Social Contagion
Principles of Complex Systems
CSYS/MATH 300, Spring, 2013 | #SpringPoCS2013

Prof. Peter Dodds
@peterdodds

Department of Mathematics & Statistics | Center for Complex Systems | Vermont Advanced Computing Center | University of Vermont

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.

Outline

Social Contagion Models
- Background
- Granovetter's model
- Network version
- Final size
- Spreading success
- Groups

References

These slides brought to you by:

Sealie & Lambie Productions

Things that spread well:

buzzfeed.com (iii):

LOL + cute + fail + wtf:

Oopsie!

BUZZFEED FELL DOWN AND WENT BOOM.

The whole lolcats thing:

:-p

Please try reloading this page. If the problem persists let us know.
Some things really stick:

wtf + geeky + omg:

Examples abound

SIR and SIRS contagion possible

[1] fashion
[2] striking
[3] smoking
[4] residential segregation
[5] ipods
[6] obesity
[7] Classes of behavior versus specific behavior: dieting
Framingham heart study:

**Evolved network stories (Christakis and Fowler):**
- The spread of quitting smoking (шейп)
- The spread of spreading (шейп)[6]
- Also: happiness (шейп)[9], loneliness, ...
- The book: Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives (шейп)

Controversy:
- Are your friends making you fat? (шейп) (Clive Thompson, NY Times, September 10, 2009).
- Everything is contagious (шейп)—Doubts about the social plague stir in the human superorganism (Dave Johns, Slate, April 8, 2010).

Social Contagion

**Two focuses for us**
- Widespread media influence
- Word-of-mouth influence

**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[10] as ‘connectors’
- The infectious idea of opinion leaders (Katz and Lazarsfeld)[16]

The hypodermic model of influence

**Why do things spread?**
- Because of properties of special individuals?
- Or system level properties?
- 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...
- Reasons for success are usually ascribed to intrinsic properties (e.g., Mona Lisa)
- System/group properties harder to understand
- Always good to examine what is said before and after the fact...
The Mona Lisa

- “Becoming Mona Lisa: The Making of a Global Icon”—David Sassoon
- Not the world’s greatest painting from the start...
- Escalation through theft, vandalism, parody, ...

The completely unpredicted fall of Eastern Europe

Timur Kuran:[17, 18] *Now Out of Never: The Element of Surprise in the East European Revolution of 1989*

The dismal predictive powers of editors...

Social Contagion

- Social Contagion
- Social Contagion
- Social Contagion
- Social Contagion

Social Contagion

- Messing with social connections
  - Ads based on message content (e.g., Google and email)
  - BzzAgent (⊕)
  - One of Facebook’s early advertising attempts: Beacon (⊕)
  - All of Facebook’s advertising attempts.

Getting others to do things for you

A very good book: *Influence*[8] by Robert Cialdini (⊕)

Six modes of influence:
1. Reciprocation: *The Old Give and Take... and Take*; e.g., Free samples, Hare Krishnas.
2. Commitment and Consistency: *Hobgoblins of the Mind*; e.g., Hazing.
3. Social Proof: *Truths Are Us*; e.g., Jonestown (⊕), Kitty Genovese (⊕) (contested).
4. Liking: *The Friendly Thief*; e.g., Separation into groups is enough to cause problems.
5. Authority: *Directed Deference*; e.g., Milgram’s obedience to authority experiment. (⊕)
6. Scarcity: *The Rule of the Few*; e.g., Prohibition.

Social Contagion

- Cialdini’s modes are heuristics that help us get through life.
- Useful but can be leveraged...

Other acts of influence:
- Conspicuous Consumption (Veblen, 1912)
- Conspicuous Destruction (Potlatch)
Social Contagion

Some important models:
- Tipping models—Schelling (1971) \[19, 20, 21\]
  - Simulation on checker boards
  - Explore the Netlogo (\(\Xi\)) online implementation (\(\Xi\))\[26\]
- Threshold models—Granovetter (1978)\[13\]
- Herding models—Bikhchandani, Hirschleifer, Welch (1992)\[2, 3\]
  - 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).

Granovetter’s Threshold model—definitions
- \(\phi^*\) = threshold of an individual.
- \(f(\phi_*)\) = distribution of thresholds in a population.
- \(F(\phi_*)\) = cumulative distribution = \(\int_{0}^{\phi_*} f(\phi_*)d\phi_*\)
- \(\phi_t\) = fraction of people ‘rioting’ at time step \(t\).
- At time \(t + 1\), fraction rioting = fraction with \(\phi_* \leq \phi_t\).

