CASCADING BEHAVIOR IN INFINITE NETWORKS

by

Alperen Yaşar Özdemir

B.S., Mathematics, Boğaziçi University, 2010

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Mathematics
Boğaziçi University

2012

ACKNOWLEDGEMENTS

I would like to thank my advisor Alp Eden for his labor and guidance. His passion has been the driving force of my research.

I owe my gratitude to Deniz Selman and Burak Gürel for their support as well as their participation in my thesis committee.

I would like to thank TÜBİTAK for providing me with scholarship.

ABSTRACT

CASCADING BEHAVIOR IN INFINITE NETWORKS

The aim of this master thesis is to analyze the underlying mathematical structure of the infinite network models of cascading behavior. Graph theoretical tools are essential to understand the structure of the network and game theoretical tools are employed for the dynamics of the model. It is tried to determine under what conditions on the structure of the graph or on the parameters of the game, cascading is possible. We also consider the optimization problem of choosing the initial set from which cascading behavior spreads through the network. For this purpose, we use the theory of submodular functions. Submodularity condition provides close approximations to the optimal value when the initial set is selected by Greedy Algorithm.

ÖZET

SONSUZ AĞLARDA SARİ DAVRANIŞ

Bu tezin amacı, sari davranışın sonsuz ağlar üzerindeki modellerinin altında yatan matematiksel yapıyı araştırmaktır. Çizge kuramsal araçlar ağ yapısını anlamak için esas olup, oyun kuramsal araçlar da modelin dinamiği için kullanılmıştır. Ağın yapısını ve oynun parametrelerini ilgilendiren hangi şartlar altında sari davranışın mümkün olduğu belirlenmeye çalışılmaktadır. Ayrıca, sari davranışa dair eniyileme problemini de gözönüne alıyoruz. Ayrıca sari davranışın ağ boyunca yayıldığı çıkış kümesini seçme probleminin eniyileme göz önünde bulunduruyouz. Bu amaçla, altmodüler fonksiyonlar kuramından faydalanıyoruz. Başlangıç kümesinin elemanları hırslı algoritma ile seçilmişse, altmodülerlik koşulu en iyi değere yakın yaklaşımlar elde etmemizi sağlıyor.

TABLE OF CONTENTS

LIST OF SYMBOLS

$\square$	End of proof
G	a graph
V(G)	The set of vertices of a graph G
E(G)	The set of edges of a graph G
X	subsets of V(G)
$\|X\|$	The cardinality of set X
$\bar{X}$	The complement of the set X in *V(G)*
2^X	The set of subsets of X
l	a labelling
$\mathcal{L}$	The set of labellings l
v_i	ith coordinate of a vertex v in a lattice $\mathbb{Z}^m$
$c(X)$	The infimum of the fraction of neighbors of vertices in a subset X.
$Prob[S]$	The probability of event S
$f(u, v)$	Influence of vertex u on vertex v
f_v	Restriction of $f(u, v)$ on the second variable
$\Gamma(v)$	The set of vertices adjacent to the vertex v
$\Gamma(X)$	The set of vertices connected to some vertex $v \in X$
$\Gamma^n(X)$	The set of vertices whose Erdös distance is less than n
$\Gamma_0(v)$	The set of adjacent vertices to a vertex playing strategy 0
$\Gamma_1(v)$	The set of adjacent vertices to a vertex playing strategy 1
$\alpha_l(n)$	The proportion of vertices in $\Gamma(v)$ with lower labels than v under labelling l
$\pi[X\|v]$	The fraction of neighbors of a vertex v in a subset X

$\Pi^p(X)$	The set of vertices for which at least p-fraction of neighbors are in X
$\Pi^p_+(X)$	The union of X and $\Pi^p(X)$
$\delta(X)$	The set of edges that have exactly one end in X
ξ	Contagion Threshold
θ_v	Threshold value for vertex v in order to adopt the desired strategy
σ	Expected number of vertices adopting the desired strategy at the end of the process

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

One of the pioneering question relevant to our topic is posed by Milgram that how many people there are in a chain of acquaintance between two random persons. Series of experiments on the structure of the network of acquaintance in the U.S. is carried out. The experiment goes in this way: At first two persons are randomly selected from U.S. population. One of them receives a document that contains the name of the other person and certain information about her. She was told to send this document to her acquaintances who are more likely than her to know the person. Then the receiver is informed about the experiment and she does the same. The experiment ends when the document is reached to the target person. The number of steps varied from 2 to 10 with median 5. It was also noted that the frequency of tightly connected communities is in direct proportion to this number. The relevant point to our topic is the demonstration of how fast a behavior may spread in large networks [1].

Two seminal works cascading behavior is discussed on theoretical ground are [2] and [3]. Banerjee prefers to use “herd behavior” instead of cascading behavior. Each member of the population sequentially chooses an asset whose return is unknown prior to decision. Before deciding on the asset, person receives a signal with some probability; and with a small probability, person is informed about the asset whose return is the largest. Then the decision is made considering the signal and the decisions of persons preceding her. It happens almost surely a cascading at an inefficient option unless the first several decision makers receive the signal of the asset with largest return [2]. In the model presented in [3], individuals decide on two options one has a payoff 1 and the other has 0, receiving a signal that can be deceptive. It is shown that the probability of cascade not to occur falls exponentially with the number of people in the population. An interesting result is that when there is a probability of a change in the payoffs of the options, cascade can switch to other option with a probability larger than the probability that the payoffs change [3]. In the thesis, we only consider deterministic models with more complex structure determining the interactions between players.

Dynamic models constructed on a large but not necessarily infinite populations where interactions are determined by a 2 by 2 coordination games are precursors to the ones with more complex network structures [4–8]. The long run Nash equilibria of these models are classified according to the distinction put forth by Harsanyi and Selten [9]. One type of Nash equilbrium is the risk dominant equilibrium where players choose the strategy that provides a payoff less susceptible to a switch in the other players' choices, in other words the best choices of players when there is no information about what others choose. The other one is the Pareto-efficient equlibrium where players choose the strategy offering at least as much as payoff as the other equilibria [9]. Throughout the thesis, we do not mention this distinction as the coordination games of the models does not distinguish between two types of equilibrium.

When the interaction rule is the matching of two random players of the population, it was shown that the constant possibilty of mutation, mutation here refers to switching strategy externally, drives population to the risk dominant equilibrium in the long run [5]. In the model, players from a finite population are matched randomly at each stage of the game. The interesting part is that the long run equilibrium does not depend on the mutation, but the payoff table of the game. Ellison develops the argument by introducing a local interaction rule that restricts players' interactions to a small group. Regardless of the interaction rule the outcome converges to the risk dominant equilibrium. In his second notable result, he derives a formula, given the magnitude of the mutation and the desired proportion of the population playing the risk dominant strategy, for the expected time to reach the desired proportion as the size of the population goes to infinity [6].

On the other hand, the payoff dominant equilibrium is shown to be achieved in the models presented in [7,8]. In the model of Galesloot and Goyal, each player is allowed to choose two actions at the same time with a learning cost for the strategy if it is the first time she plays it. At each stage, each player has a small possibility to be removed and those removed ones are replaced by new ones who choose a strategy randomly at their first round. If the cost of learning is smaller than a threshold depending on the payoffs, the game converges to the payoff dominant equilibrium in the long run [8]. Unlike the models above, In Ely's model, the interactions does not entirely follow a random matching rule. At each stage, some players are removed from the game and replaced with new ones who are allowed to choose their location in the network only once as they join in. The long run equilibrium of the game is shown to be payoff dominant equilibrium [7].

Apart from the structure of the game and the selection of the long run equilibrium, the structure of the network is another topic of investigation. In Morris' Contagion, the structure of the network is elaboratively investigated as a determinant of the cascade. Like in the earlier models mentioned, the process is deterministic. The best response dynamics lead to a long run equilibrium where a strategy initially adopted by a finite set of the population is spread to the whole population which is taken to be infinite. The spread of a strategy to the whole population is controlled by a threshold value which is defined as the maximum possible payoff of the vanishing strategy [10].

One of the result on the structure of the network as a model of a population is that the existence of tightly connected set of players increases the threshold for the cascading of a behavior among the members of a population [10]. A related phenomenon is observed earlier by Granovetter's 1983 paper in which he makes the distinction between weak ties and strong ties. Weak ties refers to acquaintance, and the strong ties stands for close friendship. Relying on empirical studies, the individuals lacking weak tie are less affected by a novelty in the population [11].

Morris also introduces two global properties for the graph to arrive at a stricter bound for the contagion threshold. First one is *low neighborhood growth*. Low neighborhood growth condition requires less than exponential growth of any chain of neighborhoods of a finite set of players. The second condition is δ-*uniformity* which refers to structural likelihood of interactions for different communities of a network [10].

The relation between graph structure and the rate of convergence to a risk dominant equilibrium is comprehensively studied by Montanari and Saberi. Best-response dynamics of their model is specified by Markov chains called Glauber dynamics for the Ising Model. The rate of convergence is determined by the hitting time to the configuration where each vertex of the graph adopts the desired strategy. The results obtained are that the convergence is faster in locally connected networks and slower in networks consisting tightly connected communities showing contrast to the ones derived from epidemic models of spread of behavior [12].

Another topic of interest is the best selection of an initial set of vertices from which the behavior spreads through the network. The problem is finding the set of given size that maximizes the number of players adopting the behavior at the end of the process. An algorithmic approach to the problem is supplied in [13]. Several models of cascading behavior in networks are used to approximate the most influential initial set. The model that is in most frequent use dealing with this problem is the Threshold Model introduced in [14]. The model describes an initial set of active players, i.e. the set of players adopting a desired strategy while the others do not. Then any inactive player is activated if the proportion of her active neighbors exceeds her threshold value. The aim is activating as many as players at the end of the process [14]. The model is adopted to more general frameworks regarding the structure of the network by Kempe and Kleinberg. It encompasses the model of [10].

In [15], the influence of an initial set is defined by a function on the subsets of the set of vertices. It gives the expected number of vertices to which the behavior spreads at the end of the process, taking the initiator set as input. The maximization problem for the influence function is shown to be NP-Hard. [15, 16]. Yet if the function is submodular, the optimal value is in close approximation. It was shown in [17] that if f is a submodular function defined on the subsets of a finite set X, at least $\left(1-\frac{1}{e}\right)$ times the optimal value among the subsets of size k is achieved by choosing the subset constructed starting from any element of X and each time adding an element that maximizes f. The influence maximization functions of the special cases of the threshold model and the cascade model are shown to be submodular in [15]. The submodularity of the cascade model is proven in [15] and the case for the threshold model is conjectured [16]. In 2007 Mossel and Roch came up with the proof of the conjecture using coupling methods [18].

