Speed versus Resilience in Contagion

Andrew Koh
MIT

Stephen Morris
MIT

European Summer Symposium on Economic Theory (Gerzensee)

July 2023
Motivation

- Threshold contagion: Infected IFF $\geq q$ proportion of my neighbours infected
  - Pretty well-studied (Econ, CS, OR, Sociology...)

But there is a natural tension between these objects
- Specific to threshold contagion
- Exception: Centola and Macy (2007): simple vs complex contagion, tension in re-wiring simulations, will discuss later

Key intuition: many faraway (weak) links means fewer local (strong) links (in proportion)

This paper: simple(st) model which channels this?
Motivation

- Threshold contagion: Infected IFF $\geq q$ proportion of my neighbours infected
  - Pretty well-studied (Econ, CS, OR, Sociology...)
- (To our knowledge) literature tends to focus on either:
  - Resilience (min seed, max threshold etc.)
  - Speed (conditional on spreading, how quickly?)
Motivation

- Threshold contagion: Infected IFF $\geq q$ proportion of my neighbours infected
  - Pretty well-studied (Econ, CS, OR, Sociology...)

- (To our knowledge) literature tends to focus on either:
  - Resilience (min seed, max threshold etc.)
  - Speed (conditional on spreading, how quickly?)

- But there is a natural tension between these objects
  - Specific to threshold contagion
  - Exception: Centola and Macy (2007): simple vs complex contagion, tension in re-wiring simulations, will discuss later
Motivation

- **Threshold contagion**: Infected IFF \( q \) proportion of my neighbours infected
  - Pretty well-studied (Econ, CS, OR, Sociology...)
- (To our knowledge) literature tends to focus on either:
  - **Resilience** (min seed, max threshold etc.)
  - **Speed** (conditional on spreading, how quickly?)
- But there is a natural tension between these objects
  - Specific to threshold contagion
  - Exception: Centola and Macy (2007): simple vs complex contagion, tension in re-wiring simulations, will discuss later

**Key intuition**: many faraway (weak) links means fewer local (strong) links (in proportion)
- This paper: **simple(st) model** which channels this?
Model

- Develop continuous network model in 1 dimension ($\mathbb{R}$)
- Analytically more tractable (c.f. global games)
- Deterministic dynamics but approximates large random graphs (Lovász)
- Link to Watts-Strogatz, Newman, Newman-Watts etc. models of random graphs (but also important differences. More on this later.)
- Other extensions: many dimensions, less structure on neighborhoods
- I am counting on network people to tell me (i) what generalizations are or aren’t important; (ii) how to test
Results

- **Tradeoff**: Contagion is faster in networks where it is harder to initiate contagion.
Results

- **Tradeoff**: contagion is *faster* in networks where it is *harder* to initiate contagion
- 3 simple propositions. Tradeoff...
  - *...is stark* when graphs ordered by FOSD: for every threshold, one is harder to initiate contagion & spreads faster;
  - *...can be absent* for fixed thresholds under ‘single-crossing’ ordering: worst of both worlds—easier to initiate contagion, but also spreads faster
  - *...always occurs* for any 2 graphs (but potentially different contagion thresholds);
Model

- infinite measure of agents indexed $i \in \mathbb{R}$ sitting on a line. Each agent has links of measure 1 (normalization).
- $i$’s links to the right given by $G_i : \mathbb{R}_+ \rightarrow [0, 1/2]$ (‘CDF’)
  - $G_i(x)$ is the weight $i$ places on agents at location $[i, i + x]$.
- Assume:
  - $G_i$ admits a density $g_i : \mathbb{R}_+ \rightarrow [0, 1]$ (‘$g(x)$ is the weight placed on agent $x$ away’)
  - $G_i$ has decreasing differences i.e., $g$ strictly decreasing on support (‘homophily’)

![Diagram showing the model and the density function $g_i(x)$]
Model

- Impose **symmetry** and **translation invariance**
  - **symmetry**: i’s links on the left and right distributed identically
  - **translation invariance**: i’s distribution of links over $[i, i + x]$ same as that of $i'$ over $[i', i' + x]$

- Work with (single) $G$ directly (sufficient statistic) rather than with explicit description of neighbourhoods
Model

- Binary action space: \( \{0, 1\} \) ‘uninfected or infected’

**q-contagion**

Player \( i \) takes action 1 if and only if \( \geq q \) proportion of her neighbours take action 1.

