ANALYSIS AND OPERATIONS IN COMPLEX NETWORKED SYSTEMS: FROM URBAN INFRASTRUCTURE TO EPIDEMICS

BY

LIQUN LU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

    Professor Yanfeng Ouyang, Chair
    Professor Ximing Cai
    Assistant Professor Hadi Meidani
    Assistant Professor Rebecca Lee Smith

Abstract

Many real-world systems run with underlying complex networked structures, and system disruptions or signals of many sorts are transmitted along these networks. This dissertation aims at investigating the dynamics and operations in different types of complex networked systems, with two main focuses: i) infrastructure cascading failures in urban interdependent infrastructure networks, and ii) infectious disease propagation dynamics and control in population contact networks. Under each topic, one or several studies are performed with different focuses to reveal insights on varies aspects of the real-world networked systems, and provide decision-making frameworks to operational problems.

First, in an urban infrastructure system, local disruptions are likely to propagate to affect a large fraction of the system due to high infrastructural interdependencies. Meanwhile, urban population's response to the disruptions may further aggravate the cascading failures when they compete with facilities for limited resources. To capture the mutual impacts between infrastructural interdependencies and population behavior, a generic game-theoretical framework is built to find the overall equilibrium among the system. Infrastructure systems are modeled as layers of interdependent networks. Infrastructural interdependencies are captured by both direct support relationships as well as indirect support via commodity flows through transportation.

On the other hand, epidemic dynamics and control in populations have raised increasing attentions using complex network models. Considering population contact networks as random networks, we investigate three topics focusing on different aspects of epidemic control: i) effects of two control approaches, quarantine and public health advisories are incorporated into an established Susceptible-Infected-Removed (SIR) dynamics model, and an optimization framework is built to find the optimal control strategy; ii) a generic system dynamics model is developed to account for vaccination during the SIR process, and a game-theoretical framework is built to investigate population voluntary vaccination behavior under vaccine-phobia during a disease outbreak; and iii) a minimum required vaccine problem to prevent an epidemic outbreak and an optimal vaccine allocation problem to mitigate local disease propagation are solved, with consideration of population heterogeneity.

Results from these studies reveal interesting insights for decision-makers in both urban infrastructure planning and public health industries. All these studies highlight the dominating importance of those infrastructures/individuals that are better connected with the others in the disruption/disease propagation processes. Moreover, self-interested gaming behavior of a population are often found to have a negative impact on system-wide utilities.

Acknowledgments

I want to express the most sincere gratitude to my advisor Dr. Yanfeng Ouyang. He saw the potential in me when I started working with him through a few small projects, and guided me over the years with great patience and inspirations. He constantly encouraged me to explore new and exciting fields on my own in the early stage of my PhD career, which eventually lead to this dissertation. His high standard set for himself in both academia and life always strongly inspire me to keep challenging and improving myself.

I also want to give profound thanks to my parents, who strive hard to provide what is beyond the best to me. They started the family at their early twenties with almost nothing in a small town in southeast China, and still managed to raise a kid well and eventually sent him to a prestigious graduate school abroad. This is actually even more amazing than it sounds considering that I as a teenager did not make it any easier for them. Yet more important is the merits such as honesty they have been trying to pass on to me over years.

I want to give special thanks to my sponsors from the Construction Engineering Research Laboratory (CERL) of the U.S. Army Corps of Engineers: Dr. George Calfas, Natalie Myers, Jeffrey Burkhalter, and Dr. Charles Ehlschlaeger. Not only my PhD career was financially sponsored by them through two consecutive projects with their firm support, I also benefited a lot learning from them, e.g., Dr. Ehlschlaeger's enthusiasm and dedication to his goal.

I greatly appreciate working with and learning from my collaborators in the Transportation and Logistic Systems research group at University of Illinois. Dr. Xin Wang, Dr. Chao Lei, Dr. Zhaodong Wang, Dr. Siyang Xie, and soon-to-be Dr. Antoine Petit, and many others are all good friends to me and brilliant academic leaders that set exceptional examples for me to follow.

Many thanks to my PhD committee, Dr. Ximing Cai, Dr. Rebecca Lee Smith, and Dr. Hadi Meidani, who provided very valuable suggestions and comments to help improve this dissertation. Many future research ideas in the last chapter of this dissertation come from their comments.

There are many others that worked together with me on different projects, helped me on my research, and supported me pursuing the PhD degree. I wish to thank all of them, if only it was possible to list all their names.

Although this dissertation was partially supported by the Engineering Research and Development Center (ERDC) - CERL, the opinions, findings, and conclusions expressed in this dissertation do not necessarily reflect the views of the ERDC-CERL.

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION

1.1 Motivation

Many real-world systems run with underlying networked structures, such as infrastructure networks and population social networks. These systems are often studied using conventional network modeling approaches, which describe the structure of the systems using *graphs* $G = \{V, E\}$, where V is a collection of *vertices* (or *nodes*) and E is the set of *edges* (or *links*). In different contexts, links can be assigned with drastically different meanings, e.g. sequential/supportive relationship between nodes in the cases of directed graphs, or connectivity/conductivity in undirected networks, etc. For instance, road networks where intersections are modeled as nodes and road segments as links, have been widely used for traffic analysis, logistics, and many other studies. These conventional network modeling, however, is not able to fully describe many sophisticated real-world systems. Two main reasons for this issue (other than that the systems themselves do not possess any underlying network structures) are that i) the system is in fact a complex system of different systems, e.g. a complex network of urban infrastructure networks, and ii) the scale of the system is too large for conventional network representation, e.g. population social networks. While infectious disease modeling and control in the human population have been intensively studied for almost a century, in the recent years people also come to realize that due to increasing infrastructural interdependencies, facility disruptions are more likely to propagate in a complex urban infrastructure system just like diseases spread in the population. To study the failure/disease propagation dynamics in these complex systems and/or to find proper operational decisions, it calls for complex network modeling tools.

1.1.1 Cascading failures in interdependent infrastructure systems

Urban infrastructure systems which the modern human society heavily rely on are usually composed of mutiple sectors of infrastructure systems, such as power, water, transportation, etc. In these systems, catastrophic cascading infrastructure failures under even minimal local disruptions have been observed in many past incidents, e.g. the 2003 Northeast blackout and the 2011 Tohoku earthquake. This is mainly because of the complex interdependencies among infrastructure

facilities (Rinaldi et al., 2001). On the other hand, urban populations, which are usually associated with great amount and high density, rely heavily on the infrastructure systems for life-essential resources. Their resource procurement behavior in the disruption events will be greatly reshaped and may further exacerbate the infrastructural disruptions when they compete with the facilities for resources. For example, over concentration of population at resource facilities increases traffic congestion, which further hinders commodity flows through transportation and thus resource replenishment of the facilities. It is important to capture the infrastructural interdependencies, population resource procurement behaviors, and the mutual impacts between these two so as to evaluate the impact of cascading failures and vulnerability of the urban systems under disruptions. Therefore, it calls for complex networks modeling to describe the multi-layer interdependent infrastructure systems (Zhang and Peeta, 2011).

1.1.2 Infectious disease modeling and control

Infectious diseases among human populations have always been among the top threats to the human beings. The difficulties in modeling and control of infectious diseases largely lie in population heterogeneity, including: i) individuals' resistance to infections and ability of recovery from diseases, ii) their interactions with the others and the environment, iii) the effectiveness and cost of disease control measures on them, and iv) their perspectives toward disease controls and responses to epidemics, etc. Many complex network modeling approaches have been adopted to address these heterogeneities in the recent years (Pastor-Satorras et al., 2015), yet there are still a lot of work that needs to be done, especially on epidemic control in large heterogeneous mixing populations.

Common disease control approaches include quarantine, health educational campaigns, vaccination, and so on. These control approaches may cause detrimental impacts to the society (e.g., disease awareness campaigns may cause panics and reduce social functionality), induce psychological sufferings to individuals (e.g., quarantine and isolation), and quite often there are very limited resources (e.g., insufficient vaccines). In order to make proper epidemic control decisions under various situations, it is imperative to use complex network modeling tools to fully understand the disease propagation process in heterogeneous mixing populations.

While enforced epidemic controls by public health agencies play an important role in mitigating disease impacts, so do autonomous responses of the population to the epidemic. It is important to study the complex behavior of individuals during an epidemic event so as to understand the development and consequence of an epidemic outbreak. For example, while vaccination has freed the human society from the fear of many infectious diseases in the modern world, opposition to vaccination has long been a non-negligible public health phenomenon resulted from people's var-

ied perceptions toward vaccination (e.g., vaccine-phobia). In addition to the risk of vaccine failures (e.g., Wu et al. (2011)), McKee and Bohannon (2016) summarized four main reasons for people of avoid vaccination: religious concerns, philosophical concerns, safety concerns due to side effects, and insufficient knowledge. As such, the idea that one may be vaccinated at some point often causes psychological sufferings. Yet the potentially disastrous outcome of infection drives them toward vaccination in the meantime. Therefore, for each individual the decisions on whether and when to take vaccine depend on the trade-off between the psychological penalty and the risk of being infected. Moreover, in a population, the infection risk also depends on the overall vaccination coverage of all individuals – if all others with whom one interacts are all vaccinated, then there is no fear about the epidemic. Therefore, the decisions of all individuals are simultaneously influenced by and influencing each other. This problem is referred to as a *vaccination game* in the literature. It is important to investigate the vaccination game among the population to understand the important factors that affect people's own decision-making processes in an epidemic event.

Furthermore, resources required to carry out epidemic control plans, such as funds, manpower, and materials are finite and often very limited. In a heterogeneous population, it is important to properly allocate the limited resources to mitigate the impacts of disease propagation utmost. Dividing the population into several subpopulations based on their social characteristics and/or spatial distribution, control strategies targeting different subpopulations may have different cost efficiencies. Given limited resources during or prior to an epidemic outbreak, it is imperative to find out the most effective resource allocation strategy.

In light of these challenges, this Ph.D. study aims at investigating the two aforementioned topics with appropriate complex network modeling approaches. For interdependent infrastructure networks, we investigate cascading failures considering the mutual impacts between infrastructural interdependencies and population gaming behavior, and evaluate the impact of system disruptions. Under the epidemic control topic, three studies are performed with different focuses: i) optimal dynamic control that considers quarantine and health educational campaigns in a multi-group population; ii) a dynamic vaccination game that investigates people's autonomous vaccination behavior under vaccine-phobia; and iii) optimal vaccine allocation in a spatially heterogeneous population.

1.2 Contributions

First, modeling the urban interdependent infrastructure system as a system of layered networks, this research proposes a mathematical framework to systematically quantify the social impacts of cascading failures in the infrastructure network system. We i) generalize various types of interdependencies among infrastructures with a layered network model, (ii) estimate entangled system failure and equilibrium community behavior as part of a game-theoretical model, and (iii) evaluate

the cascading propagation of disruptions (due to interdependencies) and the consequential societal impacts (such as demand loss and cost increase).

Moreover, we consider quarantine and health educational campaigns as epidemic control measures with detrimental influences, and solve for the optimal control to mitigate the impacts of both the disease propagation and the controls in a population of different social groups. We i) develop an epidemic dynamics model incorporating quarantine and educational campaign controls, ii) propose a latent model to evaluate the hidden dynamic system states using observations from quarantine control, and iii) establish an optimization framework to solve for the best control decisions, and adopt a heuristic solution algorithm based on value function approximation to solve for the optimal control.

Next, we investigate the voluntary vaccination behavior of a population during an epidemic outbreak, where each individual makes its own vaccination decision to minimize its expected disutility from both vaccine-phobia and the risk of infection. We i) build an epidemic dynamics model to describe the coupled disease propagation process and the vaccination dynamics, with a computation method to conveniently evaluate the final outcome of people's vaccination decisions, ii) formulate a Nash game in a large heterogeneous population into a mathematical program, to which existence and uniqueness of an equilibrium solution are investigated, and iii) propose an efficient heuristic algorithm to solve the equilibrium solution to the game.

Finally, we look for the optimal vaccine allocation in a spatially heterogeneous population, with the objective of preventing an epidemic outbreak and minimizing the impacts of local disease infections. We i) develop a method to compute the expected outbreak size given a vaccine allocation in a population with spatial heterogeneity, ii) establish an optimization framework to minimize the expected outbreak size by finding the optimal allocation, and formulate a second problem to find the minimum required vaccine resources that prevents an epidemic outbreak, and iii) propose a gradient-based heuristic solution approach to solve the optimal allocation problem, and a branch-and-bound approach to solve the minimum resource problem.

Numerical experiments are performed in each of these works on semi-realistic or hypothetical case studies to validate the modeling frameworks and solution approaches. Extensive sensitivity analysis are as well conducted to examine the important factors in these problems and to help reveal useful insights and facilitate decision-makings. The results highlight the crucial importance of nodes with higher connectivity in the networked systems in the disruption/disease propagation processes. Moreover, individuals' gaming behavior in the network are often found to be detrimental to the systems' utilities.

1.3 Outline

This dissertation is organized as follows. Chapter 2 summarizes literatures in cascading failures in interdependent infrastructure systems, and infectious disease modeling and control in population contact networks. Focusing on interdependent infrastructure systems and facility cascading failures, Chapter 3 establishes a generic game-theoretical framework to evaluate the vulnerability of urban infrastructure systems. Then with insights revealed on complex network systems, we continue to examine the key factors in epidemic modeling and control in population contact networks. In Chapter 4, we consider two commonly adopted control approaches, quarantine and educational campaigns. An optimal control framework is proposed for public health agencies to balance the trade-off between the impacts of both disease propagation and control in a population with different social groups. Another common epidemic control approach, vaccination, is studied in the following two chapters. Realizing that each individual in a population also needs to balance a trade-off between the penalties of both vaccine-phobia and infection risk, in Chapter 5 we investigate a dynamic vaccination game to understand the population's voluntary vaccination behavior during an epidemic outbreak. This chapter also examines the impacts of delayed vaccination, and the results highlight the importance of timely vaccination prior to the onset of the epidemic outbreak. This finding motivates the following topic: to find the best vaccine allocation strategy to protect the population from an upcoming epidemic outbreak given limited vaccines. Therefore, Chapter 6 studies the optimal vaccine allocation problem in a spatially heterogeneous population, and the minimum amount of vaccines required to prevent an epidemic outbreak is also found. Finally, Chapter 7 discusses potential future research topics based on the modeling framework and preliminary findings of this dissertation.

CHAPTER 2 – LITERATURE REVIEW

2.1 Infrastructure Cascading Failures under Disruptions

Cascading failures in urban infrastructure systems have raised a great amount of attentions in the past two decades, especially after the 2003 Northeast blackout incident. One noticeable cause of such fragility of the urban system under disruptions is the complex interdependencies across multiple infrastructure systems. Such complex infrastructure interdependencies have been studied by several teams of researchers. Rinaldi et al. (2001) discuss the challenge of modeling and simulating the system with a detailed analysis of four interdependency categories. Zhang and Peeta (2011) review a variety of foregoing work on addressing infrastructural interdependencies with different approaches and propose a method of using computable general equilibrium theory to generalize the interdependencies. Focusing on different types of interdependencies, Motter and Lai (2002) and Buldyrev et al. (2010) respectively perform analytical studies and prove that disruptions at a small number of components can lead to a disastrous cascading failure in the system. Moreover, efforts have been made to build a more resilient system (e.g., designing facility locations for optimal reliability) that capture the same understanding of the interdependencies. For example, Wang and Ouyang (2013) describe the demand geographical interdependencies among facilities and developed game-theoretic models to find an optimal spatial distribution of service facilities; Xie et al. (2015) propose a novel network augmentation approach that is proven to be able to decompose any infrastructure disruption interdependencies, such that the optimal design problem can be formulated and solved equivalently with only independent disruptions.

