Cascades in networks: a simple theory and applications

Graduate Summer School: Games and Contracts for Cyber-Physical Security
IPAM, UCLA
22 July 2015 (9:00-9:45)

Yongwhan Lim, Asuman Ozdaglar, Alex Teytelboym

MIT, MIT, Oxford
Competitive cascades
Motivation

- Rumors can spread quickly through social networks.
  - Examples: political rumors; firms worry about the reputations of their products.
  - In the short run, rumors are irreversible.
  - The initial seeds in the social network really matter for what rumor spreads.
Motivation

- Rumors can spread quickly through social networks.
- Examples: political rumors; firms worry about the reputations of their products.
- In the short run, rumors are irreversible.
- The initial seeds in the social network really matter for what rumor spreads.
Motivation

- Rumors can spread quickly through social networks.
- Examples: political rumors; firms worry about the reputations of their products.
- In the short run, rumors are irreversible.
- The initial seeds in the social network really matter for what rumor spreads.
Motivation

• Rumors can spread quickly through social networks.
• Examples: political rumors; firms worry about the reputations of their products.
• In the short run, rumors are irreversible.
• The initial seeds in the social network really matter for what rumor spreads.
Previous work vs. this talk

- Several recent papers on rumor spread: Moreno et al. (2004); Kostka et al. (2008); Trpevski et al. (2010); He et al. (2012); Weng et al. (2012).

- Rumor spread is similar to work on product diffusion: Goyal and Kearns (2012); Bimpikis, Ozdaglar, and Yildiz (2014) (...and Hotelling, 1929)

- Quality and seeding: Fazeli and Jadbabaie (2012a,b,c); Fazeli, Ajorlou, and Jadbabaie (2014).

- Other papers where consumers can switch products many times: Bharathi, Kempe, and Salek (2007); Alon, Feldman, Procaccia, and Tennenholtz (2010); Apt and Markakis (2011); Simon and Apt (2012); Tzoumas, Amanatidis, and Markakis (2012); Borodin, Braverman, Lucier, and Oren (2013); Apt and Markakis (2014); Mei and Bullo (2014).
Previous work vs. this talk

- Several recent papers on rumor spread: Moreno et al. (2004); Kostka et al. (2008); Trpevski et al. (2010); He et al. (2012); Weng et al. (2012).

- Rumor spread is similar to work on product diffusion: Goyal and Kearns (2012); Bimpikis, Ozdaglar, and Yildiz (2014) (...and Hotelling, 1929)

- Quality and seeding: Fazeli and Jadbabaie (2012a,b,c); Fazeli, Ajorlou, and Jadbabaie (2014).

- Other papers where consumers can switch products many times: Bharathi, Kempe, and Salek (2007); Alon, Feldman, Procaccia, and Tennenholtz (2010); Apt and Markakis (2011); Simon and Apt (2012); Tzoumas, Amanatidis, and Markakis (2012); Borodin, Braverman, Lucier, and Oren (2013); Apt and Markakis (2014); Mei and Bullo (2014).
Previous work vs. this talk

- Several recent papers on rumor spread: Moreno et al. (2004); Kostka et al. (2008); Trpevski et al. (2010); He et al. (2012); Weng et al. (2012).

- Rumor spread is similar to work on product diffusion: Goyal and Kearns (2012); Bimpikis, Ozdaglar, and Yildiz (2014) (...and Hotelling, 1929)

- Quality and seeding: Fazeli and Jadbabaie (2012a,b,c); Fazeli, Ajorlou, and Jadbabaie (2014).

- Other papers where consumers can switch products many times: Bharathi, Kempe, and Salek (2007); Alon, Feldman, Procaccia, and Tennenholtz (2010); Apt and Markakis (2011); Simon and Apt (2012); Tzoumas, Amanatidis, and Markakis (2012); Borodin, Braverman, Lucier, and Oren (2013); Apt and Markakis (2014); Mei and Bullo (2014).
We develop a tractable model of competitive cascades in networks.