The fundamental aim of the thesis is to understand the mathematical framework of [10] and treat it elaboratively. This is achieved through mostly relying on the resuls in [10, 19, 20] on deterministic models of cascading behavior in infinite networks. Then we restrict our attention to the computational problems of cascading behavior in finite networks expecting to arrive at generalizations for infinite networks. The thesis is organized as follows.

In Chapter 2, some definitions and examples for Erdös Distance, Erdös Labelling and NP-hardness are supplied.

In Chapter 3, first the model of cascading of a behavior in infinite networks is set up. Then, the results concerning the dynamics of the cascading process are presented. Finally, we discuss the conditions for the cascading to occur in relation with the game defined on the network.

Chapter 4 is about the problem of choosing the initial set from which cascading occurs through the network. In this respect different models of cascading are discussed. We give some results pertaining to the theory of submodular functions. The complexity of the problem of optimizing the initial set is the final subtopic.

2. PRELIMINARIES

We first discuss a measure of distance on graphs where the distance is defined between vertices. Secondly, a labelling rule related to the distance concept will be introduced on the set of vertices of a graph. In the remaining part, some definitions and examples belonging to the theory of NP-completeness are presented which will be referred while discussing the complexities of optimization problems on the set of vertices.

2.1. Erdös Distance

The *Erdös distance* is defined between a vertex v of graph G and a subset X of the set of vertices of G, which will be denoted by V(G), as the length of the shortest path from v to X if there exists any, otherwise is equal to ∞.

Define inductively $\Gamma^n(X)$ for the set of vertices whose Erdös distance to X is less than or equal to n.

$$\Gamma^0(X) = X, \tag{2.1}$$

$$\Gamma^{n+1}(X) = \Gamma^n(X) \cup \{v' : vv' \in E(G) \text{ for some v } \in \Gamma^n(X)\}. \tag{2.2}$$

Observe that $X \subseteq \Gamma(X) \subseteq \Gamma^2(X) \subseteq ...\Gamma^n(X) \subseteq ...$

2.2. Erdös Labelling

Definition 2.1. [10] Labelling *is a bijection from natural numbers* $\mathbb{N}$ *to the set of vertices V(G). Labelling l is an* Erdös labelling *if there exists a finite set of vertices X such that $l(m) \in \Gamma^k(X)$ and $l(n) \notin \Gamma^k(X)$ for some natural number k implies $m < n$.*

We denote the set of labellings by $\mathcal{L}$.

Define $\alpha_l(n)$ to be the proportion of adjacent vertices of the vertex $l(n)$ which are labelled with lower values, *i.e.*

$$\alpha_l(n) = \frac{|\{m : l(m)l(n) \in E(G), m < n\}|}{|\{m : l(m)l(n) \in E(G)\}|}. \tag{2.3}$$

Now, we have to find the conditions under which the existence of Erdös labelling is guaranteed. We start with an example showing that Erdös labelling is unattainable.

Example 2.2. *A graph that has no Erdös labelling.*

Each vertex has infinitely many subordinate neighbors, and exactly one superior except the one with the highest rank.

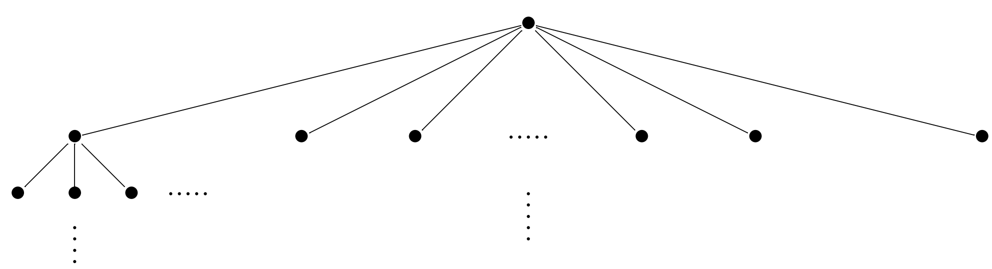

Figure 2.1. Hierarchy of infinite order.

Formally, let X_1 be the set, includes only a vertex which will be the uppermost one in the hierarchy. For all n greater than 1, let $\{X_n\}$ be a sequence of set of vertices, each consisting of countably many elements.

Now let f_n be a bijection from $\mathbb{N}$ to X_n. Partition X_n into the sets $X_n^1, ..., X_n^i, ...$ defining

$$X_n^i =: \{f_n(p_1^i), f_n(p_2^i), ..., f_n(p_m^i), ...\} \tag{2.4}$$

where p_m is the mth smallest prime number.

Define the set of edges $E_n^i =: \{f_{n-1}(i)u : f_{n-1}(i) \in X_{n-1}, u \in X_n^i\}$. Then $V(G) = \bigcup_{n=1}^{\infty} X_n$ and $E(G) = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{|X_{n-1}|} E_n^m$.

Claim 2.3. *The graph defined in Example 2.2. has no Erdös labelling.*

Proof. Let X_0 be a finite subset of $V(G)$ and l be an arbitrary labelling. Choose $x \in X_0$ that is on the lowest rank among the elements of X_0 in the sense that "hierarchy" suggests. Choose $x_2 \in \Gamma^2(X_0) \setminus \Gamma^1(X_0)$. Then under l, x_2 has a label $l(x_2) \in \mathbb{N}$. But, since $\Gamma^1(X_0)$ contains all adjacent vertices to x which are infinitely many, there exists an $x_1 \in \Gamma^1(X_0)$ such that $l(x_1) > l(x_2)$.Therefore, l is not an Erdös labelling.

□

Proposition 2.4. *An infinite graph with all of its vertices has at most finite neighbors has an Erdös labelling.*

Proof. Choose any finite subset X of $V(G)$. Label the elements of X from 1 to $|X|$ in any order. Then, since any neighbor growth is finite, i.e. $\Gamma^k(X)$ is always finite, label them accordingly to obtain an Erdös labelling.

2.3. NP-Hardness

Before defining NP-Hardness, we have to mention the class of NP problems. NP problems are the set of decision problems, which refers to the problems having "yes" or "no" answer with a given input, such that the "yes" answer to the problem is provable in polynomial time.

Definition 2.5. [21] *A problem is called* NP-hard *if it is reducible to an NP problem in polynomial time by the transformation of the inputs.*

NP-hardness does not only include decision problems, a well known example Traveling Salesman Problem, which is a search problem, is NP-hard. Yet all NP problems are not in the class of NP-Hard unless P=NP.

The problems below are NP problems that will be shown as a special case of other problems in the last chapter, from which we will deduce NP-Hardness.

Example 2.6. [21] *Set Cover Problem*

Let S be a finite set and A is a subset of the power set 2^S. And let N be a positive natural number such that $N \leq |2^S|$. The decision problem is whether there is a subset B of A with cardinality smaller than N such that

$$S \subseteq \bigcup_{X \in B} X. \tag{2.5}$$

Example 2.7. [21] *Vertex Cover Problem*

Let G be a finite undirected graph with the set of vertices $V(G)$ and set of edges $E(G)$. And let N be a positive natural number such that $N \leq |V(G)|$.

The aim of the problem is to find the smallest set of vertices in which each edge of the graph has an endpoint. The decision problem can be stated that if there is a set of vertices $X \subseteq V(G)$ such that $|X| \leq N$ and for $uv \in E(G)$, either $u \in X$ or $v \in X$.

3. LOCAL INTERACTION GAMES AND THE CASCADING BEHAVIOR

This chapter is basically concerned with the analysis of how a behavior spreads among a population and follows closely the works of [10, 16, 19]. The model posited to explain this stands at the intersection of graph theory and game theory. The graph theory part deals with the structure of the population and tries to answer how prone the population is to adopting a behavior. The population is considered to be large, which is reflected in the models with infinite graphs. On the other hand , the game theory part handles micro-level questions by imposing rules on the interactions between members of the population resulting in a change of behavior. Although, the game theoretical part lies in the core of the model, throughout the thesis we are less interested in the questions arising from it as compared to the graph theoretical ones.

In the first part, we define the game and the graph, restricting our attention to symmetric coordination games defined on infinite graphs. Then, we bring out the necessary technical information including notations and several lemmas before the cascading dynamics is introduced. The dynamics of the model is the next topic. Our main results are about a parameter called contagion threshold which is related to the parameter of the game. At the end, we present examples of local interaction system and discuss the cascading behavior on them.

3.1. On the Structure of the Game and the Graph

Definition 3.1. Local interaction game *is a pair consisting of a graph and a game defined between adjacent vertices where each player maximizes her utility having a unique strategy invariant with respect to the games played with her neighbors. More explicitly, the last* *statement of the sentence says that the player should stick to her strategy during a stage of the game and cannot change it depending on with whom she plays.*

For our purpose we need some restrictions on the graph and on the game. Let G be the simple undirected graph with possibly countably many vertices. Each vertex represents a player and it is inferred by what we have assumed on the graph that no player has interaction with herself and every interaction is reciprocal. The graph will be assumed to satisfy two properties:

Property 3.2. *Uniform Boundedness : The degree of each vertex is bounded by a natural number M. We note that if the degree of each vertex of a graph is finite, the graph is said to satisfy finite neigborhood property*

Property 3.3. *Connectedness : For any vertex u, v of the graph there exists a u, v-path.*

The game is a symmetric coordination game having a strategy space of size 2, consisting of strategies "0" and "1", for each player. More explicitly, the game has a symmetric payoff table and two pure Nash equilibria, two playing the same stategy. We first interpret this as the game does not distinguish players. Secondly, players must coordinate in order to maximize their payoff. The payoff table is of the form:
assuming that $a > c$ and $d > b$.

Table 3.1. Payoff table.

	0	1
0	a,a	b,c
1	c,b	d,d

The inequalities assumed above can be interpreted that if both of the players choose the same strategy, one of them switching to other strategy will be worse off in terms of utility. Therefore there are exactly two Nash equilibria of the game, "0"-"0" and "1"-"1".

We introduce some notations.