Urban infrastructure failures are likely to stimulate strong reactions from the population. A direct consequence of most system failure is inadequate access to life-supporting resources. A number of recent studies focus on the social impacts caused by peopleâĂŹs behavior under infrastructure failures and limited resources. Horner and Widener (2011) first consider the impact of transportation network failure on people accessing disaster relief goods after a hurricane, and then try to optimize the location of commodity distribution facilities. Many efforts have been made to model community behaviors after an emergency or a disaster. A game-theoretical model such as Nash competition has been used to capture peopleâĂŹs reaction to scarcity of resources. In partic-

ular, it is probably reasonable to assume each community/individual would try to acquire resources with the least time and monetary cost, taking into account peer reactions. Such a situation, often referred to as âĂĲequilibrium,âĂİ changes dramatically after infrastructure failures, and its quantification can help evaluate the social impact of the failures. Such game-theoretical models have been applied in related literature, e.g., within the scope of input-output models (Haimes and Jiang, 2001; Haimes et al., 2005). Zhang and Peeta (2011) further consider an equilibrium taking account of both spatial and economic factors between household and the infrastructure facilities that provide the resources. These works have not focused specifically on complex interdependencies of multiple infrastructure systems, and we suggest that population reaction to cascading failures be modeled into an equilibrium problem considering spatial distributions of facilities and population.

2.2 Epidemics in Population Contact Networks

2.2.1 Epidemic modeling

Infectious disease modeling and control have been intensively studied in the past century since the establishment of the fundamental Kermack-Mckendrick model (Kermack and McKendrick, 1927). It is one of the compartmental models, which categorize the population into several compartments based on their health status, e.g. susceptible (S), infectious/infected (I) and recovered/removed (R), etc, and investigate the transitions among these groups to study the epidemic dynamics. The susceptible population are the healthy ones that could be infected; once they are infected by the infectious population, they may also infect others until they eventually recover with immunity or decease due to illness. A system of deterministic nonlinear equations was proposed to capture the expected quantities of the three compartments over the course of the disease propagation, presented as follows.

$$\dot{S} = -rSI, \tag{2.1a}$$

$$\dot{I} = rSI - uI, \tag{2.1b}$$

$$\dot{R} = uI, \tag{2.1c}$$

where S, I and R denotes the numbers/fractions of the susceptible, infected, and removed individuals, respectively; r and u represents the infection and recovery rates, respectively. Most of the subsequent work on epidemiology are developed within or as extensions to this framework; see Anderson and May (1992) for a review. These models treat each individuals in the same compartment identically, and use a unified disease transmission rate for the whole population. Therefore, they are also known as *homogeneous mixing population models*. These models are particularly

suitable for disease transmitted via the living environment, such as water-borne (Codeço, 2001; Tien and Earn, 2010) and vector-borne (Bacaër and Guernaoui, 2006; Stoddard et al., 2009), or those transmitted among domestic animals (Smith and Gröhn, 2015; Smith, 2016).

However, these compartment models often neglect population heterogeneities and are not suitable to describe diseases transmitted via person-to-person contacts. For example, a person with more interactions with others not only has a higher infection risk himself/herself, but also poses a greater threat to all those around him/her. Many complex network models have been developed to address population heterogeneities such as different social/age groups and heterogeneous spatial distributions (Pastor-Satorras et al., 2015). In particular, facilitated by the development of random network analysis (Newman et al., 2001), numerous works focus on the disease propagation process in a population contact network, assuming only statistical information regarding node degrees is available, but microscopic topologies are unknown and random (Barthélemy et al., 2005; Volz and Meyers, 2007; Volz, 2008; Lindquist et al., 2011; Miller, 2011; Koch et al., 2013). These models are known as *heterogeneous mixing population models*. In particular, a significant part of this work is established primarily based on the configuration random network models by Newman et al. (2001), and Volz and Meyers (2007) and Miller (2011). We hereby introduce some fundamental notions, which will be used and referred to throughout this dissertation.

In a configuration random network, we have a fixed number of nodes, and the actual connections among these nodes are described probabilistically; i.e., each node is associated with a degree of *stubs*, and an edge is formed by randomly connecting two unconnected stubs. The degree of an arbitrary node follows a given probability distribution, and such *degree distribution* is expressed in the form of a probability generating function, as follows (Molloy and Reed, 1995):

$$g(x) = \sum_k p_k x^k, \tag{2.2}$$

where p_k is the probability that a random node has k degrees, and x is a dummy argument. As such, $g'(1)$ and $g''(1)$ are the expected numbers of direct and second-order neighbors of an arbitrary node, respectively.

Volz (2008) and Miller (2011) propose a system of low-dimensional equations to describe the SIR epidemic dynamics in a population with heterogeneous mixing. The population is modeled as a configuration model random network, where each individual in the population is represented by a node, and an (undirected) edge represents a direct contact between two individuals that may transmit the disease. This model tries to capture the probability for a randomly selected test node to be susceptible, based on whether or not any of its incident edges have transmitted disease from its (infected) neighbors. For convenience, we call such an event as an *infectious contact*. Two variables are defined to capture this quantity:

- $\theta(t)$: the probability that a random edge of the test node has not passed an infectious contact to the test node by time t, assuming the test node does not transmit disease to others;
- $\phi(t)$: the probability that a random edge of the test node is connected to an infected neighbor, but has not passed an infectious contact to the test node by time t, assuming the test node does not transmit disease to others.

An infectious contact along an edge and the recovery of an infected node are assumed to follow Poisson processes at a unit transmission/infection rate, denoted r, and a recovery rate, denoted u, respectively. Moreover, S, I, and R are used to represent the fractions of population in the three compartments, respectively. The system equations can be summarized as follows:

$$\dot{\theta} = -r\phi, \forall t \geq 0, \tag{2.3}$$

$$\dot{\phi} = \left(-r - u + r\frac{g''(\theta)}{g'(1)} \right) \phi, \forall t \geq 0, \tag{2.4}$$

$$\dot{R} = uI, \forall t \geq 0, \tag{2.5}$$

$$S(t) = g\left(\theta\right), \forall t \geq 0, \tag{2.6}$$

$$S(t) + I(t) + R(t) = 1, \forall t \geq 0. \tag{2.7}$$

In (2.4), the first two terms in the parentheses on the right hand side represent, respectively, the transmission of the disease along an edge that satisfies the definition of ϕ, and the removal of an infected neighbor linked by such an edge. The third term is the derivative of $-g'(\theta)/g'(1)$, where $g'(\theta)/g'(1)$ represents the probability that an arbitrary neighbor of the test node being susceptible. Therefore, the third term captures the increment of ϕ due to new infections at the neighbors of the test node (i.e., infected by their own neighbors other than the test node).

2.2.2 Epidemic controls

Quarantine and educational campaigns

The purpose of infectious disease modeling is not only for predictions of disease outbreaks, but more importantly, to serve as building blocks of mathematical frameworks for epidemic control decision-makings. There are a handful of different approaches in disease control, many of which eventually lead to two outcomes, reducing the disease transmission rate and/or increasing the removal rate (Nowzari et al., 2016) (except for vaccination control, which changes the population compartment structures directly). For this reason, many existing epidemic control literature focus on manipulating the disease parameters (Castilho, 2006; Hansen and Day, 2011; Khan et al., 2017; Xu et al., 2017). Quarantine control that screens the population and then isolates and/or treats the infected individuals can effectively increase the removal rate of those infected, and nu-

merous researches have been conducted in this regard (Yan and Zou, 2008; Xu et al., 2017). Pereira and Young (2015) study the impact of delayed isolation of infected individuals, and their results highlight the importance of timely isolation for mitigating disease propagation. Moreover, educational campaigns increase people's health awareness, improve their sanitary levels, and guide them to take precautions when making contacts with others. This control approach has also been considered in the existing literature, usually by multiplying the infection rate by a discount factor that captures the effectiveness of the campaign (Castilho, 2006; Mukandavire et al., 2009).

However, the control actions sometimes cause detrimental impacts to both individuals and the society. Through survey, Hawryluck et al. (2004) reveal that during the SARS epidemic, among the quarantined individuals in Toronto 28.9% and 31.2% people suffered from post-traumatic stress disorder and depression, respectively. Moreover, epidemic-awareness propaganda can stimulate fear and panics, and a very recent example is that the 2014 Ebola propaganda in the U.S. affected the social functionality by causing a hysteria among the population (Ahmed and Mendoza, 2014; Innes, 2015). These reports suggest that while the control approaches help reduce the impacts of disease propagation, they also induce non-negligible social costs in the meantime. Therefore, it is important to find the optimal control with these considerations.

Most of the existing work focus on homogeneous mixing population without considering population heterogeneities regarding either the networked structure or physiological/social differences. Nowzari et al. (2016) give a review over a variety of epidemic dynamics models and controls with network models that try to address such heterogeneities, yet it misses recent development on epidemic modeling on large scale random networks such as Barthélemy et al. (2005) and Miller et al. (2012).

Vaccination control

The effect of vaccination could be modeled in a variety of ways, mainly depending on its timing: i) vaccination prior to the epidemic outbreak, or immediate vaccination once a new population is born (e.g., Madar et al. (2004) and Bauch and Earn (2004)); and ii) vaccination during disease propagation when the vaccine takers have been exposed to potential risks of infection for some time (e.g., Lu et al. (2002) and Pang and Chen (2007)). It is worth noting that, if immunity obtained via vaccination does not wane, early vaccination is always more effective in mitigating disease propagation because of survival bias.

Bansal et al. (2006) and Mylius et al. (2008) compare two vaccination strategies under different vaccination timings, targeting subpopulations that have high infection risks (morbidity-based) and that more likely to decease once infected (mortality-based), respectively. Their results show that if vaccines are available prior to the onset of disease outbreak, morbidity-based vaccination should be adopted as it immunizes the high risk population and mitigates disease propagation most

effectively. However, delay in availability of vaccination has a tremendous impact on the epidemic result, and in that case mortality-based vaccination out-performs its morbidity-based counterpart.

Impacts of network topology and vaccination strategy on SIR-type disease propagation in large-scale random networks are studied in Ma et al. (2013). Stochastic simulations were conducted for various combinations of random networks (e.g., Erdos-Renyi model, scale-free, etc.) and vaccination strategies (e.g., prioritized, random, etc.). They discover that network topology has a significant impact on the spread of disease as well as the timing of vaccination. This finding highlights the importance of understanding disease transmission in realistic contact networks. The recent independent effort in Di Muro et al. (2018) incorporate dynamic vaccination into the models of Volz (2008) and Miller (2011). Despite some similarity in the research goals, there are key differences in our approaches. For example, their model does not account for population's heterogeneous vaccination behavior, which nonetheless is crucial in modeling the vaccination game in a heterogeneous mixing population.

Vaccine-phobia and vaccination game

The term "vaccination game" has been used to describe an imitation process where individuals learn about others' perspective on vaccine and follow probabilistically each other's actions (Bauch, 2005; Zhang, 2013; Han and Sun, 2016). In this research, we go back to its literal meaning of a classic game-theoretical framework, where each individual in a population is a decision maker that tries to maximize its own utility based on its knowledge on the system. A review on the impacts of typical vaccination behavior can be found in Funk et al. (2010). However, the sources of such behavior as a result of people's decision-making processes in an epidemic outbreak are still not fully understood.

Modeling the vaccination game is challenging, because it is imperative to account for heterogeneity among the individuals, such as different physical conditions and psychological belief systems. Moreover, the interactions between individuals themselves and those with the environment also vary drastically from person to person; e.g., depending on how frequently a person makes disease-transmissive contacts with the others. Consequently, different people often make different vaccination decisions, and their vaccination decisions also contribute differently to the epidemic outbreak. Cojocaru (2008) considers population heterogeneity where multiple groups of population each has a different perception on the infection risk. Nash equilibria is formulated into evolutionary variational inequalities, and their results highlight the importance of accounting for social heterogeneity in a vaccination game.

Another challenging issue associated with a vaccination game is the proper modeling of the epidemic evolution process. Since disease propagation and vaccination decisions are dynamic, epidemic dynamic models must be developed to evaluate the final or steady state of the system.

The SIR models usually demand tedious iterations over time steps to evaluate the outcome of a vaccination strategy; e.g., see Bauch et al. (2003) for an example. In the hope to avoid such difficulty, most of the existing literature on dynamic vaccination game either (i) consider susceptible-infected-susceptible problems (Reluga, 2009) or a dynamic population (Bauch and Earn, 2004; Reluga et al., 2006) such that a steady state equilibrium of the epidemic can be found, or (ii) approximate the results by using a perceived infection risk among the population (Cojocaru, 2008) or linearized system dynamics (Reluga and Galvani, 2011). A method to conveniently and accurately evaluate the outcome of a vaccination strategy during a fast-spreading SIR epidemic is yet to be found.

Moreover, to investigate the vaccination game, it is imperative to quantitatively measure people's disutility due to psychological suffering, based on not only the likelihood of vaccine uptake, but also how people perceive vaccination (e.g., the level of vaccine-phobia). Existing social studies try to find the correlation between people's perception on vaccines and their vaccine uptake decisions from field surveys, e.g. Smith et al. (2011). In light of this, we will next aim at drawing a bridge between these two aspects of vaccination, by establishing a model on how vaccine-phobia may affect peopleâĂŹs decision-making process and in turn the disease propagation.

CHAPTER 3 – VULNERABILITY OF INTERDEPENDENT INFRASTRUCTURE SYSTEM UNDER POPULATION GAMING BEHAVIOR

3.1 Introduction

In an urban infrastructure system, interdependencies exist among components (i) within the same infrastructure network, e.g. a water treatment plant provides resources to downstream water towers; and (ii) across different infrastructure networks, e.g. a power plant generates the electricity for water pumps, traffic lights, and fuel refineries. Such complex interdependencies usually leads to disastrous cascading failures in the system when local disruption takes place. Moreover, urban infrastructure failures are likely to stimulate strong reactions from the population. A direct consequence of most system failure is inadequate access to life-supporting resources. In reality, infrastructure failures and community reactions are mutually dependent, which further aggravates the infrastructure system disruptions. For example, people may have to travel through the transportation network to deliver or gain access to resources, while some infrastructural interdependencies are realized by delivering a commodity from one facility to another via transportation. When congestion increases due to peopleâĂŹs response to system failures, it grows harder for resource supply to be delivered to facilities through transportation network. Therefore, the fluidity of commodity flow may be compromised and the cascading effect could be further exacerbated.

In this chapter, we propose a mathematical framework to systematically quantify the social impacts of cascading failures in a system of infrastructure network systems. We construct the system by building different sectors of infrastructure facilities (e.g. electricity, water, transportation) into layered networks. Infrastructures, the basic system components, are dependent and jointly provide service to communities. We capture two types infrastructural interdependencies: one based on direct physical reliance, while the other via indirect resource supply through transportation. With the multi-sector network system and the interdependency structure, two failure propagation patterns are defined in correspondence to the two interdependency types. As disruption takes place in the system, the cascading failure begins to develop following the failure propagation patterns, and

people start to change the way they access resources. Eventually the disrupted system arrives at a state where the remaining infrastructures and the impacted population behaviors fall into a new equilibrium. The impacts of cascading failures are evaluated by assessing the population resource security and infrastructure functionalities under the new equilibrium. Furthermore, building upon this methodology, an extended research establishes an optimization framework for decision makers in the military to evaluate the effectiveness and efficiency of offense and control operation actions.

This rest of chapter is organized as follows. Section 3.2 establishes the model framework, including the infrastructure network, community behavior, and two types of infrastructure failure mechanisms. Then, Section 3.3 presents the solution approach. Section 3.4 applies the proposed model framework to a case study of Maiduguri, Nigeria, and presents a series of sensitivity analyses to examine the impacts of several factors.

