ECONOMICS & COMPUTATION

Harvard CS 1360, Spring 2025
Instructor: Ariel Procaccia
Scribe: Shi Feng

Social Networks 1: Cascade Models

Lecture 17

In this lecture, we introduce cascade models in social networks, which describe how ideas or behaviors spread through populations. Understanding cascade dynamics helps explain real-world phenomena, such as political movements gaining momentum, rapid adoption of technological innovations, or viral success of new products. Typically, cascades start from a small set of early adopters and propagate through social connections, influencing individuals across the network. Our goal is to study these processes to predict diffusion patterns and effectively encourage or analyze the spread of behaviors within society.

We will first introduce coordination games on graphs, characterizing their pure Nash equilibria using the concept of cohesiveness. Next, we study a progressive cascade model, proving conditions under which contagion occurs. Finally, we explore contagion thresholds in infinite graphs.

1 Coordination Games

Consider an undirected, connected graph $G = (V, E)$, with a set of players $V = \{1, \ldots, n\}$. Each player $i \in V$ chooses an action $a_i \in \{0, 1\}$, forming an action profile $\mathbf{a} = (a_1, \ldots, a_n)$.

Let N_i denote the neighborhood of player i, and let the degree of player i be $d_i = |N_i|$. For $b \in \{0, 1\}$, denote by $n_{i,b}(\mathbf{a}_{-i})$ the number of neighbors of player i choosing action b. Given a parameter $q \in [0, 1]$, the utility for player i is defined as:

$$u_i(\mathbf{a}) = \begin{cases} (1-q) \cdot n_{i,1}(\mathbf{a}_{-i}), & \text{if } a_i = 1, \\ q \cdot n_{i,0}(\mathbf{a}_{-i}), & \text{if } a_i = 0. \end{cases}$$

We first consider simultaneous-move coordination games. The best response of player i is to choose action $a_i = 1$ if and only if at least a q-fraction of neighbors choose action 1. Formally, player i's best response is 1 if and only if

$$0 \leq (1-q) \cdot n_{i,1}(\mathbf{a}_{-i}) - q \cdot n_{i,0}(\mathbf{a}_{-i}) = n_{i,1}(\mathbf{a}_{-i}) - q \cdot d_i.$$

Poll 1: How many pure Nash equilibria are guaranteed to exist in a coordination game?

Answer: Two equilibria always exist.

Proof. We will show that there are always at least two pure-strategy Nash equilibria in any coordination game defined as above. Consider two action profiles:

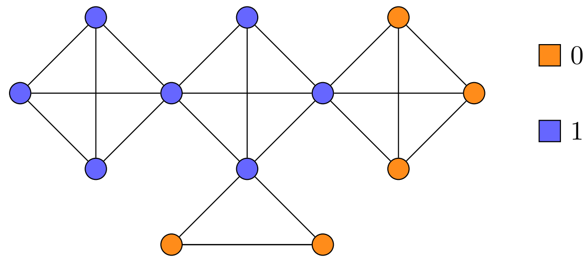
1. The profile where every player chooses $a_i = 1$: Here, each player has all neighbors choosing action 1, trivially satisfying the best-response condition $n_{i,1}(\mathbf{a}_{-i}) = d_i \geq qd_i$.
2. The profile where every player chooses $a_i = 0$: Similarly, each player sees all neighbors choosing action 0, and thus no player exceeds the threshold qd_i to switch to action 1. Hence, action 0 is a best response for all players.

Now, let's discuss why there are not necessarily three equilibria. Suppose we consider any third equilibrium. This would require some subset of players choosing 1 and another subset choosing 0 simultaneously. However, such an intermediate equilibrium is not guaranteed to exist for all graphs or all choices of q, because intermediate profiles strongly depend on graph structure and threshold q. Thus, only the two extreme equilibria ("all zeros" and "all ones") are guaranteed in general.