Action based on perceived behavior of others:
- Two states: S and I.
- \(\phi\) = fraction of contacts ‘on’ (e.g., rioting)
- Discrete time update (strong assumption!)
- This is a Critical mass model
Threshold models

Another example of critical mass model:

Example of single stable state model:

Threshold models

Chaotic behavior possible\textsuperscript{15, 14}

- Period doubling arises as map amplitude $r$ is increased.
- Synchronous update assumption is crucial

Threshold models—Nutshell

Implications for collective action theory:
1. Collective uniformity $\nrightarrow$ individual uniformity
2. Small individual changes $\Rightarrow$ large global changes

“A simple model of global cascades on random networks”

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

We’ll also explore:
- “Seed size strongly affects cascades on random networks”\textsuperscript{12}
- “Influentials, Networks, and Public Opinion Formation”\textsuperscript{24}
- “Threshold models of Social Influence”\textsuperscript{25}

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 $\phi_i$
- Individuals repeatedly poll contacts on network
- Synchronous, discrete time updating
- Individual $i$ becomes active when fraction of active contacts $\frac{\sum_j \tau_{ij} \phi_j}{k_i} \geq \phi_i$
- Individuals remain active when switched (no recovery = SI model)
Threshold model on a network

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

Snowballing

First study random networks:
- Start with $N$ nodes with a degree distribution $\rho_k$.
- Nodes are randomly connected (carefully so).
- Aim: Figure out when activation will propagate.
- Determine a cascade condition.

The Cascade Condition:
1. If one individual is initially activated, what is the probability that an activation will spread over a network?
2. What features of a network determine whether a cascade will occur or not?

Example random network structure:

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

\[ \Omega_{\text{crit}} \subset \Omega_{\text{trig}}; \Omega_{\text{crit}} \subset \Omega_{\text{final}}; \text{ and } \Omega_{\text{trig}} \subset \Omega. \]

Snowballing

Follow active links:
- An active link is a link connected to an activated node.
- If an infected link leads to at least 1 more infected link, then activation spreads.
- We need to understand which nodes can be activated when only one of their neighbors becomes active.

The most gullible

Vulnerables:
- We call individuals who can be activated by just one contact being active vulnerable.
- The vulnerability condition for node $i$:
  \[ 1/k_i \geq \phi_i \]
- Which means # contacts $k_i \leq \lfloor 1/\phi_i \rfloor$
- For global cascades on random networks, must have a global cluster of vulnerables.
- Cluster of vulnerables = critical mass.
- Network story: 1 node $\rightarrow$ critical mass $\rightarrow$ everyone.

Cascade condition

Back to following a link:
- A randomly chosen link, traversed in a random direction, leads to a degree $k$ node with probability $k \rho_k$.
- Follows from there being $k$ ways to connect to a node with degree $k$.
- Normalization:
  \[ \sum_{k=0}^{\infty} k \rho_k = \langle k \rangle \]
- So
  \[ P(\text{linked node has degree } k) = \frac{k \rho_k}{\langle k \rangle} \]
Cascade condition

Next: Vulnerability of linked node
- Linked node is vulnerable with probability
  \[ \beta_k = \int_{\phi' = 0}^{1/k} f(\phi') d\phi' \]
- If linked node is vulnerable, it produces \( k - 1 \) new outgoing active links
- If linked node is not vulnerable, it produces no active links.

Cascade condition

Putting things together:
- Expected number of active edges produced by an active edge:
  \[ R = \sum_{k=1}^{\infty} (k - 1) \cdot \beta_k \cdot \frac{kP_k}{(k)} + 0 \cdot (1 - \beta_k) \cdot \frac{kP_k}{(k)} \]
  \[ = \sum_{k=1}^{\infty} (k - 1) \cdot \beta_k \cdot \frac{kP_k}{(k)} \]

Cascade condition

So... for random networks with fixed degree distributions, cascades take off when:
\[ \sum_{k=1}^{\infty} (k - 1) \cdot \beta_k \cdot \frac{kP_k}{(k)} \geq 1. \]
- \( \beta_k \) = probability a degree \( k \) node is vulnerable.
- \( P_k \) = probability a node has degree \( k \).

Cascade condition

Cascades take off when:

Putting things together:

Two special cases:
- (1) Simple disease-like spreading succeeds: \( \beta_k = \beta \)
  \[ \beta \cdot \sum_{k=1}^{\infty} (k - 1) \cdot \frac{kP_k}{(k)} \geq 1. \]
- (2) Giant component exists: \( \beta = 1 \)
  \[ 1 \cdot \sum_{k=1}^{\infty} (k - 1) \cdot \frac{kP_k}{(k)} \geq 1. \]

Cascades on random networks

Cascades occur only if size of max vulnerable cluster \( > 0 \).
- System may be ‘robust-yet-fragile’.
- ‘Ignorance’ facilitates spreading.