3.2 Model

We focus on a geographic region Ω, such as a small country or a city. The infrastructure system is characterized by a grand network that contains seven layers. Each of these layers represents a basic type of infrastructure; i.e., water, food, electricity, fuel, transportation, healthcare, and education. We assume each layer consists of a finite number of infrastructure *nodes* as the functional unit, with specific locations (e.g., latitude and longitude) and attributes. For example, a water node can be a water tank, a fuel node can be a gas station, and a transportation node can be a road segment. In addition, we create a community layer where sub-communities (e.g. neighborhoods or housing complexes reflecting the population distribution) are served by components from the infrastructure layers. Particularly, food, water, fuel, healthcare, and education layers directly provide necessary resources to the community layer, hence are referred to as resource layers. Fuel, electricity and transportation layers serve as the backbone that supports not only the communities but also the operation of other resource layers. Hence, there are a total of seven layers in the system. We denote a generic node in the entire network by $i \in I$. We further let $r \in R$ be a resource type (i.e., water, electricity, food, healthcare) and $I_r \subseteq I$ be the corresponding set of type-r resource nodes; similarly, let $I_t \subseteq I, I_c \subseteq I$ be the respective sets of transportation and community nodes.

3.2.1 Community behavior

In communities, especially after disasters, a population may need to travel to obtain life-supporting resources via the transportation layer. (The trivial case that people get resources within the community, e.g., tap water at each household, can be incorporated as traveling a distance of zero.) Under infrastructure disruption, some resource nodes can no longer provide service, which

14

may lead to overwhelming demand concentration at the remaining nodes (and likely long queues) and longer travel detours (and likely more severe traffic congestion). Such delay and detours further increase the difficulties for people to access resources and may force some of the people to give up service. We implement a variant of traffic equilibrium model to capture this issue.

Suppose the demand of resource for each community $i \in I_c$ is independent of each other, denoted by d_i^r, and the corresponding resource access incurs a time and monetary cost c_i^r. Under infrastructure disruptions, the procurement of resources normally becomes more difficult (e.g., traveling to a well rather than using tap water from a local pipeline due to a water network disruption). This will lead to an increase of c_i^r, and hence a decrease of d_i^r, that is,

$$d_i^r = D_i^r \left(c_i^r; \hat{d}_i^r \right), \tag{3.1}$$

where $D_i^r (\cdot)$ is a monotonic decreasing demand function capturing the elasticity of demand with respect to cost c_i^r, and parameter $\hat{d}_i^r$ is the maximum demand when cost c_i^r is zero. In reality, the calculation of c_i^r is complicated. To this end, we consider the following assumptions.

First, each community i can access resource r from a collection of resource nodes $j \in I_i^r \subseteq I_r$, with a resource access cost

$$c_{ij} := c_{ij,t} + c_{ij,s}, \tag{3.2}$$

where $c_{ij,t}$ is the travel time along the shortest path from i to j via the transportation layer and $c_{ij,s}$ captures service (queueing) time at j, and I_i^r may not be the same for all i due to certain restrictions.

In particular, let $p \in P_{ij}$ be a path (i.e., $p \subseteq I_t$) connecting i and j through the transportation layer and g_k, $k \in I_t$ be the demand flow on transportation node k. Then, we have

$$c_{ij,t} = \min_{p \in P_{ij}} \sum_{k \in p} T_k (g_k), \tag{3.3}$$

where

$$T_k(g_k) = \delta_k \left(1 + \alpha_k (g_k/\gamma_k)^{\beta_k} \right) \tag{3.4}$$

is the conventional BPR function[a] (BPR, 1964) with node specific parameters $\alpha_k, \beta_k, \gamma_k, \delta_k$. Road segments (i.e., a type of infrastructure) are also represented by nodes in this paper, just like other infrastructure components. The âĂIJlink" in this model is not physically existent in the real world, but rather, it indicates the adjacency of real infrastructure components, e.g., a community node and its entry node to the transportation network are linked. This is equivalent to consider road segments as links because we compute the cost as that of traversing a node.

---

[a]More suitable transportation cost functions can also be used. To illustrate the model, we use BPR function as it is still a widely used tool for planning purposes (e.g. Bayram et al. (2015)).

Furthermore, we consider the resource procurement at node j lines up as an M/M/1 queue[b]. Hence, we have the average waiting cost

$$c_{ij,s} = 1/(\lambda_j - h_j), \tag{3.5}$$

where λ_j is the service capacity and $h_j < \lambda_j$ is the total acceptable demand flow to j. Note that this cost is not necessarily the waiting cost, but essentially a reflection on the resource scarcity. The service capacity of a facility is associated with its functionality, which will be discussed in the next section.

The second assumption presumes that, community i will choose the resource node(s) with the minimum access cost, i.e.,

$$c_i^r = \min_{j \in I_i^r} c_{ij}. \tag{3.6}$$

Then, given $f_{ij}^{p\,*}$ as the optimal demand allocation along path p, we further have

$$\sum_{j \in I_i^r, p \in P_{ij}} f_{ij}^{p\,*} = d_i^r, \; \forall i \in I, \tag{3.7}$$

$$\sum_{i \in I, p \in P_{ij}} f_{ij}^{p\,*} = h_j, \; \forall j \in I_r, \; \forall r \in R, \tag{3.8}$$

and

$$\sum_{i \in I, j \in I_i^r, p \in \{P_{ij}, k \in p\}} f_{ij}^{p\,*} = g_k, \; \forall k \in I_t. \tag{3.9}$$

(3.7)-(3.9) are flow conservation constraints at the communities, the facilities, and the transportation network nodes, respectively. Given the available resource nodes, the transportation layer reaches the above equilibrium, i.e. (3.1)-(3.9), when no community can reduce its resource access cost by unilaterally changing its resource destination or the travel path choices.

3.2.2 Disruption propagation mechanism

Now we consider the scenario where there are disrupted infrastructures. First, similar to a community node, to maintain working status, some infrastructure nodes need resource supply, which are also fulfilled via the transportation layer (e.g., a diesel generator accesses fuel from diesel tanks). Therefore, as these facilities travel through transportation layer for resource supply, the traffic congestion caused by the community behavior may significantly delay or block resource procurement as well, or even cause failures to the infrastructure. We refer to this type of failure as

---

[b]Other suitable cost functions can also be used. We adopt M/M/1 model as it properly reflects the difficulty of resource accessing due to resource scarcity (e.g. Bhaskar and Lallement (2010)).

a *resource failure*.

Second, infrastructure disruptions could also be due to direct and obvious dependencies based on physical connections. For example, a water tank with an electrical pump must be supported by a nearby power source (e.g., a power grid). Such a dependency is normally established for the long run, at a large set-up cost, and hence difficult to modify. We call this type of failure, caused by supporting infrastructure disruption, as *support failure*.

Comparing the two types of failure, we can find support failure is one major cause of cascading disruptions. In particular, a network with a tree structure can be very vulnerable âĂŞ the failure of the root node will propagate and disrupt the entire system. In contrast, resource failures may also pose a potential risk, but is more difficult to foresee in advance. However, resource failures can also bring devastating damage to the entire system, especially when it happens at some critical node that can cause consequential cascading support failures.

To model infrastructure cascading disruption, we first let $x_i \in [0, 1]$ be the working status of each node i, where the node is fully functioning if $x_i = 1$, completely failed if $x_i = 0$, and partially working otherwise. We assume the working status of a resource node linearly reflects its service capacity:

$$\lambda_j = x_j \bar{\lambda}_j + (1 - x_j)\underline{\lambda}_j, \forall j \in I_r, \forall r \in R, \tag{3.10}$$

where $\bar{\lambda}_j > 0$ and $\underline{\lambda}_j \geq 0$ are respectively the maximum and minimum service capacity of facility j.

As previously discussed, the working status of a node is a joint impact of both resource failures and support failures. For an infrastructure node $i \in I \backslash (I_t \cup I_c)$, suppose (3.1)-(3.9) also hold, and let $S_i \subseteq I \backslash (I_t \cup I_c)$ be the set of nodes that provide functional support to it. Then, we model the working status for node i due to interdependencies as the following,

$$x_i = F_i (\mathbf{y}_i, \mathbf{z}_i), \forall i \in I \backslash (I_t \cup I_c), \tag{3.11}$$

where $\mathbf{y}_i = \{x_j\}_{\forall j \in S_i}$ is a vector of states of all facilities that influence the state of facility i (i.e., via functional support), and $\mathbf{z}_i = \left(d_i^r / \hat{d}_i^r \right)_{\forall r \in R}$ captures the fulfillment of resource needs at facility i (i.e., related to resource support). We call F_i an interdependency function, and in most realistic cases it is non-decreasing with respect to its arguments, taking value from 0 to 1. We can use various forms for F_i to capture different interdependencies. An example is the Leontief form as shown in (3.12), such that the facility status is either determined by its worst supporting facility or the least satisfied resource demand:

$$x_i = \min \left(\min_{j \in S_i} x_j, \min_{r \in R} d_i^r / \hat{d}_i^r \right), \forall i \in I \backslash (I_t \cup I_c). \tag{3.12}$$

The first term on the RHS is the worst status of facility iâĂŹs supporting facilities, and the second term is the fulfilled fraction of least fulfilled resource demand; i.e., the status of i is determined by either its worst support facility or its least fulfilled resource demand.

In addition to interdependencies of infrastructure facilities within the system, we also take account of external factors. For example, a cascading failure usually start from a subset of facilities due to outside natural hazards or human actions. The status of these initially disrupted facilities is not determined by the interdependency function, but rather remain disrupted during the cascading process. We define $I_{\text{dis}} \subseteq I \backslash (I_t \cup I_c)$ to be the set of these initially disrupted facilities, and they satisfy:

$$x_i = 0, \forall i \in I_{\text{dis}}. \tag{3.13}$$

3.2.3   System equilibrium

One might have realized that the resource-seeking behavior of communities and facilities can somehow be related to the well-known Wardrop principles for network equilibrium, where all decision makers in a network choose their least cost routes to reach destinations. In this section we will first propose an equilibrium problem in an augmented network representation, and then discuss the existence, uniqueness, and solution method.

Augmented network representation

To account for elastic demand and queueing, we construct an equivalent augmented network representation with added âĂIJdummy" nodes and âĂIJvirtual" links to capture demand loss. Such treatment allows the original problem to be solved equivalently as if the demand is fixed (Sheffi, 1985). Several types of nodes, as well as corresponding virtual links, are added into the augmented network, including: sink nodes that represent demand destinations, queueing nodes at resource locations, and virtual âĂIJdummy" nodes that capture demand loss. Figure 3.1 illustrates a simple example of the new network representation. In the original network, there are trip demands from origins (i.e. resource seekers) $i_1, i_2 \in I \backslash I_t$ to resource locations $j_1, j_2 \in I_r$ as $\hat{d}_{i_1}^r, \hat{d}_{i_2}^r > 0$, and the flow travels through the transportation network. The trip demand $d_{i_1}^r, d_{i_2}^r$ varies as a function of cost based on (3.1), and due to exogenous access restrictions, i_1 may visit both j_1 and j_2 while i_2 can only visit j_1.

We will next show that by adding virtual nodes and links the augmented network can equivalently represent queueing cost, resource access restriction, and demand loss. First, for every node $i \in I \backslash I_t$ and resource type $r \in R$, we add a corresponding virtual node s_i^r to represent a virtual âĂIJsink", and set the demand from i to s_i^r be the constant $\hat{d}_i^r$. In Figure 3.1, the demands from i_1 and i_2 to j_1 and j_2 are directed into the sinks $s_{i_1}^r$ and $s_{i_2}^r$ respectively with fixed values $\hat{d}_{i_1}^r$ and
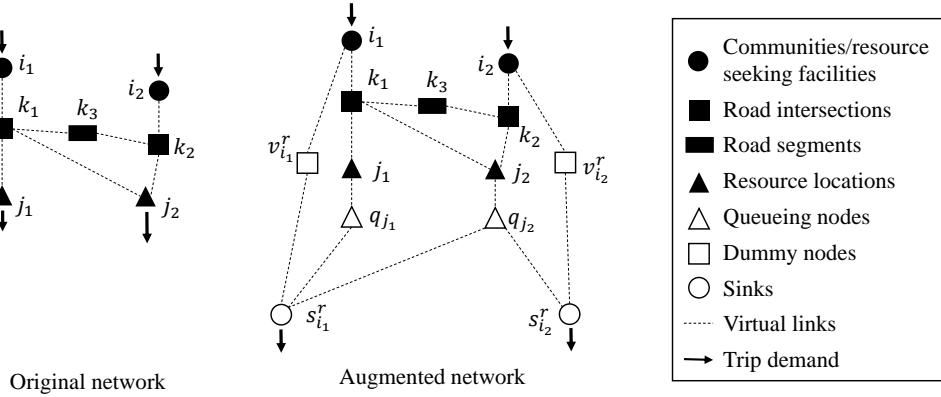
Figure 3.1: Illustration for the augmented network representation.

$\hat{d}_{i_2}^r$. Then, to allow queueing at the resource node and resource access restrictions, a virtual queueing node and two types of virtual links will be added. We add a node q_j for each resource node $j \in I_r, \forall r \in R$ to represent the queue, and add link (j, q_j); further if $j \in I_i^r$ for some $i \in I \backslash I_t$ then add link (q_j, s_i^r). This assures each node i will and can only access resource r through $j \in I_i^r$. In Figure 3.1 the queueing nodes q_{j_1} and q_{j_2} are linked to $s_{i_1}^r$, meaning i_1 can visit both j_1 and j_2 for resource r (i.e. $j_1, j_2 \in I_{i_1}^r$), while q_{j_1} is not connected to $s_{i_2}^r$, meaning i_2 is not allowed to visit j_1 (i.e. $j_2 \in I_{i_2}^r, j_1 \notin I_{i_2}^r$). The cost of travelling on q_j is

$$T_{q_j}(\lambda_j, g_{q_j}) = 1/(\lambda_j - g_{q_j}). \tag{3.14}$$

This guarantees all the demand that come to resource node j contributes to the queue and will pay the queueing cost. Hence, g_{q_j} takes exactly the value of h_j as in (3.5) and the queueing cost is converted to an equivalent transportation link cost.

Next, since the resource seekers may give up service, we add a virtual node (with proper links) as an alternative route to capture the demand loss. For all $i \in I \backslash I_t$ and $r \in R$, we add a virtual node v_i^r and two corresponding links (i, v_i^r) and (v_i^r, s_i^r). In Figure 3.1, $v_{i_1}^r$ and $v_{i_2}^r$ are added to connect i_1 to $s_{i_1}^r$ and i_2 to $s_{i_2}^r$ respectively. It allows the demand to flow from the origin to the sink via the dummy node v_i^r instead of traversing across the transportation/queuing layer, with a cost of

$$T_{v_i^r}(g_{v_i^r}) = \tilde{D}_i^r \left(\hat{d}_i^r - g_{v_i^r} \right), \tag{3.15}$$

where $\tilde{D}_i^r$ is the inverse function of D_i^r as in (3.1), and $g_{v_i^r}$ essentially represents the lost demand from node $i \in I \backslash I_t$ for resource $r \in R$. Particularly if there is no demand loss (which means $g_{v_i^r} = 0$), $T_{v_i^r}(0) = \hat{c}_i^r$ is the actual cost of traveling to s_i^r via transportation and queuing, when there is no disruption in the infrastructure system. When disruption happens, some flow will go

through v_i^r to reach s_i^r if the resource access cost becomes greater than $\hat{c}_i^r$; meanwhile, $T_{v_i^r}(g_{v_i^r})$ also increases with $g_{v_i^r}$. The value of $\hat{c}_i^r$ is computed as the cost for the flow from node i to go to sink s_i^r when $\lambda_j = \bar{\lambda}_j, \forall j \in I_r, \forall r \in R$ and $T_{v_i^r} \to \infty$ so the cost of using virtual node to reach sink (i.e., giving up service) is always greater than that of actual traveling, and demand loss will not occur. Furthermore, to differentiate the priorities of procurement of different resources for the same user, we let the cost function $T_{v_i^r}$ be more sensitive to the lost demand for resource r that with higher importance.