We conclude that at least these two equilibria exist, and the answer to the poll is indeed "2". $\square$

1.1 Coordination Games: Example

Consider the following graph, illustrating a Nash equilibrium for $q = 1/2$:



1.2 Cohesiveness: Characterization of Equilibria

The *cohesiveness* of a set $L \subseteq V$ in graph G measures how tightly connected the nodes in L are. Formally:

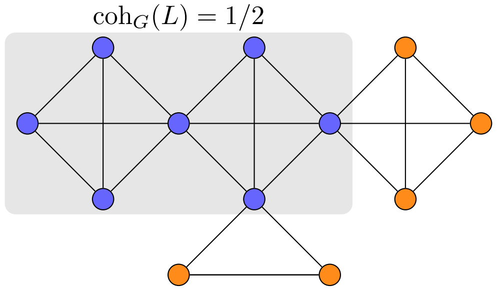
$$\mathrm{coh}_G(L) = \min_{i \in L} \left(\frac{|N_i \cap L|}{|N_i|} \right). \tag{1}$$

By convention, we define $\mathrm{coh}_G(\emptyset) = 1$.

In the previous example, consider the set L comprising the nodes playing action 1 (blue nodes). The cohesiveness is calculated as:

$$\mathrm{coh}_G(L) = 1/2$$

by considering the rightmost blue node. This means every player in L has at least half their neighbors within the set L:



Thus, cohesiveness provides a useful characterization of Nash equilibria in coordination games, indicating stable action profiles where players have no incentive to deviate.

Theorem 1. *An action profile* $\mathbf{a}$ *is a pure-strategy Nash equilibrium if and only if the sets*

$$X_0 = \{i \in V : a_i = 0\} \quad \textit{and} \quad X_1 = \{i \in V : a_i = 1\}$$

satisfy

$$coh_G(X_0) \geq 1 - q \quad \textit{and} \quad coh_G(X_1) \geq q.$$

Proof. Consider players in X_1. Their choice of action 1 is a best response if and only if at least a fraction q of their neighbors also choose action 1. Formally, for each player $i \in X_1$:

$$\frac{|N_i \cap X_1|}{|N_i|} \geq q.$$

Similarly, a symmetric argument applies to players in X_0. Each player in X_0 chooses action 0 as a best response if and only if at least a fraction $1 - q$ of their neighbors also choose action 0:

$$\frac{|N_i \cap X_0|}{|N_i|} \geq 1 - q.$$

Thus, the cohesiveness conditions exactly characterize pure-strategy Nash equilibria. □

2 Cascade Model

We now consider a progressive *cascade model*, where initially a subset of nodes (called *seeds*) adopt action 1, while all others start with action 0. This process is *progressive*, meaning that nodes can only switch from action 0 to action 1, and never back.

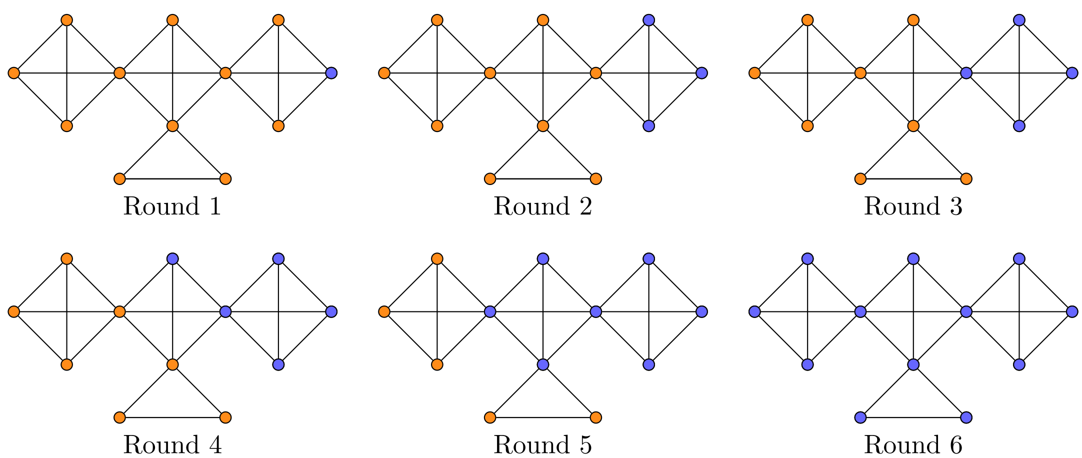
We introduce the following terminology:

- A player currently playing action 1 is called *active*.
- A player switching from action 0 to action 1 is said to be *activated*.

