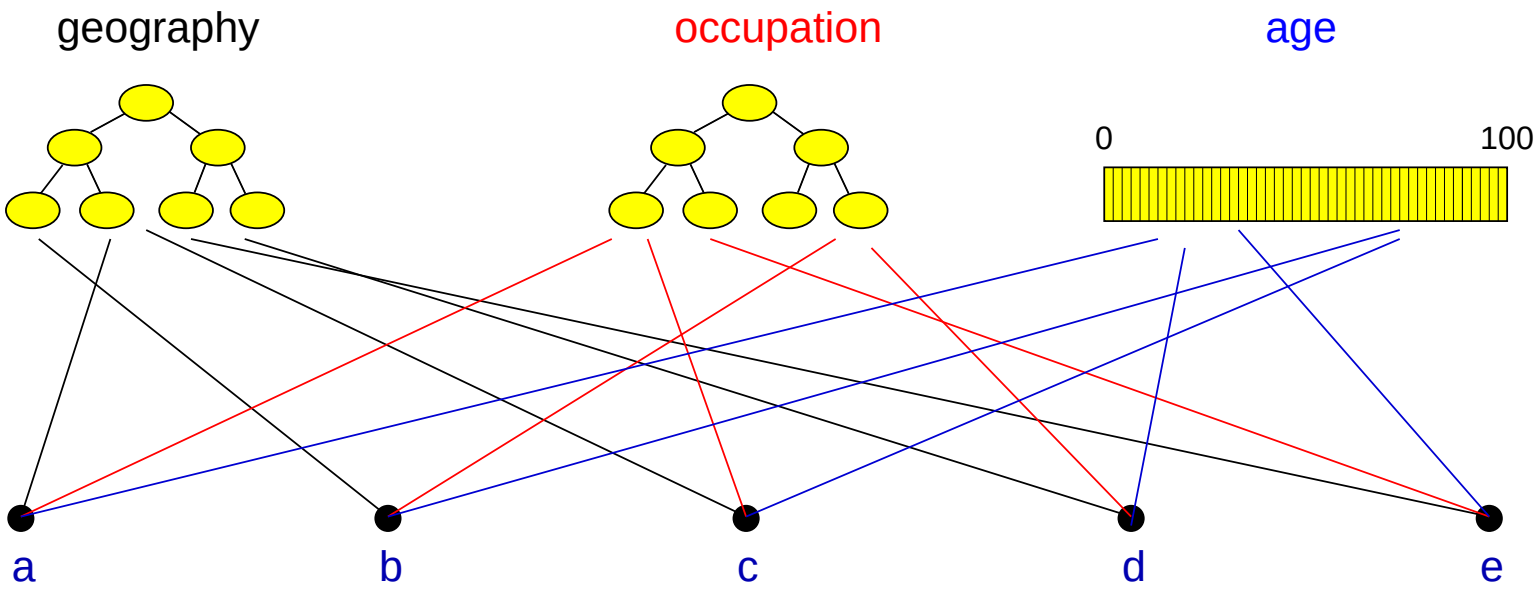


unipartite network

Context distance



Generalized affiliation model

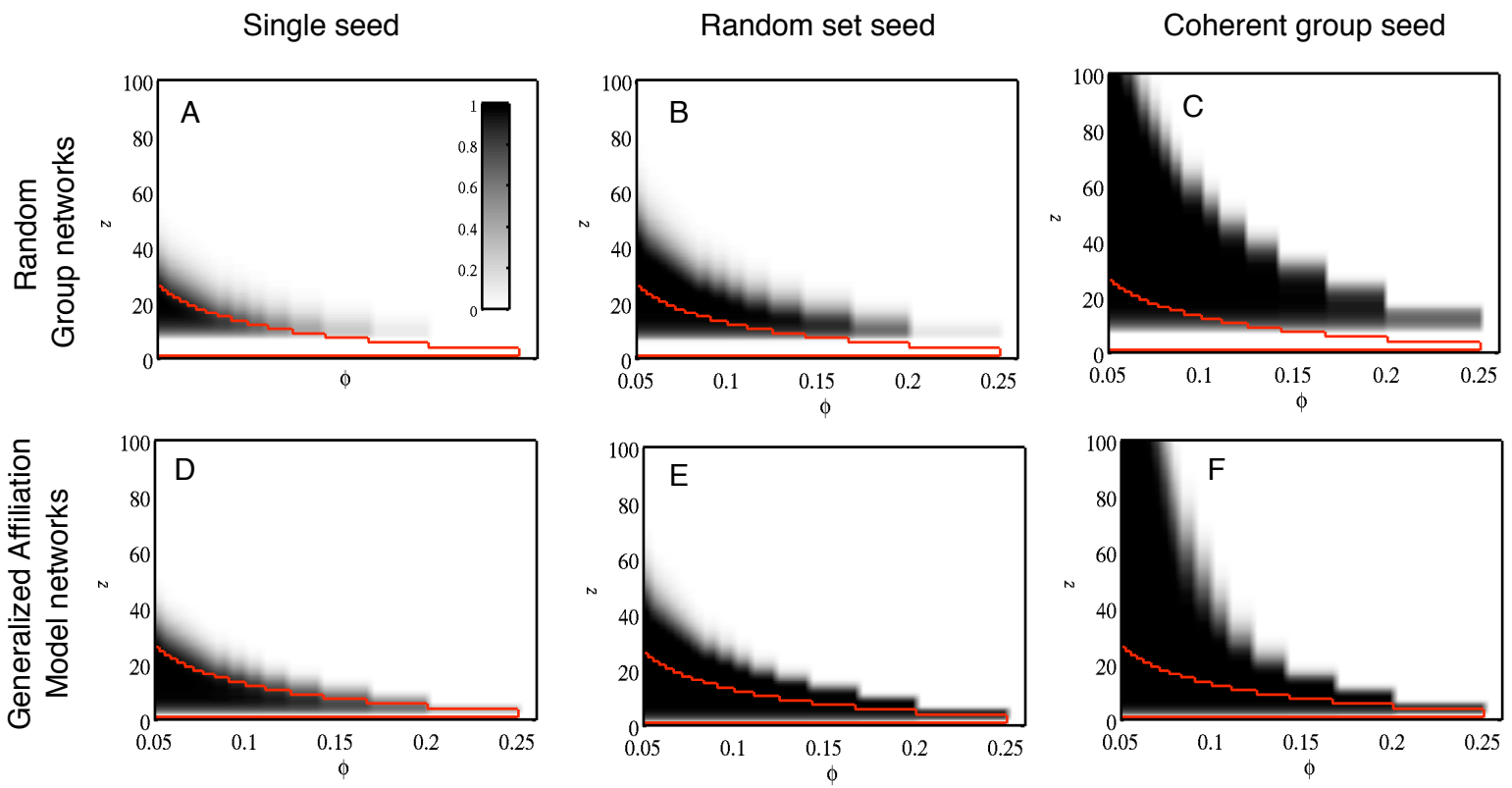


