Contagion

Last updated: 2022/08/29, 00:04:32 EDT

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 300, 303, & 394, 2022–2023 | @pocsvox

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

Outline

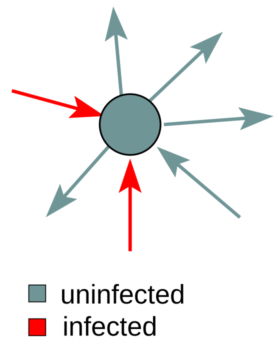
Contagion models

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?
5. What if the seed is one or many nodes?

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

Spreading mechanisms

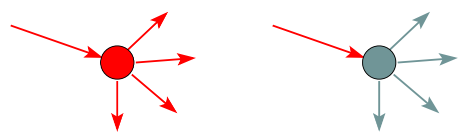


- General spreading mechanism: State of node i depends on history of i and i's neighbors' states.
- Doses of entity may be stochastic and history-dependent.
- May have multiple, interacting entities spreading at once.

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.
- First big question: for a given network and contagion process, can global spreading from a single seed occur?

Global spreading condition

- We need to find: [5]
  R = the average # of infected edges that one random infected edge brings about.
- Call **R** the gain ratio.
- Define B_{k1} as the probability that a node of degree k is infected by a single infected edge.
-

$$\mathbf{R} = \sum_{k=0}^{\infty} \underbrace{\frac{kP_k}{\langle k \rangle}}_{\substack{\text{prob. of}\\ \text{connecting to}\\ \text{a degree } k \text{ node}}} \bullet \underbrace{(k-1)}_{\substack{\#\text{ outgoing}\\ \text{infected}\\ \text{edges}}} \bullet \underbrace{B_{k1}}_{\substack{\text{Prob. of}\\ \text{infection}}}$$

$$+ \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet \underbrace{0}_{\substack{\#\text{ outgoing}\\ \text{infected}\\ \text{edges}}} \bullet \underbrace{(1-B_{k1})}_{\substack{\text{Prob. of}\\ \text{no infection}}}$$

Global spreading condition

- Our global spreading condition is then:

$$\boxed{\mathbf{R} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} > 1.}$$

- Case 1: If $B_{k1} = 1$ then

$$\mathbf{R} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) = \frac{\langle k(k-1) \rangle}{\langle k \rangle} > 1.$$

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

Global spreading condition

- Case 2: If $B_{k1} = \beta < 1$ then

$$\mathbf{R} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet \beta > 1.$$

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

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

Insert question from assignment 9

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

Global spreading condition

- Cases 3, 4, 5, ...: Now allow B_{k1} to depend on k
- Asymmetry: Transmission along an edge depends on node's degree at other end.
- Possibility: B_{k1} increases with k... unlikely.
- Possibility: B_{k1} is not monotonic in k... unlikely.
- Possibility: B_{k1} decreases with k... hmmm.
- $B_{k1} \searrow$ is a plausible representation of a simple kind of social contagion.
- The story:
  More well connected people are harder to influence.

Global spreading condition

- Example: $B_{k1} = 1/k$.
- $$\mathbf{R} = \sum_{k=1}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} = \sum_{k=1}^{\infty} (k-1) \bullet \frac{kP_k}{\langle k \rangle} \bullet \frac{1}{k}$$
  $$= \sum_{k=1}^{\infty} \frac{P_k}{\langle k \rangle} \bullet (k-1) = 1 - \frac{1-P_0}{\langle k \rangle}$$
- Since **R** is always less than 1, no spreading can occur for this mechanism.
- Decay of B_{k1} is too fast.
- Result is independent of degree distribution.

Global spreading condition

- Example: $B_{k1} = H(\frac{1}{k} - \phi)$ 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.
- $$\mathbf{R} = \sum_{k=1}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} = \sum_{k=1}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet H\left(\frac{1}{k} - \phi\right)$$
  $$= \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} (k-1) \bullet \frac{kP_k}{\langle k \rangle}$$ where $\lfloor \cdot \rfloor$ means floor.

Global spreading condition

- The uniform threshold model global spreading condition:
  $$\mathbf{R} = \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} (k-1) \bullet \frac{kP_k}{\langle k \rangle} > 1.$$
- 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.

Virtual contagion: Corrupted Blood, a 2005 virtual plague in World of Warcraft:

Social Contagion

Some important models (recap from CSYS 300)

- Tipping models—Schelling (1971) [11, 12, 13]
  - Simulation on checker boards.
  - Idea of thresholds.
- Threshold models—Granovetter (1978) [8]
- Herding models—Bikhchandani et al. (1992) [1, 2]
  - Social learning theory, Informational cascades,...