The set of neighbors of a vertex:

$$\Gamma(v) = \{v' : vv' \in E(G)\}.$$

The set of neighbors of a vertex v playing "0" :

$$\Gamma_0(v) = \{v' : vv' \in E(G) \text{ and } v\text{' chooses action } 0\}.$$

The set of neighbors of a vertex v playing "1" :

$$\Gamma_1(v) = \{v' : vv' \in E(G) \text{ and } v\text{' chooses action } 1\}.$$

The fraction of neighbors of a vertex v in a subset X of V(G):

$$\pi[X|v] = \frac{|\Gamma(v) \cap X|}{|\Gamma(v)|}.$$

The set of vertices for which at least p-fraction of neighbors are in a subset X of V(G):

$$\Pi^p(X) = \{v \in V(G) : \pi[X|v] \geq p\}.$$

3.1.1. Normalization of Payoff Table

For a vertex $v \in V(G)$ considering the games played between v and the neighbors of it, the strategy "1" is best response for v if and only if

$$\frac{(d-b)}{(a-c)+(d-b)}|\Gamma_1(v)| \geq \frac{(a-c)}{(a-c)+(d-b)}|\Gamma_0(v)|. \tag{3.1}$$

Now we can modify the pay-off table invariant to the best responses of players by the factor

$$q = \frac{(d-b)}{(a-c)+(d-b)} \tag{3.2}$$

The resulting table will be:

Table 3.2. Normalized payoff table.

	0	1
0	q,q	0,0
1	0,0	1-q,1-q

3.1.2. Best Response to a Set

Definition 3.4. [10] *Let X be a subset of $V(G)$. We call Y a* best response *to X if $Y \subseteq \Pi^q(X)$ and $\bar{Y} \subseteq \Pi^{1-q}(\bar{X})$.*

$\Pi^q(X)$ is clearly a *best response* to X, so there always exists a *best response* for any subset of V(G).

For the uniqueness, it suffices to show $V(G) = \Pi^q(X) \cup \Pi^{1-q}(\bar{X})$

Proof. Let $x \in V(G)$. If $x \notin \Pi^q(X)$, then $\pi[X|x] < p$. This means that more than $1-p$ fraction of neighbors of x are not in X. Therefore $x \in \Pi^{1-q}(\bar{X})$.

□

3.2. Cascading Behavior

We start with a local interaction game and restrict our attention to an infinite simple undirected graph and the symmetric coordination game described in the section 2.1.1. The players are allowed to choose their strategy at each round, but should stick to it for each game played with their neighbors during the round. At first round, we assume that each player in the network plays the same strategy "0", the one which has a payoff of q if chosen by both. This is clearly an equilibrium of the game, considering that if any player switches to strategy "1" will end up with 0 payoff.

Our aim is to identify the conditions under which the game switches to the other trivial equilibrium where each player chooses to play strategy "1" if a finite set of players are told to adopt strategy "1". The term "cascade" refers to this stage by stage transition to strategy "1".

Now we can define contagion threshold as:

Definition 3.5. [10] *The* contagion threshold ξ *is the largest* q *such that action "1" spreads to the whole population at some round of the repeated game from a finite subset of vertices by best response dynamics, i.e.*

$$\xi = \sup\{q \in (0,1) : \bigcup_{k\geq 1}[\Pi^q]^k(X) = V(G) \textit{ for some finite } X\} \tag{3.3}$$

But for our purpose, it does not suffice to guarantee the choice of the desired strategy by each player at varying stages of the game. We need to show the same contagion threshold as defined above also guarantees us the achievement of the ubiquitous adoption of strategy "1".

3.3. Operator Π^p_+

Referring to the last paragraph of the previous section, we define a new operator hoping it to be superior to Π^p.

$$\Pi^p_+(X) =: X \cup \Pi^p(X) \tag{3.4}$$

Observe that the operator $\Pi^p(X)$ does not necessarily include X. In fact, it may be included by X, or neither inclusion may be true. In each example, encircled vertices denote the elements of $\Pi^p(X)$, whereas elements of X are labelled by x_i's.

Example 3.6. *$X = \{x_1, x_2, x_3, x_4, x_5\}$ and take $p = \frac{1}{2}$.*

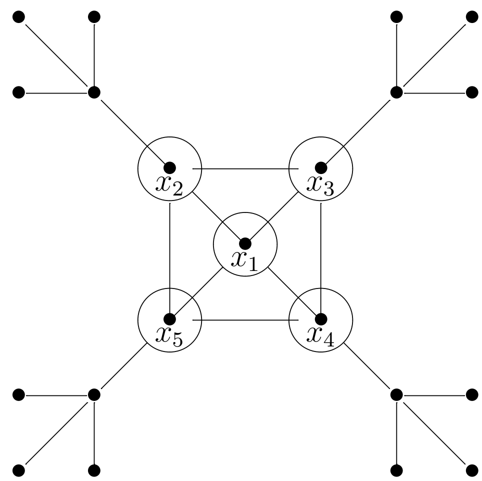

Figure 3.1. $X = \Pi^p(X)$.

Example 3.7. *$X = \{x_1, x_2\}$ and take $p = \frac{1}{3}$.*

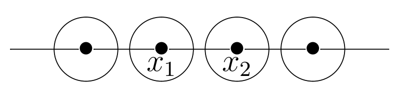

Figure 3.2. $X \subset \Pi^p(X)$.

Example 3.8. $X = \{x_1, x_2, x_3, x_4, x_5\}$ *and take* $p = \frac{1}{2}$.

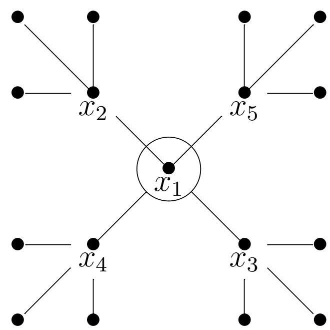

Figure 3.3. $\Pi^p(X) \subset X$.

Example 3.9. $X = \{x_1\}$ *and take* $p = \frac{1}{2}$.

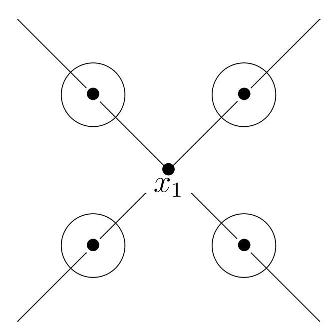

Figure 3.4. Neither $X \subseteq \Pi^p(X)$ nor $\Pi^p(X) \subseteq X$.

Example 3.10. *$X = \{x_1, x_2, x_3, x_4\}$ and take $p = \frac{2}{5}$.*

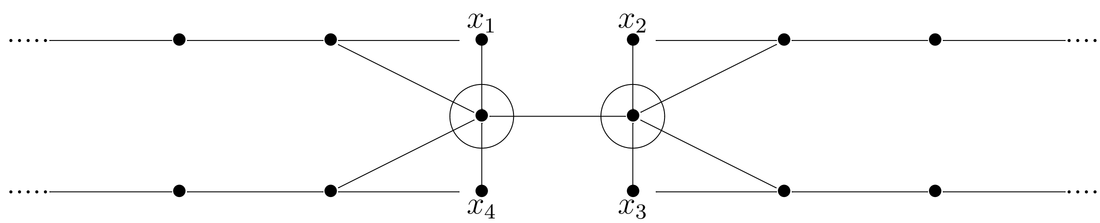

Figure 3.5. X is finite, $\bigcup_{k\geq 1}[\Pi^p]^k(X)$ is finite, but $\bigcup_{k\geq 1}[\Pi_+^p]^k(X)$ is infinite.

Observe that $\Pi^{\frac{2}{5}}[\Pi^{\frac{2}{5}}(X)] = \Pi^{\frac{2}{5}}(X)$ consists of the two encircled elements above, so is $\bigcup_{k\geq 1}[\Pi^{\frac{2}{5}}]^k(X)$. But $\Pi_+^{\frac{2}{5}}[\Pi_+^{\frac{2}{5}}(X)]$ also includes two pairs of vertices, one is left to x_1 and x_4, the other is right to x_2 and x_3. From then on, action "1" spreads along the graph and $\bigcup_{k\geq 1}[\Pi_+^{\frac{2}{5}}]^k(X) = V(G)$.

Lemma 3.11. *(i) If $X \subseteq Y$, then $\Pi^p(X) \subseteq \Pi^p(Y)$ and $\Pi_+^p(X) \subseteq \Pi_+^p(Y)$.*

(ii) If $X_k \nearrow X$, then $\Pi^p(X) = \bigcup_{k\geq 1} \Pi^p(X_k)$ and $\Pi_+^p(X) = \bigcup_{k\geq 1} \Pi_+^p(X_k)$.

(iii) For $p < r$, then $\Pi^r(X) \subseteq \Pi^p(X)$ and $\Pi_+^r(X) \subseteq \Pi_+^p(X)$.

(iv) If $p_k \nearrow p$, then $\Pi^{p_k}(X) \searrow \Pi^p(X)$ and $\Pi_+^{p_k}(X) \searrow \Pi_+^p(X)$.

Proof. (i) If $X \subseteq Y$, then $\pi[X|v] \leq \pi[Y|v]$ for every $v \in V(G)$. From which it follows that $\Pi^p(X) \subseteq \Pi^p(Y)$ and $\Pi_+^p(X) \subseteq \Pi_+^p(Y)$.

(ii) By (1), $\bigcup_{k\geq 1} \Pi^p(X_k) \subseteq \Pi^p(X)$.

For the other inclusion, for any $v \in V(G)$, there exists a k such that $\Gamma(v) \cap X \subseteq X_k$ by finite neighborhood property, (here $\Gamma(v)$ could be equal to X, and X could be infinite) so $\Pi^p(X) = \Pi^p(X_k)$ for some k. This implies $\Pi^p(X) = \bigcup_{k\geq 1} \Pi_+^p(X_k)$.

Now $\Pi_+^p(X) = X \cup \Pi^p(X) = \bigcup_{k\geq 1}[X_k \cup \Pi^p(X_k)] = \bigcup_{k\geq 1} \Pi_+^p(X_k)$.

(iii) $\pi[X|v] \geq r$ and $r > p$ implies that $\pi[X|v] \geq p$. So if $r > p$, $\Pi^r(X) \subseteq \Pi^p(X)$ and $\Pi^r_+(X) = \Pi^r(X) \cup X \subseteq \Pi^p(X) \cup X = \Pi^p_+(X)$.