- We study the game on the network using cascade centrality.
- We characterize pure-strategy Nash equilibrium, price of anarchy and budget multiplier.
Outline of this talk

- We develop a tractable model of competitive cascades in networks.
- We study the game on the network using cascade centrality.
- We characterize pure-strategy Nash equilibrium, price of anarchy and budget multiplier.
Outline of this talk

- We develop a tractable model of competitive cascades in networks.
- We study the game on the network using cascade centrality.
- We characterize pure-strategy Nash equilibrium, price of anarchy and budget multiplier.
Model
Preliminaries

- Simple, undirected graph $G(V, E)$.
- A spreading threshold for agent $i$ is a random variable $\Theta_i$ drawn from a probability distribution with support $[0, 1]$.
- The associated multivariate probability distribution for all the agents in the graph is $f(\theta)$.
- Each agent is $i \in V$ assigned a threshold $\theta_i$. Let’s define the threshold profile of agents as $\theta := (\theta_i)_{i \in V}$. A network $G_\theta$ is a graph endowed with a threshold profile.
Preliminaries

- Simple, undirected graph $G(V, E)$.
- A spreading threshold for agent $i$ is a random variable $\Theta_i$ drawn from a probability distribution with support $[0, 1]$.
- The associated multivariate probability distribution for all the agents in the graph is $f(\theta)$.
- Each agent is $i \in V$ assigned a threshold $\theta_i$. Let’s define the threshold profile of agents as $\theta := (\theta_i)_{i \in V}$. A network $G_\theta$ is a graph endowed with a threshold profile.
Preliminaries

- Simple, undirected graph \( G(V, E) \).
- A spreading threshold for agent \( i \) is a random variable \( \Theta_i \) drawn from a probability distribution with support \([0, 1]\).
- The associated multivariate probability distribution for all the agents in the graph is \( f(\theta) \).
- Each agent is \( i \in V \) assigned a threshold \( \theta_i \). Let’s define the threshold profile of agents as \( \theta := (\theta_i)_{i \in V} \). A network \( G_\theta \) is a graph endowed with a threshold profile.
Preliminaries

- Simple, undirected graph $G(V, E)$.
- A *spreading threshold* for agent $i$ is a random variable $\Theta_i$ drawn from a probability distribution with support $[0, 1]$.
- The associated multivariate probability distribution for all the agents in the graph is $f(\theta)$.
- Each agent is $i \in V$ assigned a threshold $\theta_i$. Let’s define the threshold profile of agents as $\theta := (\theta_i)_{i \in V}$. A *network* $G_\theta$ is a graph endowed with a threshold profile.
Seeding

- Two firms: $A$ spreads rumor $a$ and $B$ spreads rumor $b$. Rumors are incompatible.
- The state of agent $i$ at time $t$ is denoted $x_i(t) \in \{0, a, b\}$.
- Denote by $S^A_t(G_\theta)$ and $S^B_t(G_\theta)$ the sets of new spreaders of rumor $a$ and $b$ in network $G_\theta$ at time $t$ resp.
- At time $t = 0$, $x_i(0) = 0$ for all $i$, and each firm simultaneously chooses one agent $S^A_0, S^B_0 \in V$ as a seed for their rumor. Overlap in seed sets resolved randomly.
Seeding

- Two firms: A spreads rumor $a$ and $B$ spreads rumor $b$. Rumors are incompatible.
- The state of agent $i$ at time $t$ is denoted $x_i(t) \in \{0, a, b\}$.
- Denote by $S^A_t(G_\theta)$ and $S^B_t(G_\theta)$ the sets of new spreaders of rumor $a$ and $b$ in network $G_\theta$ at time $t$ resp.
- At time $t = 0$, $x_i(0) = 0$ for all $i$, and each firm simultaneously chooses one agent $S^A_0, S^B_0 \in V$ as a seed for their rumor. Overlap in seed sets resolved randomly.
Seeding

