Enabling and Controlling Diffusion Processes in Networks

Zhifeng Sun

May 4, 2011

Abstract

Diffusion processes are important models for many real-world phenomena, such as the spread of disease or rumors. We propose to study different aspects of diffusion processes in networks, focusing on designing efficient distributed algorithms for positive diffusion processes and good intervention strategies to control harmful diffusions.

First, we design and analyze various distributed algorithms for diffusion processes. We want to devise efficient distributed algorithms, which are easy to implement, to help the spreading of positive/useful information. We refer to these processes as positive diffusions. Earlier work has studied this for a variety of models, mainly based on static networks. The major point that separates our research with previous work is that we consider dynamically changing networks, which extends previous models to a larger range of real-life situations. Depending on the ways that networks are altered, we propose to study diffusion processes over the following two types of dynamic networks: (1) networks are changed due to individuals' decisions or behaviors; (2) networks are controlled by an adversary.

Secondly, we study how to devise good intervention strategies to control diffusion processes. This problem is crucial when we deal with harmful information like human diseases or computer viruses. We refer to these processes as harmful diffusions. We distinguish between centralized and decentralized intervention strategies. In centralized intervention strategies, there is a controller who has a limited amount of intervention resources (e.g. vaccinations or antidotes in the case of diseases). We study the problem of allocating these limited resources among the network agents so that the spread of the diffusion process is minimized. In decentralized intervention strategies, each individual in the network makes their own decision on protecting themselves, based on their individual utility and local knowledge. In such settings, we are interested in questions such as: is there a stable set of intervention strategies? What's the cost of decentralized solutions compared with an optimal centralized one? Lastly, we augment our studies of intervention strategies with the consideration about individual behavior changes which would lead to a new kind of network dynamics. Earlier work has shown that the combination of behavior change and intervention failure (e.g. failed vaccination) can lead to perverse outcomes where less (intervention resources) is more (effective). However, the extent of the perversity and its dependence on network structure as well as the precise nature of the behavior change has remained largely unknown.

1 Introduction

A number of real world applications can be modeled as diffusion processes over networks. Some prominent examples include diseases transmitted among humans, viruses transmitted over computer networks, information/ideas spread over contact networks, and creation of friendships through social networks. Despite the diversity among these applications, there are lots of fundamental similarities in the mathematical models. Understanding the dynamics of these mathematical models can help us anticipate, exploit, and control the propagation processes.

A lot of interesting questions could be asked about the dynamics. Consider human disease or computer virus transmission for example. Will it become an epidemic? How much time does it take to become an epidemic? Who will get infected? What's the social cost of the epidemic? Once we understand all these, we could design interventions to control the dynamics. For instance, how do we vaccinate or quarantine the population so that the epidemic is controlled? How do we secure computers to enhance the network resilience? What polices should be applied with budget constraints (limited vaccines or anti-virus software licenses), how should we distribute resources, and how much can we reduce the social cost? Often these interventions can be translated into voluntary directives from government, like take vaccines or stay at home. However, people usually don't adhere to such recommendations. Instead, they make decisions based on their specific utilities and objectives. Such decisions happen in a decentralized manner, which makes game theory a natural approach to study these problems. Moreover, people alter their contacts dynamically. For example, a vaccinated person may increase his/her contacts with friends, due to perceived secure feelings. These behavioral changes have a huge impact on the dynamics and the effectiveness of these interventions, so that "good" intervention strategies might be ineffective, depending on the behavioral changes. All these make the analysis of dynamically changing networks more interesting and challenging.

We classify our research work into the following two categories. First, we design and analyze efficient distributed algorithms for positive diffusion processes over dynamically changing networks. We introduce our work in Section 1.1, and state detailed results and proposed research in Section 3. Secondly, we study intervention strategies for harmful diffusion processes. This is extremely important when we deal with diseases, viruses, or other harmful information. Not only do we need to know when there is an epidemic, but also we should be able to design good intervention strategies to reduce our loss. We introduce this line of work in Section 1.2, and state detailed results and proposed research in Section 4.

1.1 Distributed algorithms for positive diffusions

In peer-to-peer (P2P), wireless, and sensor networks, how to spread information efficiently in a distributed manner is an important problem. Thus, designing simple (easy to implement and deploy) distributed algorithms, that provide a good guarantee on the spreading time, is crucial for the success of such systems. Indeed, a huge body of research work has been devoted to this area [57, 39, 89, 23, 35, 59, 43, 98]. However, the vast majority study and analyze the diffusion processes defined on static networks. In many P2P, wireless, and sensor networks, however, the underlying communication links may be broken temporarily, and users/nodes may join and leave over time. Only recently has there been some work considering diffusion over dynamic adversarial networks [4, 67]. Therefore, we propose and analyze several distributed algorithms over different kinds of dynamically changing networks. First we consider the case where networks are changed by the diffusion process itself. A canonical example of such a network is resource discovery problem, which is introduced in Section 1.1.1. The second dynamic model we consider is the adversary networks. We study token dissemination problem under such dynamics, which is introduced in Section 1.1.2.

1.1.1 Resource discovery

When a P2P network is formed, the first task for a node is to discover all the other available nodes on the network. We refer such problem as *resource discovery.* Designing good distributed algorithms for resource discovery can improve the efficiency and scalability of these P2P systems.

We model the network as a graph. Each node represents a computer/user in the network. If node u knows the existence of node v (namely node u discovered node v), then there is an edge between u and v. Each node can only discover new nodes through the nodes he already knows. Every time u discovers a new node in the network, he creates an edge to this node. We want to understand how much time it takes for every node in the network to discover all the nodes.

We propose two "simple" distributed algorithms for the resource discovery problem, *triangulation* and *2-hop random walk.* We run simulations for both algorithms on various families of graphs, and observe that both of them have good running time (nearly linear time). Thus, in Section 3.1, we focus on proving the following conjecture: the running time for triangulation and 2-hop random walk algorithms is $O(n \log n)$ on graphs with n nodes.

1.1.2 Token dissemination

The token dissemination problem is one of the most popular and well studied problems in distributed computing. At a high level, it can be stated as follows. Every node in the network has a piece of information she wants to spread. The goal is to design an distributed algorithm that spreads each piece of information to all the other nodes as efficiently as possible. Most of the studies on the token dissemination problem have been restricted to static networks, which do not model the real world very well, especially in the case of wireless and sensor networks, where the underlying links are not reliable and nodes can join and leave over time.

Inspired by the works [4, 67], we propose several efficient distributed algorithms for adversary networks. If the adversary can disconnect the graph, then it is imposible to disseminate tokens to all nodes. Thus, we require the graph to be connected at each round of communication. The adversary can decide how the nodes are connected. Our adversary model is very general; if we can devise good distributed algorithms under such a dynamic model, then we can apply them to more restricted dynamic models. We conjecture that the *randomized token forwarding* algorithm completes in $O(n \log n)$ time on all graphs.