(Blau & Schwartz, Simmel, Breiger)

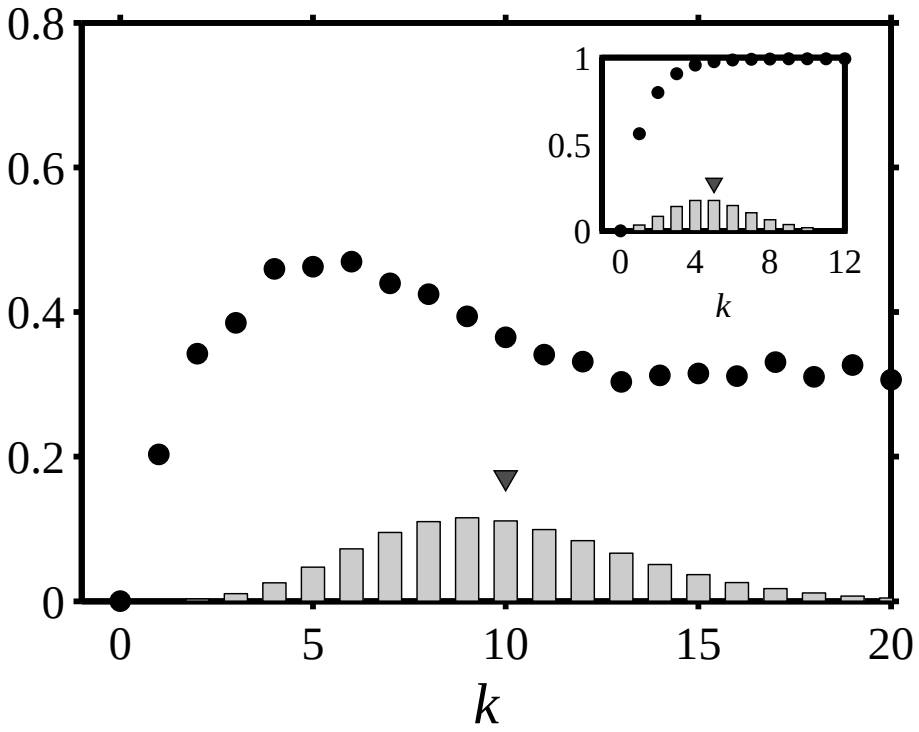
Generalized affiliation model networks with triadic closure