- Two firms: A spreads rumor $a$ and $B$ spreads rumor $b$. Rumors are incompatible.
- The state of agent $i$ at time $t$ is denoted $x_i(t) \in \{0, a, b\}$.
- Denote by $S_t^A(G_\theta)$ and $S_t^B(G_\theta)$ the sets of new spreaders of rumor $a$ and $b$ in network $G_\theta$ at time $t$ resp.
- At time $t = 0$, $x_i(0) = 0$ for all $i$, and each firm simultaneously chooses one agent $S_0^A, S_0^B \in V$ as a seed for their rumor. Overlap in seed sets resolved randomly.
Seeding

- Two firms: A spreads rumor $a$ and B spreads rumor $b$. Rumors are incompatible.
- The state of agent $i$ at time $t$ is denoted $x_i(t) \in \{0, a, b\}$.
- Denote by $S_t^A(G_\theta)$ and $S_t^B(G_\theta)$ the sets of new spreaders of rumor $a$ and $b$ in network $G_\theta$ at time $t$ resp.
- At time $t = 0$, $x_i(0) = 0$ for all $i$, and each firm simultaneously chooses one agent $S_0^A, S_0^B \in V$ as a seed for their rumor. Overlap in seed sets resolved randomly.
Linear threshold process dynamics

- Any agent who has not spread any rumor by some period \( t \), decides to spread a rumor in time period \( t + 1 \) iff
  \[
  \frac{\# \text{ friends who adopted } a + \# \text{ friends who adopted } b}{\# \text{ friends}} \geq \theta_i
  \]
  i.e. Granovetter’s linear threshold model.

- If the threshold is reached, the probability of spreading rumor \( a \) is
  \[
  \frac{\# \text{ friends who adopted } a \text{ at } t}{\# \text{ friends who adopted } a \text{ at } t + \# \text{ friends who adopted } b \text{ at } t}
  \]

- Agents use the latest spreaders to select the rumor, but total spreaders to decide to spread.

- Once an agent spreads rumor \( a \), he remains in state \( a \) in all subsequent periods.

- This process converges to a random set: eventual spreaders \( S^A \) of rumor \( a \) and \( S^B \) of rumor \( b \).
Linear threshold process dynamics

- Any agent who has not spread any rumor by some period $t$, **decides to spread a rumor** in time period $t + 1$ iff
  \[
  \frac{\# \text{ friends who adopted } a + \# \text{ friends who adopted } b}{\# \text{ friends}} \geq \theta_i
  \]
  i.e. Granovetter’s linear threshold model.

- If the threshold is reached, the **probability of spreading rumor** $a$ is
  \[
  \frac{\# \text{ friends who adopted } a \text{ at } t}{\# \text{ friends who adopted } a \text{ at } t + \# \text{ friends who adopted } b \text{ at } t}
  \]

- Agents use the **latest** spreaders to select the rumor, but **total** spreaders to decide to spread.

- Once an agent spreads rumor $a$, he remains in state $a$ in all subsequent periods.

- This process converges to a random set: eventual spreaders $S^A$ of rumor $a$ and $S^B$ of rumor $b$. 
Linear threshold process dynamics

- Any agent who has not spread any rumor by some period $t$, decides to spread a rumor in time period $t + 1$ iff
  \[ \frac{\# \text{ friends who adopted } a + \# \text{ friends who adopted } b}{\# \text{ friends}} \geq \theta_i \]
i.e. Granovetter’s linear threshold model.

- If the threshold is reached, the probability of spreading rumor $a$ is
  \[ \frac{\# \text{ friends who adopted } a \text{ at } t}{\# \text{ friends who adopted } a \text{ at } t + \# \text{ friends who adopted } b \text{ at } t} \]

- Agents use the latest spreaders to select the rumor, but total spreaders to decide to spread.

- Once an agent spreads rumor $a$, he remains in state $a$ in all subsequent periods.