1.2 Intervention strategies to prevent diffusions

When studying diffusion processes of harmful information (such as human diseases, computer viruses, gossips, etc.), an important task is to design good intervention strategies to prevent the diffusion. In the rest of this proposal, we often use human disease as an example of harmful information, and vaccination as an intervention example. Thus, when we talk about diseases and vaccinations, they refer to general harmful information and interventions.

We propose to study intervention strategies in three different settings. First, we look at centralized intervention strategies, where there is a centralized entity that is in charge of distributing vaccinations to control diseases. Second, we study decentralized intervention strategies, where individuals decide whether to vaccinate themselves in a decentralized manner. Lastly, we augment our models with behavior changes; when an individual is vaccinated, he may apply certain behavior changes due to the perceived secure feelings, which will alter the structure of the underlying contact network and hence have an impact on the diffusion process. We take into account the impact of change in the behavior of vaccinated individuals in conjunction with potential failure of the vaccines.

1.2.1 Centralized intervention strategies

For many diseases, such as influenza, vaccinations are a commonly used strategy in controlling the spread. A fundamental question in mathematical epidemiology is to determine what fraction of the population needs to be vaccinated in order to eradicate the disease, and how to allocate a limited supply of vaccines. The key point here is to identify a set of critical vertices in the graph and secure them, which could in turn cut the transmission path of diseases. However, such problems are often NP-hard to solve. Thus, finding efficient algorithms with good approximation ratios is desired. A lot of research work has been devoted to solving this problem [3, 12, 69, 102, 91, 66].

We generalize the mathematical model proposed in [3]. Undirected graph G represents the contact network between people/computers. For each node in G, there is an intervention cost and infection cost. These costs can vary between nodes. When a node takes intervention, it cannot be infected any more. Thus, we remove this node from G. A node is chosen randomly according to some arbitrary distribution to start the infection. Every other nodes that are in the same connected component will be infected. The goal is to minimize the sum of each individual's cost. We improve [3]'s $O(\log^{1.5} n)$ approximation ratio to $O(\log n)$. We also back up our theoretical results with comprehensive simulations over a number of families of graphs, which shows that, in practice, our approximation algorithm gives a much better guarantee. These results can provide guidance for government and network administrators to optimize resource allocation.

1.2.2 A game-theoretic study of decentralized intervention decisions

Often times, individuals make decisions based on their specific utilities and objectives in a decentralized manner. For example, in the case of disease transmission, individuals decide whether to secure themselves based on information including perceived infection cost, prices of vaccines (or antidotes), and decisions of other people. We use a game theoretic framework for analyzing decentralized intervention.

In our research, we found out that the spread of a disease and the intervention strategies crucially depend on the amount and quality of information available to the individuals. The amount of information available can be characterized by its locality: the distance d within the network up to which information is available to the vertex (d-local). Our results in [68] suggest that d has a big influence on the existence and structure of Nash equilibria [82]. More specifically, our research answers the following questions. Under what conditions do pure Nash equilibria exist for intervention games in which the information available to each node is d-local? What is the complexity of finding these equilibria? How good are these equilibria compared with social optimum in terms of costs (a.k.a. price of anarchy [65])?

1.2.3 Impact of behavioral changes

Most vaccines have very limited efficacy (typically 30% in the case of influenza). However, people are not very well aware of this limitation, and often over-estimate the efficacy of vaccines. Indeed, the perceived protection from infection might cause behavior changes, leading to an increase in contact by a vaccinated individual. In a series of important papers [17, 16], Sally Blower and her collaborators demonstrated risk behavior change, in the context of HIV vaccination, could lead to perverse outcomes.

We study the impact of behavior changes on disease dynamics in networks and observe (through simulation) a rich and complex behavior dependent both on the underlying network characteristics as well as the "sidedness" of the risk behavior change. The contact network is an undirected graph with each edge having a certain probability of disease transmission. We consider both uniform random vaccination (where each node is vaccinated independently with the same probability) as well as targeted vaccination (where nodes are vaccinated based on their degree of connectivity). Vaccines are assumed to fail uniformly and randomly. We model risk behavior change by an increase in the disease transmission probability. A significant aspect of our work is the consideration of "sidedness" of risk behavior change. We classify diseases as 1-sided or 2-sided based on whether the increase of disease transmission probability requires an increase in risk behavior of both the infector and the infectee or just the infector. As example: influenza (H1N1) may be modeled as 1-sided disease since a vaccinated individual may be motivated to behave more riskily (going to crowded places, traveling on planes, etc.), thus increasing the chance of infecting all he comes in contact with; whereas AIDS (HIV) may be modeled as 2-sided disease since the increase in disease transmission probability requires both the individuals participating in the interaction to engage in risky behavior. Of course, these examples are simplistic and most diseases have elements of both 1-sided and 2-sided risk behaviors.

With the existence of risk behavior, what would be good intervention strategies? Is targeted vaccination alway better than random vaccination? With more intervention resources, are we guaranteed to have less people infected? In Section 4.4, we present our findings, and propose to prove rigorously what we observed in the simulation.

2 Related work

In this section, we review related work that is most relevant to our research. We have classified previous related work into areas of interests (epidemiology, social networks, economics, and network security). We also give a brief overview of some of the mathematical and computational techniques and models that have been developed in previous work.

Epidemiology. The study of spread of diseases and the ways to control them has been a major research topic for centuries. One of the earliest documented scientific studies of vaccination policy was due to Bernoulli, who analyzed smallpox morbidity and mortality data to demonstrate the effectiveness of vaccination [15]. He presented a mathematical model for calculating the cost and benefits of smallpox vaccination, and presented a convincing mathematical analysis for universal vaccination. Modern epidemiological analysis is largely based on an elegant class of models, called SIR (susceptible-infected-recovered). This model was first formulated by Reed and Frost in 1920s, and has been developed over the decades. In the SIR model, the population is divided into three groups: susceptible (S), infected (I), and recovered (R), as shown in Figure 1. A *susceptible* individual becomes *infected* at a certain rate, β, when contacting with other infected individuals. Once infected, an individual may either recover and receive lifelong immunity or die at a certain rate, ν. In either case, it moves to *recovered* group. The lifelong immunity assumption is suitable for most common childhood diseases (measles, mumps, rubella, etc.). Another well studied model is the SIS (susceptible-infected-susceptible) model, shown in Figure 2. In SIS model, when an individual recovers from an infection, it moves back to the susceptible group [99]. Because recovered individuals can be infected again, this model is good to model fast mutated diseases, like seasonal flu.

Figure 1: SIR model.

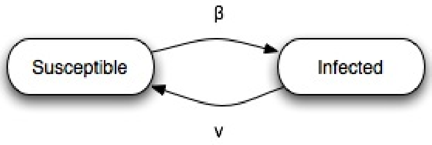

Figure 2: SIS model.

The SIR model and its variants have been highly influential in the study of epidemics [102, 73, 75, 56, 49, 74]. These models, however, do not attempt to capture the rich and structure of the contact network, along which interactions occur. Network structure has a direct effect on both the spread of diseases as well as the nature of interactions, which has been observed by a number of researchers, e.g. [83, 53]. In the emerging area of contact network epidemiology, an underlying contact graph captures the patterns of interactions which leads to the transmission of a disease [72, 77, 78, 84]. Many studies have predicted the spread of diseases through networks using mathematical analysis or simulations, e.g. estimating the size of an epidemic, or determining the basic reproductive rate (see [77] for some pointers).