Threshold model on a network

Original work:

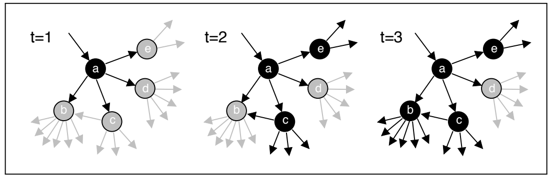
"A simple model of global cascades on random networks"
Duncan J. Watts,
Proc. Natl. Acad. Sci., **99**, 5766–5771, 2002. [15]

- Mean field Granovetter model → 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 number of active contacts $a_i \geq \phi_i k_i$
- Activation is permanent (SI)

Threshold model on a network



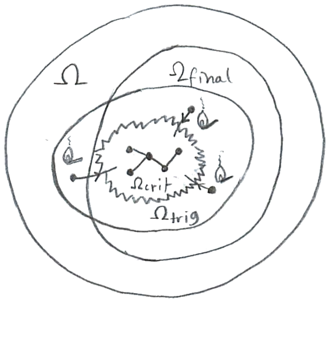
- All nodes have threshold $\phi = 0.2$.

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 spreading events (cascades) on random networks, must have a *global component of vulnerables* [15]
- For a uniform threshold ϕ, our global spreading condition tells us when such a component exists:
  $$\mathbf{R} = \sum_{k=1}^{\lfloor \frac{1}{\phi} \rfloor} \frac{kP_k}{\langle k \rangle} \bullet (k-1) > 1.$$

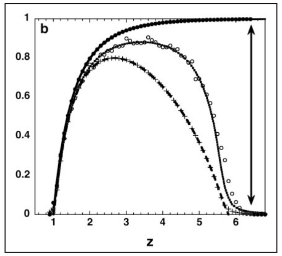
Example random network structure:



- Ω_{crit} = critical mass = global vulnerable component
- Ω_{trig} = triggering component
- Ω_{final} = potential extent of spread
- Ω = entire network

$\Omega_{\text{crit}} \subset \Omega_{\text{trig}}$; $\Omega_{\text{crit}} \subset \Omega_{\text{final}}$; and $\Omega_{\text{trig}}, \Omega_{\text{final}} \subset \Omega$.

Global spreading events on random networks [15]

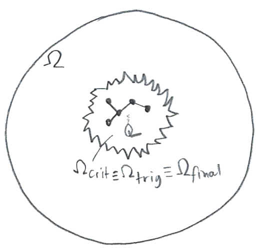


$z = \langle k \rangle$

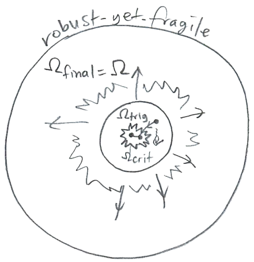
- Top curve: final fraction infected if successful.
- Middle curve: chance of starting a global spreading event (cascade).
- Bottom curve: fractional size of vulnerable subcomponent. [15]

- Global spreading events occur only if size of vulnerable subcomponent > 0.
- System is robust-yet-fragile just below upper boundary [3, 4, 14]
- 'Ignorance' facilitates spreading.

Cascades on random networks

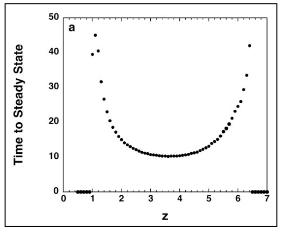


- Above lower phase transition



- Just below upper phase transition

Cascades on random networks

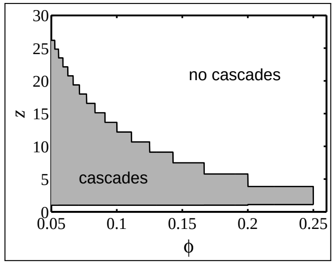


- Time taken for cascade to spread through network. [15]
- Two phase transitions.

(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.

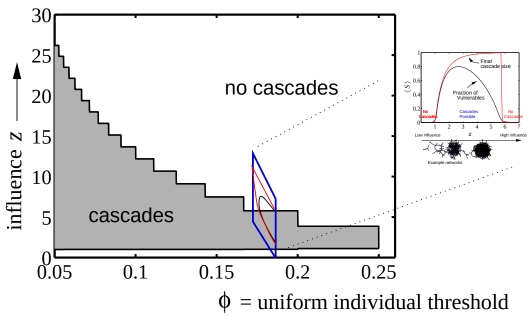
Cascade window for random networks



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

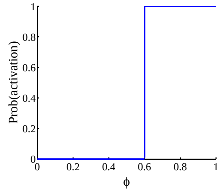
- Outline of cascade window for random networks.

