

DIFFUSIONS OF INNOVATIONS ON DETERMINISTIC TOPOLOGIES

Ercan Yildiz^{*} Daron Acemoglu[‡] Asu Ozdaglar[†] Anna Scaglione^{*†}

^{*} MIT LIDS, meyildiz@mit.edu

[‡] MIT ECON, daron@mit.edu

[†] MIT LIDS, asuman@mit.edu

^{*†} UC-Davis ECE, ascaglione@ucdavis.edu

ABSTRACT

In this paper, we are interested in modeling diffusion of innovations on social networks. We focus on a scenario where innovation emerges at a small number of nodes in the society, and each individual needs certain portion of his neighbors (thresholds) to adopt the innovation before he does so. We analyze the dynamics of the diffusion process under deterministic topologies and threshold values. We show that the diffusion process depends on both the topology and threshold values through so called *coherent* sets. We investigate several interesting topologies, and utilize *coherent* set argument to determine the behavior of the process on these topologies.

1. INTRODUCTION

Modeling diffusion of innovations through societies has been proven to be a challenging task. In this work, we focus on a particular scenario where the innovation emerges at a collection of nodes (seeds) at time $k = 0$, and propagates through the network via local interactions among neighbors. Our local interaction rule will be based on the *linear threshold model* by Granovetter [1], where each individual needs certain portion of his neighbors to adopt the innovation before he does so. Once an individual decides to adopt the innovation, he will stay as an adopter, *i.e.*, he can not abandon the innovation.

Linear threshold models have been discussed in detail by Granovetter in his study on collective behavior of societies [1]. In his seminal work, Granovetter discusses that threshold models can be used to explain several aggregate level behaviors such as diffusion of innovations, spread of rumors and diseases, riot behaviors, and migration. For instance, he focuses on a society with 100 people, where there exists one individual with threshold 0, one with threshold 1, one with threshold 2, and so on up to the last individual with threshold 100. He discusses that, under the linear threshold model, we would observe a *domino* effect in this society, *i.e.*, the individual with threshold 0 initiates the riot and activates the individual with threshold 1, then the individual with threshold 2 is activated, so on and so forth. At the end of the day, we will end up with a riot of 100 people. Granovetter discusses that the aggregate behavior of this particular model coincides with what sociologists observe in general: While individuals do not engage in certain activities in isolation, as a part of the society they might do so. Therefore, he concludes that threshold models in general might be a suitable model for explaining collective behaviors of the society in terms of the individuals' preferences. Similar arguments have also been made by Schelling where he discusses how personal motives may result in unintended aggregate level behaviors [2].

The dynamics of the linear threshold model has been studied in [3] by Watts, where the author utilizes branching process to model the behavior of the diffusion on random networks and under random thresholds, and determine the phase transitions as well as expected size of the final adopters. While the results are quite interesting, the branching based analysis does not capture collective effects in one's neighborhood, since it is based on the fact that each node has a single neighbor (parent) that exerts influence on him. The linear threshold models have also been discussed by Kleinberg *et al.* in [4, 5] under deterministic topologies and random thresholds. The authors focus on determining the most influential nodes under this particular model and show that the optimization problem is submodular.

In this paper, we will study the diffusion characteristics of the linear threshold model on deterministic topologies and deterministic thresholds values. We will propose a topology based rule for determining the final adopter set for a given network, threshold values, and initial seed set of the innovation. We will investigate the behavior of the algorithm on certain network topologies.

2. LINEAR THRESHOLD MODEL

In this study, we consider a society S and its corresponding network representation $G(\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of individuals (agents, vertices, nodes) and \mathcal{E} is the set of links (edges) among these individuals. The underlying network is deterministic. We define the neighbor set of agent $i \in \mathcal{V}$ as $\mathcal{N}_i = \{j | (i, j) \in \mathcal{E}\}$. We note that due to undirected nature of the edges, $j \in \mathcal{N}_i$ if and only if $i \in \mathcal{N}_j$.

The state of the model at iteration $k \geq 0$ is given by a function $x(k) : \mathcal{V} \rightarrow \mathcal{S}^{|\mathcal{V}|}$, where \mathcal{S} is the set of possible discrete types. Since there are two possible types, *i.e.*, adopter $\{1\}$, and non-adopter $\{0\}$, the set $\mathcal{S} = \{1, 0\}$. Moreover, each agent i has a threshold (tolerance) value $\phi_i \in (0, 1]$, where individual i needs at least ϕ_i portion of his neighbors to be adopters before adopting the innovation. Denote $\Phi(k)$ as the set of agents who adopt the product at iteration k .