Over the past decade, a number of researchers have analyzed the effectiveness of control strategies in a game-theoretic setting [10, 11, 9, 30]. Such studies enable the comparison between Nash equilibria strategies driven by self-interest [3, 42] and those centralized strategies proposed by health agencies.

A major focus of the thesis is on how intervention strategies may in turn cause risky behavior and affect network dynamics, which significantly affect the equilibrium and resulting perverse social outcomes. The counterintuitive impact of vaccination owing to risky behavior was first discovered by Blower and McLean [17, 16] in the case of HIV. Subsequently, several independent studies have confirmed this phenomenon and have offered potential explanations [94, 13, 100, 97]. However, no formal models have been proposed for studying the interplay between interventions, network dynamics, and individual behaviors.

Computer viruses and networking. Several researchers have analyzed network security problems and the spread of viruses in computer networks [3, 12, 43, 50, 69, 98]. Another direction of work is based on SIS models for the worm spread, e.g. the n-intertwined model [87]. In this model, nodes are in two states: susceptible or infected. Each infected node spreads the infection to its neighbors with some probability.

Non-cooperative game theory has been used in analyzing a number of problems in traffic and communication networks, e.g. routing [92], topology control and network formation [34, 81], and security [50, 87]. The basic question of interest have usually been about the existence, the structure of Nash equilibria, and the price of anarchy, which is the worst case cost of Nash equilibrium to the social optimum [65]. See [86] for a good introduction on the use of game theoretic techniques for networking applications.

Social networks and economics. The spread of contagion, influence, or behavior, in social networks has been extensively studied. This has spawned the research area of *network games*, starting from the seminal work of Jackson and Wolinsky [54], and Bala and Goyal [5] on how networks are formed when individuals choose to add or sever links so as to maximize their influence. A variety of game-theoretic models have been developed to study diverse applications including the very formation of networks [54, 5, 55, 37], provision of public goods [25], and research collaboration among firms [48]. (See [51, 47, 40] for excellent overviews of this research area.) Most relevant to this project is prior work on the diffusion of social and economic behavior in networks, e.g., purchase of a product or adoption of a technology [41], decision to undertake criminal activity [6]. While much of prior research has assumed that perfect information is available to all players, more recent work has developed models for incomplete information [40, 52]. There are several fundamental differences between our proposed project and this line of previous research. First, the focus of our work is on the *impact of intervention strategies on the spread of contagion*, rather than analyzing how contagions spread. Second, we take into account *changes in the network* in the course of the diffusion process. Third, we consider a much *richer strategy space* which takes into account temporal issues (when an intervention is adopted) and the strength of the interventions; in contrast, previous work assumes binary strategy spaces.

There is a huge body of work on *gossiping*, applied to both the spread of ideas in a social network as well as routing or broadcast of data in a communication network [57, 39, 89, 23]. Several analyses have been given for gossip-based routing protocols [24, 35, 59]. Gossiping has also proved effective in aggregate computations in distributed networks [58, 32]. Recent work has considered gossiping in dynamic adversarial networks [4, 67]. Combinatorial optimization aspects of influence spread are explored in [60, 61] where the goal is to pick an initial set in a stochastic model with maximal expected influence. This model is extended further in [14] to a competitive setting within the stochastic framework where different players compete (sequentially) to maximize their expected influence.

Models and techniques. The SIR, SIS, and related models in epidemiology can be characterized in terms of differential equations, which captures the rate of change of the susceptible, infected, and recovered individuals.

Contact network epidemiology enhance these models with contact graphs. Contact graphs are also important aspects of models in network security and social networks. A number of random graph models have been considered for modeling contacts (e.g. [78, 46, 76]). The most basic random graph model is due to Erdös and Rényi [36] and Gilbert [45], who defined the $G(n,m)$ and $G(n,p)$ models. In the former, a graph is chosen uniformly at random from all n-vertex m-edge graphs, while the latter is an n-vertex graph in which each edge appears independently at random with probability p. A number of alternative random graph models have been suggested in an attempt to yield random graphs with more general type of degree distribution. Molloy and Reed's model is based on considering random graphs of fixed order with a given arbitrary degree distribution [79, 80, 85]. Later, Chung and Lu proposed a model to generate random graphs with expected degree distribution [28, 29]. Other models motivated by specific real-world networks include Kleinberg's small world random graphs [62, 63], and Barabási-Albert preferential attachment graph [7]. The preferential attachment model is a mechanism for generating random scale-free graphs [7, 21], which have been observed to capture important properties of several real world networks, including the Internet AS-level graph [38], the Worldwide Web [1, 64], human sexual contact network [71], scientific collaborations network [8].

The spread of diseases/viruses in networks has been analyzed by a variety of methods: model-based simulations [2, 88, 31]; percolation theory techniques using differential equations and generating functions that yield asymptotic analysis [95, 84, 101]; rigorous probabilistic techniques that apply to finite random graph models [19, 20, 12]; and game-theoretic analysis that studies uniqueness, complexity, and the structure of equilibria [26, 86, 52, 40].

3 Distributed algorithms for positive diffusions

Designing efficient distributed algorithms for positive diffusions is crucial for the success of P2P, wireless, and sensor networked systems. Indeed, a lot of research work has been devoted to this problem [57, 39, 89, 23, 35, 59, 44, 98, 22, 70]. A major limitation in this line of work, however, is they only consider diffusion processes defined on static networks. Only recently, has there been some work on diffusion over dynamic adversarial networks [4, 67].

In our proposed research, we focus on dynamically changing networks. We consider the following two types of dynamics. First, in Section 3.1, we look at the *resource discovery* problem, where the underlying dynamic network is changed by the diffusion process itself. Second, in Section 3.2, we look at the *token dissemination* problem, where the underlying dynamic network is controlled by an adversary.

3.1 Resource discovery

When a node joins a P2P system, it first discovers all the other available nodes on the network. Such a problem is referred as resource discovery problem. We model the P2P system by a graph. Each node represents a resource (or a computer) in the network. If node u knows node v, we assume that u can communicate with v, and hence there is an edge between u and v. For example, if u knows v's IP address, u can send packages to v. Thus, as a node discovers more and more other nodes, the graph changes dynamically. Such dynamic is induced by the diffusion process itself.

We designed two distributed algorithms for resource discovery problem, which are very easy to implement and deploy. The first algorithm is the following: in each round, every node in the graph chooses two of its neighbors randomly and "introduces" them to each other. This algorithm can be analyzed by the follow *triangulation process.*

Triangulation process. Given a connected graph $G = (V, E)$, in each round, for all $v \in V$, vertex v picks 2 random neighbors u and w, add an edge (u, w) to graph G. The question is how many rounds this process needs to take in order to make G a complete graph K_n.