- We will study the **evolution** of the set of infected agents.
- Let \( I_t \subseteq \mathbb{R} \) be the set of agent infected at \( t \)
- \( i \) is infected at time \( t + 1 \) if and only if

\[
\int_{-\infty}^{+\infty} 1[i - x \in I_t] g(x) \, dx \geq q
\]

- Kind of unwieldy carrying an infinite-dimensional object around
Key simplification $\rightarrow$ interval contagion

- Key simplifications: $I_0$ interval $+$ $g$ strictly decreasing
- Implies $\{I_t\}_t$ are intervals (immediately from induction)

$$G(\mu(I_0) + x) - G(x) = q$$

Discussion:
- $I_0$ is not minimal to induce contagion. (but can get LB)
- Homophily (at least in expectation) seems natural.
Defining resilience & speed

- Let’s keep track of measure of infected agents rather than sets.
  - Define \( m_t = \mu(I_t) \) ← measure of infected agents at \( t \)
  - Define \( a_t = m_t - m_{t-1} \) ← change in measure from \( t - 1 \) to \( t \)

**Definition (Contagion occurring)**

Contagion occurs if \( \lim_{t \to +\infty} m_t = +\infty \)

**Definition (Resilience)**

\( m_0(G, q) := \inf m_0 \quad s.t. \quad \text{contagion occurs} \)

- Note: could also fix \( m_0 \) and look at max threshold \( q \)

**Definition ((Limit) Speed)**

\( a_\infty(G, q) := \lim_{t \to \infty} a_t \)
Expression for $m_0$

- Tuple $(G, m_0, q)$ sufficient to pin down contagion dynamics.

**Lemma**

Contagion occurs from $(G, m_0, q)$ if and only if $G(m_0) > q$

- Sketch: (⇐ = ) Take $I_0 = [x, x + m_0]$. WLOG b/c translation invariance. Let $a_1/2$ be additional measure infected at $t = 1$ on the left of $I_0$ and conjecture that $> 0$
- We know that the guy $x - a_1/2$ must be ‘indifferent’:
  
  $G(a_1/2 + m_0) - G(a_1/2) = q < G(m_0)$

If $x$ strictly prefers to take 1,
then by dec. differences someone to the left must also take 1
Expression for $m_0$

**Lemma**

Contagion occurs from $(G, m_0, q)$ if and only if $G(m_0) > q$

- Sketch: (⇒) Consider any $i < x$ and define $\epsilon := x - i$. If the condition on $G$ doesn’t hold,

  $$q \geq G(m_0) > G(\epsilon + m_0) - G(\epsilon)$$

  because $g$ is decreasing.

- But then any $i < x$ is not infected in period 1. So $m_{t+1} \leq m_t$ and $\lim_{t \to \infty} m_t \leq m_0 < +\infty$

- **Useful expression from the lemma:**

  $$m_0(G, q) = G^{-1}(q)$$

  (Min-Seed)

- Interpretation: **Local links** matter for resilience
  - $G(x)$ is the ‘CDF’ of links: ‘what proportion of my links are less than $x$ distance away from me?’
  - $G^{-1}(q)$: ‘what’s the distance of guys around me which needs to be infected before $q$ proportion of my neighbours are infected?’
Expression for $a_\infty$

- Recall: $a_t = m_t - m_{t-1}$
- Define $\bar{G} := 1/2 - G$ as the anti-CDF
- If infection occurs, $a_t$ solve the nonlinear diff. eqn.