(iv) By (3), $\Pi^{p_k}(X)$ is a decreasing sequence of sets and $\Pi^p(X) \subseteq \Pi^{p_k}(X)$ for all k. Now take $v \in \bigcap_{k\geq 1} \Pi^{p_k}(X)$, then $\pi[X|v] \geq p_k$ for all k, which implies $x \in \Pi^p(X)$. Hence, $\Pi^{p_k}(X) \searrow \Pi^p(X)$ and $\Pi^{p_k}(X) = (X \cup \Pi^{p_k}_+(X)) \searrow (X \cup \Pi^p(X)) = \Pi^p_+(X)$.

$\square$

Lemma 3.12. *The followings are equivalent*

(i) $\bigcup_{k\geq 1}[\Pi^p_+]^k(X)$ *is co-finite, for some finite* X*.*

(ii) $[\Pi^p_+]^k(X) \nearrow V(G)$*, for some finite* X*.*

(iii) $[\Pi^p]^k(X) \nearrow V(G)$*, for some finite* X*.*

Proof. Suppose $\bigcup_{k\geq 1}[\Pi^p_+]^k(X)$ is co-finite, for some finite X.
Let $Y = X \cup (\overline{\bigcup_{k\geq 1}[\Pi^p_+]^k(X)})$, then Y is finite.
By (3), $\bigcup_{k\geq 1}[\Pi^p_+]^k(X) \subseteq \bigcup_{k\geq 1}[\Pi^p_+]^k(Y)$.
By monotonicity of Y, the result follows.

$\square$

Proposition 3.13. [10]

$$\xi = \sup\{q \in (0,1) : \bigcup_{k\geq 1}[\Pi^q_+]^k(X) = V(G) \textit{ for some finite } X\} \tag{3.5}$$

Proof. The proof is by induction on k,

$$[\Pi^p_+]^k(X) = X \cup \Pi^p([\Pi^p_+]^{k-1}(X)). \tag{3.6}$$

For $k = 1$ it is true by the definition of Π_+. Now,

$$\begin{aligned}
[\Pi^p_+]^{k+1}(X) &= \Pi^p_+([\Pi^p_+]^k(X)) && (3.7)\\
&= [\Pi^p_+]^k(X) \cup \Pi^p([\Pi^p_+]^k(X)) && (3.8)\\
&= X \cup \Pi^p([\Pi^p_+]^{k-1}(X)) \cup \Pi^p([\Pi^p_+]^k(X)) && (3.9)\\
&= X \cup \Pi^p([\Pi^p_+]^k(X)) && (3.10)\\
X &\subseteq Y \to \Pi^p(X) \subseteq \Pi^p(Y). && (3.11)
\end{aligned}$$

Suppose that X is finite and $\bigcup_{k\geq 1}[\Pi^p_+]^k(X) = V(G)$. Let $Y = X \cup \Gamma(X)$. Since X is finite implies that Y is finite by finite neighborhood property of the graph. Now, choose K such that $Y \subseteq [\Pi^p_+]^K(X)$ and therefore $Y \subseteq [\Pi^p_+]^k(X)$ for every $k \geq K$. Since $\Gamma(X) \subseteq Y \subseteq [\Pi^p_+]^K(X)$, $X \subseteq \Pi^p(\Pi^p_+)^k(X)$ for every $k \geq K$.

Now by the statement proved above,

$$[\Pi^p_+]^{k+1}(X) = X \cup \Pi^p([\Pi^p_+]^k(X)) = \Pi^p([\Pi^p_+]^k(X)) \tag{3.12}$$

for every $k \geq K$. Then,

$$[\Pi^p]^k([\Pi^p_+]^K(X)) = [\Pi^p_+]^{K+k}(X) \tag{3.13}$$

for every $k \geq K$. Thus,

$$\bigcup_{k \geq 1} [\Pi^p]^k([\Pi^p_+]^K(X)) = V(G) \tag{3.14}$$

$[\Pi^p_+]^K(X)$ is the desired finite subset.

$\square$

3.4. Contagion Threshold

Theorem 3.14. [16] *The contagion threshold is at most* $\frac{1}{2}$.

Proof. Assume for a contradiction that $\xi > \frac{1}{2}$ and there exists a finite set X_0 from which action "1" spreads to the whole population. We can assume that no player switches from strategy "1" to strategy "0" by Proposition 3.13.

Now, let X_k be the set of players adopting strategy "1" at the *kth* stage of the game. And let $\delta(X_k)$ be the set of edges that have one end in X_k and the other end in $\bar{X}_k$. $\delta(X_k)$ is the interface where the action "1" spreads. To arrive at a contradiction, it has to be shown that the size of $\delta(X_k)$ decreases strictly as k increases, so that action "1" ceases to spread.

Define $S_k = X_{k+1} \setminus X_k$. For an arbitrary $s_k \in S_k$, $\Gamma_1(s_k)$ is strictly larger than $\Gamma_0(s_k)$ provided that $q > \frac{1}{2}$. So $|\bigcup_{s_k \in S_k} \Gamma_1(s_k)| > |\bigcup_{s_k \in S_k} \Gamma_0(s_k)|$.

Now, we identify the set $\delta(X_{k+1})$. By the assumption that no player switches from strategy "1" to strategy "0" $\bigcup_{s_k \in S_k} \Gamma_1(s_k) \cap \delta(X_{k+1}) = \emptyset$.

Then for any uv in the set of edges of S_k, if u or v is in X_k, then $uv \in \delta(X_k) \setminus \bigcup_{s_k \in S_k} \Gamma_1(s_k)$. Otherwise, $uv \in \Gamma_0(s_k)$. Therefore, $\delta(X_{k+1}) \subseteq (\delta(X_k) \cup \Gamma_0(s_k)) \setminus \Gamma_1(s_k)$.

By the inequality above, $|\delta(X_{k+1})| < |\delta(X_k)|$.

Since $\delta(X_k)$ is finite for any k by uniform boundedness property and decreases in terms of cardinality as k increases, will terminate the cascade at some finite point. Action "1" is not spread to the whole population, we arrive at a contradiction.

Hence, ξ must be smaller than or equal to $\frac{1}{2}$.

□

Definition 3.15. [10] Cohesion *is a property of subsets X of $V(G)$ and is defined as*

$$c(X) = \inf_{v \in X} \pi[X|v]. \tag{3.15}$$

X is called $p-$*cohesive* if $c(X) \geq p$. Equivalently,

$$c(X) = \sup\{p \in (0,1) : X \subset \Pi^p(X)\}. \tag{3.16}$$

Cohesion measures how tightly a community is connected within a network. Those communities having a large cohesion is referred to as "clusters" in the literature [16]. One result which will be presented next is that large cohesion is an obstacle for contagion.

Lemma 3.16. *There exists an $\epsilon \geq 0$ such that $\bigcup_{k \geq 1} [\Pi_+^r]^k(X)$ is $(1-p)$ cohesive for every subgraph H of G and $r \leq p + \epsilon$.*

Proof. Consider

$$F(M) = \{\alpha \in (0,1) : \alpha = \frac{n}{m} \text{ for some integers m,n with } 0 < m \leq M \text{ and } 0 \leq n \leq m\}. \tag{3.17}$$

Given p, choose $\epsilon > 0$ such that $F(M) \cap (p, p+\epsilon) = \emptyset$. By uniform boundedness assumption,

$$\frac{|X \cap \Gamma(x)|}{|\Gamma(x)|} \in F(M) \text{ for every } x \in V(G) \text{ and for every } X \subseteq V(G). \tag{3.18}$$

Since $F(M) \cap (p, p+\epsilon) = \emptyset$,

$$\Pi^r(X) = \Pi^{r'}(X) \text{ for every } r, r' \in (p, p+\epsilon). \tag{3.19}$$

So for any $X \subseteq V(G)$ let

$$Y = \bigcup_{k \geq 1} [\Pi^r_+]^k(X)) \text{ for every } r \in (p, p+\epsilon). \tag{3.20}$$

Now for every $x \in Y$ and $r \in (p, p+\epsilon)$, $\pi(Y|x) > 1-r$, so
$\pi(Y|x) \geq 1-p$ for every $x \in Y$.

□

Theorem 3.17. [19] *The action "1" spreads contagiously from a finite set X if and only if $\bar{X}$ does not contain a $(1-p)$-cohesive subset Y for some $p < \xi$.*

Proof. For the first part, consider an infinite graph G that the strategy "1" spreads contagiously with threshold ξ from a finite subset X. Assume for a contradiction that $\bar{X}$ contains a subset of cohesion larger than $(1-\xi)$. Let t be the first stage of the game that action “1” spreads to Y, and let $v \in Y$ that adopts strategy “1” at the tth stage. Now, we will show that v could not have enough neighbor choosing strategy “1” at the $t-1^{st}$ stage and arrive at a contradiction.

At the $t-1^{st}$ stage, no vertex belonging to Y chooses strategy “1”, that is how we assumed. So more than $(1-\xi)$ fraction of neighbors of v were still playing “0” by cohesiveness of Y larger than $(1-\xi)$. But then, the fraction neighbors of v having chosen “1” could not exceed ξ. Therefore, v must have chosen “0” at the tth, which is a contradiction. Hence, contrary to our assumption $V(G)-X$ does not contain a subset of cohesion larger than $(1-\xi)$.

For the necessity part, suppose that action “1” is initiated to a finite subset X of a graph G and ceases to spread at some stage of the game. Let Y be the subset of players that still playing “0” after this stage. We will show that Y is more than $(1-\xi)$-cohesive.

Let v be an arbitrary element of Y, since v chooses strategy “0”, then more than $(1-\xi)$ - fraction of neighbors of v chooses “0”, so are in Y. Therefore, $\pi[Y|v] > (1-\xi)$ for $v \in Y$, from which it follows that $c(Y) \geq (1-\xi)$, hence Y is $(1-\xi)$-cohesive.

Yet we need stricter result, for this purpose we employ uniform boundedness assumption. Let M be the natural number that bounds the number of neighbors one player has. Then, there exists a natural number n such that

$$\frac{n}{M} > (1-\xi) \geq \frac{(n-1)}{M} \tag{3.21}$$

for $1 \leq n \leq M$.

Since $\pi[Y|v] > (1-\xi)$ for $v \in Y$, the infimum of $\pi[Y|v]$ over the elements of Y

is greater than or equal to $\frac{n}{M}$. This implies that Y is more than $(1-\xi)$ cohesive.