Cascade window for random networks



Social Contagion

Granovetter's Threshold model—recap



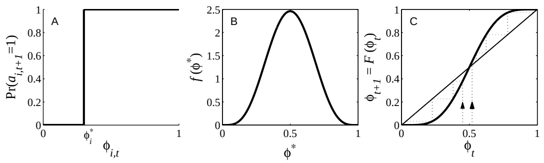
- Assumes deterministic response functions
- ϕ_* = threshold of an individual.
- $f(\phi_*)$ = distribution of thresholds in a population.
- $F(\phi_*)$ = cumulative distribution = $\int_{\phi_*'=0}^{\phi_*} f(\phi_*')\mathrm{d}\phi_*'$
- ϕ_t = fraction of people 'rioting' at time step t.

Social Sciences—Threshold models

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

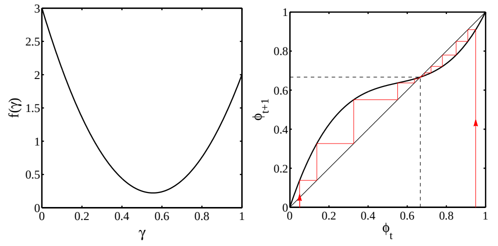
Social Sciences—Threshold models

Action based on perceived behavior of others.



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

Social Sciences—Threshold models



- Example of single stable state model

Social Sciences—Threshold models

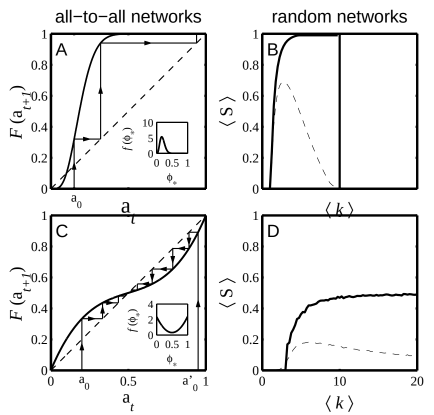
Implications for collective action theory:

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

Next:

- Connect mean-field model to network model.
- Single seed for network model: $1/N \to 0$.
- Comparison between network and mean-field model sensible for vanishing seed size for the latter.

All-to-all versus random networks

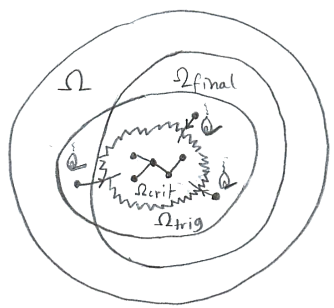


Threshold contagion on random networks

Three key pieces to describe analytically:

1. The fractional size of the largest subcomponent of vulnerable nodes, S_{vuln}.
2. The chance of starting a global spreading event, $P_{\text{trig}} = S_{\text{trig}}$.
3. The expected final size of any successful spread, S.
   - n.b., the distribution of S is almost always bimodal.

Example random network structure:



- $\Omega_{\text{crit}} = \Omega_{\text{vuln}}$ = critical mass = global vulnerable component
- Ω_{trig} = triggering component
- Ω_{final} = potential extent of spread
- Ω = entire network

$\Omega_{\text{crit}} \subset \Omega_{\text{trig}}$; $\Omega_{\text{crit}} \subset \Omega_{\text{final}}$; and $\Omega_{\text{trig}}, \Omega_{\text{final}} \subset \Omega$.

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 monotonic threshold distribution $f(\phi)$, a degree k node is vulnerable with probability

$$B_{k1} = \int_0^{1/k} f(\phi)\mathrm{d}\phi\,.$$

Threshold contagion on random networks

- We now have a generating function for the probability that a randomly chosen node is vulnerable and has degree k:

$$F_P^{(\text{vuln})}(x) = \sum_{k=0}^{\infty} P_k B_{k1} x^k.$$

- The generating function for friends-of-friends distribution is similar to before:

$$F_R^{(\text{vuln})}(x) = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} B_{k1} x^{k-1}$$

$$= \frac{\frac{\mathrm{d}}{\mathrm{d}x} F_P^{(\text{vuln})}(x)}{\frac{\mathrm{d}}{\mathrm{d}x} F_P(x)|_{x=1}} = \frac{\frac{\mathrm{d}}{\mathrm{d}x} F_P^{(\text{vuln})}(x)}{F_R(1)}$$

- Detail: We still have the underlying degree distribution involved in the denominator.

Threshold contagion on random networks

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

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

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

- Can now solve as before to find

$$S_{\text{vuln}} = 1 - F_\pi^{(\text{vuln})}(1).$$

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:

$$F_\pi^{(\text{trig})}(x) = xF_P\left(F_\rho^{(\text{vuln})}(x)\right)$$

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

- Solve as before to find $P_{\text{trig}} = S_{\text{trig}} = 1 - F_\pi^{(\text{trig})}(1)$.