$$\bar{G}(a_t/2) - \bar{G}(a_t/2 + m_{t-1}) = q$$

mass of infected neighbours:

$$\bar{G}(a_t/2) - \bar{G}(a_t/2 + m_{t-1})$$

infected at start of $t$ (mass $m_{t-1}$)  |  newly infected at time $t$ (mass $a_t/2$)
Expression for $a_\infty$

$$\bar{G}(a_t/2) - \bar{G}(a_t/2 + m_{t-1}) = q$$

- Path $\{a_t\}_t$ will, in general, depend on $G$.
- Observe that if contagion occurs, $a_t$ will be increasing.
  - from decreasing differences since $m_{t-1} = \sum_{0}^{t-1} a_s$ is growing
- But will converge to a limit:

  $$a_\infty := \lim_{t \to \infty} a_t$$

  $$= 2(\bar{G})^{-1}(q) \quad \text{since} \quad \lim_{m_{t-1} \to \infty} \bar{G}(a_t/2 + m_{t-1}) = 0.$$  
  (Lim-Speed)

- Interpretation: **Distant links** matter for speed.
  - $\bar{G}^{-1}(q)$ is the distance ‘from the interval stretching to infinity’ required to have $q$ proportion of neighbours
  - Contrast with **Min-Seed**: $m_0(G, q) = G^{-1}(q)$
Proposition

If $G, G' \in \mathcal{G}$ are such that $G \leq G'$, then for all $q \in (0, 1/2)$,

(i) $G$ is more resilient than $G'$ i.e., $m_0(G, q) \geq m_0(G', q)$; and

(ii) $G$ has a quicker limit speed than $G'$ i.e., $a_\infty(G, q) \geq a_\infty(G', q)$. 

Tradeoff is **stark** when FOSD-ordered
Tradeoff is **stark** when FOSD-ordered

- Moving from $G$ to $G'$ as if we’re ‘shifting mass’ from nearby links (closer than $x$) further away (beyond $\bar{x}$).

![Diagram showing the tradeoff between $G$ and $G'$, with $x$ and $\bar{x}$ as reference points.](image-url)
Tradeoff is **stark** when FOSD-ordered

- **Proof immediate from expressions...**
- **Resilience:** Fix any $q \in (0, 1/2)$. $G \leq G'$ implies $G^{-1} \leq G'^{-1}$. Hence from our expression for $m_0$

  $$m_0(G, q) = G^{-1}(q) \leq G'^{-1}(q) = m_0(G', q).$$

- **Speed:** If contagion doesn’t occur, speed is identically zero. If it does, $G \leq G'$ implies $\overline{G} \geq \overline{G'}$ and so $(\overline{G})^{-1} \geq (\overline{G}')^{-1}$ and from our expression for $a_\infty$,

  $$a_\infty(G, q) = 2(\overline{G})^{-1}(q) \geq 2(\overline{G'})^{-1}(q) = a_\infty(G', q).$$
Example: normal distributions

- Suppose that $i$’s links are ‘normally distributed’ ($i$ is sitting in the middle so $i$’s links are $\sim N(i, \sigma^2)$)
- $G_\sigma(x) = \Phi(x/\sigma) - \frac{1}{2}$
- Invert and rearrange...

$$m_0(G_\sigma, q) = \sigma \cdot \Phi(q + \frac{1}{2})$$
$$a_\infty(G_\sigma, q) = 2\sigma \cdot \Phi(1 - q)$$

- High $\sigma$: more mass on faraway links.
Other ways to shift mass...

- Increasing the s.d. of normal is quite special...
- We’ve seen that local links matter for resilience; faraway links matter for speed.
- These two things can coexist by shifting ‘middle links’:
  - Closer: more local links, less resilient
  - Further: more tail links, quicker speed

**Proposition**

For $G, G' \in \mathcal{G}$, suppose that there exists some $\bar{x} \in (0, +\infty)$ such that for all $x' \leq \bar{x}$, $G(x') \leq G'(x')$ and for all $x'' \geq \bar{x}$, $G(x'') \geq G'(x'')$.