$\square$

Corollary 3.18. ***Upper Bound****- If every co-finite set of vertices contains an infinite, $(1-p)$-cohesive, set of vertices, then $\xi \leq p$.*

Theorem 3.19. [10] $\xi = sup_{l\in\mathcal{L}}\Big(\liminf_n \alpha_l(n)\Big)$.

Proof. We first show that for any labelling l, $\xi \geq \liminf_n \alpha_l(n)$. Let l be an arbitrary labelling, X be a set that action "1" spreads contagiously and N be a natural number that is larger than the greatest label among elements of X. Now define the set $Y =: \{l(1), l(2), ..., l(N)\}$, Y is a superset of X. By Lemma 3.12., an inductive argument would reveal that Y is also contagious. Since Y is contagious, $\Pi^p(Y)$ is non-empty for every $p < \xi$. Then, for every p, there exists an v such that $\pi[Y|v] \geq p$. From which we can deduce by the construction of Y that $\alpha_l(n) \geq p$ where n is the label of v under l. Therefore, for every labelling l and every sufficiently large N, $inf_{n\geq N}\alpha_l(k)) \geq p$ for every $p < \xi$. The desired result is obtained.

Now, we show that the limit cannot be larger than ξ to obtain maximality. By the definition of ξ, there exists a finite set X such that $\bigcup_{k\geq 1}[\Pi^{\xi}_{+}]^k(X) = V(G)$. Define the set $X_n = [\Pi^{\xi}_{+}]^n(X) \cap \overline{[\Pi^{\xi}_{+}]^{n-1}(X)}$ for $n \geq 1$ and $X_0 =: X$, so X_n is the set of vertices adopts strategy "1" beginning from the nth stage of the game. Then label the elements of X_n from $|\bigcup_{i=0}^{n-1} X_i| + 1$ to $|\bigcup_{i=0}^{n} X_i|$ in an arbitrary order, call the labelling l_0. Then for any v in X_N for an $N \geq 1$ with label n_v,

$$\xi \leq \Pi[X_{n-1}|v] = \frac{|X_{n-1} \cap \Gamma(v)|}{|\Gamma(v)|} \leq \frac{|\{m : l_0(m)l_0(n_v) \in E(G) \text{ and } m < n_v\}|}{|\Gamma(v)|} = \alpha_{l_0}(n_v). \tag{3.22}$$

Therefore, $\xi \leq lim_{N\to\infty}(inf_{n\geq N}\alpha_{l_0}(n))$ which implies by the result above that ξ is the maximum.

$\square$

Corollary 3.20. ***Lower Bound****- If there exists a labelling l with $\alpha_l(k) \geq p$ for all sufficiently large k, then $\xi \geq p$.*

The following definition introduces a stricter condition regarding to neighborhood growth.

Definition 3.21. [10] *An infinite graph G is said to satisfy* low neighborhood growth *if for any $\gamma > 1$ $\lim_{n\to\infty} \gamma^{-n}|\Gamma^n(X)| = 0$ for all finite subsets X of $V(G)$.*

Another graph property, this time related to labelling, is:

Definition 3.22. [10] *An infinite graph G satisfies δ* - uniformity *if there exists an Erdös labelling l such that for all sufficiently large K,*

$$\max_{k,k'\geq K} |\alpha_l(k') - \alpha_l(k)| \leq \delta. \tag{3.23}$$

Theorem 3.23. [10] *If a local interaction system satisfies low neighbour growth and δ-uniformity, then $\xi \geq \frac{1}{2} - \delta$.*

Proof. Let l be an Erdös labelling on the graph that satisfies δ-uniformity. Then there exists an α such that

$$\alpha - \delta \leq \frac{|\{m : l(m)l(n) \in E(G) \text{ and } m < n\}|}{|\{m : l(m)l(n) \in E(G)\}|} \leq \alpha, \tag{3.24}$$

for all $n > N$ where N is sufficiently large. By Corollary 3.20, $\xi > \alpha - \delta$. So we need to show $\alpha \geq \frac{1}{2}$ to prove the theorem. Assume for a contradiction that $\alpha < \frac{1}{2}$.

Now, there exists a set X such that $l(m) \in \Gamma^k(X)$ and $l(n) \notin \Gamma^k(X)$ implies $m < n$, since l is an Erdös labelling. Define $X_0 = X$ and $X_n = \Gamma^n(X) \setminus \Gamma^{n-1}(X)$. By 3.23., for a sufficiently large n we obtain,

$$\Big| \bigcup_{l(i) \in X_n} \{j : l(j)l(i) \in E(G) \text{ and } j < i\} \Big| \leq \alpha \Big| \bigcup_{l(i) \in X_n} \{j : l(j)l(i) \in E(G)\} \Big| \tag{3.25}$$

$$\Big| \bigcup_{l(i) \in X_n} \{j : l(j)l(i) \in E(G) \text{ and } j < i\} \Big| \leq (\frac{\alpha}{1-\alpha}) \Big| \bigcup_{l(i) \in X_n} \{j : l(j)l(i) \in E(G) \text{ and } j > i\} \Big| \tag{3.26}$$

By the definition of X_n and the above inequality we have

$$\begin{aligned} &\big|\{uv \in E(G) : u \in X_{n-1} \text{ and } v \in X_n\}\big| + \big|\{uv \in E(G) : u \in X_n \text{ and } v \in X_n\}\big| \leq \\ &\Big(\frac{\alpha}{1-\alpha}\Big)\big|\{uv \in E(G) : u \in X_n \text{ and } v \in X_{n+1}\}\big| + \big|\{uv \in E(G) : u \in X_n \text{ and } v \in X_n\}\big|. \end{aligned} \tag{3.27}$$

By uniform boundedness property for some natural number M and from the assumption that $\alpha < \frac{1}{2}$ we have

$$|X_n| \geq \frac{\left|\{uv \in E(G) : u \in X_{n-1} \text{ and } v \in X_n\}\right|}{M} \geq \Big(\frac{1-\alpha}{\alpha}\Big)^n |X_0|. \tag{3.28}$$

$$|\Gamma^n(X)| = \sum_{i=0}^{n} |X_n| \geq \sum_{i=0}^{n} \Big(\frac{1-\alpha}{\alpha}\Big)^n |X_0|. \tag{3.29}$$

Since $\alpha < \frac{1}{2}$, $\left(\frac{1-\alpha}{\alpha}\right) > 1$. Now for $\gamma \in \left(1, \left(\frac{1-\alpha}{\alpha}\right)\right)$,

$$\gamma^{-n}|\Gamma^n(X)| \geq \sum_{i=0}^{n} \Big(\frac{1-\alpha}{\alpha}\Big)^n \gamma^{-n}|X_0| \geq \Big(\frac{1-\alpha}{\alpha}\Big)^n \gamma^{-n}|X_0| \to \infty. \tag{3.30}$$

But this contradicts the low neighborhood growth assumption. Therefore $\alpha \geq \frac{1}{2}$ and the theorem is proven.

$\square$

3.5. Uniform Boundedness Assumption

Example 3.24. *An infinite graph satisfying finite neighborhood that does not necessarily have a contagion threshold.*

Let X_1 be the set, includes only a vertex which will be the uppermost one in the hierarchy. Inductively for all n greater than 1, let X_n be a set of vertices consisting of $n|X_{n-1}|$ elements.Partition X_n into the sets $X_n^1, ..., X_n^{n-1}$ each having exactly n members. For an arbitrary labelling of X_{n-1}, define the set of edges $E_n^i = \{v_i u : v_i \in X_{n-1}, u \in X_n^i\}$.

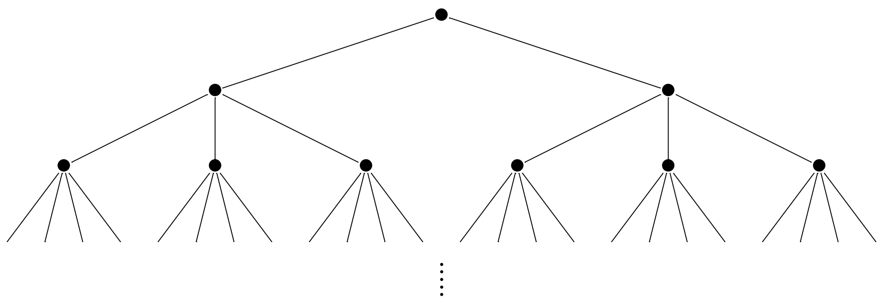

Figure 3.6. Hierarchy of increasing order.

Let $V(G) = \bigcup_{n=1}^{\infty} X_n$ and $E(G) = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{|X_{n-1}|} E_n^m$. Then for any $q > 0$, there exists a natural number N such that $\frac{1}{N} < q$. So for all $n > N$, the action "1" does not spread contagiously from a vertex of X_n to subordinate ones. Hence, the contagion is unattainable for any $q \in (0,1)$.

Proposition 3.25. *The contagion threshold ξ exists for any infinite graph satisfying uniform boundedness assumption.*

Proof. Let m be a natural number such that for any $v \in V(G)$, $|\Gamma(v)| < M$. Then for any $q < \frac{1}{M}$, $\Pi^q(v) = \Gamma(v)$.

Now take any finite subset X of $V(G)$. For any $q < \frac{1}{M}$, $\Pi^q(X) = \Gamma(X)$ and $[\Pi^q]^n(X) = \Gamma^n(X)$. But since the graph is connected,

$$\bigcup_{n \geq 1} [\Pi^q]^n(X) = lim_{n\to\infty} \Gamma^n(X) = V(G). \tag{3.31}$$

$\square$

3.6. Examples

Example 3.26. [10] ***Interaction on a line***

Each vertex has exactly two neighbors, one to her left and one to her right.

Figure 3.7. A graph with vertices of integers.

Now, we show that the contagion threshold is $\frac{1}{2}$ for this example. Take $q < \frac{1}{2}$. Then any vertex having a neighbor playing "1" is better off by switching to "1". So, initiating the srategy "1" to a pair of adjacent vertices, it spreads to whole population. Hence $\xi \geq \frac{1}{2}$.

If $q > \frac{1}{2}$, then no player switches to "1" unless both of the neighbors play "1". But since any finite set has a farmost left and a farmost right element having negihbors playing "0", the strategy "1" does not spread. Therefore, $\xi = \frac{1}{2}$.

Example 3.27. [10] ***n-max distance interaction in m-dimensions***

The vertices are put in an infinite m dimensional lattice, $\mathbb{Z}^m$. u is a neighbor of v if $max_{i \leq m}|u_i - v_i| \leq n$.