We assume that at iteration $k = 0$, a subset of individuals $\Phi(0) \subset \mathcal{V}$ is selected as the seed of the innovation. In the next iteration, a node $i \in \mathcal{V} \setminus \Phi(0)$ will adopt the product if at least ϕ_i percent of his neighbors are adopters, *i.e.*,

$$\frac{|\Phi(0) \cap \mathcal{N}_i|}{|\mathcal{N}_i|} \geq \phi_i \Rightarrow i \in \Phi(1). \quad (1)$$

For a given $k \geq 1$, the above discussion can be generalized as fol-

Anna Scaglione thanks NSF CCF-1011811 grant.

lows: a node $i \in \mathcal{V} \setminus \bigcup_{l=0}^{k-1} \Phi(l)$ adopts the product at time k if,

$$\frac{|\bigcup_{l=0}^{k-1} \Phi(l) \cap \mathcal{N}_i|}{|\mathcal{N}_i|} \geq \phi_i, \quad (2)$$

where $\{\mathcal{V} \setminus \bigcup_{l=0}^{k-1} \Phi(l)\}$ denotes the set of individuals who have not yet adopted the innovation before iteration k , and $\bigcup_{l=0}^{k-1} \Phi(l)$ denotes the set of adopters at beginning of iteration k . We note that nodes satisfying (2) form the set $\Phi(k)$.

In this work, we are interested in the relationship between the behavior of the society, *i.e.*, evolutions of $\Phi(k)$ as well as $\bigcup_{l=0}^k \Phi(l)$, and the underlying connectivity \mathcal{G} , threshold values ϕ in the time limit, $k \rightarrow \infty$.

We will utilize the following assumptions in the rest of the paper.

- A1) The number of nodes in the network $|\mathcal{V}| = N$ is finite.
- A2) The underlying topology G is strongly connected, *i.e.*, there exists a path (not necessarily single hop) between any given node pairs.
- A3) The seed set $\Phi(0)$ is non-empty.

We note that the second assumption simply states that *no man is an island*, and the last assumption states that the innovation will be initiated at a non-empty subset of the society.

3. MAIN RESULTS

In this section, we will investigate the behavior of the diffusion process in terms of the network characteristics and threshold values.

We first introduce the following corollary which follows from the diffusion model given in (1)-(2).

Corollary 1. *The diffusion model in (1)-(2) will terminate at iteration $k \geq 0$, if and only if either the set $\{\mathcal{V} \setminus \bigcup_{l=0}^{k-1} \Phi(l)\}$ is empty, or the set $\Phi(k)$ in (2) is empty.*

The corollary simply states that the diffusion process stops when either the whole society adopts the behavior, or thresholds of the non-adopted individuals are sufficiently high. As a result of this corollary, we can also conclude the following:

Corollary 2. *The diffusion model in (1)-(2) will terminate in at most N steps.*

The corollary simply follows from the fact that if the new adopter set at time k , $\Phi(k)$, is empty, the diffusion will terminate by Corollary 1. Since there are N nodes in the network \mathcal{G} , $\Phi(k)$ can be non-empty for at most N iterations.

While the diffusion process is guaranteed to terminate in finite time $K^* \leq N$, the limiting diffusion process we defined in Section 2 $\{\Phi(k)\}_{k=0}^{\infty}$ is still well defined since $\{\Phi(k)\}_{k=K^*+1}^{\infty} = \emptyset$.

In the following, we will first introduce two definitions, and then, a necessary and sufficient condition for fixed points in terms of the underlying graph properties and the threshold values.

Definition 1. *For a given society \mathcal{G} and threshold values ϕ ; the adopter set Φ^* is defined to be a fixed point if;*

$$\Phi(0) = \Phi^* \Rightarrow \Phi(k) = \emptyset, \quad \forall k > 0. \quad (3)$$

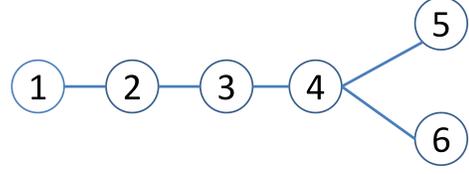


Fig. 1. A sample network with $\phi_i = \frac{1}{2} - \epsilon, \forall i \in \{1, 2\}$ and $\phi_i = \frac{1}{2} + \epsilon, \forall i \in \{3, 4, 5, 6\}$.