Then

(i) for sufficiently low values of $q$, $G$ is both more resilient than $G'$ as well as has slower limit speeds; and

(ii) for sufficiently high values of $q$, $G$ is both less resilient than $G'$ as well as has quicker limit speeds.
Other ways to shift mass...

Here \( G' \geq G \) before \( \bar{x} \), and the opposite after \( \bar{x} \). Implies that \( G' \) has more local links but also fatter tails.

Note that \( q \) controls ‘how local’ and ‘how far away’ the links need to be for them to matter for resilience and speed.

- lower \( q \) → more remote tails matter, more local neighbourhoods matter
Other ways to shift mass...

**Sketch:** Choose \( q' = G(\overline{x}) \) and note that by the condition of single crossing at \( \overline{x} \) in the proposition, for any \( q \leq q' \),

\[
m_0(G, q) = G^{-1}(q) \geq G'^{-1}(q) = m_0(G', q).
\]

with the reverse equality for \( q \geq q' \). Choose \( q'' = \overline{G}(\overline{x}) \) and for \( q \leq q'' \) by the condition in the proposition,

\[
a_\infty(G, q) = 2\overline{G}^{-1}(q) \leq 2\overline{G}'^{-1}(q) = a_\infty(G', q)
\]

with the reverse equality for \( q \geq q'' \). Part (i) follows for thresholds \( q \leq q' \land q'' \); part (ii) follows for thresholds \( q \geq q' \lor q'' \).
Other ways to shift mass...

- Cauchy: \( G_{C,\gamma}(x) = \frac{1}{\pi} \arctan(x/\gamma) \)
  - \( m_0(G_{C,\gamma}, q) = \gamma \cdot \tan(q\pi) \), \( a_\infty(G_{C,\gamma}, q) = 2\gamma \cdot \tan((\frac{1}{2} - q)\pi) \)

- Cauchy tails decay polynomially \((\propto 1/x^2)\) hence \( a_\infty \simeq 1/q^2 \).
  - Subgaussian: \( a_\infty \lesssim (\log(1/q))^{1/2} \)

- More generally, always have freedom to control:
  - tails: sub-exponential, heavy tailed, polynomial decay etc.
  - local ‘peakedness’
Tradeoff obtains for any pair of networks

- Note: need \( q < 1/2 \) for contagion to occur; same logic as Morris (2000)

**Proposition**

For \( G, G' \in \mathcal{G} \), if \( G \neq G' \) then there exists \( q, q' \in (0, 1/2) \) such that one is more resilient than the other under \( q \), but has a quicker limit speed than the other under \( q' \).

- Any two graphs exhibit the tradeoff for some contagion thresholds
- E.g., can find \( q = 0.3, \ q' = 0.1 \) so that
  - \( m_0(G, 0.1) > m_0(G', 0.1) \) \( \leftarrow \) G is more resilient than G'
  - \( a_\infty(G, 0.3) > a_\infty(G', 0.3) \) \( \leftarrow \) G spreads faster than G'
Tradeoff obtains for **any pair** of networks

**Sketch pf.**: let’s assume WLOG that $G(x) < G'(x)$ for some $x \in [0, \infty)$. This implies that there exists $y \in (G(x), G'(x))$ such that $G^{-1}(y) > x > G'^{-1}(y)$. Now set $q = y < 1/2$ and by the expressions

$$m_0(G, q) = G^{-1}(q) = G^{-1}(y)$$

$$> G'^{-1}(y) = G'^{-1}(q) = m_0(G', q).$$

Next, recall we defined $\overline{G} = 1/2 - G$. There exists $z \in (\overline{G}'(x), \overline{G}(x))$ such that $\overline{G}'^{-1}(z) < x < \overline{G}^{-1}(z)$ and setting $q' = z < 1/2$, we have

$$a_{\infty}(G, q') = \overline{G}^{-1}(q') = \overline{G}^{-1}(z)$$

$$> \overline{G}'^{-1}(z) = \overline{G}'^{-1}(q') = a_{\infty}(G', q').$$
Taking stock