Physical derivation of possibility and probability of global spreading:

- Possibility: binary indicator of phase. Global spreading events are either possible or can never happen.
- For random networks, global spreading possibility is understood as meaning a giant component of vulnerable nodes exists.
- Next: what's the probability that a randomly infected node will cause a global spreading event?
- Call this P_{trig}.
- As usual, it's all about edges and we need to first determine the probability that an infected edge leads to a global spreading event.
- Call this Q_{trig}.
- Later: Generalize to more complex networks involving assortativity of all kinds.

Probability an infected edge leads to a global spreading event:

- Q_{trig} must satisfying a one-step recursion relation.
- Follow an infected edge and use three pieces:
  1. Probability of reaching a degree k node is $Q_k = \frac{kP_k}{\langle k \rangle}$.
  2. The node reached is vulnerable with probability B_{k1}.
  3. At least one of the node's outgoing edges leads to a global spreading event = 1 - probability no edges do so = $1 - (1 - Q_{\text{trig}})^{k-1}$.
- Put everything together and solve for Q_{trig}:

$$Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[1 - (1 - Q_{\text{trig}})^{k-1}\right].$$

Good things about our equation for Q_{trig}:

$$Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[1 - (1 - Q_{\text{trig}})^{k-1}\right] = f(Q_{\text{trig}}; P_k, B_{k1})$$

- $Q_{\text{trig}} = 0$ is always a solution.
- Spreading occurs if a second solution exists for which $0 < Q_{\text{trig}} \leq 1$.
- Given P_k and B_{k1}, we can use any kind of root finder to solve for Q_{trig}, but ...
- The function f increases monotonically with Q_{trig}.
- We can therefore use an iterative cobwebbing approach to find the solution: $Q_{\text{trig}}^{(n+1)} = f(Q_{\text{trig}}^{(n)}; P_k, B_{k1})$.
- Start with a suitably small seed $Q_{\text{trig}}^{(1)} > 0$ and iterate while rubbing hands together.

- Global spreading is possible if the fractional size S_{vuln} of the largest component of vulnerables is "giant".
- Interpret S_{vuln} as the probability a randomly chosen node is vulnerable and that infecting it leads to a global spreading event:

$$S_{\text{vuln}} = \sum_k P_k \bullet B_{k1} \bullet \left[1 - (1 - Q_{\text{trig}})^k\right] > 0.$$

- Amounts to having $Q_{\text{trig}} > 0$.
- Probability of global spreading differs only in that we don't care if the initial seed is vulnerable or not:

$$P_{\text{trig}} = S_{\text{trig}} = \sum_k P_k \bullet \left[1 - (1 - Q_{\text{trig}})^k\right]$$

- As for S_{vuln}, P_{trig} is non-zero when $Q_{\text{trig}} > 0$.

Connection to generating function results:

- We found that $F_\rho^{(\text{vuln})}(1)$—the probability that a random edge leads to a finite vulnerable component—satisfies

$$F_\rho^{(\text{vuln})}(1) = 1 - F_R^{(\text{vuln})}(1) + 1 \cdot F_R^{(\text{vuln})}\left(F_\rho^{(\text{vuln})}(1)\right).$$

- We set $F_\rho^{(\text{vuln})}(1) = 1 - Q_{\text{trig}}$ and deploy $F_R^{(\text{vuln})}(x) = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} B_{k1} x^{k-1}$ to find

$$1 - Q_{\text{trig}} = 1 - \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} B_{k1} + \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} B_{k1} \left(1 - Q_{\text{trig}}\right)^{k-1}.$$

- Some breathless algebra it all matches:

$$Q_{\text{trig}} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[1 - \left(1 - Q_{\text{trig}}\right)^{k-1}\right].$$

Fractional size of the largest vulnerable component:

- The generating function approach gave $S_{\text{vuln}} = 1 - F_\pi^{(\text{vuln})}(1)$ where

$$F_\pi^{(\text{vuln})}(1) = 1 - F_P^{(\text{vuln})}(1) + 1 \cdot F_P^{(\text{vuln})}\left(F_\rho^{(\text{vuln})}(1)\right).$$

- Again using $F_\rho^{(\text{vuln})}(1) = 1 - Q_{\text{trig}}$ along with $F_P^{(\text{vuln})}(x) = \sum_{k=0}^{\infty} P_k B_{k1} x^k$, we have:

$$1 - S_{\text{vuln}} = 1 - \sum_{k=0}^{\infty} P_k B_{k1} + \sum_{k=0}^{\infty} P_k B_{k1} \left(1 - Q_{\text{trig}}\right)^k.$$

- Excited scrabbling about gives us, as before:

$$S_{\text{vuln}} = \sum_{k=0}^{\infty} P_k B_{k1} \left[1 - \left(1 - Q_{\text{trig}}\right)^k\right].$$