For $m = 2$ and $n = 1$, we have the following graph,

We utilize Theorem 3.17. to obtain a bound the contagion threshold. Take any finite set of vertices X of $V(G)$ and consider the corresponding co-finite set $\bar{X}$. Now, for the first coordinate of the vertices in the lattice define $\alpha =: \max_{v \in X} v_1$. Take a vertex $u \in \bar{X}$ such that $u_1 = \alpha + 1$. Then u has exactly $(2n+1)^{m-1} - 1$ neighbors with first coordinate

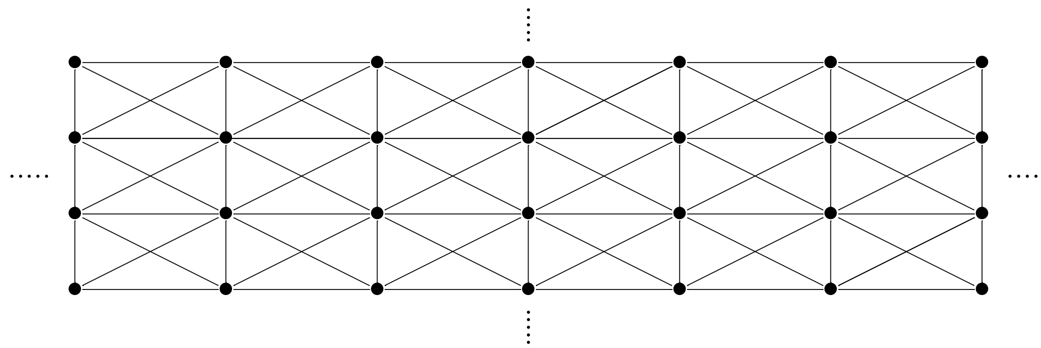

Figure 3.8. m-dimensional lattice.

$\alpha+1$. And u has $n(2n+1)^{m-1}$ neighbors with first coordinate greater than $\alpha+1$. So u has at least $(n+1)(2n+1)^{m-1}-1$ neighbors within the set $\bar{X}$. Since total neighbors of any vertex is $(2n+1)^m-1$ many, $\bar{X}$ is $\frac{((n+1)(2n+1)^{m-1}-1)}{((2n+1)^m-1)}$ is cohesive.

By the theorem, $\xi \leq \frac{n(2n+1)^{m-1}}{((2n+1)^m-1)}$.

Example 3.28. [10] ***Regions***
The set of vertices consist of infinite complete graphs of n vertices. Each complete graph has two adjacent complete graphs, here adjacency is connectedness of each vertex of a graph with exactly one vertex from an adjacent graph. The following figure illustrates the case for $n=3$.

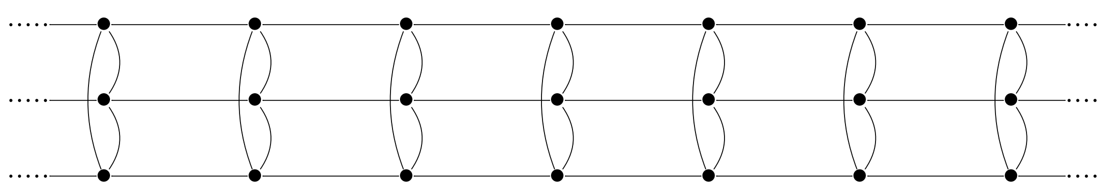

Figure 3.9. Regions of order three.

For the cohesiveness of the graph, consider the subset X of vertices consisting of two adjacent complete graphs. Each vertex v of X has $n+1$ neighbors, where $n-1$ of them are in the same complete graph with v, so is in X. One of the remaining two is in the adjacent complete graph included by X, and the other is in $\bar{X}$. Hence X is $\frac{n}{n+1}$ cohesive. But since any subset of vertices Y where $\bar{Y}$ is finite includes a pair of adjacent complete graphs and $\frac{n}{n+1}$ is not exceedable unless $Y = V(G)$, by Theorem 3.17. $\xi \leq \frac{1}{n+1}$.

Moreover, $\max_{v \in V(G)} |\Gamma(v)| = n+1$ implies $\xi \geq \frac{1}{n+1}$. Therefore $\xi = \frac{1}{n+1}$.

Example 3.29. [10] ***Hierarchy***

Let X_1 be the set, includes only a vertex which will be the uppermost one in the hierarchy. For all n greater than 1, let $\{X_k\}$ be a sequence of set of vertices, each consisting of n^k elements.

Now let f_k be a bijection from $\{1, 2, ..., n^k\}$ to X_k. Partition X_k into the sets $X_k^1, ..., X_n^i, ...$ defining

$$X_k^i =: \{f_k(in+1), ..., f_k(i(n+1))\} \tag{3.32}$$

Now define the set of edges $E_k^i =: \{f_{k-1}(i)u : f_{k-1}(i) \in X_{k-1}, u \in X_k^i\}$.

Let $V(G) = \bigcup_{k=1}^{\infty} X_k$ and $E(G) = \bigcup_{k=2}^{\infty} \bigcup_{m=1}^{|X_{k-1}|} E_k^m$.

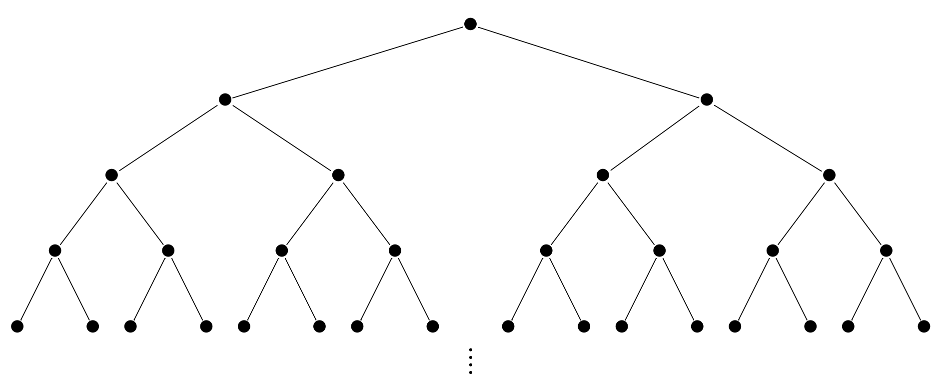

Figure 3.10. Hierarchy of order two.

Take any subset X of the set vertices consisting of all subordinates of a vertex v. X is $\frac{n}{n+1}$ cohesive, since only vertex having a neighbor in $\bar{X}$ is v, and it has exactly one neighbor in $\bar{X}$. Since any subset of vertices Y where $\bar{Y}$ is finite contains a subset of the form above, by Theorem 2.17. $\xi \leq \frac{1}{n+1}$. Moreover, $\max_{v \in V(G)} |\Gamma(v)| = n+1$ implies $\xi \geq \frac{1}{n+1}$. Therefore $\xi = \frac{1}{n+1}$.

For δ-uniformity, consider an Erdös labelling where the uppermost vertex $\bar{v}$ is labelled by "1", i.e. $\bar{v}$ is $l(1)$. Since $X_k = \Gamma(\bar{v})$, by Definition 1.1. $l(m) \in X_k$ and $l(n) \notin X_k$ for some k implies $m < n$. Therefore for any vertex $l(i) \in V(G)$, $l(i)$ has exactly one neighbor with a lower label. So, $\alpha_l(i) = \frac{1}{n}$ for all $i > 1$. Hence 0-uniformity is satisfied.

4. EXTENSIONS

This chapter is devoted to a question arising from the cascading model of the previous chapter which is dealt extensively in [15], [16], [22], [17], [20] and [18]. In the model, the condition for cascading to occur is the existence of an initial contagious set. However, almost nothing has mentioned about this set yet. From now on, we deal with the problem of finding a good set to start with. The approaches developed in the mentioned works are helpful to some extent, yet they are not completely satisfactory. The difficulty is due to cardinality. The optimization problem for the best initial set is transformed into the maximization of the number of vertices converted by the initial set, but in case of infinite graphs the maximum is unattainable. In this chapter, instead of restricting ourselves to infinite graphs, we deviate a bit and mention several results related to the initial set selection problem in finite graphs which bear hope of generalization.

First, we introduce other models of cascading which focus on the initial set selection problem rather than the conditions for cascading to occur. Then, we compare the models and apply possible transformations. Next, the optimization problem is introduced and the function related to this problem is expected to satisfy a property, namely the submodularity. The submodularity is shown to hold for each model presented. Finally, the complexity of this optimization problem for different models is examined.

4.1. General Models of Cascading Behavior

So far we have considered the cascading behavior specifically on local interaction systems with infinite vertices and certain simplifications such as indistiguishable players, reciprocal relations etc. Now, we present several cascading behavior models that capture more general settings of diffusion of a behavior in general networks.

4.1.1. Linear Threshold Model

In this model, we allow players to distinguish their neighbors in terms of their influence on the adoption of a new behavior. Also, it will be ascribed seperate contagion thresholds to each player. Just as before, there are two strategies, call them A and B and we are looking for the strategy A to overcome B when A is initiated to a set of vertices. The model is as follows:

Let G be possibly an infinite undirected graph. On the set $V(G) \times V(G)$ we define a non-negative function f requiring $f(u, v) = 0$ if u and v are not neighbors and $\Sigma_{u \in \Gamma(v)} f(u, v) \leq 1$. So player v attains seperate weights on her neighbor's influence whose sum does not exceed 1. Secondly, each player v has a threshold θ_v, uniformly distributed between 0 and 1. This is the required fraction of neighbors adopting the cascading strategy.

At the beginning of the game, each player adopts the strategy B. Then at the first round, possibly infinite set of players are told to choose strategy A. Then, at the kth round, player v, still adopting the strategy B, switches to strategy A if

$$\Sigma_{u \in \Gamma_B(v)} f(u, v) \geq \theta_v. \tag{4.1}$$

In the cascading model we discussed in the previous chapter, we simply define f for two adjacent vertices u and v as $f(u, v) = \frac{1}{|\Gamma(v)|}$. The threshold $\theta(v) = q$ for each $v \in V(G)$ where q is the normalized payoff of the initial strategy if they both adopt it.

4.1.2. General Threshold Model

One generalization of the model above is defining the influence function on the subsets of vertices through which we allow group influence on players. But in the case of infinite graphs, the domain of the function has the same cardinality with $2^{\mathbb{N}}$. So it is more convenient to deal with distinct functions f_v defined on the set $2^{\Gamma(v)}$ for $v \in V(G)$.