The second algorithm is, in each round, every node in the graph choose a random neighbor, and ask this neighbor to tell him a random node that he knows. This algorithm can be analyzed by the following *2-hop random walk process.*

2-hop random walk process. Given a connected graph $G = (V, E)$, in each round, for all $v \in V$, vertex v takes a 2-hop random walk and reaches vertex u. We add an edge (v, u) to graph G. The question is how many rounds this process needs to take in order to make G a complete graph K_n.

3.1.1 Our results

We have done a comprehensive simulations for both triangulation process and 2-hop random walk process, over different families of graphs. We observe that both processes complete quickly, in nearly linear time. Thus, we conjecture that this is true across all families of graphs. We will focus on proving this conjecture in future research.

3.1.2 Proposed research

We plan to address the following open questions. We mainly focus on proving rigorously what we observed in the simulation.

1. We conjecture that the completion time for the triangulation process and 2-hop random walk process is $O(n \log n)$ in undirected graphs. We will work on resolving this conjecture.
2. We conjecture that the completion time for the 2-hop random walk process in directed graphs is $O(n^2)$. We also conjecture that it is lower bounded by $\Omega(n^2)$. We will work on resolving these conjectures.

3.2 Token dissemination

Inspired by [4, 67], we study the token dissemination problem in adversarial networks. The adversary model is important for many applications in wireless and P2P networks. Such networks are highly dynamic, as nodes join, leave, and move around, and as links appear and disappear. Such dynamic changes are not under the control of users in the network.

3.2.1 Definition

Let V be the set of users in the contact/communication network, and let $|V| = n$. There are k different pieces of information assigned to a set of users, which we will refer as *tokens*. Each user can hold different tokens. The goal is to diffuse all k tokens to all the users on the network. This information diffusion problem is often refereed to as the k-token dissemination problem.

We consider synchronous communication in this model. The diffusion process is divided into rounds. At the beginning of each round, the *adversary* constructs the edges/links in the contact/communication network. Namely, he decides which users can talk to whom directly. Then, every user is allowed to exchange 1 token (or constant number of tokens) with each of its neighbors in the contact network. In our model, we assume the message size to be $\Theta(\log n)$. Observe that, if the adversary is allowed to make the contact network disconnected, then diffusing all k tokens to all users is impossible to complete. Therefore, we enforce the *connectivity* constraint on the adversary. This means the adversary can arrange edges whichever way he wants, but he has to make sure the contact network is connected in each round.

3.2.2 Our results

Based on the model stated in Section 3.2.1, we proposed 2 distributed algorithms. One is PriorityForward, shown in Algorithm 1. The other is RandomForward, shown in Algorithm 2. We have shown the running time lower bound for PriorityForward is $\Omega(kn)$.

Algorithm 1 PriorityForward (G)

```
Require: Dynamic network $G = (V, E)$.
Ensure: Nodes receive all $k$ tokens.
  Each node $v$ executes the following procedure:
  for all $u$ such that $u$ is $v$'s neighbor in $G$ do
      $v$ sends $u$ a token with lowest possible priority such that $u$ doesn't have this token yet.
  end for
```

Algorithm 2 RandomForward (G)

```
Require: Dynamic network $G = (V, E)$.
Ensure: Nodes receive all $k$ tokens.
  Each node $v$ executes the following procedure:
  for all $u$ such that $u$ is $v$'s neighbor in $G$ do
      $v$ sends $u$ a token randomly chosen from the tokens that $v$ has and $u$ doesn't have yet.
  end for
```

3.2.3 Proposed research

We propose the following research problems in the adversarial network domain.

1. We conjecture that RandomForward algorithm completes in $O(n \log k)$ time. We will work on resolving the conjecture.

2. In RANDOMFORWARD algorithm line 2, we didn't state the time complexity for v to choose a token which u doesn't have at random. This problem can be formulated as follows. Nodes A and B both have an n-bit vector. A doesn't know the values of B's bit vector, and vice versa. Each time one can send the other a message of size $O(\log n)$, to query the values of its bit vector. How many messages have to be exchanged for A to identify a random bit position among all positions that are set in B's vector but not in A's vector.

4 Intervention strategies to prevent diffusions

In this section, we study how to design good intervention strategies to control harmful diffusion processes. We characterize intervention strategies into two categories, centralized and decentralized. Both kinds of strategies share a common mathematical framework, which we define formally in Section 4.1. Then we state our results and propose research for centralized and decentralized intervention strategies in Section 4.2 and 4.3, respectively. Lastly, in Section 4.4, we enrich our intervention strategy designs with the consideration of individual behavior changes.

4.1 Definition and framework

All three areas of our research on intervention strategies (centralized intervention strategy, decentralized intervention strategy, behavior changes) have a common model framework. Thus, we briefly present the framework first, and then state our results and proposed research in subsequent sections.

Let V denote the set of users/individuals/computers (henceforth, referred to as *nodes*). Let $G = (V, E)$ denote the underlying contact graph over V; an edge $e = (u, v) \in E$ indicates that the diffusion can spread from u to v. We let $p(e)$ or $p(u, v)$ denote the probability that diffusion spread from node u to node v. For $S \subset V$, we let $G[S]$ denote the subgraph of G induced by set S. We assume infection is initiated at a node chosen from V according to some arbitrary probability distribution. Let w_v denote the probability that node v is chosen as the initial infection point, and let $w(S) = \sum_{v \in S} w_v$ for $S \subset V$. In most of this proposal, we use the SIR model for epidemics, in which each infected node u infects each susceptible neighbor v independently with probability $p(u, v)$ (see [33, 84, 83] for additional information). But, our framework and associated questions are relevant for a broader class of diffusion processes. Nodes which are successfully vaccinated are assumed not to spread the infection. An important aspect of our framework is that interventions, like vaccines, are only partially efficacious, and succeed only with probability p_s. However, nodes do not know whether the intervention has succeeded, and might alter their behaviors.

The strategy for each node v consists of two decisions:

1. Whether to adopt an intervention, such as vaccination or quarantine (modeled node deletion). Let $a_v \in [0, 1]$ denote the probability that node v gets vaccinated. Let $\vec{a}$ denote the strategy vector for all nodes. We need the notation of *attack graph*, denoted by $G_{\vec{a}}$, which is the subgraph of G in which each node is deleted with probability a_v [3].

2. Behavioral changes that alter connections: such behavior changes could happen in many ways, and we abstract it by considering $p_b(u, v)$, which denotes the modified transmission probability on edge (u, v). While the behavior changes can be manifested as fairly general vectors $\vec{p_b}$, we focus on an important special case where these changes occur only when applied interventions. We consider two specific models. In the *1-sided model*, vertex u after vaccination (if failed) increases the transmission probability to p_b on all incident edges (u, v), while in the *2-sided model*, vertex u after vaccination (if failed) increases the transmission probability only on edges (u, v) for which vertex v is also vaccinated.