For notational convenience, we define A as the set of all nodes in the augmented network, such that

$$A := I \bigcup \{s_i^r, v_i^r\}_{\forall i \in I \setminus I_t, r \in R} \cup \{q_j\}_{\forall j \in I_r, r \in R}.$$

In so doing, all resource demands at i are constant regardless of the resource procurement cost or working status of infrastructures (overflows or lost demand are "captured" via virtual links), and the demand destinations are fixed regardless of the choice of resource locations. Moreover, we can write (3.11) as a function of demand lost instead of demand satisfaction to exploit the convenience of such a network representation:

$$x_i = F_i(\mathbf{y}_i, \mathbf{z}_i), \forall i \in I \setminus (I_t \cup I_c), \tag{3.16}$$

where $\mathbf{y}_i = \{x_j\}_{\forall j \in S_i}$, $\mathbf{z}_i = \left\{ \left(\hat{d}_i^r - g_{v_i^r}\right)/\hat{d}_i^r \right\}_{\forall r \in R}$.

Equilibrium definition

With the augmented network representation, we are now ready to analyze system equilibrium. System equilibrium is defined as a state where the interdependency constraints (3.10) and (3.16) are satisfied, and all the communities and facilities find their best resource-seeking strategy in the network. That means that (i) given the facilitiesâĂŹ status and capacity, the network flow reaches an equilibrium where none of the resource consumers can unilaterally reduce its access cost, and (ii) based on such a flow pattern, the facility status is consistent with the resource support and functional support it receives. To show a more rigorous definition, we first propose the following convex optimization problem that captures the resource-seeking strategy of consumers, which is in

the form of a user equilibrium:

$$\min_{f,g} \sum_{k \in I_t} \int_0^{g_k} T_k(x)dx + \sum_{\substack{j \in I_r \\ r \in R}} \int_0^{g_{q_j}} T_{q_j}(x)dx + \sum_{\substack{i \in I \setminus I_t \\ r \in R}} \int_0^{g_{v_i^r}} T_{v_i^r}(x)dx \tag{3.17}$$

$$\text{s.t.} \sum_{p \in P_{is_i^r}} f_{is_i^r}^p = \hat{d}_i^r, \; \forall i \in I, \forall r \in R, \tag{3.18}$$

$$\sum_{\substack{p \in P_{is_i^r}:q_j \in p \\ i \in \{i \in I \setminus I_t : j \in I_i^r\}}} f_{is_i^r}^p = g_{q_j}, \; \forall j \in I_r, \; \forall r \in R, \tag{3.19}$$

$$f_{is_i^r}^{p:v_i^r \in p} = g_{v_i^r}, \; \forall i \in I \setminus I_t, \forall r \in R, \tag{3.20}$$

$$\sum_{\substack{p \in P_{is_i^r}:k_t \in p \\ i \in I \setminus I_t, r \in R}} f_{is_i^r}^p = g_{k_t}, \forall k_t \in I_t, \tag{3.21}$$

$$f_{is_i^r}^p \geq 0, \forall p \in P_{is_i^r}, \forall i \in I, \forall r \in R, \tag{3.22}$$

$$g_{q_j} + \varepsilon_j \leq \lambda_j, \forall j \in I_r, \forall r \in R. \tag{3.23}$$

Here, $T_k(g_k)$ is defined by (3.4), $T_{v_i^r}(g_{v_i^r})$ by (3.15), and conditional on λ_j, $T_{q_j}(g_{q_j})$ only depends on g_{q_j}:

$$T_{q_j}(g_{q_j}) = 1/(\lambda_j - g_{q_j}). \tag{3.24}$$

The objective function (3.17) is standard for equilibrium problems, where we consider three components for transportation nodes, queueing nodes, and lost-demand nodes respectively; Constraints (3.18)-(3.21) are for flow conservation, and (3.22) for flows non-negativity. Moreover, Constraints (3.23) enforce that the arrival rates at the queues cannot be greater than the service capacities, and ε_j is a sufficiently small positive number, in light of (3.24), to obtain a closed feasibility region; in addition, as the queueing function (3.24) is not well-defined when $x_i = 0$ and $\underline{\lambda}_j = 0$, we let $\underline{\lambda}_j' = \begin{cases} \underline{\lambda}_j, \text{if } \underline{\lambda}_j \geq \varepsilon_j \\ \varepsilon_j, \text{if } \underline{\lambda}_j < \varepsilon_j \end{cases}$. be the truncated minimum service capacity, i.e., now $\lambda_j = x_j \bar{\lambda}_j + (1 - x_j)\underline{\lambda}_j'$. The following two lemmas are proposed to show that such numerical treatment does not affect the equilibrium solution.

Lemma 1. *For sufficiently small ε_j, the truncated minimum service capacity $\underline{\lambda}_j'$ only brings a negligible error to the optimal solution to (3.17)-(3.24).*

Proof. see Appendix. □

Lemma 2. *For sufficiently small ε_j, the inclusion of ε_j in (3.23) does not affect the optimal solution to (3.17)-(3.24).*

Proof. see Appendix. □

We are now ready to define the system equilibrium:

Definition 1. *Given the collection of all system parameters*

$$\theta := \left\{ \alpha_k, \beta_k, \gamma_k, \delta_k, \bar{\lambda}_j, \lambda_{-j}, \hat{d}_i^r \right\}_{k \in I_t, i \in I, j \in I_r, r \in R},$$

the collection of system variables

$$\Gamma(\theta) := \left\{ \begin{array}{l} x_i, \lambda_j, d_i^r, c_i^r, f_{is_i^r}^p, g_a : \\ s.t. Constaints\ (3.10)\ and\ (3.16), and\ \left(f_{is_i^r}^p, g_a \right)\ solves\ (3.17)\text{-}(3.24) \end{array} \right\}$$

form a collection of equilibrium states. With boundary conditions (3.13), the state

$$\sigma := \{ \sigma' \in \Gamma(\theta) : x_i = 0, \forall i \in I_{dis} \}$$

describes equilibrium of both the infrastructure system and the community.

Particularly, we let $\sigma^{\varnothing} := \{ \sigma' \in \Gamma(\theta) : x_i = 1,\ \forall i \in I \}$ be the system equilibrium in the normal scenario without disruptions; i.e., $I_{\text{dis}} = \varnothing$.

Social impact evaluation

In this subsection, we quantitatively evaluate the social impact of system failure. In general, the social impacts can be multi-faceted. First, from an economic perspective, the resource procurement cost significantly increases; meanwhile, the disrupted infrastructure may need significant investment to enable recovery. Second, from a humanitarian perspective, the reduction of community demand may reflect the extent of population suffering under the disruption(s). To this end, we consider the following three measures as proxies of social impacts: (i) total demand loss, (ii) average procurement cost increase, and (iii) total infrastructure failures, for each type of resource.

Given the initial disrupted nodes I_{dis}, the system equilibrium moves from the normal scenario $\sigma^{\varnothing}$ to σ. Hence, for each resource $r \in R$, we have the total demand loss

$$\Delta_d^r = \sum_{i \in I_c} \left(\hat{d}_i^r - d_i^r |_{\sigma} \right), \tag{3.25}$$

the average procurement cost increase

$$\overline{\Delta}_c^r = \sum_{i \in I_c} \left(c_i^r |_{\sigma} - \hat{c}_i^r \right) / |I_c|, \tag{3.26}$$

22

and the total infrastructure failures

$$\Delta_x^r = \sum_{i \in I_r} (x_i|_{\sigma^\varnothing} - x_i|_\sigma). \tag{3.27}$$

3.3 Equilibrium Analysis and Solution Approach

With the augmented network representation, we have converted the equilibrium problem into a modified traffic assignment problem with link interactions. In particular, the parameters of each link performance function depend on not only the demand flow on its own link, but also those on other links. For example, the queueing cost at a resource location is affected by the facility status, which further depends on its resource support, i.e. the amount of resource it can obtain. The following proposition provides the conditions for the existence of such an equilibrium.

Proposition 1. *(sufficient condition) There exists an equilibrium that satisfies the boundary condition* (3.13)*, if the interdependency function* $F_i, \forall i \in I \backslash (I_t \cup I_c)$*is continuous, concave, and non-decreasing w.r.t.* $\mathbf{y}_i$*, and* $\mathbf{z}_i$*.*

Proof. see Appendix. □

It is intuitive to have a strictly increasing demand-loss penalty and non-decreasing interdependency function. The concavity of the interdependency function means facility status is less sensitive to disturbance in the system when it still has relatively sufficient functional and resource support. In other words, the facilities are more robust to disruption when the other facilities are relatively well-functioning, and the functionality reduces faster when the other facilities are more disrupted.

Although we have shown the existence of a system equilibrium, it is still challenging to find. The aforementioned traffic equilibrium problems with asymmetric link interactions usually cannot be formulated into typical convex optimization problems. In the remainder of this section, we propose an iterative algorithm to compute the equilibrium. Noticing that when we fix the facility service capacity in (3.24), the problem can be reduced to a standard traffic assignment problem. We hereby consider a diagonalization method (Sheffi, 1985) to solve it. In particular, we first ignore link interactions and solve a degenerated traffic equilibrium sub-problem, then update the link cost functions based on the degenerated traffic equilibrium, and then repeat such a process until the equilibrium converges. To improve converging speed, we take affine combinations of both facility status and flow across consecutive iterations.

Diagonalization Algorithm

We use Frank-Wolfe algorithm based diagonalization method to obtain σ. We iteratively update facilitiesâĂŹ status and resource capacity given the flow using the interdependency functions, and

compute new flow assignment based on the updated link cost functions with Frank-Wolfe algorithm. The iterations are terminated when maximum iteration number is reached, or the updated flow shows little difference from the previous step. Note that the queueing nodes have strict capacity constraints; in the Frank-Wolfe algorithm we use a linear projection for the cost function when the flow approaches and exceeds the capacity, so the constraints are removed by Lagrangian relaxation.

We now discuss the solution uniqueness and algorithm convergence. First, for each facility j, we define its âĂIJresource-dependent upstream support" V_j as a collection of facilities which has positive resource demand and provide functional support directly or indirectly through other facilities. In particular,

$$
V_j = \left\{
\begin{array}{l}
i: \\
\exists r \in R, \text{s.t. } \hat{d}_i^r > 0, \text{ and} \\
\exists i_1, i_2, ..., i_K \in I, \text{s.t. } i \in S_{i_1}, i_K \in S_j, \text{and } i_k \in S_{i_{k+1}}, \forall 1 \le k < K
\end{array}
\right\},
$$

Then, we have the following proposition.

Proposition 2. *The diagonalization method gives the unique equilibrium point with guaranteed global convergence if either one of the following two conditions is satisfied:*

1. *The facility status is not sensitive to resource failure such that $\left| \partial F_j\left(\mathbf{y}_j, \mathbf{z}_j\right)/\partial g_{v_j^{\tilde{r}}} \right|$ is sufficiently close to 0 $\forall j \in I_r, \forall r, \tilde{r} \in R$;*

2. *The resource demand is inelastic enough such that $\left| dT_{v_i^r}\left(g_{v_i^r}\right)/dg_{v_i^r} \right|$ is sufficiently large, $\forall i \in I \backslash I_t, r \in R$. Particularly, when the interdependency functions take the form of (3.12), the requirements are:*

$$
T_{v_i^r}' \ge \left(\overline{\lambda}_j/\varepsilon_j \hat{d}_i^r \right)^2, \forall i \in V_j, r \in R, \forall j \in I_{\widehat{r}}, \widehat{r} \in R.
$$

Proof. see Appendix. □

The conditions in Proposition 2 are sufficient conditions for solution uniqueness and algorithm convergence. While these conditions are seemingly restrictive for application, there are many real-world scenarios where they are satisfied. For example, some facilities may be regulated to serve the civilians even when it is difficult for them to access resource supply. Furthermore, in some test numerical experiments, the algorithm converges even when the conditions are not satisfied, indicating a weaker condition for algorithm convergence exists. Therefore, the conditions stated in this proposition, while not exhausting all the possible scenarios for solution uniqueness and algorithm convergence, provide valuable insights on the problem.

24

3.4 Numerical Experiments

In this section, we implement our model to analyze the social impact of infrastructure failures based on a combination of real-world and inferred data from the city of Maiduguri, Nigeria. Maiduguri is the capital city of Borno State ($11°51'$N, $13°05'$E), with a total estimated population of 1.2 million. Concurrently with urban growth, the local government has been facing additional severe challenges. First, natural hazards such as drought and floods cause significant adverse effects (Odihi, 1996). Next, both active military events and terrorist attacks threaten peopleâĂŹs daily life and the security of urban infrastructure. In addition, large numbers of internally displaced persons have fled into Maiduguri after terrorist attacks, exhausting the resources in the city and resulting further pressure on the system. From this angle, our model aims to (i) better quantify, understand and evaluate these pressing social concerns, and (ii) provide policy insights into how to best protect critical infrastructure, reserve emergency supplies, and avoid humanitarian disasters.
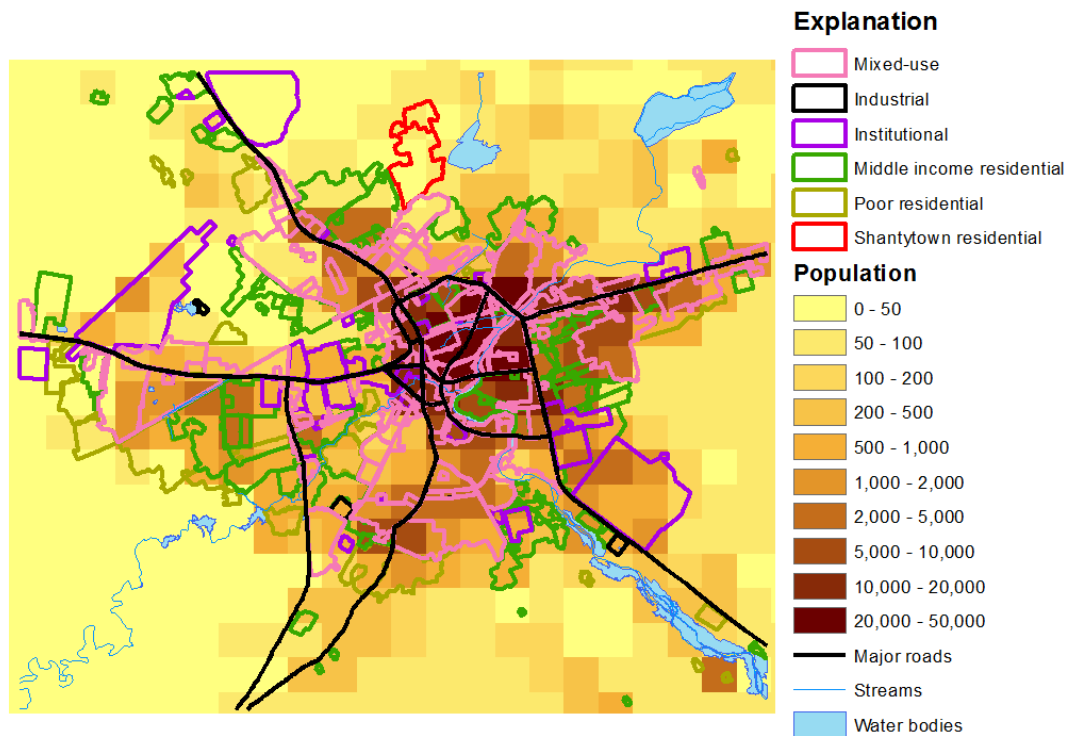

Figure 3.2: Community and population distribution.