The function f_v is a restricted form of the function f above and has further restrictions. In accordance with the model above, f_v will be non-negative and smaller than 1. In addition, the function will be monotone, i.e. if $X \subseteq Y \subseteq \Gamma(v)$, then $f_v(X) \leq f_v(Y)$. The interpretation is that there is no negative influence.

The threshold θ_v is as defined on the previous model and at any stage of the game the action A is adopted by player v who chose B before if,

$$f_v(\Gamma_A(v)) \geq \theta_v. \tag{4.2}$$

4.1.3. Independent Cascade Model

In the local interaction for this model, we once again assume there are two strategies for each player. In the beginning, each player choose strategy A, then a set of players is initiated to strategy B as in every model we have investigated. Each player adopting strategy B is given a single chance to introduce the strategy one by one to her neighbors. The process is independent from the order in which a player is tried to be influenced by her neighbors and from the history of the game. It terminates until no more diffusion of strategy B is possible.

For this purpose, we define the function p on the set $V(G) \times V(G)$, which takes values between 0 and 1. We will call $p(u, v)$ the probability for u to initiate strategy B to v. If the graph is undirected, then $p(u, v) = p(v, u)$.

4.1.4. General Cascade Model

This model is a generalization of the one above. The distinction is that now the probability for a player to influence one of her neighbors could possibly depend on the set of players that have already tried and failed to do so. But as in the general threshold model, the probability of influence functions will be defined distinctly.

For any vertex v of G, p_v is defined from $\Gamma(v) \times 2^{\Gamma(v)}$ to [0,1]. The value $p_v(u, X)$ refers to the probability that u succeeds influencing v to adopt the strategy B, given that the set X has already tried and failed. What we infer is $p_v(u, X)$ is always 0, if $u \in X$. Furthermore, we still expect the functions to be order-independent, i.e.if a set of neighbors of v try to influence v, the order in which they try does not affect the overall success that v adopts B.

4.1.5. Equivalence of the Models

Now, it will be shown that the two general models introduced above are in fact equivalent in the sense that given the influence functions of one, the functions of the other can be expressed in terms of them.

Firstly, consider a vertex v of a graph and the corresponding influence function f_v. Now, v tries to influence u given that vertices of X, a subset of $\Gamma(u)$, have already tried and failed. Then, by Bayes' formula the conditional probability that u influences v while X could not is:

$$p_v(u, X) = \frac{f_v(X \cup u) - f_v(X)}{1 - f_v(X)}. \tag{4.3}$$

In this way, we obtain order-independent functions which can be shown by taking an arbitrary set X of neighbors of a vertex v and calculating the overall success that v adopts

B through the attempts of the elements of X. Regardless of the orders of the attempts this calculation will give $f_v(X)$.

Conversely, take a vertex v in the general cascade model and a subset of $\Gamma(v)$, call it X and $X = \{u_1, u_2, ..., u_k\}$. Now, our aim is to calculate the probability of success of the set X converting v to strategy B. Since, p_v is order-independent, the calculation will be independent of the order of u_i's.

Let $X_i = \{u_1, u_2, ..., u_i\}$ and $X_0 = \emptyset$. Then the probability that v still chooses A after the attempts of the players in X is $\Pi_{i=1}^{k}(1 - p_v(u_i, X_{i-1}))$. So by order-independence, the influence function of the general threshold model is well-defined as:

$$f_v(X) = 1 - \Pi_{i=1}^{k}(1 - p_v(u_i, X_{i-1})). \tag{4.4}$$

4.2. Submodularity of the Influence Function

In this chapter, we introduce a new property for the influence functions of the models defined above. Assuming this property on the functions, it is aimed to achieve an approximation to the optimal solution of the selecting most influential vertices problem.

Let G be a graph and V(G) be the set of its vertices.

Definition 4.1. [17,18] *The function $f : 2^{V(G)} \to \mathbb{R}$ is* submodular *if for all $X, Y \subseteq V(G)$*

$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y). \tag{4.5}$$

The equivalent condition is if $X \subset Y \subset V(G)$ and $v \in V(G)$, then

$$f(X \cup v) - f(X) \geq f(Y \cup v) - f(Y). \tag{4.6}$$

This expression has a clearer interpretation: the effect of adding a vertex decreases as the set increases. That is why the condition is classified as a diminishing returns property.

Now we come back to our problem, we try to find an initial subset of vertices with a limited cardinality that has the most influence, in other words we try to choose such a small set of vertices as the first adopters of a strategy that at the end of the process the most possible number of adoption of the strategy is achieved.

Definition 4.2. [17,18] *Let G be a graph and X_0 be a subset of $V(G)$. X_n denotes the set of players adopting the new strategy n rounds after X_0 is introduced to it. The* influence function $\sigma : 2^V(G) \to \mathbb{R}^+$ *is defined as*

$$\sigma(X) = \mathbb{E}[X_N] \tag{4.7}$$

where $X \subseteq V(G)$ and N is the cardinality of $V(G)$ for finite graphs, otherwise a predetermined finite number.

The main result of the chapter is

Theorem 4.3. [17] *Let σ be a monotone submodular function, and let X^* be the n-element set for which σ is maximized. On the other hand, let X be an n-element set obtained by each time adding an element that maximizes σ (this corresponds to the greedy algorithm below). Then*

$$\sigma(X) \geq (1 - \frac{1}{e})\sigma(X^*). \tag{4.8}$$

Algorithm 4.4. *Greedy Algorithm.*

- $X = \emptyset$
- ***for*** $i = 1$ ***to*** n
- *Let* v_i *solves the problem* $max_{v \in V(G)} \sigma(X \cup v) - \sigma(X)$.
- $X \longleftrightarrow X \cup v_i$
- *end for*

Example 4.5. *The influence function* σ *is not submodular.*

Let G consist of five vertices. Suppose that the condition for one vertex to adopt a strategy is that all of its neighbors play the strategy, i.e. $\theta_v = 1$ for all $v \in V(G)$; and f be the expected value function of number of vertices adopting the desired strategy at the end of the game. Consider the figure below. $X = \{x_1, x_2\}$, $Y = \{x_1, x_2, x_3\}$ *and* v

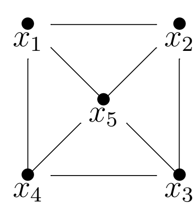

Figure 4.1. A graph that on which a non-submodular function can be defined.

Claim 4.6. f *is not submodular.*

Since adding x_4 to X only increases the set of vertices playing the strategy by 1 throughout the game,

$$f(X \cup x_4) - f(X) = 1. \tag{4.9}$$

On the other hand, adding x_4 to Y guarantees x_5 to play the same strategy with the others.

$$f(Y \cup x_4) - f(Y) = 2. \tag{4.10}$$

But $X \subset Y$, hence f is not submodular.

Now, we will show that the influence functions derived from the models described above are in fact submodular. First we start with two lemmas.

Lemma 4.7. *If $f_1, f_2, ..., f_n$ are submodular functions, and $c_1, c_2, ..., c_n$ are non-negative real numbers, then $f = \Sigma_{i=1}^n c_i f_i$ is also submodular.*

Proof. For $X, Y \subseteq V(G)$

$$\begin{aligned} f(X) + f(Y) &= \Sigma_{i=1}^n c_i f_i(X) + \Sigma_{i=1}^n c_i f_i(Y) &(4.11)\\ &\leq \Sigma_{i=1}^n c_i f_i(X \cup Y) + \Sigma_{i=1}^n c_i f_i(X \cap Y) &(4.12)\\ &= f(X \cup Y) + f(X \cap Y). &(4.13) \end{aligned}$$

□

Lemma 4.8. *Let $C_1, C_2, ..., C_n$ be a collection of sets and $X \subseteq \{1, 2, ..., n\}$. Define $f(X) = |\bigcup_{i \in X} C_i|$. Then f is submodular.*

Proof. For $X, Y \subseteq 1, 2, ..., n$

$$|\bigcup_{i \in X \cup Y} C_i| = |\bigcup_{i \in X} C_i| + |\bigcup_{i \in Y} C_i| - |\bigcup_{i \in X \cap Y} C_i|. \tag{4.14}$$

Hence, f is submodular.

□

Theorem 4.9. [15, 16] *The influence function σ derived from the Independent Cascade Model is submodular.*

Proof. A new approach will be developed for this and the following proofs. The cascading process is considered as a network formation game. The influence probability $p(u, v)$ for $u, v \in V(G)$ is taken to be the probability of edge uv to be formed. The game does not take place stage by stage, but all at once every edge is determined to be "active" or "inactive", as if coins of biases $p(u, v)$'s are thrown beforehand and noted, then the game unfolds deterministically stage by stage. This model is equivalent to the Independent Cascade Model in terms of analysis and helps us to resolve computational intricacies of stage by stage approach.

Let α be one of the $2^{|E(G)|}$ possible outcomes of the game. Let $\sigma_\alpha(X)$ denote the vertices playing strategy B at the end of the game when strategy B is first initiated to the set $X \subseteq V(G)$, For a vertex v, the set P_α^v denotes the set of vertices that are connected to v through active edges of event α. Since at the beginning, only players adopting B belongs to X,

$$\sigma_\alpha(X) = |\bigcup_{v \in X} P_\alpha^v|. \tag{4.15}$$

By Lemma 4.7, σ_α is monotone for every α in the set of events.

Finally, we express the influence function σ of the Independent Cascade Model as,

$$\sigma(X) = \sum_\alpha \sigma_\alpha(X) Prob(\alpha). \tag{4.16}$$

Hence, by Lemma 4.6, σ is submodular.

□

Theorem 4.10. [15,16] *The influence function σ derived from the Linear Threshold Model is submodular.*

Proof. [20] In order to prove this, it will be shown that the Linear Threshold Model is equivalent to the network formation model above in terms of the expected number of vertices playing B at a certain stage of the game. The claim is that for any $X' \subseteq X''$ the probability that exactly X' consists of vertices playing B at tth stage and X'' at $t+1^{st}$ stage is the same in both the Linear Threshold model and the network formation game.

The proof is by induction on t. For $t = 0$, choosing $X = \emptyset$ and $X = X_0$, where X_0 is the initial set playing B, the probability is 1; otherwise 0 for both cases.