The cascade evolves in discrete rounds. In each round, every inactive player who has at least a fraction q of active neighbors becomes activated and thus switches to action 1.

Consider an example of the cascade model with threshold $q = 1/3$. We illustrate how activation propagates through the network:

- *Round 1:* Initially, one node (seed) is active, and all other nodes are inactive.
- *Round 2:* Nodes having at least 1/3 of neighbors active become activated.
- *Round 3–5:* Activation progressively spreads as each inactive node reaches the threshold due to previously activated neighbors.
- *Round 6:* Eventually, all nodes become active, demonstrating full contagion (every agent is active) through the network.



2.1 Contagion

We say *contagion* occurs if eventually every agent in the network becomes active.

Theorem 2. *A seed set S causes contagion in a graph G if and only if $coh_G(L) \leq 1-q$, for every $L \subseteq V \setminus S$.*

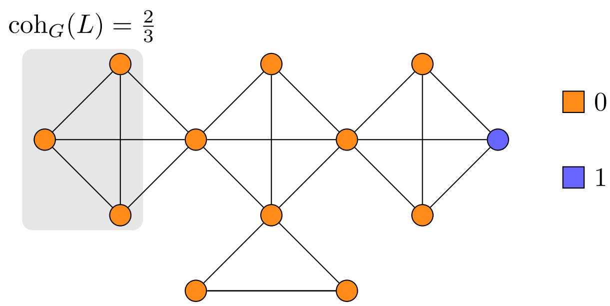
Proof. We prove both directions explicitly:

($\Rightarrow$) Suppose that the seed set S causes contagion. Assume for contradiction that there exists a subset $L \subseteq V \setminus S$ such that $\text{coh}_G(L) > 1 - q$. By definition of cohesiveness, this implies that each node in L has strictly more than a fraction $1 - q$ of its neighbors within L, and therefore fewer than a fraction q neighbors outside L. Thus, no node in L can ever meet the threshold of active neighbors to become activated first. This contradicts the assumption of contagion. Therefore, we must have $\text{coh}_G(L) \leq 1-q$ for every subset $L \subseteq V \setminus S$.

($\Leftarrow$) Suppose that $\text{coh}_G(L) \leq 1 - q$ for every $L \subseteq V \setminus S$. Consider any intermediate step of the cascade where some nodes remain inactive. Let $X_0 \subseteq V \setminus S$ be the set of currently inactive nodes. Since $\text{coh}_G(X_0) \leq 1-q$, there exists at least one node in X_0 that has at most a fraction $1-q$ of its neighbors still inactive, hence at least a fraction q of its neighbors are active. Thus, this node meets the activation threshold and becomes activated in this round. Consequently, the activation process continues at every step until no inactive node remains, resulting in full contagion.

This completes the proof. □

Consider the following example, with an initial seed node colored blue (action 1): The subset L, highlighted below, has cohesiveness: $\text{coh}_G(L) = 2/3$, and no subset of $V \setminus S$ has higher cohesiveness. Thus, contagion will occur starting from this single seed if and only if: $q \leq 1/3$. If $q > 1/3$, contagion does not fully propagate through the network.



2.2 Infinite Graphs

We now extend our analysis to infinite graphs. Assume the set of nodes V is countably infinite and the degree of each node d_i is bounded.

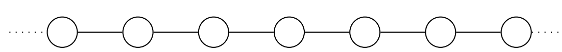
A finite set of seeds $S \subseteq V$ is called *contagious* if contagion occurs, meaning eventually all nodes become active starting from this set. Intuitively, contagion becomes easier as the threshold q is small. Thus, we define the *contagion threshold* of a graph G as the maximum value of q for which there exists a finite contagious set:

$$\text{contagion threshold of } G = \max\{q : \exists \text{ finite contagious set}\}.$$

We consider two examples of infinite graphs to understand contagion thresholds clearly.

2.2.1 Example 1: Infinite Path

Consider an infinite path graph where each node has exactly two neighbors (except the ends, which extend infinitely).



Poll 2: What is the contagion threshold of G?