Definition 2. *A nonempty subset $\mathcal{M} \subset \mathcal{V}$ is called a coherent set if;*

$$\frac{|\mathcal{M} \cap \mathcal{N}_i|}{|\mathcal{N}_i|} > 1 - \phi_i, \quad \forall i \in \mathcal{M}. \quad (4)$$

Definition 2 simply states that slightly greater than $1 - \phi_i$ portion of the neighbors of node i are inside the set \mathcal{M} , for all $i \in \mathcal{M}$. In other words, members of a coherent set satisfies certain threshold on the relative frequency of links among set members compared to non-members. Definition 2 is closely related with p -cohesive set definition by Morris [6]. In his seminal work, Morris used p -cohesive sets, where at least p portion of the interactions are within the set members, to analyze the behavior of the local interaction game on a given network.

Fig. 1 is a sample topology with six individuals. We assume that threshold values are equal to $0.5 - \epsilon$ for the first two agents, and are equal to $0.5 + \epsilon$ for the last four agents, where $0 < \epsilon \ll 1$. In other words, each agent $i \in \{1, 2\}$ needs slightly less than half of his neighbors to adopt the innovation before doing so, while agents $i \in \{3, 4, 5, 6\}$ need strict majority in their neighborhood before adopting. The sample topology has multiple coherent subsets, *e.g.*, $\{3, 4, 5, 6\}$, $\{4, 5, 6\}$, $\{1, 2, 3, 4, 5\}$, $\{1, 2, 3, 4, 6\}$. For instance, for agents $\{4, 5, 6\}$, the ratio in (4) is equal to $\{2/3, 1, 1\} > 0.5 - \epsilon$.

Lemma 1. *For a given society \mathcal{G} and threshold values ϕ ; the adopter set Φ^* is a fixed point if and only if $\mathcal{V} \setminus \Phi^*$ is a coherent set or an empty set.*

Proof. We will first prove the if part of the lemma. Let's assume that $(\Phi^*)^c \triangleq \mathcal{V} \setminus \Phi^*$ is a coherent set. Therefore, by the definition of a coherent set;

$$\frac{|(\Phi^*)^c \cap \mathcal{N}_i|}{|\mathcal{N}_i|} = \frac{|\mathcal{N}_i| - |\Phi^* \cap \mathcal{N}_i|}{|\mathcal{N}_i|} > 1 - \phi_i, \quad \forall i \in (\Phi^*)^c.$$

However, above inequality implies that:

$$\frac{|\Phi^* \cap \mathcal{N}_i|}{|\mathcal{N}_i|} < \phi_i, \quad \forall i \in (\Phi^*)^c. \quad (5)$$

By (2), above equation implies that Φ^* is a fixed point.

If $(\Phi^*)^c = \emptyset$, the whole society has already adopted the innovation, thus (Φ^*) is also a fixed point.

The only if part of the proof simply follows from the definition of a fixed point and equation (2). \square

The above lemma shows that possible fixed points of the diffusion process depends on the both the underlying connectivity \mathcal{G} and the threshold values ϕ through so called *coherent sets*.

The society in Fig. 1 has several fixed points, e.g., $\{1, 2\}$, $\{1, 2, 3\}$, $\{4, 5, 6\}$, $\{5\}$, $\{6\}$ and $\{1, 2, 3, 4, 5, 6\}$. We note that for each of these sets (except the whole society), the complement set is coherent.

While above lemma introduces topological conditions for existence of fixed points for the linear threshold model, it does not clarify what would be the fixed point for a given initial seed $\Phi(0)$. The following lemma characterizes the limiting behavior of the diffusion process for a given $\Phi(0)$.

Lemma 2. For a given society \mathcal{G} , threshold values ϕ and the seed $\Phi(0)$, one of the followings holds:

1. If there exists no Φ^* such that $\Phi(0) \subseteq \Phi^*$ and $(\Phi^*)^c$ is coherent, then the innovation diffuses to the whole society.
2. If there exists a unique Φ^* such that $\Phi(0) \subseteq \Phi^*$ and $(\Phi^*)^c$ is coherent, then the set Φ^* adopts the innovation.
3. If there exists multiple subsets, i.e., $K > 1$ and $\{\Phi_k^*\}_{k=1}^K$, such that $\Phi(0) \subseteq \{\Phi_k^*\}_{k=1}^K$ and $(\{\Phi_k^*\}_{k=1}^K)^c$ are coherent, then the set:

$$\bigcap_{k=1}^K \Phi_k^*, \quad (6)$$

adopts the innovation.