- Connect nodes with probability $\propto \exp^{-\alpha d}$
  where
  α = homophily parameter
  and
  d = distance between nodes (height of lowest common ancestor)
- τ_1 = intergroup probability of friend-of-friend connection
- τ_2 = intragroup probability of friend-of-friend connection

Cascade windows for group-based networks



Assortativity in group-based networks



- Very surprising: the most connected nodes aren't always the most influential
- Assortativity is the reason

Social contagion

Summary

- 'Influential vulnerables' are key to spread.
- Early adopters are mostly vulnerables.
- Vulnerable nodes important but not necessary.
- Groups may greatly facilitate spread.
- Seems that cascade condition is a global one.
- Most extreme/unexpected cascades occur in highly connected networks
- 'Influentials' are posterior constructs.
- Many potential influentials exist.

Social contagion

Implications

- Focus on the influential vulnerables.
- Create entities that can be transmitted successfully through many individuals rather than broadcast from one ‘influential.’
- Only simple ideas can spread by word-of-mouth.
  (Idea of opinion leaders spreads well...)
- Want enough individuals who will adopt and display.
- Displaying can be passive = free (yo-yo’s, fashion), or active = harder to achieve (political messages).
- Entities can be novel or designed to combine with others, e.g. block another one.

References I


S. Bikhchandani, D. Hirshleifer, and I. Welch.
A theory of fads, fashion, custom, and cultural change as informational cascades.
J. Polit. Econ., 100:992–1026, 1992.


S. Bikhchandani, D. Hirshleifer, and I. Welch.
Learning from the behavior of others: Conformity, fads, and informational cascades.
J. Econ. Perspect., 12(3):151–170, 1998. pdf (⊞)


J. Carlson and J. Doyle.
Highly optimized tolerance: A mechanism for power laws in design systems.
Phys. Rev. Lett., 60(2):1412–1427, 1999. pdf (⊞)

References II

J. Carlson and J. Doyle.
Highly optimized tolerance: Robustness and design in complex systems.
Phys. Rev. Lett., 84(11):2529–2532, 2000. pdf (⊞)

M. Gladwell.
The Tipping Point.
Little, Brown and Company, New York, 2000.

J. P. Gleeson and D. J. Cahalane.
Seed size strongly affects cascades on random networks.
Phys. Rev. E, 75:Article # 056103, 2007. pdf (⊞)

M. Granovetter.
Threshold models of collective behavior.
Am. J. Sociol., 83(6):1420–1443, 1978. pdf (⊞)

References III


E. Katz and P. F. Lazarsfeld.
Personal Influence.
The Free Press, New York, 1955.


T. Schelling.
Dynamic models of segregation.
J. Math. Sociol., 1:143–186, 1971.


T. C. Schelling.
Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities.
J. Conflict Resolut., 17:381–428, 1973.


T. C. Schelling.
Micromotives and Macrobehavior.
Norton, New York, 1978.

References IV

D. Sornette.
Critical Phenomena in Natural Sciences.
Springer-Verlag, Berlin, 2nd edition, 2003.

D. J. Watts.
A simple model of global cascades on random networks.
Proc. Natl. Acad. Sci., 99(9):5766–5771, 2002.
pdf (⊞)