Triggering probability for single-seed global spreading events:

- Slight adjustment to the vulnerable component calculation.
- $S_{\text{trig}} = 1 - F_\pi^{(\text{trig})}(1)$ where

$$F_\pi^{(\text{trig})}(1) = 1 \cdot F_P\left(F_\rho^{(\text{vuln})}(1)\right).$$

- We play these cards: $F_\rho^{(\text{vuln})}(1) = 1 - Q_{\text{trig}}$ and $F_P(x) = \sum_{k=0}^{\infty} P_k x^k$ to arrive at

$$1 - S_{\text{trig}} = 1 + \sum_{k=0}^{\infty} P_k \left(1 - Q_{\text{trig}}\right)^k.$$

- More scruffing around brings happiness:

$$S_{\text{trig}} = \sum_{k=0}^{\infty} P_k \left[1 - \left(1 - Q_{\text{trig}}\right)^k\right].$$

Connection to simple gain ratio argument:

- Earlier, we showed the global spreading condition follows from the gain ratio $\mathbf{R} > 1$:

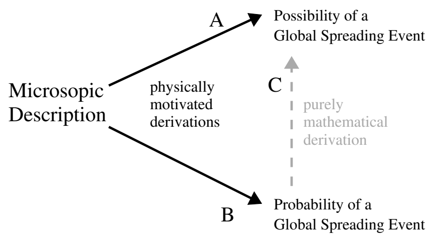
$$\mathbf{R} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} > 1.$$

- We would very much like to see that $\mathbf{R} > 1$ matches up with $Q_{\text{trig}} > 0$.
- It really would be just so totally awesome.
- Must come from our basic edge triggering probability equation:

$$Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[1 - (1 - Q_{\text{trig}})^{k-1}\right].$$

- When does this equation have a solution $0 < Q_{\text{trig}} \leq 1$?
- We need to find out what happens as $Q_{\text{trig}} \to 0$. [9]

What we're doing:



- For $Q_{\text{trig}} \to 0^+$, equation tends towards

$$Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[\cancel{1} + \left(\cancel{1} + (k-1)Q_{\text{trig}} + \ldots\right)\right]$$

$$\Rightarrow Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet (k-1)Q_{\text{trig}}$$

$$\Rightarrow 1 = \sum_k \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1}$$

- Only defines the phase transition points (i.e., **R** $= 1$).
- Inequality?

- Again take $Q_{\text{trig}} \to 0^+$, but keep next higher order term:

$$Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[\cancel{1} + \left(\cancel{1} + (k-1)Q_{\text{trig}} - \binom{k-1}{2} Q_{\text{trig}}^2\right)\right]$$

$$\Rightarrow Q_{\text{trig}} = \sum_k \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[(k-1)Q_{\text{trig}} - \binom{k-1}{2} Q_{\text{trig}}^2\right]$$

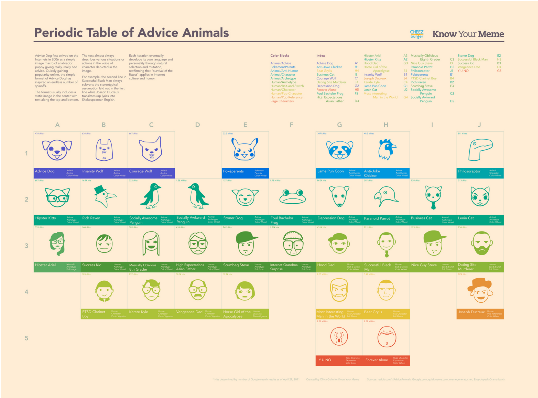
$$\Rightarrow \sum_k \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} = 1 + \sum_k \frac{kP_k}{\langle k \rangle} B_{k1} \binom{k-1}{2} Q_{\text{trig}}$$

- We have $Q_{\text{trig}} > 0$ if $\sum_k \frac{kP_k}{\langle k \rangle} \bullet (k-1) \bullet B_{k1} > 1$.
- Repeat: Above is a mathematical connection between two physically derived equations.
- From this connection, we don't know anything about a gain ratio **R** or how to arrange the pieces.