Proof. 1. We first note that the seed $\Phi(0)$ has to be a subset of the final adopters. Therefore, if there does not exist any Φ^* such that $\Phi(0) \subseteq \Phi^*$ and $(\Phi^*)^c$ is coherent, then the only fixed point of the diffusion process is the whole society \mathcal{V} by Lemma 1.

2. Given the set Φ^* such that $\Phi(0) \subseteq \Phi^*$ and $(\Phi^*)^c$ is coherent, it is clear that Φ^* is a fixed point for the algorithm by Lemma 1. Moreover, since $(\Phi^*)^c$ is the unique coherent set, the only other fixed point of the process is the whole society \mathcal{V} .

We first note that if the diffusion process $\{\Phi(k)\}_{k=0}^\infty$, initiated at the seed $\Phi(0)$, hits the set Φ^* , i.e.,

$$\exists k^* \in [0, \infty), \text{ such that } \bigcup_{k=0}^{k^*} \Phi(k) = \Phi^*,$$

then the diffusion process will stop at k^* with Φ^* being the final adopter set. Let's assume that such k^* does not exist, i.e., the diffusion process $\{\Phi(k)\}_{k=0}^\infty$ hits the whole society \mathcal{V} without hitting the coherent set Φ^* . Such a claim implies that, $\exists k_1 \in [0, \infty)$ such that:

$$\begin{aligned} \bigcup_{k=0}^{k_1-1} \Phi(k) \cap (\Phi^*)^c &= \emptyset \\ \bigcup_{k=0}^{k_1} \Phi(k) \cap (\Phi^*)^c &\neq \emptyset. \end{aligned}$$

In other words, k_1 is the first iteration at which an individual $i \in (\Phi^*)^c$ adopts the innovation. We note that, $(\Phi^*)^c$ is a coherent set, therefore, at least $1 - \phi_i$ portion of the neighbors are inside the set for each individual $i \in (\Phi^*)^c$. Since $\bigcup_{k=0}^{k_1-1} \Phi(k) \cap (\Phi^*)^c = \emptyset$, i.e., no individual in $(\Phi^*)^c$ has adopted the innovation on or before time $k - 1$, there does not exist any agent in $(\Phi^*)^c$ for whom (2) can hold at time

k_1 . This is a contradiction, therefore the diffusion process has to hit Φ^* .

3. We first note that since $\Phi(0) \subseteq \{\Phi_k^*\}_{k=1}^K$, then $\Phi(0) \subseteq \bigcap_{k=1}^K \Phi_k^*$ holds true. Therefore, the intersection in (6) is non-empty. Moreover, the intersection set is also a fixed point since,

$$\left(\bigcap_{k=1}^K \Phi_k^* \right)^c = \bigcup_{k=1}^K (\Phi_k^*)^c,$$

and finite unions of coherent sets are also coherent. At this point, we need to show that the diffusion process $\{\Phi(k)\}_{k=0}^\infty$, initiated at the seed $\Phi(0)$, hits the intersection set before hitting any other fixed points. The proof of our claim follows from the exact same argument discussed in part 2, i.e., $\nexists k_1 \geq 0$ such that individual $i \in \left(\bigcap_{k=1}^K \Phi_k^* \right)^c$ adopts the innovation at time k_1 . Therefore, the process has to hit the intersection before hitting any other fixed points. □

If we choose the seed set as $\Phi(0) = \{1\}$ for our example in Fig. 1, we are in case 3 since there exists multiple fixed points (other than the whole set) of which the seed set is a subset, i.e., $\Phi(0) = \{1\} \subseteq \{1, 2\}, \{1, 2, 3\}, \{1, 2, 5\}, \{1, 2, 6\}$. Therefore, the spread of innovation stops once it reaches the intersection of these sets, i.e., the first two individuals.