Now let X_t denote the set of vertices to which strategy B is spread at the $t - 1^{st}$ stage of the game. Take a vertex v such that v still plays A at the tth stage of the game. We need to show that the probability that v plays B at the $t+1^{st}$ stage of the game is the same in either model.

For the Linear Threshold Model, that v still plays A at the tth stage of the game implies $\theta_v \geq \sum_{u \in X_{t-1}} f(u,v)$. So θ_v is taken to be uniformly distributed in the interval $(\sum_{u \in X_{t-1}} f(u,v), 1]$. Therefore, the probability that v adopts strategy B is

$$\frac{\sum_{u \in X_t \setminus X_{t-1}} f(u,v)}{1 - \sum_{u \in X_{t-1}} f(u,v)}. \tag{4.17}$$

For the network formation game, v plays B at $t+1^{st}$ stage if and only if its active edge comes from $X_t \setminus X_{t-1}$. The probability that v's active edge does not come from X_{t-1} is $1 - \sum_{u \in X_{t-1}} f(u,v)$. On the other hand, it comes from $X_t \setminus X_{t-1}$ with probability $\sum_{u \in X_t \setminus X_{t-1}} f(u,v)$. Hence, the conditional probability that v plays B at the $t+1^{st}$ stage is the same with the Linear Threshold Model's.

So for any X, the probability that exactly the vertices of X plays B at certain stage of the game is the same under both models. We obtain

$Prob$[exactly X' plays B at t and exactly X'' plays B at $t+1$]
$= \sum_X$ $Prob$[exactly X plays B at $t-1$ and exactly X' plays B at t] $\cdot$ $Prob$[exactly $X'' \setminus X'$ plays B at t — X' plays B at t].

Hence, by induction, the expected number of vertices playing B at a certain stage of the game is the same for both. By Theorem 4.9., σ is submodular.

$\square$

Theorem 4.11. [22] *The influence function σ derived from the General Cascade Model where the probabilty functions are non-increasing, is monotone and submodular.*

Theorem 4.12. [18] *The influence function σ derived from the General Threshold Model where all the threshold functions are monotone submodular, is monotone and submodular.*

Now the proof of Theorem 4.3 is presented, we start with a lemma.

Lemma 4.13. *Let σ be a monotone, submodular function and X, Y be subsets of the set of vertices. Denote $\sigma(X \cup \{v\}) - \sigma(X)$ by $s_v(X)$ for $v \in V(G)$. Then,*

$$\sigma(Y) \leq \sigma(X) + \sum_{v \in Y \setminus X} s_v(X) \tag{4.18}$$

Proof. Start with labelling the elements of $Y \setminus X$ as $y_1, y_2, ..., y_n$. Since σ is monotone,

$$\sigma(Y) \leq \sigma(X) + \sum_{y_i \in Y \setminus X} \sigma(X \cup \{y_1, ..., y_i\}) - \sigma(X \cup \{y_1, ..., y_{i-1}\}). \tag{4.19}$$

And by submodularity,

$$\sigma(Y) \leq \sigma(X) + \sum_{y_i \in Y \setminus X} \sigma(X \cup y_i) - \sigma(X). \tag{4.20}$$

Proof of Theorem 4.3 [20] Let X^t be the set obtained at the tth iteration of the Algorithm 4.4. and X^* be the optimal set for the influence maximization problem. Denote $\sigma(X^{i+1}) - \sigma(X^i)$ by s_i. Then,

$$\sigma(X^t) = \sum_{i=0}^{t-1} s_i. \tag{4.21}$$

Since $|X^* - X^t| \leq N$ where N is the cardinality of $V(G)$, we obtain by the lemma above,

$$\sigma(X^*) \leq \sigma(X^t) + N s_t \tag{4.22}$$

for all t smaller than the number of iterations.

Then we have,

$$\sigma(X^{t+1}) \geq \sigma(X^t) + \frac{1}{N}\big(\sigma(X^*) - \sigma(X^t)\big). \tag{4.23}$$

Now the claim is

$$\sigma(X^t) \geq \Big(1 - \Big(1 - \frac{1}{N}\Big)^t\Big)\sigma(X^*). \tag{4.24}$$

The proof is by induction on t. For the case $t = 0$, $\sigma(X^t) \geq 0$ which holds.For the inductive step by the above inequality we have,

$$\begin{aligned} \sigma(X^{t+1}) &\geq \Big(1 - \frac{1}{N}\Big)\sigma(X^t) + \frac{1}{N}\sigma(X^*) && (4.25)\\ \sigma(X^{t+1}) &\geq \Big(1 - \frac{1}{N}\Big)\Big(1 - \Big(1 - \frac{1}{N}\Big)^t\Big)\sigma(X^*) + \frac{1}{N}\sigma(X^*) && (4.26)\\ & && (4.27) \end{aligned}$$

by the induction hypothesis. We obtained the desired result.

Since $\left(1 - \left(1 - \frac{1}{N}\right)^t\right) \geq \left(1 - \frac{1}{e}\right)$ for $t \leq N$, the theorem is proven.

□

4.3. Complexity of the Influence Maximization Problems

Theorem 4.14. [15, 16] *The influence maximization problem is NP-hard for the independent cascade model.*

Theorem 4.15. [15, 16] *The influence maximization problem is NP-hard for the linear threshold model.*

Proof. The proof is by showing that the vertex cover problem in Example 2.7. is a special case of the influence maximization problem of the linear threshold model, so that the influence maximization problem is as hard as the vertex cover problem which suffices to show NP-hardness.

Now, let N be the desired largest size of a subset of $V(G)$ such that a set of vertices X of smaller cardinality than N includes an endpoint of every edge in $E(G)$. If there exists such a subset X we want to show $\sigma(X) = |V(G)|$ where σ is the influence function. For this purpose we assume,

$$\sum_{u \in \Gamma(v)} f(u, v) \leq 1 \tag{4.28}$$

for any $v \in V(G)$ in the linear threshold model 4.1.1.

Then, if there exists such an X, among vertices those who does not belong to X have all their neighbors in X. But by the equation above, they will adopt the desired strategy at the first round, if X is chosen to be initial set. Therefore, $\sigma(X) = |V(G)|$.

Theorem 4.16. [15,16] *If submodularity is not satisfied for the general cascade model and the general threshold model, it is NP-hard to approximate the optimal value of the influence function within a factor of* $n^{1-\epsilon}$ *for any* $\epsilon > 0$ *where* n *is the number of vertices.*

5. CONCLUSION

In this thesis, we presented only a limited perspective on the cascading behavior in networks relying on a game theoretical perspective to the problem.

On the other hand, the models used in evolutionary biology to describe the cascading behavior are remarkably different. The difference is due to the fact that the agents in the models are completely exposed to a certain behavior, as they are infected with disease unlike the models described in the thesis where the agents adopt a behavior in accordance with best response dynamics.

Another perspective is the structure of the graph and its relation with the cascading behavior both in finance and biology. The type of connectedness of the graph can also play a significant role in the models we have described. This point of view needs a further scrutiny. Also similar studies on finite graphs with special structures ought to be performed.

As a last remark, introducing randomness to the decisions of players in the models has interesting results but we restrained ourselves only to deterministic models.

REFERENCES

1. Milgram, S., "The Small World Problem", *Psychology Today*, Vol. 1, pp. 60–67, 1967.
2. Banerjee, A., "Sign Language Tutoring Tool", *Proceedings of eNTERFACE 2007, The Summer Workshop on Multimodal Interfaces*, Vol. 21, 2007.
3. Bikhchandani, S., D. Hirshleifer and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, Vol. 100, pp. 72–83, 1992.
4. Blume, L., "The Statistical Mechanics of Best-Response Strategy Revision", *Games and Economic Behavior*, Vol. 11, pp. 111–145, 1995.
5. Kandori, M., G. Mailath and R. Rob, "Learning, Mutation and Long Run Equilibria in Games", *Econometrica*, Vol. 61, pp. 29–56, 1993.
6. Ellison, G., "Learning, Local Interaction, and Coordination", *Econometrica*, Vol. 62, pp. 1047–1071, 1993.
7. Ely, J., "Local Conventions", *Advances in Theoretical Economics*, Vol. 2, pp. 1–30, 2002.
8. Galesloot, B. M. and S. Goyal, "Costs of Flexibility and Equilibrium Selection", *Journal of Mathematical Economics*, Vol. 28, pp. 249–264, 1997.
9. Harsanyi, J. C. and R. Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA, USA, 1988.
10. Morris, S., "Contagion", *Review of Economic Studies*, Vol. 67, pp. 57–78, 2000.

11. Granovetter, M. S., "The Strength of Weak Ties", *American Journal of Sociology*, Vol. 78, pp. 1360–1380, 1973.

12. Montanari, A. and A. Saberi, "Convergence to Equilibrium in Local Interaction Games", *50th Annual IEEE Symposium on Foundations of Computer Science*, pp. 303–312, 2009.

13. Richardson, M. and P. Domingos, "Mining Knowledge-Sharing Sites for Viral Marketing", *Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 61–70, 2002.

14. Granovetter, M. S., "Threshold Models of Collective Behavior", *American Journal of Sociology*, Vol. 83, pp. 1420–1443, 1978.

15. Kempe, D., J. Kleinberg and E. Tardos, "Maximizing the Spread of Influence Through a Social Network", *In Proceedings 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, 2003.

16. Kleinberg, J., *Algorithmic Game Theory*, Cambridge University Press, New York, NY, USA, 2007.

17. Nemhauser, G., L. Wolsey and M. Fisher, "An Analysis of the Approximations for Maximizing Submodular Set Functions", *Mathematical Programming*, Vol. 14, pp. 265–294, 1978.

18. Mossel, E. and S. Roch, "Submodularity of Influence in Social Networks: From Local to Global", *SIAM Journal on Computing*, Vol. 39, pp. 2176–2188, 2010.

19. Easley, D. and J. Kleinberg, *Networks, Crowds and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, New York, NY, USA, 2010.

20. Kempe, D., *Structure and Dynamics of Information in Networks*, 2011, `http://www-bcf.usc.edu/ dkempe/teaching/structure-dynamics.pdf`, accessed at May 2012.

21. Garey, M. R. and D. S. Johnson, *Computers and Intractability : A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, CA, USA, 1979.

22. Kempe, D., J. Kleinberg and E. Tardos, "Influential Nodes in a Diffusion Model for Social Networks", *In Proceedings 32nd International Colloqium on Automata, Languages and Programming*, pp. 1127–1138, 2005.