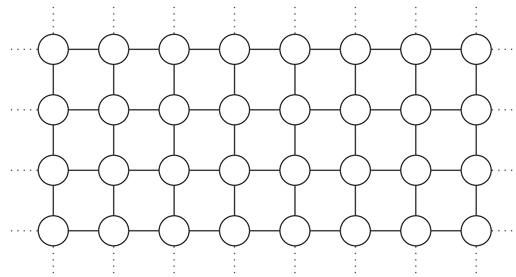
Answer: $\frac{1}{2}$

Proof. In an infinite path, each internal node has degree $d_i = 2$. To initiate contagion from a finite seed set, at least one node must become activated initially. This requires at least half of its neighbors to be active, hence $q \leq \frac{1}{2}$. Indeed, starting with a single node active, activation spreads sequentially to adjacent nodes as each inactive node encounters exactly half (1 out of 2) neighbors active.

If $q > \frac{1}{2}$, then no finite set can trigger contagion, as initially, the node next to the leftmost seed always sees at most half their neighbors active, not strictly greater. Thus, the contagion threshold is exactly $\frac{1}{2}$. □

2.2.2 Example 2: Infinite Grid (2-dimensional lattice)

Now, consider an infinite 2-dimensional lattice graph (infinite grid). Each node has exactly four neighbors.



Poll 3: What is the contagion threshold of G?

Answer: $\dfrac{1}{4}$.

Proof. Each node in the infinite grid has 4 neighbors. A node becomes active when at least $q \cdot 4$ of its neighbors are active. Consider the following:

- When $q = \frac{1}{4}$, a single active neighbor (since $\frac{1}{4} \cdot 4 = 1$) is sufficient to eventually trigger the activation of a neighboring node. By arbitrarily selecting an appropriate seed node set (for example, a single node or a small cluster), the activation will spread outward.
- If $q > \frac{1}{4}$, then a node would require at least 2 active neighbors (because $q \cdot 4 > 1$) to activate. Imagine a rectangle around the set of seeds. The contagion cannot expand beyond this rectangle because nodes on the boundary will not reach the higher activation requirement.

Thus, the maximum value of q that allows a finite seed to cause infinite contagion is $q = \frac{1}{4}$. This proves that the contagion threshold for the infinite 2D grid is exactly $\frac{1}{4}$. □

Contagion Threshold

We have seen an example of a graph whose contagion threshold is $\frac{1}{2}$. A natural question arises:

Does there exist a graph with contagion threshold $> \frac{1}{2}$?

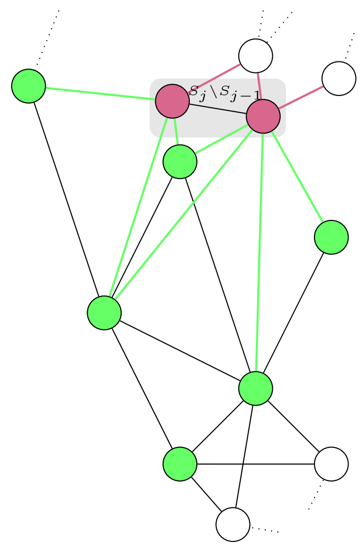
The following theorem provides a negative answer.

Theorem 3. *The contagion threshold of any graph G is at most $\frac{1}{2}$.*

Proof. Let $q > \frac{1}{2}$ and a finite seed set $S \subseteq V$. Furthermore, let S_j denote the set of active nodes at the end of round j. For any set $X \subseteq V$, define

$$\delta(X) \;=\; \{\, (u,v) \in E \;\mid\; u \in X,\; v \notin X \} \quad \text{(the } \textit{boundary} \text{ of } X\text{)}.$$

We prove that if $S_{j-1} \neq S_j$, then $|\delta(S_j)| < |\delta(S_{j-1})|$. To see why, consider a newly activated node $i \in S_j \setminus S_{j-1}$. By assumption, i was activated when it observed that more than half of its neighbors (because $q > \frac{1}{2}$) were already in S_{j-1}. Hence, i has strictly more edges going *into* S_{j-1} than into $V \setminus S_j$. As a result, when i is added to the active set, the total boundary $\delta(S_j)$ decreases by at least one edge compared to $\delta(S_{j-1})$.



Because $\delta(S)$ for a finite set S is finite, and each round strictly reduces the boundary size whenever new nodes join, we cannot keep activating new nodes indefinitely (the boundary cannot shrink below zero). Hence, the process must stabilize at a finite set of active nodes. $\square$