Cascade window for random networks

‘Cascade window’ widens as threshold \( \phi \) decreases.
- Lower thresholds enable spreading.
Cascade window for random networks

All-to-all versus random networks

Threshold contagion on random networks

Cascade window—summary

Expected size of spread

For our simple model of a uniform threshold:

1. Low $\langle k \rangle$: No cascades in poorly connected networks.
   No global clusters of any kind.
2. High $\langle k \rangle$: Giant component exists but not enough vulnerables.
3. Intermediate $\langle k \rangle$: Global cluster of vulnerables exists. Cascades are possible in "Cascade window."

Next: 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:

Developed further by Gleeson in "Cascades on correlated and modular random networks," Phys. Rev. E, 2008.\textsuperscript{[11]}

References

Idea:

- Randomly turn on a fraction $\phi_0$ 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 = $\phi_0$)
  - $t = 1$: $i$ 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

- $t = 0$, $\phi = 1/3$
Expected size of spread

Notes:
- Calculations are possible if nodes do not become inactive (strong restriction).
- 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 $k$ switching 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.

References

Models

Social Contagion

First connect $\theta_0$ to $\theta_1$:
- $\theta_1 = \phi_0 + (1 - \phi_0) \sum_{k=0}^{\infty} k P_k \sum_{j=0}^{k-1} \binom{k-1}{j} \theta_j^t (1 - \theta_0)^{k-1-j} B_{kj}$
- $k P_k / k = R_k = Pr$ (edge connects to a degree $k$ node).
- $\sum_{j=0}^{k-1} \theta_j^t$ piece gives Pr(degree node $k$ activates) of its neighbors $k - 1$ incoming neighbors are active.
- $\phi_0$ and $(1 - \phi_0)$ terms account for state of node at time $t = 0$.
- See this all generalizes to give $\theta_{i+1}$ in terms of $\theta_i$...
Expected size of spread

Two pieces: edges first, and then nodes

1. \( \theta_{t+1} = \frac{\phi_0}{\sum_{k=1}^{\infty} \frac{k^P}{k!} \sum_{j=0}^{k-1} \binom{k-1}{j} \theta_j^j (1 - \theta_j)^{k-1-j} B_{kj} } \)

\( \text{exogenous} \)

\( + (1 - \phi_0) \sum_{k=1}^{\infty} \frac{k^P}{k!} \sum_{j=0}^{k-1} \binom{k-1}{j} \theta_j^j (1 - \theta_j)^{k-1-j} B_{kj} \)

\( \text{social effects} \)

with \( \theta_0 = \phi_0 \).

2. \( \phi_{t+1} = \frac{\psi_0}{\sum_{k=1}^{\infty} \frac{k^P}{k!} \sum_{j=0}^{k-1} \binom{k-1}{j} \theta_j^j (1 - \theta_j)^{k-1-j} B_{kj} } \)

\( \text{exogenous} \)

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

\( \text{social effects} \)

Expected size of spread:

- 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:
  \[ G(0; \phi_0) = \sum_{k=1}^{\infty} \frac{k^P}{k!} \cdot B_{k0} > 0. \]
  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 if
  \[ G'(0; \phi_0) = \sum_{k=0}^{\infty} \frac{k^P}{k!} \cdot (k - 1) \cdot B_{k1} > 1. \]

Expected size of spread:

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 \( \phi_0 \), so as we change \( \phi_0 \), we also change \( G \).

General fixed point story:

- Given \( \theta_0(= \phi_0) \), \( \theta_\infty \) 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 \( \phi_0 \).
- So choice of \( \phi_0 \) dictates both \( G \) and starting point—can’t start anywhere for a given \( G \).

Early adopters—degree distributions

The multiplier effect:

- Fairly uniform levels of individual influence.
- Multiplier effect is mostly below 1.
**The multiplier effect:**

- Skewed influence distribution example.

**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**

“A few harmless flakes working together can unleash an avalanche of destruction.”

**Special subnetworks can act as triggers**

- \( \phi = 1/3 \) for all nodes
Context distance

Generalized affiliation model

Generalized affiliation model networks with triadic closure

- Connect nodes with probability $\propto \exp^{-kd}$
- where
  - $\alpha$ = homophily parameter
  - $d$ = distance between nodes (height of lowest common ancestor)
- $\tau_1$ = intergroup probability of friend-of-friend connection
- $\tau_2$ = intragroup probability of friend-of-friend connection
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.

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

References II

References III

References IV
Social Contagion

References V

Threshold models of diversity: Chinese restaurants, residential segregation, and the spiral of silence.

Threshold models of interpersonal effects in consumer demand.

Personal Influence.


References VI


Dynamic models of segregation.

Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities.

Micromotives and Macrobehavior.

References VII

[22] D. Sornette.
Critical Phenomena in Natural Sciences.

[23] D. J. Watts.
A simple model of global cascades on random networks.

Influentials, networks, and public opinion formation.

References VIII

Threshold models of social influence.

Netlogo segregation model.