- This process converges to a random set: eventual spreaders $S^A$ of rumor $a$ and $S^B$ of rumor $b$. 
Linear threshold process dynamics

- Any agent who has not spread any rumor by some period $t$, decides to spread a rumor in time period $t + 1$ iff
  \[
  \frac{\# \text{ friends who adopted } a + \# \text{ friends who adopted } b}{\# \text{ friends}} \geq \theta_i
  \]
  i.e. Granovetter’s linear threshold model.
- If the threshold is reached, the probability of spreading rumor $a$ is
  \[
  \frac{\# \text{ friends who adopted } a \text{ at } t}{\# \text{ friends who adopted } a \text{ at } t + \# \text{ friends who adopted } b \text{ at } t}
  \]
- Agents use the latest spreaders to select the rumor, but total spreaders to decide to spread.
- Once an agent spreads rumor $a$, he remains in state $a$ in all subsequent periods.
- This process converges to a random set: eventual spreaders $S^A$ of rumor $a$ and $S^B$ of rumor $b$. 
Linear threshold process dynamics

- Any agent who has not spread any rumor by some period $t$, **decides to spread a rumor** in time period $t + 1$ iff
  \[
  \frac{\# \text{ friends who adopted } a + \# \text{ friends who adopted } b}{\# \text{ friends}} \geq \theta_i
  \]
  i.e. Granovetter’s linear threshold model.

- If the threshold is reached, the **probability of spreading rumor** $a$ is
  \[
  \frac{\# \text{ friends who adopted } a \text{ at } t}{\# \text{ friends who adopted } a \text{ at } t + \# \text{ friends who adopted } b \text{ at } t}
  \]

- Agents use the **latest** spreaders to select the rumor, but **total** spreaders to decide to spread.

- Once an agent spreads rumor $a$, he remains in state $a$ in all subsequent periods.

- This process converges to a random set: eventual spreaders $S^A$ of rumor $a$ and $S^B$ of rumor $b$. 
Expected number of spreaders

- Fixing seeds $S^A_0$ and $S^B_0$ and a graph $G$, and re-run the process by drawing the agents’ thresholds from $f(\theta)$ each time.
- Denoting the probability of any agent spreading rumor $a$ is

$$P^A_i(G, S^A_0, S^B_0) = \int_{\mathbb{R}^n} |S^A(G_\theta, S^A_0, S^B_0) \cap \{i\}| f(\theta) d\theta$$

- Expected number of spreaders of rumor $a$ is

$$E[S^A(G, S^A_0, S^B_0)] = \int_{\mathbb{R}^n} |S^A(G_\theta, S^A_0, S^B_0)| f(\theta) d\theta$$

$$= \sum_{i=1}^{n} P^A_i(G, S^A_0, S^B_0)$$
Expected number of spreaders

- Fixing seeds $S_A^0$ and $S_B^0$ and a graph $G$, and re-run the process by drawing the agents’ thresholds from $f(\theta)$ each time.
- Denoting the probability of any agent spreading rumor $a$ is

$$P^A_i(G, S_A^0, S_B^0) = \int_{\mathbb{R}^n} |S^A(G_\theta, S_A^0, S_B^0) \cap \{i\}| f(\theta) d\theta$$

- Expected number of spreaders of rumor $a$ is

$$\mathbb{E}[S^A(G, S_A^0, S_B^0)] = \int_{\mathbb{R}^n} |S^A(G_\theta, S_A^0, S_B^0)| f(\theta) d\theta$$

$$= \sum_{i=1}^{n} P^A_i(G, S_A^0, S_B^0)$$
Expected number of spreaders

- Fixing seeds $S_0^A$ and $S_0^B$ and a graph $G$, and re-run the process by drawing the agents’ thresholds from $f(\theta)$ each time.
- Denoting the probability of any agent spreading rumor $a$ is

$$P_i^A(G, S_0^A, S_0^B) = \int_{\mathbb{R}^n} |S^A(G_\theta, S_0^A, S_0^B) \cap \{i\}| f(\theta) d\theta$$

- Expected number of spreaders of rumor $a$ is

$$\mathbb{E}[S^A(G, S_0^A, S_0^B)] = \int_{\mathbb{R}^n} |S^A(G_\theta, S_0^A, S_0^B)| f(\theta) d\theta$$

$$= \sum_{i=1}^{n} P_i^A(G, S_0^A, S_0^B)$$
Uniform thresholds

Assumption

For any $G_\theta$ and every $i \in V$, $\Theta_i \sim U(0, 1)$ and independent.
Cascade centrality