The rationale behind the behavioral change models is that after vaccination, it makes sense for an individual to take advantage of the social contact network to enhance their economic value. We formalize this aspect below. The 1- and 2-sided models signify two potential ways in which the contagion spread could be affected: whether it requires increased contact from any one endpoint, or both endpoints. There are a number of natural hybrid models that we will also consider in our study. Since increased contact comes with increased risks, we sometimes refer to this behavior change as "risky behavior".

The information used by nodes is an important aspect of the decision making process, and we quantify this in terms of a parameter d, which denotes the maximum number of hops to which a node can get information about infections and/or vaccinations. Now we consider a generalized non-cooperative game formulation $\mathcal{G}(G, d)$ by defining individual costs, which involve the following components: (i) C_v denotes the cost of node v getting vaccinated. We assume this is independent of the disease dynamics; (ii) B_v denotes the "benefit" of increased contact (e.g., in the form of information and centrality of the node in the social organization); and (iii) L_v denotes the cost for node v resulting from an infection. We assume that node v only has information about the graph within its d-neighborhood; let $p_v(\vec{a}, \vec{p_b}, d)$ denote the probability that node v becomes infected, given the decision vectors $\vec{a}$ and $\vec{p_b}$. Then, the cost to node v is defined as

$$\text{cost}_v\left(\bar{a}, \vec{p_b}, d\right) = C_v a_v + B_v p_b(v) + (1 - a_v) L_v \cdot p_v\left(\bar{a}, \vec{p_b}, d\right). \tag{1}$$

A Nash equilibrium [82] (henceforth, NE) is a strategy vector $\vec{a}$ such that no node v has any incentive to switch his strategy, if all other nodes' strategies are fixed. In most of the proposal, we focus on pure NE. We consider variants in which the term $p_v(\vec{a}, \vec{p_b}, d)$ using only partial information about the graph and individual decisions within the d-neighborhood; as we discuss later, information has a significant impact on the dynamics and structure of NE in these games.

The total social cost of a strategy profile is the sum of the individual costs, which is $\text{cost}\left(\bar{a}, \vec{p_b}, d\right) = \sum_{v=1}^{n} \text{cost}_v\left(\bar{a}, \vec{p_b}, d\right)$. A socially optimal strategy is a pair of vectors $\vec{a}, \vec{p_b}$ that minimizes this cost - this is not necessarily (and is not usually) a pure NE. Therefore, the cost of a pure NE relative to the social cost is an important measure; the maximum of this ratio (i.e., over all possible pure NE) is also known as the *price of anarchy* [65].

4.2 Centralized intervention strategies

In this section, we study centralized intervention strategies. The goal is to minimize the social cost define in Section 4.1. Here we restrict our graphs to be static, and assume there is no behavior change. Therefore, the problem boils down to find a strategy vector $\vec{a}$ such that

$$\sum_v \text{cost}_v\left(\bar{a}, d\right) = \sum_v \left[C_v a_v + (1 - a_v) L_v \cdot p_v\left(\bar{a}, d\right)\right]$$

is minimized.

4.2.1 Our results

In [68], we have designed an algorithm that is an $O(\log n)$-approximation for the $d = \infty$ case (recall d is locality parameter), and a $2d$-approximation for $d < \infty$. A logarithmic-approximation for the $d = \infty$ case was also obtained independently by [26].

4.2.2 Proposed research

If time permits, I'd like to try the following problem.

- Consider the same problem in the setting where arbitrary disease transmission probability and behavior change are allowed. One hurdle for such problem is, with arbitrary disease transmission probability, even calculating the cost of an individual is hard. We plan to pursue two directions. One is by using the notion of cut-based decompositions of networks. Thanks to an elegant result of Räcke that approximates arbitrary networks by trees [90] (note that vaccination strategy is equivalent to vertex cut in graphs). The other is by building on work of [27] for percolation in arbitrary finite graphs. Their results imply that for constant-degree networks, the cost of an intervention strategy is asymptotically dominated by those clusters (obtained after applying the intervention) that have high second order average degree, defined as $\sum_v d(v)^2/(\sum_v d(v))$, where $d(v)$ denotes the degree of node v in the cluster. It is still challenging to extend these ideas to derive efficient approximation algorithms.

4.3 Decentralized intervention decisions

In this section, we analyze diffusion processes using the non-cooperative game model defined in Section 4.1. More specifically, we consider the games with static connections, which means we assume there are no behavioral changes. Therefore, by Equation 1, the cost function for v can be simplified as

$$\text{cost}_v(\bar{a}, d) = C_v a_v + (1 - a_v) L_v \cdot p_v(\bar{a}, d)$$

As defined in Section 4.1, $p_v(\bar{a}, d)$ denotes the probability that v gets infected, if the disease starts at a random initial node in the d-neighborhood of v.

Under the non-cooperative game model, we are interested in answering the following questions:

1. Under what conditions do NE exist, when information (about the individual strategies and disease states) is available only within the d-neighborhood? If NE exist, when are they unique and how robust are they to perturbations in the network or locality of information? What is the computational complexity of finding NE or reaching NE via best-response dynamics? What is the effect of the underlying network (e.g., the degree sequence and conductance)? Can we characterize the kinds of nodes that tend to get vaccinated in any NE?

2. Price of anarchy: Find a vaccination strategy $\vec{a}$ that minimizes the total cost. What is the maximum ratio of the cost of a social optimum and cost of NE (referred to as the price of anarchy)? What are practical strategies that cause convergence to a socially desirable NE?

4.3.1 Our results

The results in this section are published in [68]. There we have considered the special case where $p(e) = 1$ for all e and intervention are perfect. We find that the parameter d has a significant role in the structure of the resulting games.

1. For $d = 1$, a pure NE always exists and can be found by best response dynamics; that is, every sequence of best response steps by the individual players converges to a pure NE. Finding the NE with least cost is NP-complete. The price of anarchy is $\Delta + 1$, where Δ is the maximum degree in the contact graph.

2. For $d = \infty$, a pure NE always exists and can be found by best response dynamics. The price of anarchy is $O(1/\alpha(G))$, where $\alpha(G)$ is the vertex expansion of graph G.

3. For $d \in (1, \infty)$, there always exist instances which have no pure NE. Further, deciding if a pure NE exists is NP-complete.

4.3.2 Open problems

I may try to think about the following open problems if time allows.

1. Solve the general model with arbitrary disease transmission probability. This problem is much more challenging to handle. One of the difficulties we face in terms of best response function is that with probabilistic transmission, even the calculation of the cost for an individual node can be shown to be #P-hard [96]. Efficient approximations (using Monte Carlo simulations) can be computed, however, leading to the question: do "approximate" NE exist and can they be efficiently found?

2. Explore more on the structures of NE and its implications on public policy. In the simulation using the preferential attachment graph model [68], we have observed that high degree nodes tend to get vaccinated in NE (when they exist). It would be nice to have some theoretical proofs about such observations.

3. Analyze Stackelberg strategies, in which the choices of some nodes can be control by the policy planer, and the goal would be to design low cost strategies that cause the remaining nodes to reach a "good" NE.