- Tradeoff...
  - ...is stark when graphs ordered by FOSD: for every threshold, one is harder to initiate contagion & spreads faster;
  - ...can be absent for fixed thresholds under ‘single-crossing’ ordering: worst of both worlds—easier to initiate contagion, but also spreads faster
  - ...always occurs for any 2 graphs (but potentially different contagion thresholds);

- Extension 1: What is the link with random (discrete) graphs?

- Extension 2: Higher dimensions vs 1D
  - analytically quite ugly, but some results go through. We may simulate the rest....
(Informal) Link to finite mass population random discrete graphs

- So far we worked on $\mathbb{R}$. Allows us to speak of ‘limit speed’, work with canonical full-support distributions etc.
- Now: Unit measure of agents $i \in [0, 1]$.
  - Allows us to link contagion results to discrete random graphs: sample uniformly from $[0, 1]$ (see Lovász (2012))
- Define $G$ as was our space of graphs on $\mathbb{R}$. Define $G^T$ as the space of graphs over the unit circle $[0, 1]$ with similar conditions (density exists, homophily etc.)
- **Goal**: define a transformation $G \rightarrow G^T$ which preserves contagion dynamics (& tradeoffs) studied in $\mathbb{R}$.
### Link to random discrete graphs

**Goal:** find map \( \psi_s : \mathcal{G} \rightarrow \mathcal{G}^T \) which ‘preserves contagion dynamics’

Here’s the map we use:

Let \( \psi_s \) be this map, where \( s \) controls the shrinkage factor: (in 1st step \( G_s(sx) = G(x) \))

\[
\lim_{t \to \infty} \lim_{s \to 0} a_t(\psi_s(G), q) = s \cdot a_\infty(G, q) \\
\lim_{s \to 0} m_0(\psi_s(G), q) = s \cdot m_0(G, q)
\]
Link to random discrete graphs

- For small but finite \( s \), exhibit the same tradeoffs.
  - Now with avg speed (up to full infected) rather than limit speed.
  - Could work directly on \([0, 1]\) but uglier.

Taking stock:

\[ \infty \text{ measure} \xrightarrow{\text{now}} \text{Unit measure} \xrightarrow{\text{Random graphs}} \]

- Graphs in \( G^T \) are graphons: \( W : [0, 1]^2 \rightarrow [0, 1] \).
- \( W(i, j) \) : weight that \( i \) puts on \( j \). In our setting: \( W(i, j) = g(|i - j|) = W(j, i) \).
- Graphons approximate random graphs:
  - Sample \( S \) from \([0, 1]\) uniformly at random.
  - Let’s say \( S = \{i, j\} \). Then on the random graph, \( i \) and \( j \) are connected with probability \( W(i, j) \).

Recent paper in JET by Erol, Parise, and Teytelboym (2020): contagion on graphons approximate contagion on sampled graph
Higher dimensions

- General idea is to work on the Euclidian ball in $\mathbb{R}^n$. Analog of translation invariance and symmetry s.t. graph can once again be summarized by a single CDF
- But now the dimension, size of the ball, and distance all matter!

$$p^{n=2}(r, d, x)$$ gives the length of the arc of the dotted ball within the solid ball divided by the circumference of the dotted ball
Conclusion

- **Tradeoff**: contagion is faster in networks where it is harder to initiate contagion
Conclusion

- **Tradeoff**: contagion is *faster* in networks where it is *harder* to initiate contagion
- clear empirical implications
- special model but clearly generalizes: what is the right way to do so to make it compelling to network theorists?
Thanks!