**Definition**

Cascade centrality of node $i$ in graph $G$ is the expected number of spreaders of rumor $a$ in that graph given $i$ is the seed and firm $A$ is a monopolist, namely

$$C_i(G) := E[S^A(G, \{i\})] = 1 + \sum_{j \in V \setminus \{i\}} P_j^A(G, \{i\}) = 1 + \sum_{j \in V \setminus \{i\}} \sum_{P \in P_{ij}} \frac{1}{\chi_P}$$
The cascade centrality of any node $i$ in $G$ is:

$$C_i(G) = 1 + d_i - \sum_{j \in V \setminus \{i\}} \sum_{L \in \mathcal{L}_{ij}} \frac{1}{\chi_L}$$

where $\chi_L$ is the degree sequence product along a loop.
Game: uniform thresholds

- Action space of firms $A$ and $B$: $\Sigma := \Sigma_A \times \Sigma_B := V \times V$
- Action profile $\sigma := (\sigma_A, \sigma_B)$ is simply a pair of nodes.
- Payoff profile: $\pi := (\pi_A(\sigma), \pi_B(\sigma))$ is the expected number of spreaders of rumors $a$ and $b$. 
Game: uniform thresholds

- Action space of firms $A$ and $B$: $\Sigma := \Sigma_A \times \Sigma_B := V \times V$
- Action profile $\sigma := (\sigma_A, \sigma_B)$ is simply a pair of nodes.
- Payoff profile: $\pi := (\pi_A(\sigma), \pi_B(\sigma))$ is the expected number of spreaders of rumors $a$ and $b$. 
Game: uniform thresholds

- Action space of firms $A$ and $B$: $\Sigma := \Sigma_A \times \Sigma_B := V \times V$
- Action profile $\sigma := (\sigma_A, \sigma_B)$ is simply a pair of nodes.
- Payoff profile: $\pi := (\pi_A(\sigma), \pi_B(\sigma))$ is the expected number of spreaders of rumors $a$ and $b$. 
Game: uniform thresholds

- For $i \neq j$, let us denote $\Xi(i, j)$ as the set of all paths that begin at $i$ and include (but do not necessarily end) at $j$.

**Proposition**

The expected number of spreaders of rumor $a$ (i.e. firm A’s payoff) is

$$
\pi_A(\sigma_A, \sigma_B) = \begin{cases} 
\frac{C_{\sigma_A}}{2} & \text{if } \sigma_A = \sigma_B \\
C_{\sigma_A} - \epsilon(\sigma_A, \sigma_B) & \text{if } \sigma_A \neq \sigma_B
\end{cases}
$$

where

$$
\epsilon(i, j) = \sum_{P \in \Xi(i, j)} \frac{1}{\chi_P}
$$

It’s a constant-sum game.
The game is defined as \( \Gamma := (\Sigma, \pi) \).

**Definition**

A profile of actions \( \sigma^* := (\sigma_A^*, \sigma_B^*) \in \Sigma \) is a pure-strategy Nash equilibrium if:

- \( \pi_A(\sigma_A^*, \sigma_B^*) \geq \pi_A(\sigma_A, \sigma_B^*) \) for all actions \( \sigma_A \in \Sigma_A \)
- \( \pi_B(\sigma_A^*, \sigma_B^*) \geq \pi_B(\sigma_A^*, \sigma_B) \) for all actions \( \sigma_B \in \Sigma_B \)

Define \( \Sigma^* \) as the set of all pure-strategy Nash equilibria.
The game is defined as $\Gamma := (\Sigma, \pi)$.