3.4.1 Illustrative case study

Communities within the urban area of Maiduguri, Nigeria are categorized into four groups based on income and land use characteristics: mixed-use commercial/residential, middle-income residential, poor residential, and shantytown (AGC, 2015), and their geographic distributions are shown in Figure 3.2. The population as represented by the LandScan2010 dataset (ORNL, 2010) is clustered near the city center (mostly in the mixed-use areas) and gradually diminishes toward the periphery while clustering along the major roads into the city. For simplicity in the numerical experiment, we assume the population within each category of community is identical and estimated as 7,500, 750, 250, and 500 for mixed-use, middle-income, poor-income, and shantytown, respectively. We further estimate the resource demand for community $i \in I_c$ by

$$
\hat{d}_i^r = \left\{
\begin{array}{ll}
d_i^P, & \text{if } r \in \{\text{water, food}\} \\
0.4 d_i^P, & \text{if } r \in \{\text{education, healthcare}\} \\
0, & \text{otherwise}
\end{array}
\right. ,
$$

where d_i^P is the population of community i. The factor 0.4 reflects the resource demand for education and healthcare is relatively less critical than water and food. The interdependency functions take the form of (12), such that the status of each facility is either determined by the critical support failure or the most devastating resource failure. In addition, we consider an inverse demand function of resource r for community i as

$$
D_i^r \left(c_i^r; \hat{d}_i^r \right) = \left\{
\begin{array}{ll}
\hat{d}_i^r, & \text{if } c_i^r < \hat{c}_i^r \\
\hat{d}_i^r - (c_i^r - \hat{c}_i^r)/m_i^r, & \text{otherwise}
\end{array}
\right. , \tag{3.28}
$$

where $\hat{c}_i^r$ is the resource procurement cost under the normal scenario and m_i^r is a constant that satisfies $m_i^r \geq \left(\overline{\lambda}_j / \varepsilon_j \hat{d}_i^r \right)^2, \forall i \in V_j, r \in R, \forall j \in I_{\hat{r}}, \hat{r} \in R$, such that the second condition in Proposition 2 is met and convergence is guaranteed.

Representative infrastructure locations are presented in Figure 3.3, where each point indicates an infrastructure node, and the type/layer is differentiated by the marker shape. The structure of the seven modeled infrastructure layers are summarized here as follows: (i) the electricity layer supports the water layer, and they jointly support education and healthcare layers; (ii) the transportation layer neither is supported nor supports other layers, but purely provides mobility to the system; (iii) the food layer only provides food to the community, where the downstream food market node seeks supply from the upstream food processing facilities; and (iv) the fuel layer only provides fuel to the infrastructure nodes such as those in the electricity layer. The resource demand of these infrastructure nodes normally do not generate huge traffic, hence we simply assign $\hat{d}_i^r = 1$ for each infrastructure node seeking resources (e.g., fuel and food); the inverse demand function

of an infrastructure i is the same as (3.28). We should also note that some infrastructure nodes within a resource layer, such as water treatment plants, the central fuel depot, and food processing facilities, might not directly provide resources but support other downstream resource nodes. The maximum resource capacity $\bar{\lambda}_j$ is computed by assigning all the resource seekers to their nearest resource locations and computing the demand at j, and $\bar{\lambda}_j$ is set to be 1.1 times the demand, so that there are sufficient resources in normal scenario; furthermore, if the computed demand at a resource facility is lower than a certain threshold $\hat{\lambda}_j$, the threshold value is assigned to $\bar{\lambda}_j$ for realism. In other words, let the $I_{\text{near } j} \subseteq I\backslash I_t$ be the set of nodes that need resource r and their closest resource location is $j \in I_r$, then $\bar{\lambda}_j = \begin{cases} \kappa_{cap} \sum\limits_{i \in I_{\text{near } j}} \hat{d}_i^r, & \text{if } \sum\limits_{i \in I_{\text{near } j}} \hat{d}_i^r \geq \hat{\lambda}_j \\ \kappa_{cap}\hat{\lambda}_j, & \text{o.w.} \end{cases}$ where $\kappa_{cap} = 1.1$.
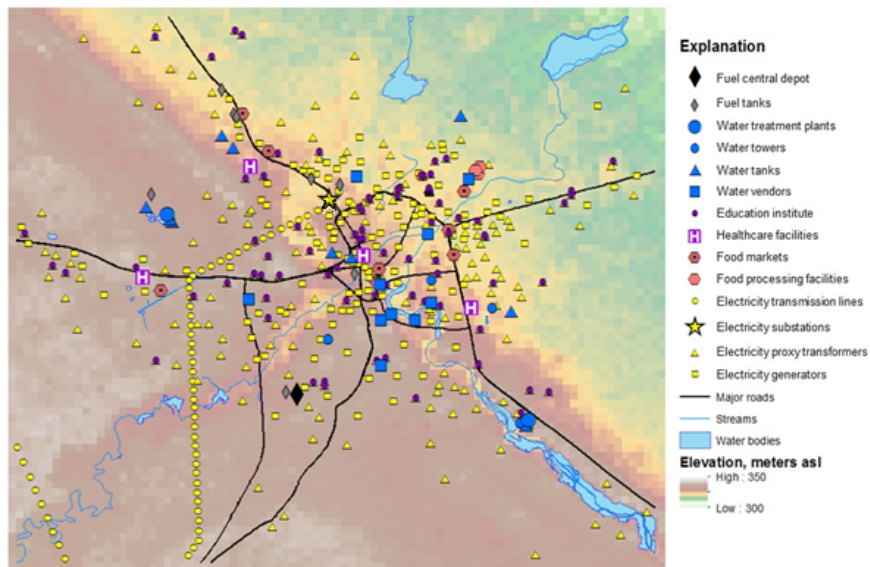

Figure 3.3: Infrastructure distribution.

The rest of parameters are set as follows: $\alpha_k = 0.15$ min, $\beta_k = 4$, $\gamma_k = 10,000$ vehicle per lane per day, δ_k is computed based on actual distance between nodes at a free-flow speed of 35 mph.

We consider a scenario that disruption initiates at the main power substation in the center of the city (the star in Figure 3.3), and investigate the failure propagation and social impacts. The electricity substation is a critical node that supports a variety of other infrastructure nodes, including water treatment plants and a large number of local transformers. After implementing our algorithm, the system converges to an equilibrium. In the last few iterations of the diagonalization algorithm, the change in the total infrastructure failures Δ_x^r is no greater than $2 \cdot 10^{-3}$ between consecutive iterations, and the change continues to diminish. This demonstrates the convergence of the result. The resulting infrastructure status is shown in Figure 3.4, where a component having a working status less than 0.5 is represented by a hollow symbol to indicate either complete or

partial loss of functionality. As expected, all electricity transformers are shut down, the electricity network is disabled, while only local electricity generators can work based on fuel, providing limited power supply to nearby communities. Furthermore, the water network is also disrupted, while only local commercial water vendors that can pump water from wells using generators still function. As a result, education institutions and healthcare facilities are all disrupted due to the shortage of power and water.
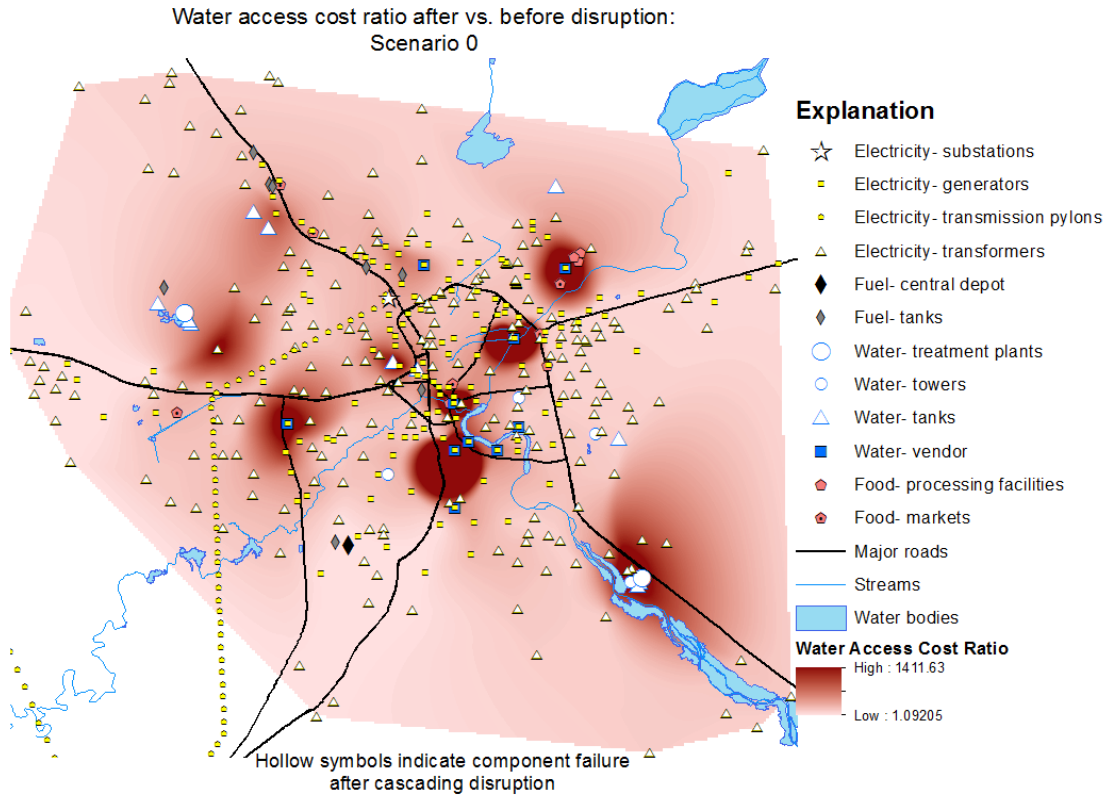

Figure 3.4: Community water access cost ratio after versus before disruption.

The social impacts are estimated as follows. In total, 62.4% of the water nodes, 0.0% of the food nodes, 0.0% of the fuel nodes, and 100.0% of education and healthcare nodes are disrupted (i.e., $\Delta_x^{water} = 17.5$, $\Delta_x^{food} = 0.0$, $\Delta_x^{fuel} = 0.0$, $\Delta_x^{education} = 84.0$, and $\Delta_x^{healthcare} = 4.0$). There is no demand loss for water and food, but all demand for education and healthcare are lost. Interestingly, after disruption, the average access cost for food even decreases from 0.35 to 0.32. This is mainly due to the loss of education and healthcare needs which relieve traffic congestion in some parts of the city. The overall average access cost for water increases from 0.25 to 1.42, and 4.3% of population lost access to water, which is caused by the increase of queueing cost at remaining water facilities. The spatial distribution of changes in communitiesâĂŹ water access

costs (Figure 3.4) shows that costs increase significantly around several of the commercial water providers as competition increases for those limited resources.

3.4.2 Sensitivity analysis

We denote the above case study as a benchmark Scenario 0, and now examine several representative scenarios under different parameter settings. Scenario 1 considers only transportation cost to examine the assumption that there is no queueing cost at resource locations; scenario 2 and 3 investigated the model sensitivity to facility service capacity; scenario 4, 5, and 6 were selected to test the reliability of the system when other critical components were disrupted. The corresponding societal impacts are summarized in Table 3.1. The augmented network contains 3266 nodes and 15319 links, with 1105 O-D pairs . Due to the giant network size, it took the algorithm from 20 minutes to approximately 1 day to evaluate each of the six disruption scenarios. First, in Scenario 1, we remove the consideration of queueing cost at resource location. To do that, we set the maximum resource capacity $\overline{\lambda}_j$ of each facility $j \in I_r, r \in R$ to be infinity, and let the interdependency function take a binary value, such that:

$$
x_i = \begin{cases} 1, \text{ if } \min\left(\min_{j \in S_i} x_j, \min_{r \in R} d_i^r / \hat{d}_i^r \right) \geq \delta_i, \forall i \in I \setminus (I_t \cup I_c) \\ 0, \text{ otherwise} \end{cases},
$$

where δ_i is a constant that represents a threshold value. In this case we let $\delta_i = 0.9$. Moreover, the removal of queueing makes the transportation cost the only resource procurement. cost, which is far less sensitive to the flows, hence the elastic demand function of facilities is changed into:

$$
D_i^r \left(c_i^r; \hat{d}_i^r \right) = \begin{cases} \hat{d}_i^r, \text{ if } c_i^r < \max\left\{ 0.6 \cdot \max_{j \in I_{r'}} \hat{c}_j^r, 1.2 \hat{c}_i^r \right\}, \\ 0, \end{cases} \tag{3.29}
$$

where $\hat{c}_j^r$ represents the resource procurement cost for resource r of facility j under no disruption, and $I_{r'}$ is the set of facilities of the same type as facility i (i.e. $i \in I_{r'}$). The factors "0.6" and "1.2" are chosen based on past experiences and status quo, such that facility has a relatively competitive ability to change its resource locations compared to other facilities of the same type, and the facilities have an overall low flexibility to change their resource locations. Since the transportation cost is the total access cost, the reduction of traffic leads to a more noticeable decrease in the food access cost. There is no significant change in the water access cost after disruption, which means the negative effect of water facilities failure is cancelled out by the reduced traffic congestion.

Scenario 2 and 3 examines the impact of facility maximum resource capacity $\overline{\lambda}_j$. It is expected

Table 3.1: Sensitivity analysis results of different disruption scenarios and different parameters.

Scenarios	Food			Water		
	Access cost increase	Demand loss	# Node failed (total 11)	Access cost increase	Demand loss	# Node failed (total 28)
0: Case Study	-7%	0.00%	0	458%	4.30%	17.5
1: No Queueing	-36%	0.00%	0	0.50%	0.00%	18
2: $\kappa_{cap} = 5.0$	-49%	0.00%	0	-25%	0.00%	17.5
3: $\kappa_{cap} = 10.0$	-54%	0.00%	0	-35%	0.00%	17.5
4: Init. Water	-27%	0.00%	0	20%	2.90%	9.3
5: Init. Fuel	22%	1.30%	0.5	157%	5.90%	10.6
6: Water and Fuel	-2%	0.00%	0.3	860%	8.70%	16.2

that an infrastructure system with higher maximum resource capacity would have greater resiliency to system disruption. In Scenario 0, we let the resource users visit their nearest resource locations, and compute the resource capacity of the resource locations based on the demand by multiplying a factor κ_{cap} of 1.1. In Scenario 2 and 3, κ_{cap} is changed to 5.0 and 10.0 respectively to consider an infrastructure system with moderate and high resource capacity. Similar to what we observe in Scenario 1, as the queueing cost reduces, the transportation cost takes a greater portion in the overall access cost, hence the alleviation of traffic congestion results in a greater percent reduction in access cost. We observe a decrease of 49% and 25% respectively in food and water access cost in Scenario 2 and 54% and 35% in Scenario 3.

Scenario 4 explores a different initial disruption at a water treatment plant. Counter-intuitively, although water treatment plant is the key infrastructure in the water layer, shutting it down does not significantly affect the entire system. Although 4.0 water towers (due to support failures), 3.0 water tanks and 2.3 water vendors (due to resource failures) are disrupted, the water access cost only slightly increases by 20%, and there is 2.9% water demand loss. This is mainly because of the high inelasticity of resource demand of facilities makes resource failures very unlikely to happen. Further, we find no disruptions in the food layer, and there is a decrease of 27% in food procurement cost.

In Scenario 5, we assume the initial disruption occurs at the central fuel depot, which directly cut down the fuel supply to the entire system. As a result, ten local water vendors are completely disrupted (due to resource failures), and the status of several water tanks are impaired due to increased congestion. Similarly, the food markets are also affected due to congestion, with a total status reduction of 0.5. The community can still survive well without fuel supply, whereas the access cost for water and food have a moderate increase. Therefore, the fuel layer does not play a critical role in the infrastructure system if the power grid and water network still function well.