4.4 Behavioral changes

In this section, we study non-cooperative games with connection-altering interventions (a.k.a. behavior changes). As before, locality is parameterized by d. Among all possible models for behavior changes, we focus on the 1-sided and 2-sided model (defined in Section 4.1). We are interested in studying the following questions. How are the properties of NE affected by the behavior change parameter p_b and the vaccine efficacy? Are there threshold phenomena involving these parameters, so that an increase in p_b beyond the threshold value p_b^* changes the dynamic significantly? How does price of anarchy change? Furthermore, what kind of nodes tend to choose to get vaccinated? Can we achieve equilibria that are close to the social optimum through a public policy that provides certain amount of vaccines (subject to a budget) and provides appropriate local and global information to individual players?

4.4.1 Our results

As a first step, we will focus on symmetric strategies, in which the intervention strategy a_v and p_b are the same for all nodes. We want to understand how diffusion dynamics (e.g. the average outbreak size, and the time to peak), and the total utility as defined in Equation 1 vary as a function of a_v and p_b. Furthermore in order to answer the public policy issues would require understanding the interplay between the locality parameter d and the behavior change parameter p_b, which may depend significantly on the vaccine efficacy and could be very different in the 1-sided and 2-sided models.

In our preliminary work, we have studied the social cost and differences between 1-sided and 2-sided models. We assume that interventions failed with probability p_f. Through simulations, we found that both forms of behavior changes would lead to perverse outcomes across a wide range of contact networks. 1-sided behavior change leads to perverse outcomes at low levels of intervention, in which the average outbreak increases with a_v, up to a point, as shown in Figure 3. 2-sided behavior change leads to perverse outcomes at high levels of intervention, in which the average outbreak size starts increasing beyond a threshold value of a_v.

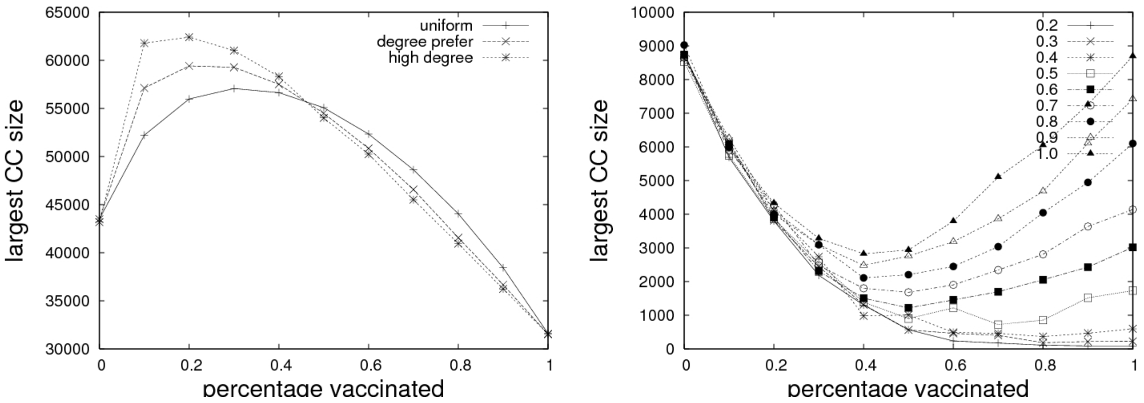

Figure 3: Variation in the giant component size with a_v for the 1-sided and 2-sided models of behavior changes in scale-free graphs. On the left is the comparison of uniform vaccination with two kinds of degree-preferred vaccination policies in the 1-sided model, for $p = 0.2$, $p_b = 0.9$, $p_f = 0.4$. On the right are the curves for different values of behavior change probability p_b in the 2-sided model, for $p = 0.1$, $p_f = 0.2$. The giant component size is a good estimate for the expected epidemic size.

To answer those questions in which we are interested, a significant fraction of our efforts will be devoted to developing analytical tools. Our initial foray has been on Erdös-Rényi networks that have formed the basic underlying network in traditional SIR models. A major technical hurdle we face is that when the behavioral models are incorporated, even homogeneous Erdös-Rényi networks are transformed to heterogeneous networks. One powerful tool for analyzing percolation on heterogeneous networks is that of *multi-type branching processes.* We plan to build on recent results of Söderberg [95] and Bollobás et al [19] in this regard.

Consider a random graph model denoted by $\mathcal{G}(N, K, \mathbf{r}, \mathbf{c})$, where (i) K is a positive integer, (ii) $\mathbf{r} = \{r_1, \ldots, r_K\}$ is a probability vector, (iii) $\mathbf{c} = (c_{ij})$ is a $K \times K$ matrix, (iv) each node $j = 1, \ldots, N$, is assigned a type $i \in \{1, \ldots, K\}$ with probability r_i, and (v) each pair of nodes i, j are connected by an edge with probability $p(i,j) = c_{ij}/N$. Söderberg [95] and Bollobás et al. [19] establish the following: (i) if the eigenvalues of the matrix $\{c_{ij}r_j\}$ are all less than 1, it is sub-critical (i.e., has no giant component), and (ii) if some eigenvalue is larger than 1, it is super-critical (i.e., has a giant component) with asymptotically $r_i(1-f_i)N$ nodes of type i, where f_i satisfies the coupled set of equations: $f_i = exp\left(\sum_j c_{ij} r_j (f_j - 1)\right)$. Using the above threshold result, we have given a rigorous proofs for the diverse non-monotonicity phenomena of Figure 3 in the case of Erdös-Rényi random graphs.

Empirically, we have observed that the phenomenon is widespread across a range of networks including preferential attachment networks. We are also applying traditional percolation theory techniques to the class of locally-finite infinite networks [18] as a proof of concept. Our findings have implications for public policy and the distribution of vaccines. More surprisingly we observe that targeted vaccination can be strictly worse than random vaccination for the same level of vaccine coverage and this phenomenon occurs both for one-sided as well as two-sided risk behavior change (as shown in Figure 3). Given the prior work on targeting vaccine distributions this finding flies in the face of intuition that expects that vaccinating highly connected individuals would always confer greater benefits.

4.4.2 Proposed research

We propose to study the following open problems.

1. We have a rigorous proof for the existence of non-monotonicity in Erdös-Rényi random graphs. Can we capture more detailed features of the non-monotonicity curves shown in Figure 3 (e.g. do the curves have unique maximum/minimum point)?.
2. Extend rigorous proofs of Erdös-Rényi random graphs to more general networks, like preferential attachment graphs or locally-finite infinite graphs, to explain the diverse non-monotonicity phenomena shown in Figure 3.
3. In the simulation, we observe targeted vaccination can be strictly worse than random vaccination for some levels of vaccine coverage. Can we mathematically prove this?
4. Develop simulations to cover more families of graphs to confirm our findings.

5 Proposed timeline and plan

May-July, 2011	Proposal defense, work on open problems in Section 3.1 and 3.2 .
June-August, 2011	Work on open problems in Section 4.4. If time permits, I will try open problems in Section 4.2 and Section 4.3.
August-November, 2011	Consider any manageable open questions that come up while writing the final thesis.
December	Tentative defense date.