4. DISCUSSIONS

In this section, we will focus on specific topologies and try to infer the behavior of the diffusion process in the light of Lemma 2. Our first example is fully connected network in Fig. 2(a), where each individual can talk to every other agent in the society. Moreover, we assume that all the nodes have a common threshold value $p \in (0, 1]$. We first note that, for a given seed set $\Phi(0)$, the complement set $(\Phi(0))^c$ will be coherent unless at most $1 - p$ portion of the neighbors are in $(\Phi(0))^c$ for each node $i \in (\Phi(0))^c$. This, in return, implies that if at least p portion of the nodes are in the seed set $\Phi(0)$, the innovation will diffuse through the whole society. Otherwise, $\Phi(0)$ is a fixed point by Lemma 1 and the innovation gets stuck in the initial set by Lemma 2. In other words, for fully connected network and initial seed set $\Phi(0)$, $p = |\Phi(0)|/N$ is the transition point for diffusion through the whole society.

Our second example is tree network where each node has two children in Fig. 2(b). If we focus on the scenario where the parent node is selected as the initial seed, then for $p \in (1/3, 1)$, the rest of the network will form a coherent set. However, for any $p \in (0, 1/3]$, we will be in case 1 of Lemma 2, thus the innovation will spread through the society. Therefore, for this particular case, $p = 1/3$ is the transition point.

Our last example will be a line network in Fig. 2(c). We will focus on the scenario where node 1 is selected as the initial seed. In this particular example, for $p \in (0, 1/2]$, the innovation will spread through the society, while, for $p \in (1/2, 1]$, only the seed will adapt the innovation. For this particular case, $p = 1/2$ is the transition point.

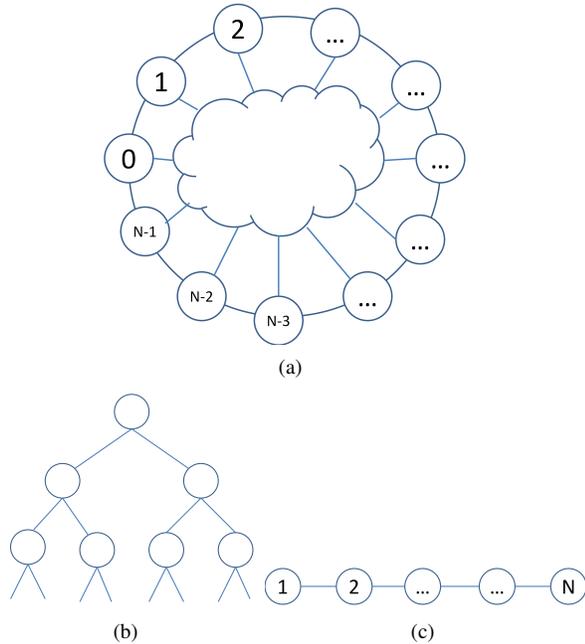


Fig. 2. Three different topologies.

If we revisit the classical example by Granovetter that we have briefly discussed in Section 1, this particular scenario is equivalent to a society with fully connected network of N nodes, where for each individual $i \in \{1, 2, \dots, N - 1\}$ the threshold value is i/N , and individual 0 is the seed. In this case, there does not exist a coherent subset \mathcal{M} of \mathcal{V} to which node 0 does not belong. The argument is simple, since node identities are integers in $[0, N]$, the set \mathcal{M} has a unique minimum (in terms of the node identities), *i.e.*, $k \leq N$. However, this implies that nodes $\{0, 1, \dots, k - 1\} \in (\mathcal{M})^c$, since, otherwise node k would not be the minimum in \mathcal{M} . Therefore, (4) does not hold for k , and \mathcal{M} is not coherent. If we repeat the argument for all $k > 0$, it will be clear that we will be in case 1 of Lemma 2.

5. CONCLUSION

In this paper, we focused on linear threshold model to model the diffusion of an innovation through a given society. Assuming the topology and individual threshold's are deterministic, we proposed a topology based method for determining the set of final adopters. The topology based condition is based on so called *coherent* sets, and is a function of both the underlying connectivity \mathcal{G} and the threshold values ϕ . We investigated the behavior of the process on certain graphs, and determine the phase transition values for thresholds at which the diffusion process spreads to the whole society from given initial seeds.

6. REFERENCES

[1] M. Granovetter, "Threshold models of collective behavior," *The American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, May 1978.

[2] T. C. Schelling, *Micromotives and Macrobehavior*. W.W. Norton and Company, 1978.

[3] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 9, pp. 5766–5771, April 2002.

[4] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003.

[5] D. Kempe, J. Kleinberg, and va Tardos, "Influential nodes in a diffusion model for social networks," in *IN ICALP*. Springer Verlag, 2005, pp. 1127–1138.

[6] S. Morris, "Contagion," *Review of Economic Studies*, no. 67, pp. 57–78, 2000.