In Scenario 6, the water treatment plant and the fuel depot are both failed in the initial disruption. The result turns out to be more devastating than the simple summation of the results where

the two facilities are disrupted respectively. The water access cost is increased by 860%, with a population of 8.7% lost water access. The demand loss appears to be low while there is a dramatic increase in access cost, because both of the inelasticity of resource demand and that there is still relatively sufficient resource supply in the infrastructure system.

The results could be sensitive w.r.t. the cost function parameters as well as the elasticity of facilitiesâĂŹ resource demand. Through the numerical experiments, the result is observed to be sensitive w.r.t. the facilitiesâĂŹ maximum resource capacity, especially when the capacity is at the low end. Moreover, the solution can also be sensitive w.r.t. the elasticity of facilitiesâĂŹ resource demand. To be specific, the demand elasticity affects the facilitiesâĂŹ resiliency against resource failures, and the resiliency has a significant impact on the resource seeking behavior of the communities. However, the demand elasticity of the population does not have such a remarkable impact on the solution âĂŞ change in the demand elasticity only means that the population can endure higher/lower resource access cost, and it only brings in a negligible change to the traffic.

CHAPTER 4 – SIR EPIDEMIC CONTROL IN A MULTI-GROUP POPULATION WITH QUARANTINE AND EDUCATIONAL PROGRAMS

4.1   Introduction

The previous chapter focuses on interdependent infrastructure systems and facility failure propagation. Starting from this chapter, the following parts of this dissertation focus on population contact networks and infectious disease propagation among the population. We inherit the notation and the formulas from Section 2.2.1 for the rest of this dissertation.

A critical issue in epidemic modeling and control is social heterogeneity among populations. First, individuals differ from each other in their ability to resist potential infections and that to recover/survive if infected. The reasons for such differences include physiological features such as age and gender, as well as social characteristics such as career and income level. Moreover, the contact networks of people also highly depend on their physiological/social traits; e.g., people that share similar characteristics often interact more among themselves than with others. Furthermore, the effectiveness of epidemic control measures on different subpopulations can be different. For example, it may require less efforts to monitor and control disease propagation among school children than that among working class adults.

On the other hand, the detrimental impacts of disease control measures are non-negligible, and it is important to balance the trade-off between the impact of disease propagation and that of the control. For instance, quarantine and isolation of infected individuals help provide timely treatment to them and protect the susceptible population to mitigate disease propagation, yet it may cause psychological sufferings to the isolated personnels and might raise political concerns. Moreover, the intensity and timing of an educational campaign can greatly affect the disease spreading process as well as the consequential social impacts due to societal panics toward the disease.

To address such challenges, this chapter proposes an optimization framework to find the optimal control strategies during an epidemic in a multi-group and networked population with the aforementioned considerations. We incorporate the effects of two control approaches, educational

program and quarantine into an existing multi-group epidemic dynamics model. A system state evaluation model is developed based on observations from quarantine control to help decision makers understand the disease propagation status. Then an optimization framework is established and a value function approximation approach is adopted to solve for the optimal control.

The rest of this chapter is organized as follows. Section 4.2 develops an epidemic dynamics model with educational programs and quarantine controls, and proposes a latent model to estimate the dynamic system states. Next, Section 4.3 establishes a control framework and introduces a solution approach to solve the optimal control problem. Then in Section 4.4, a semi-realistic case study is performed to validate the modeling framework and the solution approach, and sensitivity analysis results are presented to reveal some key factors that affect the operational decisions and the impacts of disease propagation.

4.2 Epidemic Dynamics and Control Strategies

4.2.1 Epidemic dynamics model

We categorize the population into several groups based on their physical and social characteristics, such as age, gender, career, etc. We let the set of groups be $\mathcal{I}$, and each group $i \in \mathcal{I}$ is associated with a subpopulation size N_i. In a multi-group population, the degree distribution generating function for a group-i node is given as follows:

$$g_i(\mathbf{x}) = \sum_{k_j, j \in \mathcal{I}} P_i\left(\ldots, k_j, \ldots\right) \prod_{j \in \mathcal{I}} x_{ij}^{k_j}, \forall i \in \mathcal{I}, \tag{4.1}$$

where k_j is the degree of this group-i node to group j nodes, $P_i\left(\ldots, k_j, \ldots\right)$ is the probability that this node has degree k_j to group j, $\forall j \in \mathcal{I}$, and x_{ij} is a dummy variable.

Miller and Volz (2013) extend their earlier model (2.3) - (2.7) to incorporate population spatial heterogeneity, which could also be generalized to address a much broader spectrum of population heterogeneities. The system dynamics equations of the multigroup heterogeneous mixing epidemic dynamics model are quite similar to those of the single-group model Eq (2.3) - (2.7), presented as

follows:

$$\dot{\theta}_{ij} = -r_{ij}\phi_{ij}, \forall i, j \in \mathcal{I}, \tag{4.2}$$

$$\dot{\phi}_{ij} = (-r_{ij} - u_j)\phi_{ij} - \frac{1}{\partial g_j(\mathbf{1})/\partial x_i} \frac{d}{dt}(\partial g_j(\boldsymbol{\theta}_j)/\partial x_i), \forall i, j \in \mathcal{I}, \tag{4.3}$$

$$\dot{R}_i = u_i I_i, \forall i \in \mathcal{I}, \tag{4.4}$$

$$S_i(t) = g_i(\boldsymbol{\theta}_i(t)), \forall i \in \mathcal{I}, t \geq 0, \tag{4.5}$$

$$I_i(t) = 1 - S_i(t) - R_i(t), \forall i \in \mathcal{I}, t \geq 0, \tag{4.6}$$

where $\theta_{ij}(t)$ denotes the fraction of edges between group i and j that has not transmitted disease from group j to i at time t, $\phi_{ij}(t)$ is the fraction of edges between group i and j that has not transmitted disease from group j to i at time t where the j node is infectious, and $\boldsymbol{\theta}_i = (\theta_{ij})_{\forall j \in \mathcal{I}}$. Moreover, r_{ij} and u_i are the Poisson process rates of an infectious contact from group j to i and the recovery of a group-i infected node, respectively. In (4.3), the first part on the right hand side means the decrease of ϕ_{ij} due to (i) disease transmission along the (j, i) edges, and (ii) the natural removal of infected group-j nodes. The second part means the increase of ϕ_{ij} due to the susceptible group-j neighbors of the group-i test node becoming infected by their own neighbors (other than the test node as infection from the test node is prohibited).

Moreover, whether an infected individual recovers or deceases has no impact on the disease propagation process; however, they have completely different impacts and can lead to different control objectives and decisions. In particular, we separate the recovered population from the deceased, and denote u_{ic} and u_{id} as the natural recovery rate and death rate of infected group-i nodes, respectively, such that $u_i = u_{ic} + u_{id}, \forall i \in \mathcal{I}$; then we let R_{ic} and R_{id} be the corresponding population fractions, respectively:

$$\dot{R}_{ic} = u_{ic} I_i, \forall i \in \mathcal{I}, \tag{4.7a}$$

$$\dot{R}_{id} = u_{id} I_i, \forall i \in \mathcal{I}. \tag{4.7b}$$

4.2.2 Disease control strategies

There exists a public health agency that takes two control actions, educational programs and random screening-based quarantine, to mitigate the disease propagation process. There might be limited medical resources for the agency, and that the agency's control actions may cause detrimental impacts to the society, such that the it needs to optimize its course of actions when making control decisions. We examine each of them over their effects in the dynamic process, their resource requirements, and detrimental impacts, respectively.

Caution contact advisory

When an epidemic outbreak happens, the public health agency can improve the people's awareness to the disease, e.g. by posting sanitary advisories to the population. In doing so, precautions are taken when individuals interact with the others, and people's personal sanitary levels increases. Consequently, the transmission rates of the infectious disease are reduced.

When the public agency posts an advisory to group i, the transmission rates from other groups to group i and those from group i to others are reduced at the same proportion. To be specific, we define a discount factor $\gamma_i(t)$ to capture the effect of advisory to group i, where $\gamma_i(t) \in \left[\underline{\gamma}_i, 1\right]$, and $\underline{\gamma}_i > 0$ is a constant lower bound, for $i \in \mathcal{I}$. Then the reduced transmission rate from group $j \in \mathcal{I}$ to group i, is $\gamma_i(t)\gamma_j(t)r_{ij}$, and that of the opposite direction is $\gamma_i(t)\gamma_j(t)r_{ji}$. In reality, the effect of an advisory may gradually wear off, and we assume that the discount factor approaches 1 exponentially over time:

$$\dot{\gamma}_i = \eta_i \left(1 - \gamma_i\right), \forall i \in \mathcal{I}, \tag{4.8}$$

where $\eta_i > 0$ is the diminishing rate of the advisory effect.

There will be detrimental impacts caused by the educational programs, such as social panics and functionality loss. These impacts will be accumulated over time, while the instantaneous impact of an advisory diminishes with the advisory effects. We use a function $f_i(\gamma_i)$ to capture the instantaneous impact. It can take arbitrary form that is non-increasing and satisfies $f_i(1) = 0$. An example is as follows:

$$f_i\left(\gamma_i\right) = -a_i \log_{b_i}\left(\gamma_i\right), \forall i \in \mathcal{I}. \tag{4.9}$$

where $a_i > 0, b_i > 1$ are constant parameters. Then the total impact from $t = 0$ to the end of the control horizon T, is given by $\int_0^T f_i\left(\gamma_i(t)\right) dt$.

Moreover, constantly changing warning levels in a short time frame might damage the agency's authority and thus the effectiveness of advisories. Therefore, the agency must decide the best combination of warning levels and timings. In particular, we denote $\tilde{\gamma}_i$ as the agency's decision variable on the warning level of the advisory to group i, and $\tilde{t}_i$ as the time of posting the warning to group i. Then the discount factor γ_i can be expressed as follows:

$$\gamma_i(t) = \begin{cases} 1, & \text{if } t \leq \tilde{t}_i, \\ \left(\tilde{\gamma}_i - 1\right) e^{-\eta_i\left(t - \tilde{t}_i\right)} + 1, & \text{o.w.} \end{cases}, \forall i \in \mathcal{I}, \tag{4.10}$$

By definition, $\tilde{\gamma}_i$ takes value between a lower bound $\underline{\gamma}_i$ and a upper bound of 1:

$$\tilde{\gamma}_i \in \left[\underline{\gamma}_i, 1\right], \forall i \in \mathcal{I}. \tag{4.11}$$

Random-screening-based quarantine

Quarantine controls are usually carried out by screening the population randomly; e.g., by setting up checkpoints and testing the passersby. It then provides treatment to and/or isolates the detected infected individuals, such that the infected ones are removed from the disease transmission network. We assume an infected group-i individual is detected by the agency at *detection rate* $\mu_i(t)$ following Poisson distribution.[a] Since the detection and natural removal of the infected nodes are independent processes, the overall removal rate at time t of an infected group-i node is $\mu_i(t) + u_i$, $\forall i \in \mathcal{I}$. Therefore, by re-writing (4.2), (4.3), and (4.4) with the discounted transmission rates and overall removal rates, the new system dynamics of the spread of disease under both control strategies are presented as follows:

$$\dot{\theta}_{ij} = -r_{ij}\gamma_i\gamma_j\phi_{ij}, \forall i, j \in \mathcal{I}, \tag{4.12}$$

$$\dot{\phi}_{ij} = (-r_{ij}\gamma_i\gamma_j - u_j - \mu_j)\phi_{ij} - \frac{1}{\partial g_j(\mathbf{1})/\partial x_i}\frac{d}{dt}\left(\partial g_j(\boldsymbol{\theta}_j)/\partial x_i\right), \forall i, j \in \mathcal{I}. \tag{4.13}$$

$$\dot{R}_i = (u_i + \mu_i)I_i, \forall i \in \mathcal{I}. \tag{4.14}$$

Due to varying health-related characteristics, subpopulation sizes, and degree distributions of different groups, the effectivenesses of quarantine can differ drastically among groups. We let $\mathcal{L}$ be the set of screen actions the agency can take, each of which has a different effect on various groups. We let $\omega_l(t)$ be the intensity of action $l \in \mathcal{L}$ at time t, namely how much effort the agency put into action l, then

$$\mathbf{A}(t)\boldsymbol{\omega}(t) = \boldsymbol{\mu}(t), \forall t \geq 0. \tag{4.15}$$

where $\boldsymbol{\omega} = (\omega_l)_{\forall l \in \mathcal{L}}$, $\boldsymbol{\mu} = (\mu_i)_{i \in \mathcal{I}}$, and $\mathbf{A}(t) \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{L}|}$ is a matrix that maps action intensity to the detection rate in each group at time t. In other words, the elements in $\mathbf{A}$ reflect the actual detection rate in each group per unit of intensity for each action. The entries in $\mathbf{A}$ could be increasing over time, as when some population are deceased or isolated and the potential quarantine subjects are reduced, higher detection rate could be achieved with the same level of action intensity.

Moreover, the agency needs to properly allocate different types of resources $q \in \mathcal{Q}$ to each screen action. To be specific, we denote c_{ql} as the amount of resource q required to achieve 1 unit of intensity of action l, the resource consumption at t due to screening actions is $\mathbf{c}_q^T\boldsymbol{\omega}(t)$, where $\mathbf{c}_q = (c_{ql})_{\forall l \in \mathcal{L}}$:

$$\mathbf{c}_q^T\boldsymbol{\omega}(t) \leq B_q(t), \forall q \in \mathcal{Q}, t \geq 0, \tag{4.16}$$

$B_q(t)$ is the available amount of type-q non-accumulative resource at t. In addition, the action

---

[a]Some infected people may actively seek for help from the authority, and such consideration can be easily incorporated into the natural removal rate of the associated group.

intensities and detections rates can only take non-negative values:

$$\boldsymbol{\mu}(t), \boldsymbol{\omega}(t) \geq 0, \forall t \geq 0. \tag{4.17}$$

Quarantine as well induces negative social impacts among population. We use another penalty function $h_i\left(\mu_i(t)I_i(t)\right)$ to reflect the impact caused by quarantine. h_i can take arbitrary forms that is non-decreasing and satisfies that $h_i(0) = 0$. For example,

$$h_i\left(\mu_i I_i\right) = c_i \left(e^{d_i \mu_i I_i} - 1\right), \forall i \in \mathcal{I}, \tag{4.18}$$

where $c_i, d_i > 0$ are constant parameters.

Now we have presented the epidemic dynamics along with effects of the agency's controls. In order to make optimal control decisions, the agency needs complete knowledge of all system state variables, including some unobservable quantities such as $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. In the following subsection, we propose a latent model to infer the hidden state variables based on observations obtained by the quarantine.

4.2.3 System state evaluation

Although the system state is not completely observable, certain observations can be made on the system over time. They are compared with the decision-maker's current perception of the system state, such that the perception can be adjusted along the way. We first set up the latent model framework, and then discuss how observations are fed into the model.

First, as the dynamic system (4.5), (4.6), (4.12)-(4.14) is deemed as deterministic in statistical senses, once the initial conditions and the controls are known, the trajectory of the system can be computed. Therefore, the system state at any time can be derived based on merely the initial condition and the controls. The initial condition of the system can be approximated as follows:

$$\left(\boldsymbol{\theta}_i(0), \boldsymbol{\phi}_i(0), S_i(0), I_i(0), R_i(0)\right) \approx \left(\mathbf{1}, \mathbf{e}_i, 1, 0, 0\right), \forall i \in \mathcal{I}. \tag{4.19}$$

The fractions of infected and removed populations are assumed to be small and thus approximated by 0; naturally, the susceptible fraction is approximated by 1, as well as $\boldsymbol{\theta}_i(0), \forall i \in \mathcal{I}$. These approximations only introduce a negligible error into the dynamics. However, it has been observed that the system is quite sensitive to the initial condition of $\boldsymbol{\phi}$, namely $\mathbf{e} = (\mathbf{e}_i)_{\forall i \in \mathcal{I}}$. Ideally, the agency should be able to assess the initial condition $\mathbf{e}$ based on the early endemic phase disease propagation. We assume the agency has a prior distribution for the initial condition $\mathbf{e}$, which is a Gaussian distribution with mean $\bar{\mathbf{e}}$ and variance Σ^2, i.e., the estimation of $\mathbf{e}$ has prior distribution

37

$\mathcal{N}\left(\bar{\mathbf{e}}, \Sigma_{\mathbf{e}}^2\right)$. However, such perception might well be biased, and the agency needs to validate and adjust its perception on e based on available observations.