Threshold contagion on random networks

- Third goal: Find expected fractional size of spread.
- Not obvious even for uniform threshold problem.
- Difficulty is in figuring out if and when nodes that need ≥ 2 hits switch on.
- Problem solved for infinite seed case by Gleeson and Cahalane: "Seed size strongly affects cascades on random networks," Phys. Rev. E, 2007. [7]
- Developed further by Gleeson in "Cascades on correlated and modular random networks," Phys. Rev. E, 2008. [6]

Meme species:



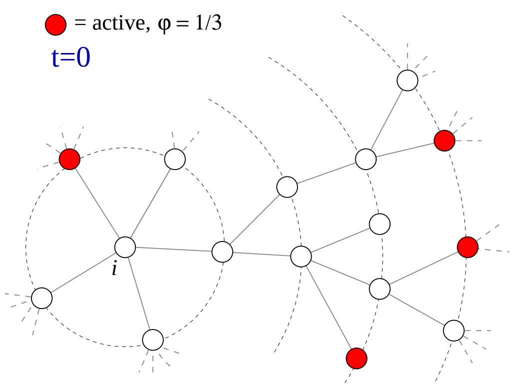
- More here at http://knowyourmeme.com

Expected size of spread

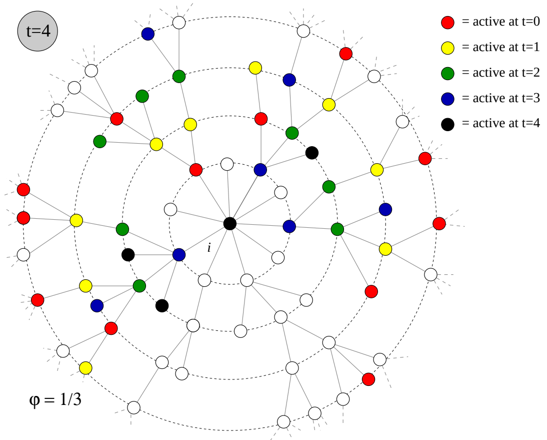
Idea:

- Randomly turn on a fraction ϕ_0 of nodes at time $t = 0$
- Capitalize on local branching network structure of random networks (again)
- Now think about what must happen for a specific node i to become active at time t:
  - $t = 0$: i is one of the seeds (prob = ϕ_0)
  - $t = 1$: i was not a seed but enough of i's friends switched on at time $t = 0$ so that i's threshold is now exceeded.
  - $t = 2$: enough of i's friends and friends-of-friends switched on at time $t = 0$ so that i's threshold is now exceeded.
  - $t = n$: enough nodes within n hops of i switched on at $t = 0$ and their effects have propagated to reach i.

Expected size of spread



Expected size of spread



Expected size of spread

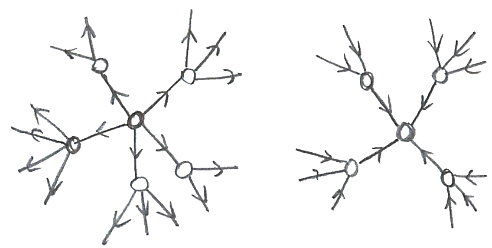
Notes:

- Calculations presume nodes do not become inactive (strong restriction, liftable)
- Not just for threshold model—works for a wide range of contagion processes.
- We can analytically determine the entire time evolution, not just the final size.
- We can in fact determine **Pr**(node of degree k switches on at time t).
- Even more, we can compute: **Pr**(specific node i switches on at time t).
- Asynchronous updating can be handled too.

Expected size of spread

Pleasantness:

- Taking off from a single seed story is about expansion away from a node.
- Extent of spreading story is about contraction at a node.



Expected size of spread

- Notation: $\phi_{k,t}$ = **Pr**(a degree k node is active at time t).
- Notation: B_{kj} = **Pr** (a degree k node becomes active if j neighbors are active).
- Our starting point: $\phi_{k,0} = \phi_0$.
- $\binom{k}{j}\phi_0^j(1-\phi_0)^{k-j}$ = **Pr** (j of a degree k node's neighbors were seeded at time $t=0$).
- Probability a degree k node was a seed at $t=0$ is ϕ_0 (as above).
- Probability a degree k node was not a seed at $t=0$ is $(1-\phi_0)$.
- Combining everything, we have:

$$\phi_{k,1} = \phi_0 + (1-\phi_0)\sum_{j=0}^{k}\binom{k}{j}\phi_0^j(1-\phi_0)^{k-j}B_{kj}.$$

Expected size of spread

- For general t, we need to know the probability an edge coming into a degree k node at time t is active.
- Notation: call this probability θ_t.
- We already know $\theta_0 = \phi_0$.
- Story analogous to $t=1$ case. For specific node i:

$$\phi_{i,t+1} = \phi_0 + (1-\phi_0)\sum_{j=0}^{k_i}\binom{k_i}{j}\theta_t^j(1-\theta_t)^{k_i-j}B_{k_ij}.$$

- Average over all nodes with degree k to obtain expression for ϕ_{t+1}:

$$\phi_{t+1} = \phi_0 + (1-\phi_0)\sum_{k=0}^{\infty}P_k\sum_{j=0}^{k}\binom{k}{j}\theta_t^j(1-\theta_t)^{k-j}B_{kj}.$$

- So we need to compute θ_t... massive excitement...