**Definition**

A profile of actions $\sigma^* := (\sigma^*_A, \sigma^*_B) \in \Sigma$ is a pure-strategy Nash equilibrium if:

- $\pi_A(\sigma^*_A, \sigma^*_B) \geq \pi_A(\sigma_A, \sigma^*_B)$ for all actions $\sigma_A \in \Sigma_A$
- $\pi_B(\sigma^*_A, \sigma^*_B) \geq \pi_B(\sigma^*_A, \sigma_B)$ for all actions $\sigma_B \in \Sigma_B$

Define $\Sigma^*$ as the set of all pure-strategy Nash equilibria.
Theorem

Consider a duopoly with unit budgets $\Gamma$. Then $\Gamma$ admits at least one PSNE if and only if either:

1. There exists $i \in V$ such that, for any $j \in V \setminus \{i\}$:
   \[
   \frac{C_i}{C_j} \geq 2 - 2 \cdot \left( \frac{\epsilon(j, i)}{C_j} \right)
   \]
   then there exists a $\sigma^* = (i, i)$ PSNE, or...
PSNE: existence characterization

**Theorem**

Consider a duopoly with unit budgets $\Gamma$. Then $\Gamma$ admits at least one PSNE if and only if either Condition 1 is satisfied or

2. There exist $i, j \in V$ such that, $C_i \geq C_j$ and for any $k \in V \setminus \{i, j\}$:

\[
\frac{C_i}{C_k} \geq 1 + \frac{\epsilon(i, j) - \epsilon(k, j)}{C_k}
\]

\[
\frac{C_j}{C_k} \geq 1 + \frac{\epsilon(j, i) - \epsilon(k, i)}{C_k}
\]

\[
\frac{1}{2} + \frac{\epsilon(i, j)}{C_j} \leq \frac{C_i}{C_j} \leq 2 - 2 \cdot \left(\frac{\epsilon(j, i)}{C_j}\right)
\]

in which case there exists a $\sigma^* = (i, j)$ (and $\sigma^* = (j, i)$ by symmetry) PSNE.
Budget multiplier

**Definition**
For arbitrary integer budgets $\mathcal{B}_A$ and $\mathcal{B}_B$, the budget multiplier is defined as:

$$BM(\Gamma) := \max_{\sigma \in \Sigma^*} \frac{\pi_A(\sigma)/\pi_B(\sigma)}{\mathcal{B}_A/\mathcal{B}_B}$$

**Theorem**
*For any $\Gamma$ that admits at least one PSNE,*

$$1 \leq BM < 2$$
Price of anarchy

Social planner’s objective: \( Y(\sigma) := \pi_A(\sigma) + \pi_B(\sigma) \) (i.e. firms’ total payoffs).

**Definition**

Price of Anarchy is defined as:

\[
\text{PoA}(\Gamma) = \frac{\max_{\sigma \in \Sigma} Y(\sigma)}{\min_{\sigma \in \Sigma^*} Y(\sigma)}
\]

**Theorem**

For any \( \Gamma \) that admits at least one PSNE,

\[ 1 \leq \text{PoA}(\Gamma) < 1.5 \]
Proposition

Suppose $G$ is a tree. Then $\Gamma$ admits at least one PSNE.
Trees

We provide necessary and sufficient conditions only on the largest and second largest degree of the trees such that:

- all PSNEs are efficient
- no PSNEs are efficient
- at least one PSNE is efficient
- at least one PSNE is inefficient
- there is at least one efficient and one inefficient PSNE
Conclusions

- Using a new notion of *cascade centrality*, we analyzed a tractable cascade process on general networks.
- We applied these insights to studying competitive diffusion.
- The competition model can be extended in a few ways (different strength of rumors, sequential entry).
Conclusions

- Using a new notion of *cascade centrality*, we analyzed a tractable cascade process on general networks.
- We applied these insights to studying competitive diffusion.
- The competition model can be extended in a few ways (different strength of rumors, sequential entry).
Conclusions

- Using a new notion of *cascade centrality*, we analyzed a tractable cascade process on general networks.
- We applied these insights to studying competitive diffusion.
- The competition model can be extended in a few ways (different strength of rumors, sequential entry).
Future research questions

- Multiple seeds
  - Hard because you need to figure out the optimal seeding problem first.

- Sequential entry
  - A little easier, but less obvious how to set up.
Future research questions

- Multiple seeds
  - Hard because you need to figure out the optimal seeding problem first.

- Sequential entry
  - A little easier, but less obvious how to set up.
References I