In practice, there might be many possible methods of making different observations on the system. In this paper, we utilize the observations that made available by the random screening to demonstrate the latent model, but the model can also be easily generalized if other observations become available. We assume that, once detected, an infected individual's social group information will be revealed to the agency, which allows the agency to keep track of the number of detected infectious individuals in each group. We denote $\mathcal{T}_t$ as the set of time instants when observations are made from the beginning to time t, and denote the fraction of group-i individuals that are recently detected at time t' as $\delta_i(t')$, $\forall i \in \mathcal{I}, t' \in \mathcal{T}_t$. Moreover, the observations may contain errors because, for example, the disease detection approach is not perfectly accurate. We assume δ_i follows Gaussian distribution with mean $\mu_i I_i + \epsilon_i$ and variance $\Sigma_{\delta_i}^2$, where ϵ_i is the bias of the observation, i.e., $\delta_i \sim \mathcal{N}\left(\mu_i I_i + \epsilon_i, \Sigma_{\delta_i}^2\right)$.

Based on the current estimation of e, denoted as $\hat{\mathbf{e}}^{(t)}$, the system state at any time t' can be computed following the system equations (4.5), (4.6), and (4.12)-(4.14), and we denote $\hat{I}_i\left(t', \hat{\mathbf{e}}^{(t)}\right)$ as the prediction of infected fraction of group i population I_i given $\hat{\mathbf{e}}^{(t)}$ at time t', $\forall i \in \mathcal{I}, t, t' \geq 0$. At time t, we solve the following optimization problem, which is essentially a maximum a posteriori (MAP) model, to obtain the best estimation on the system initial condition:

$$\textbf{SEP1}_t \quad \max_{\mathbf{0} \leq \hat{\mathbf{e}}^{(t)} \leq \mathbf{1}} P\left(\frac{\delta_i(t') - \epsilon_i}{\mu_i(t')} = \hat{I}_i\left(t', \hat{\mathbf{e}}^{(t)}\right), \forall i \in \mathcal{I}, t' \in \mathcal{T}_t \middle| \hat{\mathbf{e}}^{(t)}\right) P\left(\hat{\mathbf{e}}^{(t)}\right) \quad (4.20)$$
$$\text{s.t. } (4.5), (4.6), \text{and } (4.12) - (4.14).$$

The objective (4.20) is to find the "most likely" value of the initial condition e that may yield the observations δ_i. In a special case where the observation is unbiased and the observation randomness is negligible, i.e., $\epsilon_i, \Sigma_{\delta_i}^2 = 0, \forall i \in \mathcal{I}$, we have:

$$P\left(\frac{\delta_i(t') - \epsilon_i}{\mu_i(t')} = \hat{I}_i\left(t', \hat{\mathbf{e}}^{(t)}\right), \forall i \in \mathcal{I}, t' \in \mathcal{T}_t \middle| \hat{\mathbf{e}}^{(t)}\right) = \begin{cases} 1, & \text{if } \frac{\delta_i(t')}{\mu_i(t')} = \hat{I}_i(t'), \forall i \in \mathcal{I}, t' \in \mathcal{T}_t, \\ 0, & \text{o.w.} \end{cases}$$
$$(4.21)$$

In this case, the problem $\textbf{SEP1}_t$ is equivalent to the following problem by observation:

$$\textbf{SEP2}_t \quad \max_{\mathbf{0} \leq \hat{\mathbf{e}}^{(t)} \leq \mathbf{1}} P\left(\hat{\mathbf{e}}^{(t)}\right) \quad (4.22)$$

$$\text{s.t. } \frac{\delta_i(t')}{\mu_i(t')} = \hat{I}_i(t'), \forall i \in \mathcal{I}, t' \in \mathcal{T}_t, \quad (4.23)$$
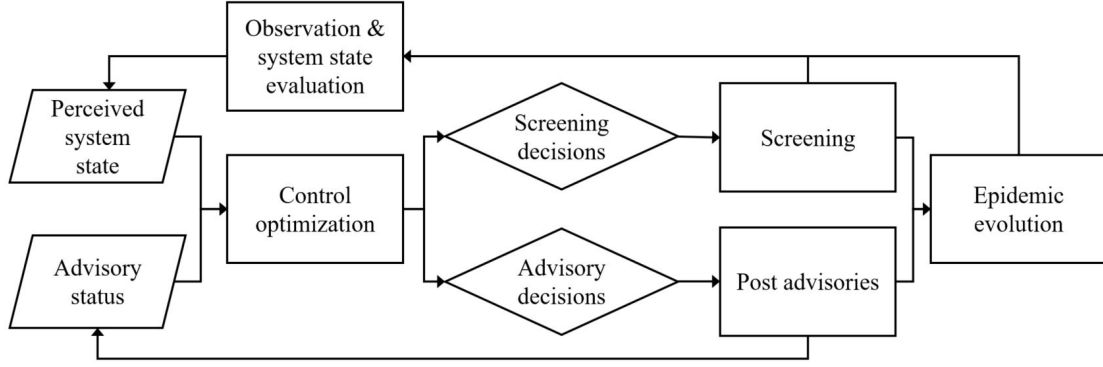
$$\text{and } (4.5), (4.6), \text{and } (4.12) - (4.14).$$

Figure 4.1: Control scheme illustration.

The nonlinear system equations (4.5), (4.6), (4.12)-(4.14), and (4.23) defines the feasible set of **SEP2**$_t$, whose size naturally decreases as observations are accumulated over time. We conjecture that there exists a $t > 0$, such that the feasible set of **SEP2**$_t$ contains one unique solution, which is the ground-truth value of the initial condition e.

4.3 Optimal Control and Solution Approach

So far we have discussed the control strategies and presented the system state evaluation model, we are will now present the optimal control problem, and propose a solution approach to solve the problem.

4.3.1 Optimal control problem

The continuous time system dynamics and controls is discretized into small time intervals. We choose a *period* with moderate length such as a day as a suitable time step for discretization of the dynamic equations, and an equal or larger interval such as a week, as a plausible length of a *decision epoch* in reality. We denote the length of an epoch in terms of Δ periods, and consider a control horizon of K epocs, which corresponds to a time span of $T = K\Delta$ periods. We also denote $t = 0$ as time of the initial state of the system. By the end of each period, the number of detected infected individuals in each group will be reported to the agency for system state evaluation, i.e., the observation time instants $\mathcal{T}_t = \{1, \ldots, t\}, \forall t = 1, \ldots, T$.

The overall control scheme is illustrated in Figure 4.1. At the current epoch, the agency solves an optimal control problem based on its knowledge on the system state, and determines the best control decisions for the current epoch as well as the future; then it carries out the optimal decisions to intervene the natural evolution of the epidemics; through screening, the agency is able to obtain

new observations to calibrate its perception on the initial condition and the current state of the system; in the next decision epoch, the agency uses the updated perceived system state to solve another optimal control problem, and repeat this process until the end of the control horizon.

Although variables $\tilde{t}_i$ and $\tilde{\gamma}_i$ are time-invariant decisions as the agency can only post advisories to each group for once, in order to unify the notions with the other decision variables, they are still treated as time-dependent variables and expressed as $\tilde{t}_{i,t}$ and $\tilde{\gamma}_{i,t}$, respectively. Correspondingly, constraints must be added to regulate these decision variables to avoid multiple warning-posts. To this end, we define a binary indicator $y_{i,t}$ to capture the status of the post:

$$y_{i,t} = \mathbb{1}_{\tilde{t}_{i,t} \leq t}, \forall i \in \mathcal{I}, t = 0, \ldots, T, \tag{4.24}$$

If $y_{i,t-1} = 0$, a warning has not been posted to group i by the beginning of period t, and yes otherwise. Without loss of generality, we denote $y_{i,0} = 0, \forall i \in \mathcal{I}$. Then the constraints on $\tilde{t}_{i,t}$ and $\tilde{\gamma}_{i,t}$ are presented as follows:

$$t(1 - y_{i,t-1}) + \tilde{t}_{i,t-1} y_{i,t-1} \leq \tilde{t}_{i,t} \leq T(1 - y_{i,t-1}) + \tilde{t}_{i,t-1} y_{i,t-1}, \forall i \in \mathcal{I}, t = 1, \ldots, T, \tag{4.25}$$

$$\underline{\gamma}_i (1 - y_{i,t-1}) + (\gamma_{i,t-1} + \eta_i(1 - \gamma_{i,t-1})) y_{i,t-1} \leq \tilde{\gamma}_{i,t} \leq 1 - y_{i,t-1} + (\gamma_{i,t-1} + \eta_i(1 - \gamma_{i,t-1})) y_{i,t-1},$$
$$\forall i \in \mathcal{I}, t = 1, \ldots, T. \tag{4.26}$$

Constraints (4.25) state that once a warning is posted, $\tilde{t}_{i,t}$ remains the same as the actual posting time $\tilde{t}_i$, $\forall t \geq \tilde{t}_i$; and Constraints (4.26) regulate that only when the warning has not been posted, $\tilde{\gamma}_{i,t}$ can take values other than that naturally deteriorated from the previous period. Without loss of generality, we let $\tilde{t}_{i,0} = T$ and $\gamma_{i,0} = 1$. Moreover, for simplicity we assume the advisories can only be posted at the beginning of each epoch, i.e.,

$$\tilde{t}_{i,t} \in \{(k-1)\Delta + 1 | k = 1, \ldots, K\}, \forall i \in \mathcal{I}. \tag{4.27}$$

At the beginning of the epoch k, namely period $t = (k-1)\Delta + 1$, the agency solves the following optimization problem in a discretized dynamic system. For notational brevity, we use the subscript k interchangeably with $t = (k-1)\Delta + 1$. Note that when solving the following problem at epoch k, decision variables with subscript $t' \leq (k-1)\Delta$ belong to the past, and thus

are no longer decisions to make in the current epoch.

$$
\textbf{OCP}_k \min_{\substack{\boldsymbol{\omega}_{t'},\boldsymbol{\mu}_{t'},\boldsymbol{\gamma}_{t'}, \\ \tilde{\boldsymbol{\gamma}}_{t'},\tilde{\boldsymbol{t}}_{t'},\boldsymbol{y}_{t'}, \\ \forall t'=t,\ldots,T}} \sum_{t'=t}^{T} \sum_{i\in\mathcal{I}} \alpha_1 N_i \sum_{j\in\mathcal{I}} \frac{\partial g_i(\boldsymbol{\theta}_{i,t'-1})}{\partial x_j} \gamma_{i,t'} \gamma_{j,t'} r_{ij} \phi_{ij,t'-1} \tag{4.28}
$$

$$
+ \alpha_2 N_i u_{id} I_{i,t'-1} + \alpha_3 f_i\left(\gamma_{i,t'}\right) + \alpha_4 h_i\left(\mu_{i,t'} I_{i,t'-1}\right)
$$

$$
\text{s.t. } \theta_{ij,t'} = \theta_{ij,t'-1} - r_{ij} \gamma_{i,t'} \gamma_{j,t'} \phi_{ij,t'-1}, \forall i,j \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.29}
$$

$$
\phi_{ij,t'} = \left(1 - \gamma_{i,t'} \gamma_{j,t'} r_{ij} - u_j - \mu_{j,t'}\right) \phi_{ij,t'-1}
$$

$$
+ \frac{\sum_{m\in\mathcal{I}} \gamma_{j,t'} \gamma_{m,t'} r_{jm} \phi_{jm,t'-1} \frac{\partial^2 g_j(\boldsymbol{\theta}_{j,t'-1})}{\partial x_i \partial x_m}}{\partial g_j(\mathbf{1})/\partial x_i}, \forall i,j \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.30}
$$

$$
S_{i,t'} = g_i\left(\boldsymbol{\theta}_{i,t'}\right), \forall i \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.31}
$$

$$
R_{ic,t'} = R_{ic,t'-1} + u_{ic} I_{i,t'-1}, \forall i \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.32}
$$

$$
R_{id,t'} = R_{id,t'-1} + u_{id} I_{i,t'-1}, \forall i \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.33}
$$

$$
R_{i,t'} = R_{i,t'-1} + \left(u_i + \mu_{i,t'}\right) I_{i,t'-1}, \forall i \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.34}
$$

$$
I_{i,t'} = 1 - S_{i,t'} - R_{i,t'}, \forall i \in \mathcal{I}, t' = 1,\ldots,T, \tag{4.35}
$$

$$
\gamma_{i,t'} = \tilde{\gamma}_{i,t'} y_{i,t'-1} + \left(1 - y_{i,t'-1}\right) \left(\mathbb{1}_{\tilde{t}_{i,t'}=t'} \tilde{\gamma}_{i,t'} + 1 - \mathbb{1}_{\tilde{t}_{i,t'}=t'}\right), \forall t' = t,\ldots,T, i \in \mathcal{I}, \tag{4.36}
$$

$$
\mathbf{A}'_t \boldsymbol{\omega}_{t'} = \boldsymbol{\mu}'_t, t' = t,\ldots,T, \tag{4.37}
$$

$$
\mathbf{c}_q^T \boldsymbol{\omega}_{t'} \leq B_{q,t'}, \forall q \in \mathcal{Q} \backslash \tilde{\mathcal{Q}}, t' = t,\ldots,T, \tag{4.38}
$$

$$
\boldsymbol{\mu}_{t'}, \boldsymbol{\omega}_{t'} \geq 0, \forall t' = t,\ldots,T, \tag{4.39}
$$

$$
\boldsymbol{\omega}_{t'} = \boldsymbol{\omega}_{k'}, \forall (k'-1)\Delta + 1 \leq t' \leq k'\Delta, k' = k,\ldots,K, \tag{4.40}
$$

$$
\left(\boldsymbol{\theta}_{i,0}, \boldsymbol{\phi}_{i,0}, S_{i,0}, I_{i,0}, R_{i,0}\right) = \left(\mathbf{1}, \hat{\mathbf{e}}_i^{(k-1)}, 1, 0, 0\right), \forall i \in \mathcal{I}, \tag{4.41}
$$

and (4.9), (4.18), (4.24) − (4.27).

The objective contains four parts: (i) the total number of predicted infections occurred by time T, (ii) the total predicted death tolls caused by the disease, (iii) the social impact caused by quarantine of infected individuals, and (iv) the impact caused by the warnings. Different weight factors $\alpha_m, m = 1,\ldots,4$ are assigned to these terms respectively to reflect different objectives. Constraints (4.36) states that if the warning has not been posted and the decision is to post the warning later than t', $\gamma_{i,t'} = 1$, otherwise it equals $\tilde{\gamma}_{i,t'}$. Constraints (4.40) regulate that the quarantine control must be consistent throughout a period. Moreover, Constraints (4.41) are the perception of the initial condition at current stage, where $\hat{\mathbf{e}}_i^{(k-1)}$ is the optimal solution to $\textbf{SEP1}_{(k-1)\Delta}$.

Proposition 3. *If the detection rate $\mu_{i,k'}$ is always bounded by a finite number M, $\forall k' = k,\ldots,K, i \in$*

$\mathcal{I}$, then the objective (4.28) is bounded, and the solution to **OCP**$_k$ exists for a finite $K > 0$ and $k = 1, \ldots, K$.