Expected size of spread

First connect θ_0 to θ_1:

- $\theta_1 = \phi_0 +$

$$(1-\phi_0)\sum_{k=1}^{\infty}\frac{kP_k}{\langle k\rangle}\sum_{j=0}^{k-1}\binom{k-1}{j}\theta_0^j(1-\theta_0)^{k-1-j}B_{kj}$$

- $\frac{kP_k}{\langle k\rangle} = Q_k$ = **Pr** (edge connects to a degree k node).
- $\sum_{j=0}^{k-1}$ piece gives **Pr** (degree node k activates if j of its $k-1$ incoming neighbors are active).
- ϕ_0 and $(1-\phi_0)$ terms account for state of node at time $t=0$.
- See this all generalizes to give θ_{t+1} in terms of θ_t...

Expected size of spread

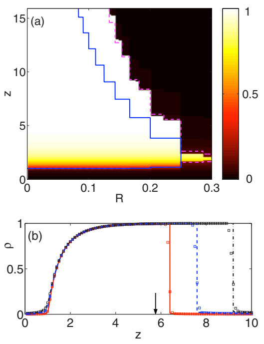
Two pieces: edges first, and then nodes

1. $$\theta_{t+1} = \underbrace{\phi_0}_{\text{exogenous}} + (1-\phi_0)\underbrace{\sum_{k=1}^{\infty}\frac{kP_k}{\langle k\rangle}\sum_{j=0}^{k-1}\binom{k-1}{j}\theta_t^j(1-\theta_t)^{k-1-j}B_{kj}}_{\text{social effects}}$$

   with $\theta_0 = \phi_0$.

2. $$\phi_{t+1} = \underbrace{\phi_0}_{\text{exogenous}} + (1-\phi_0)\underbrace{\sum_{k=0}^{\infty}P_k\sum_{j=0}^{k}\binom{k}{j}\theta_t^j(1-\theta_t)^{k-j}B_{kj}}_{\text{social effects}}.$$

Comparison between theory and simulations



From Gleeson and Cahalane [7]

- Pure random networks with simple threshold responses
- R = uniform threshold (our ϕ_*); z = average degree; $\rho = \phi$; $q = \theta$; $N = 10^5$.
- $\phi_0 = 10^{-3}$, 0.5×10^{-2}, and 10^{-2}.
- Cascade window is for $\phi_0 = 10^{-2}$ case.
- Sensible expansion of cascade window as ϕ_0 increases.

Notes:

- Retrieve cascade condition for spreading from a single seed in limit $\phi_0 \to 0$.
- Depends on map $\theta_{t+1} = G(\theta_t;\phi_0)$.
- First: if self-starters are present, some activation is assured:

$$G(0;\phi_0) = \sum_{k=1}^{\infty}\frac{kP_k}{\langle k\rangle}\bullet B_{k0} > 0.$$

meaning $B_{k0} > 0$ for at least one value of $k \geq 1$.

- If $\theta = 0$ is a fixed point of G (i.e., $G(0;\phi_0) = 0$) then spreading occurs for a small seed if

$$G'(0;\phi_0) = \sum_{k=0}^{\infty}\frac{kP_k}{\langle k\rangle}\bullet(k-1)\bullet B_{k1} > 1.$$

Insert question from assignment 10

Notes:

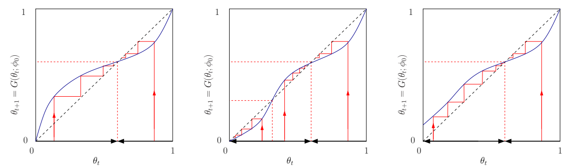
In words:

- If $G(0;\phi_0) > 0$, spreading must occur because some nodes turn on for free.
- If G has an unstable fixed point at $\theta = 0$, then cascades are also always possible.

Non-vanishing seed case:

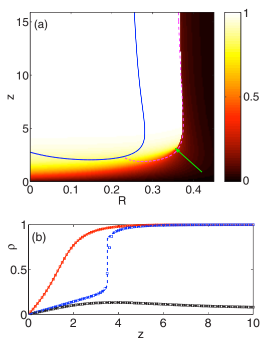
- Cascade condition is more complicated for $\phi_0 > 0$.
- If G has a stable fixed point at $\theta = 0$, and an unstable fixed point for some $0 < \theta_* < 1$, then for $\theta_0 > \theta_*$, spreading takes off.
- Tricky point: G depends on ϕ_0, so as we change ϕ_0, we also change G.

General fixed point story:



- Given $\theta_0 (= \phi_0)$, θ_∞ will be the nearest stable fixed point, either above or below.
- n.b., adjacent fixed points must have opposite stability types.
- Important: Actual form of G depends on ϕ_0.
- Important: ϕ_t can only increase monotonically so ϕ_0 must shape G so that ϕ_0 is at or above an unstable fixed point.
- First reason: $\phi_1 \geq \phi_0$.
- Second: $G'(\theta;\phi_0) \geq 0$, $0 \leq \theta \leq 1$.