5.1 Plan

Existing, published results (the majority of which are listed in the "Our results" sections of this proposal) will make up the core of my thesis. There are some open problems in Section 3.1 and 3.2 that I think I can solve between May and July. I plan to first work on proving both triangulation process and 2-hop random walk process complete in $O(n \log n)$ time. Then I will work on proving RandomForward algorithm (defined in Section 3.2) can complete in $O(n \log n)$ time. Along with the analysis on PriorityForward algorithm (also defined in Section 3.2), this will give us some insight on the role of randomness in adversary networks. If time permits, I'd like to try some other open problems in the same section, like communication complexity lower bound for Line 2 in Algorithm 2.

Next, I'd like to spend some time focusing on open problems in Section 4.4. We have some game theoretic results on intervention strategies (in Section 4.2 and 4.3), but little has been done on the effects of behavior changes. We already observed that the perverse outcome (shown in Figure 3) is widespread across a range of networks; it would be nice if we can nail down the rigorous proofs for some general families of graphs, e.g. preferential attachment graphs and locally-finite infinite graphs. We also observe in the simulation that randomly picked intervention can be better than targeted strategies in preferential attachment graphs. I will think about how to mathematically prove our observations. And I will take some time to develop more simulations to cover a wider range of parameters and different families of graphs, to confirm our findings in more broader settings. If I still have time left, I will look at the open problems in Section 4.2 and 4.3.

Finally, I've left three months for following up on loose ends and assembling the thesis by the end of November.

References

[1] R. Albert, H. Jeong, and A.L. Barabási. Diameter of the world wide web. *Nature*, 1999.

[2] R. Albert, H. Jeong, and A.L. Barabási. Error and attack tolerance of complex networks. *Nature*, 2000.

[3] J. Aspnes, K.L. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *J. Comput. Syst. Sci.*, 2006.

[4] Chen Avin, Michal Koucký, and Zvi Lotker. How to explore a fast-changing world (cover time of a simple random walk on evolving graphs). In *ICALP (1)*, pages 121–132, 2008.

[5] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[6] C. Ballester, A. Calv-Armengol, and Y. Zenou. Who's who in networks. Wanted: The key player. *Econometrica*, 74:1403–1417, 2006.

[7] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509512, 1999.

[8] A.L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica*, 2002.

[9] C. Bauch. Imitation dynamics predict vaccinating behaviour. *The Royal Society B*, 2005.

[10] C. Bauch, A.P. Galvani, and D.J.D. Earn. Group interest versus self-interest in smallpox vaccination policy. *PNAS*, 2003.

[11] C.T. Bauch and D.J.D. Earn. Vaccination and the theory of game. *PNAS*, 2004.

[12] N. Berger, C. Borgs, J.T. Chayes, and A. Saberi. On the spread of viruses on the Internet. *Proceedings of SODA*, 2005.

[13] D. Bezemer, F. Wolf, M.C. Boerlijst, A. Sighem, T.D. Hollingsworth, M. Prins, R.B. Geskus, L. Gras, R. Goutinho, and C. Fraser. A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. *AIDS*, 2008.

[14] Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In *WINE*, pages 306–311, 2007.

[15] S. Blower. Daniel Bernoulli's "an attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. 1766". *Rev Med Virol*, 14:275–88, 2004.

[16] S. Blower, K. Koelle, and J. Mills. Health policy modeling: epidemic control, HIV vaccines and risky behavior. In *Quantitative Evaluation of HIV Prevention Programs*. Yale University Press, 2002.

[17] S.M. Blower and A.R. McLean. Prophylactic vaccination, risk behavior change, and the probability of eradicating HIV in San Francisco. *Science*, 1994.

[18] B. Bollobas. *Percolation Theory*. Cambridge University Press, 2006.

[19] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 2007.

[20] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Mathematics*, 2004.

[21] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Struct. Alg.*, 2001.

[22] C. Borgs, J.T. Chayes, A. Ganesh, and A. Saberi. How to distribute antidotes to control epidemics. *Random Structures and Algorithms*, 2010.

[23] Stephen P. Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[24] David Braginsky and Deborah Estrin. Rumor routing algorithm for sensor networks. In *WSNA*, pages 22–31, 2002.

[25] Y. Bramoullé and R. Kranton. Strategic experimentation in networks. *Journal of Economic Theory*, 135:478–494, 2007.

[26] Po-An Chen, Mary David, and David Kempe. Better vaccination strategies for better people. In *ACM Conference on Electronic Commerce*, pages 179–188, 2010.

[27] F. Chung, P. Horn, and L. Lu. The giant component in a random subgraph of a given graph. *WAW*, 2009.

[28] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, 2002.

[29] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math.*, 2003.

[30] M.G. Cojocaru. Dynamic equilibria of group vaccination strategies in a heterogeneous population. *Journal of Global Optimization*, 2008.

[31] D. M. Conforth, T. C. Reluga, E. Shim, C. T. Bauch, A. P. Galvani, and L. A. Meyers. Erratic flu vaccination emerges from short-sighted behavior in contact networks. Under review, August 2010.

[32] Supratim Deb, Muriel Médard, and Clifford Choute. Algebraic gossip: a network coding approach to optimal multiple rumor mongering. *IEEE Transactions on Information Theory*, 52(6):2486–2507, 2006.

[33] Rick Durrett. *Random Graph Dynamics.* Cambridge University Press, New York, NY, USA, 2006.

[34] Stephan Eidenbenz, V. S. Anil Kumar, and Sibylle Zust. Equilibria in topology control games for ad hoc networks. *MONET*, 11(2):143–159, 2006.

[35] Robert Elsässer and Thomas Sauerwald. On the runtime and robustness of randomized broadcasting. *Theor. Comput. Sci.*, 410(36):3414–3427, 2009.

[36] P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math.*, 1960.

[37] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H. Papadimitriou, and Scott Shenker. On a network creation game. In *PODC '03*, pages 347–351, New York, NY, USA, 2003. ACM Press.

[38] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Proceedings of SIGCOMM*, 1999.

[39] Uriel Feige, David Peleg, Prabhakar Raghavan, and Eli Upfal. Randomized broadcast in networks. In *SIGAL International Symposium on Algorithms*, pages 128–137, 1990.

[40] A. Galeotti, S. Goyal, M. Jackson, F. Vega-Rodondo, and L. Yariv. Network games. *Review of Economic Studies*, 77:218–244, 2010.

[41] Andrea Galeotti. Consumer networks and search equilibria. Tinbergen Institute Discussion Paper 2004-75, 2004.

[42] A.P. Galvani, T.C. Reluga, and G.B. Chapman. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. *PNAS*, 2007.

[43] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *Proceeding of INFOCOM 2005*, 2005.

[44] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. *Proceedings of INFOCOM*, 2005.

[45] E.N. Gilbert. Random graphs. *Ann. Math. Statist.*, 1959.

[46] J. Gómez-Gardenes, V. Latora, Y. Moreno, and E. Profumo. Spreading of sexually transmitted diseases in heterosexual populations. *PNAS*, 2007.

[47] S. Goyal. *Connections: An Introduction to the Economics of Networks.* Princeton University Press, New Jersey, 2007.