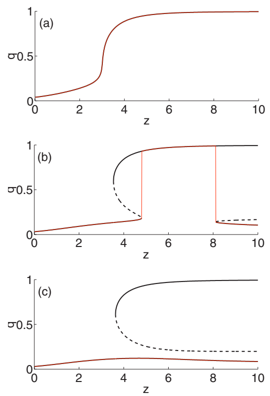
Interesting behavior:



From Gleeson and Cahalane [7]

- Now allow thresholds to be distributed according to a Gaussian with mean R.
- R = 0.2, 0.362, and 0.38; $\sigma = 0.2$.
- $\phi_0 = 0$ but some nodes have thresholds ≤ 0 so effectively $\phi_0 > 0$.
- Now see a (nasty) discontinuous phase transition for low $\langle k\rangle$.

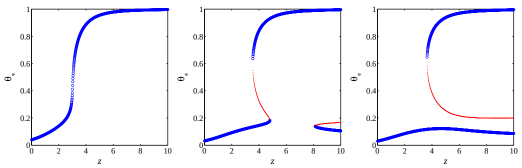
Interesting behavior:



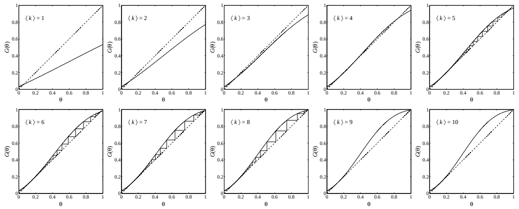
- Plots of stability points for $\theta_{t+1} = G(\theta_t; \phi_0)$.
- n.b.: 0 is not a fixed point here: $\theta_0 = 0$ always takes off.
- Top to bottom: R = 0.35, 0.371, and 0.375.
- Saddle node bifurcations appear and merge (b and c).

From Gleeson and Cahalane [7]

What's happening:



- Fixed points slip above and below the $\theta_{t+1} = \theta_t$ line:



Time-dependent solutions

Synchronous update

- Done: Evolution of ϕ_t and θ_t given exactly by the maps we have derived.

Asynchronous updates

- Update nodes with probability α.
- As $\alpha \to 0$, updates become effectively independent.
- Now can talk about $\phi(t)$ and $\theta(t)$.

Nutshell:

- Solid dive into understanding contagion on generalized random networks.
- Threshold model leads to idea of vulnerables and a critical mass. [16, 8]
- Generating function approaches provided first breakthroughs and gave possibility and probability of spreading. [10, 16]
- Later: A probabilistic, physical method solved the whole story for a fractional seed—final size, dynamics, ... [7, 6]
- Much can be generalized for more realistic kinds of networks: degree-correlated, modular, bipartite, ...
- The single seed contagion condition and triggering probability can be fully developed using a physical story. [5, 9]
- Many connections to other kinds of models: Voter models, Ising models, ...

References I

[1] 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.

[2] 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

[3] J. M. Carlson and J. Doyle.
Highly optimized tolerance: A mechanism for power laws in designed systems.
Phys. Rev. E, 60(2):1412–1427, 1999. pdf

References II

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

[5] P. S. Dodds, K. D. Harris, and J. L. Payne.
Direct, phyisically motivated derivation of the contagion condition for spreading processes on generalized random networks.
Phys. Rev. E, 83:056122, 2011. pdf

[6] J. P. Gleeson.
Cascades on correlated and modular random networks.
Phys. Rev. E, 77:046117, 2008. pdf

References III

[7] J. P. Gleeson and D. J. Cahalane.
Seed size strongly affects cascades on random networks.
Phys. Rev. E, 75:056103, 2007. pdf

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

[9] K. D. Harris, J. L. Payne, and P. S. Dodds.
Direct, physically-motivated derivation of triggering probabilities for contagion processes acting on correlated random networks.
http://arxiv.org/abs/1108.5398, 2014.

References IV

[10] M. E. J. Newman, S. H. Strogatz, and D. J. Watts.
Random graphs with arbitrary degree distributions and their applications.
Phys. Rev. E, 64:026118, 2001. pdf

[11] T. C. Schelling.
Dynamic models of segregation.
J. Math. Sociol., 1:143–186, 1971. pdf

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

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

References V

[14] D. Sornette.
Critical Phenomena in Natural Sciences.
Springer-Verlag, Berlin, 1st edition, 2003.

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

[16] D. J. Watts, P. S. Dodds, and M. E. J. Newman.
Identity and search in social networks.
Science, 296:1302–1305, 2002. pdf