[48] S. Goyal and J. L. Moraga-Gonzalez. R&D networks. *Rand Journal of Economics*, 32:686–707, 2001.

[49] N.C. Grassly and C. Fraser. Mathematical models of infectious disease transmission. *Nature*, 2008.

[50] J. Grossklags, N. Christin, and J. Chuang. Secure or insure? A game-theoretic analysis of information security games. In *World Wide Web Conference (WWW)*, 2008.

[51] M. Jackson. *Social and Economic Networks.* Princeton University Press, New Jersey, 2008.

[52] M. Jackson and L. Yariv. Diffusion of behavior and equilibrium properties in network games. In *American Economic Review (Papers and Proceedings)*, volume 97, pages 92–98, 2007.

[53] M. Jackson and L. Yariv. Diffusion, strategic interaction, and social structure. In *The Handbook of Social Economics*, 2010.

[54] Matthew Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.

[55] Ramesh Johari, Shie Mannor, and John N. Tsitsiklis. A contract-based model fordirected network formation. *Games and Economic Behavior*, 56(2):201–224, 2006.

[56] E.H. Kaplan, D.L. Craft, and L.M. Wein. Emergency response to a smallpox attack: The case for mass vaccination. *PNAS*, 2002.

[57] Richard M. Karp, Christian Schindelhauer, Scott Shenker, and Berthold Vöcking. Randomized rumor spreading. In *FOCS*, pages 565–574, 2000.

[58] David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-based computation of aggregate information. In *FOCS*, pages 482–491, 2003.

[59] David Kempe and Jon M. Kleinberg. Protocols and impossibility results for gossip-based communication mechanisms. In *FOCS*, pages 471–480, 2002.

[60] David Kempe, Jon M. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *KDD*, pages 137–146, 2003.

[61] David Kempe, Jon M. Kleinberg, and Éva Tardos. Influential nodes in a diffusion model for social networks. In *ICALP*, pages 1127–1138, 2005.

[62] J. Kleinberg. Navigation in a small world. *Nature*, 2000.

[63] J. Kleinberg. The small world phenomenon: An algorithmic perspective. *Proceedings of STOC*, 2000.

[64] J. Kleinberg, S.R. Kumar, P. Raphavan, S. Rajagopalan, and A. Tomkins. The web as a graph: Measurements, models and methods. *Proceedings of the International Conference on Combinatorics and Computing*, 1999.

[65] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *STACS '99*, pages 404–413, March 1999.

[66] C. Kuhlman, V.S. Anil Kumar, M. Marathe, S.S. Ravi, and D. Rosenkrantz. Finding critical nodes for inhibiting diffusion of complex contagions in social networks. In *ECML-PKDD*, 2010.

[67] Fabian Kuhn, Nancy A. Lynch, and Rotem Oshman. Distributed computation in dynamic networks. In *STOC*, pages 513–522, 2010.

[68] V.S. Anil Kumar, Rajmohan Rajaraman, Zhifeng Sun, and Ravi Sundaram. Existence theorems and approximation algorithms for generalized n etwork security games. In *Proceedings of IEEE ICDCS*, 2010.

[69] M. Lelarge and J. Bolot. Economic incentives to increase security in the Internet: The case for insurance. In *IEEE Infocom*, 2009.

[70] T. M. Liggett. Stochastic interacting systems: contact, voter and exclusion processes. *Springer-Verlag*, 1999.

[71] F. Liljeros, C. Edling, L. Amaral, H. Stanley, and Y. Aberg. The web of human sexual contacts. *Nature*, 2001.

[72] A. L. Lloyd and R. M. May. Epidemiology. how viruses spread among computers and people. *Science*, page 1316, 2001.

[73] C.F. Manski. Vaccination with partial knowledge of external effectiveness. *PNAS*, 2010.

[74] A.R. McLean and S.M. Blower. Imperfect vaccines and herd immunity to HIV. *Proceedings of the Royal Society of London*, 1993.

[75] J. Medlock and A.P. Galvani. Optimizing influenza vaccine distribution. *Science*, 2009.

[76] S. Meloni, A. Arenas, and Y. Moreno. Traffic-driven epidemic spreading in finite-size scale-free networks. *PNAS*, 2009.

[77] L.A. Meyers. Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. *American Mathematical Society*, 2006.

[78] L.A. Meyers, B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, and R.C. Brunham. Network theory and SARS: Predicting outbreak diversity. *Theoretical Biology*, 2005.

[79] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 1995.

[80] M. Molloy and B. Reed. The size of the largest component of a random graph on a fixed degree sequence. *Combinatorics, Probability and Computing*, 1998.

[81] A. Nahir and A. Orda. Topology design and control: A game-theoretic perspective. In *IEEE INFOCOM*, 2009.

[82] J.F. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[83] M. Newman. The structure and function of complex networks. *SIAM Review*, 45, 2003.

[84] M.E. Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 2002.

[85] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their applications. *Physical Review E*, 2001.

[86] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani. *Algorithmic Game Theory.* Cambridge University Press, 2008.

[87] J. Omic, A. Orda, and P. Mieghem. Protecting against network infections: A game theoretic perspective. In *IEEE Infocom*, 2009.

[88] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 2000.

[89] Boris Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47(1):213–223, 1987.

[90] Harald Räcke. Optimal hierarchical decompositions for congestion minimization in networks. In *STOC*, pages 255–264, 2008.

[91] T. Reluga and J. Medlock. Resistance mechanisms matter in SIR models. *Math. Biosciences*, 2007.

[92] T. Roughgarden and E. Tardos. How bad is selfish routing? *J. ACM*, page 236259, 2002.

[93] Tim Roughgarden. Stackelberg scheduling strategies. *SIAM J. Comput.*, 33(2):332–350, 2004.

[94] R. Smith and S. Blower. Could disease-modifying HIV vaccines cause population-level perversity? *Infectious Diseases*, 2004.

[95] B. Söderberg. A general formalism for inhomogeneous random graphs. *Phys. Rev. E*, 2002.

[96] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.

[97] J.X. Velasco-Hernandez, H.B. Gershengorn, and S.M. Blower. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? *Infectious Diseases*, 2002.

[98] Yang Wang, Deepayan Chakrabarti, Chenxi Wang, and Christos Faloutsos. Epidemic spreading in real networks: An eigenvalue viewpoint. In *In SRDS*, pages 25–34, 2003.

[99] G. H. Weiss and M. Dishon. On the asymptotic behavior of the stochastic and deterministic models of an epidemic. *Math. Biosci.*, 11:261–265, 1971.

[100] D.P. Wilson, P.M. Coplan, M.A. Wainberg, and S.M. Blower. The paradoxical effects of using antiretroviral-based microbicides to control HIV epidemics. *PNAS*, 2008.

[101] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karonski and H.J. Proemel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.

[102] Y. Yang, J.D. Sugimoto, M.E. Halloran, N.E. Basta, D.L. Chao, L. Matrajt, G. Potter, E. Kenah, and I.M. Longini. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science*, 2009.