Proof. It is straightforward to see that the first two terms in the objective (4.28) is bounded, as $S_i(T) \geq 0$. Furthermore, as μ_i is bounded and $I_i \leq 1, \forall i \in \mathcal{I}$, we can always find a linear function $\bar{h}_i(x) = A_i x$, such that $\bar{h}_i(x) \geq h_i(x), \forall x \in [0, M]$. Then we have

$$\sum_{t'=t}^{T} h_i\left(\mu_{i,t'} I_{i,t'-1}\right) \leq \sum_{t'=t}^{T} \bar{h}_i\left(\mu_{i,t'} I_{i,t'-1}\right) = \sum_{t'=t}^{T} A_i \mu_{i,t'} I_{i,t'-1} \leq A_i\left(R_{i,T} - R_{ic,T} - R_{id,T}\right) \leq A_i.$$

Moreover, since γ_i's are bounded by $\underline{\gamma}_i$ as indicated by (4.26), $\sum_{t'=t}^{T} f_i(\gamma_i)$ is as well bounded, and the objective is well-defined and bounded. Furthermore, the solution set defined by Constraints (4.9), (4.18), (4.27) and (4.29) - (4.41) is bounded and closed, the solution to problem **OCP**$_k$ exists. □

4.3.2 Approximate dynamic programming

Problem **OCP**$_k$ is a multistage decision problem in a highly nonlinear dynamic system, and it appears especially formidable as the agency needs to solve a similar problem at each epoch with the updated perception on the system states. We use approximate dynamic programming (ADP) (Powell, 2007) to solve the discretized optimal control problem. In particular, we use value function approximation (VFA) is utilized to help make decisions at each stage, which is obtained from iteratively performing value iteration.

To elaborate, we first define *state* as the collection of variables that fully captures the systems status at the beginning of stage k', and denote it by

$$\boldsymbol{\psi}_{k'} := \left(\boldsymbol{\theta}_{(k'-1)\Delta}, \boldsymbol{\phi}_{(k'-1)\Delta}, \mathbf{R}_{(k'-1)\Delta}, \boldsymbol{\gamma}_{(k'-1)\Delta}, \mathbf{y}_{(k'-1)\Delta}\right), \forall k' = k, \ldots, K, k = 1, \ldots, K, \quad (4.42)$$

where $\mathbf{y}_{(k'-1)\Delta}$ is the collection of $y_{i,(k'-1)\Delta}, \forall i \in \mathcal{I}$. We also denote the *action*, i.e. the collection of decision variables at stage k' by $\boldsymbol{\Omega}_{k'} = \left(\boldsymbol{\omega}_{k'}, \boldsymbol{\mu}_{k'}, \boldsymbol{\gamma}_{k'}, \tilde{\boldsymbol{\gamma}}_{k'}, \tilde{\mathbf{t}}_{k'}, \mathbf{y}_{k'}\right)$. At epoch k, for the current and future stages $k' = k, \ldots, K$, the agency makes control decisions by solving the following subproblem to achieve optimal solution to **OCP**$_k$:

$$\boldsymbol{\Omega}_{k'}^* = \operatorname*{argmin}_{\boldsymbol{\Omega}_{k'}} V_{k'+1}\left(\boldsymbol{\psi}_{k'+1}\left(\boldsymbol{\psi}_{k'}, \boldsymbol{\Omega}_{k'}\right)\right) + \sum_{t''=t'}^{k'\Delta} \sum_{i \in \mathcal{I}} \alpha_1 N_i \sum_{j \in \mathcal{I}} \frac{\partial g_i(\boldsymbol{\theta}_{i,t''-1})}{\partial x_j} \gamma_{i,t''} \gamma_{j,t''} r_{ij} \phi_{ij,t''-1}$$

$$+ \alpha_2 N_i u_{id} I_{i,t''-1} + \alpha_3 f_i\left(\gamma_{i,t''}\right) + \alpha_4 h_i\left(\mu_{i,t''} I_{i,t''-1}\right),$$

$$(4.43)$$

$$\text{s.t. } (4.9), (4.18), (4.24) - (4.27), \text{ and } (4.29) - (4.41),$$

where value function $V_{k'+1}(\boldsymbol{\psi}_{k'+1})$ is defined as:

$$
\begin{aligned}
V_{k'+1}(\boldsymbol{\psi}_{k'+1}) = \min_{\substack{\boldsymbol{\Omega}_{k''}, \\ k''=k'+1,\dots,K}} \sum_{t''=t'}^{T} \sum_{i \in \mathcal{I}} \alpha_1 N_i \sum_{j \in \mathcal{I}} \frac{\partial g_i(\boldsymbol{\theta}_{i,t''-1})}{\partial x_j} \gamma_{i,t''} \gamma_{j,t''} r_{ij} \phi_{ij,t''-1} \\
+ \alpha_2 N_i u_{id} I_{i,t''-1} + \alpha_3 f_i(\gamma_{i,t''}) + \alpha_4 h_i(\mu_{i,t''} I_{i,t''-1}), \\
\forall k' = k, \dots, K-1, k = 1, \dots K-1,
\end{aligned}
\tag{4.44}
$$

and

$$V_{K+1}(\boldsymbol{\psi}_{K+1}) = 0, \forall \boldsymbol{\psi}_{K+1}, \tag{4.45}$$

where $\boldsymbol{\psi}_{K+1}$ is the final state of the system. The value functions, by definition, should follow the backward Bellman equation:

$$
\begin{aligned}
V_{k'}(\boldsymbol{\psi}_{k'}) = \min_{\boldsymbol{\Omega}_{k'}} \sum_{t''=t'}^{k'\Delta} \sum_{i \in \mathcal{I}} \left[\alpha_1 N_i \sum_{j \in \mathcal{I}} \frac{\partial g_i(\boldsymbol{\theta}_{i,t''-1})}{\partial x_j} \gamma_{i,t''} \gamma_{j,t''} r_{ij} \phi_{ij,t''-1} \right. \\
\left. + \alpha_2 N_i u_{id} I_{i,t''-1} + \alpha_3 f_i(\gamma_{i,t''}) + \alpha_4 h_i(\mu_{i,t''} I_{i,t''-1}) \right] \\
+ V_{k'+1}(\boldsymbol{\psi}_{k'+1}(\boldsymbol{\psi}_{k'}, \boldsymbol{\Omega}_{k'})), \forall k' = k, \dots, K, k = 1, \dots K,
\end{aligned}
\tag{4.46}
$$

In conventional dynamic programming approach, to obtain the value function requires solving the Bellman equation (4.46) for all possible states, which is computationally intractable for high dimension continuous state spaces. We use value function approximation (VFA) techniques to address such an issue. To be specific, we use an approximation to $V_{k'}(\boldsymbol{\psi}_{k'})$, denoted as $\hat{V}_{k'}(\boldsymbol{\psi}_{k'})$, to help solve the subproblems (4.43). By iteratively taking sample paths and evaluating the "value" of sample states at each stage, we try to obtain an increasingly accurate VFA, which known as *value iteration*. We propose the following approach for value iteration, and use feed-forward neural network (FNN) to approximate the value functions, for its theoretical applicability to fit arbitrary shapes of functions. First, we denote the number of hidden nodes in the FNN as N_0, and let s be a parameter to control the batch size of the sample paths in each iteration.

To obtain a good approximation to the value functions stills require a large number of sample paths in each iteration, which is due to the intrinsic complexity of the problem. However, while generating sample paths, random feasible solutions can be easily obtained, which significantly reduces the time on optimization using likely poorly trained VFAs as in conventional value iteration approaches. Moreover, each sample path do not interfere with the others, which enables parallel programing to further reduce the computation time.

Algorithm 1 Approximate dynamic programing

Step 1.1 Initialize the training set $\mathcal{V}_k = \emptyset, \forall k = 1, \ldots, K$, set $\hat{V}_{K+1} = 0$;

Step 1.2 For $k = K, \ldots, 2$, do:

- Let $N_k = N_0 s(K - k + 1)$, and generate a set of N_k different initial conditions $\mathcal{E}_k = \{\mathbf{e}_l\}_{l=1,\ldots,N_k}$;
- For $l = 1, \ldots, N_k$, do:
  - Let $\boldsymbol{\psi}_0 = \{\mathbf{1}, \mathbf{e}_l, \mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{P}_q\}$, let $k' = 1$;
  - While $k' \leq k - 1$: i) generate a random feasible solution $\boldsymbol{\Omega}_{k'}$, ii) compute $\boldsymbol{\psi}_{k'+1} = \boldsymbol{\psi}_{k'+1}(\boldsymbol{\psi}_{k'}, \boldsymbol{\Omega}_{k'})$; iii) let $k' \leftarrow k' + 1$;
  - While $k' \leq K$: i) solve Bellman equation (4.46) with the approximated value function $\hat{V}_{k'+1}$, to obtain the objective $V_{k'}(\boldsymbol{\psi}_{k'})$ and the corresponding new system state $\boldsymbol{\psi}_{k'+1}(\boldsymbol{\psi}_{k'}, \boldsymbol{\Omega}_{k'}^*)$; ii) let $\mathcal{V}_k \leftarrow \mathcal{V}_k \cup \{(V_{k'}(\boldsymbol{\psi}_{k'}), \boldsymbol{\psi}_{k'})\}$; iii) let $k' \leftarrow k' + 1$;
- For $k' = k, \ldots, K$, (re)train the FNN using $\mathcal{V}_k$ to obtain $\hat{V}_{k'}$.

4.4 Numerical Experiment

In this section, we present the numerical results obtained from a semi-realistic case study. We first validate the applicability of the latent model, and then demonstrate the results of the optimization framework. Moreover, through extensive numerical experiments, we also examine the sensitivity of the results with respect to the model parameters and input data.

4.4.1 Case study

We consider a college town whose population is composed of mostly faculties and students, i.e., $\mathcal{I} = \{s, f\}$, and a hypothetical disease which spreads quickly and diminishes relatively slowly. Moreover, we assume a students/faculties ratio of 20, which is close to that of a public university, and the population in each group is $N_s = 20000$ and $N_f = 1000$, respectively. Based on existing researches including Kossinets and Watts (2006) and Srinivasan et al. (2006), real-life contact networks are more suitably described by degree distributions without heavy tailed property. Poisson distribution is chosen as the degree distribution function for that it is not heavy-tailed and its wide applications in the existing literature (Newman et al., 2001; Barthélemy et al., 2005; Volz, 2008). We assume the following parameters for the mean degrees: $\lambda_{s,s} = 7.5$, $\lambda_{s,f} = 0.6$, and $\lambda_{f,f} = 4.5$, where $\lambda_{i,j}$ represents the average degrees of a group i node connecting to group j. Naturally, we have $\lambda_{f,s} = 12$ due to degree conservation. The overall average degrees in the network is 8.5. The disease parameters are given as follows: $r_{s,s} = 1.5 \times 10^{-2}$ per day, $r_{s,f} = 5 \times 10^{-3}$ per day, $r_{f,s} = 5 \times 10^{-3}$ per day, and $r_{f,f} = 8 \times 10^{-3}$ per day; moreover, $u_i = 8 \times 10^{-3}$ per day with

$u_{i,d} = 5 \times 10^{-3}$ and $u_{i,c} = 3 \times 10^{-3}$ for $i \in \mathcal{I}$. Therefore, the students have higher connectivities among themselves compared to that among faculties, and they are much easier to be infected by other students. However, an infected faculty may infect more people than an infected student.

We consider 1 day as a period and $\Delta = 14$ days as an epoch, and $K = 6$ epochs in total. Therefore, the control horizon is $T = 84$ days. For the hypothetical epidemic with the presented parameters, such a control horizon is appropriate. The observations on the infected fraction are accumulated over each period to feed into the latent model.

The other parameters in the objective and in the solution algorithm are given as follows: $\eta_i = 0.04, \underline{\gamma}_i = 0.01, \forall i \in \mathcal{I}$; the action intensity-detection rate matrix A is a 2×2 identity matrix, i.e., there are two independent screening actions, each of which quarantines one group. We consider one resource type, and the cost for the screening actions $c_{q,s} = 20$ and $c_{q,f} = 3$ units of resource per unit intensity for the two groups, respectively, with a total resource $B_{q,t} = 0.08$ units, $\forall t$. Furthermore, the objective weights are $\alpha_1 = \$1.0$, $\alpha_2 = \$10.0$, $\alpha_3 = \$2.0$, and $\alpha_4 = \$0.5$. The impact penalty function parameters are given as follows: $a_s = 1000, a_f = 50$, and $b_i = 150, c_i = 300, d_i = 150, \forall i \in \mathcal{I}$. Finally, the batch size control parameter $s = 50$.

In the latent model, the mean of the prior distribution, namely $\bar{\mathbf{e}}$, is randomly generated from a uniform distribution with mean $2.5 \cdot 10^{-3}$. Each element of the variance matrix $\Sigma_{\mathbf{e}}^2$ is as well randomly generated from another uniform distribution with mean 0.05. The ground-truth value of the initial condition $\mathbf{e}$ is given as follows: $\phi_{s,s}(0) = 1 \cdot 10^{-3}$, $\phi_{s,f}(0) = 2 \cdot 10^{-3}$, $\phi_{f,s}(0) = 1 \cdot 10^{-3}$, and $\phi_{f,f}(0) = 0.5 \cdot 10^{-3}$. Moreover, we assume the observations are unbiased, such that the formulation of $\mathbf{SEP2}_t$ can be adopted in the latent model.

The results of the natural disease process is presented in Figure 4.2. By the end of the 84-days control horizon, the susceptible fraction of each group decreases to 0.3430 for students and 0.3834 for faculties, respectively, and a total accumulated infected population of 13,756 people. Moreover, there are death tolls of 1097 and 50 people among the students and faculties, respectively. It leads to a total cost of $25,226 in total, due to both infections and deaths. Therefore, without proper infectious disease control, the result of the epidemic can be devastating.

The major computation burden is on obtaining the training data for VFAs, which requires solving a large amount of nonlinear optimization problems. With the aforementioned parameter settings, it takes approximately 1 hour on a 4-core CPU (i7-4800MQ CPU @ 2.70 GHz) laptop with parallel computing to obtain the approximated value functions. Once the VFA is completed, solving the latent model and the optimal control problem following the framework as shown in Figure 4.1 is rather efficient and can be finished within 5 minutes.

After the VFAs are obtained, we run a total of 50 realizations of the control scheme to validate the latent model as well as the control framework. In each simulation, a different prior distribution of the initial condition is generated. Figure 4.3 demonstrates the good performance of the latent
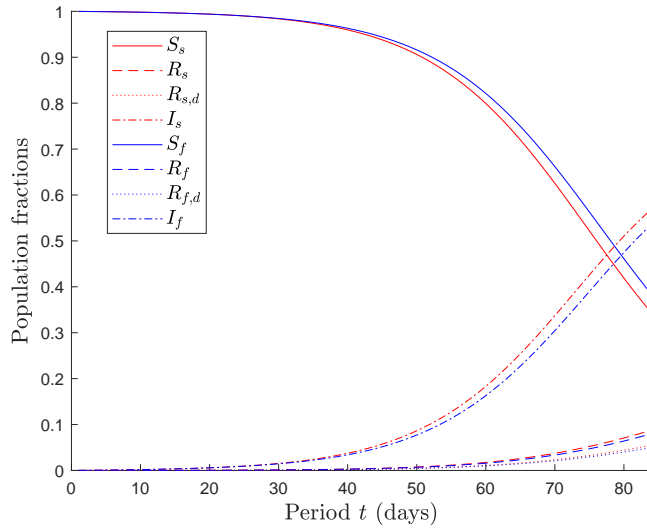
Figure 4.2: The epidemic dynamics of the system without disease control measures.

model. The relative error is computed as $\frac{|\hat{\mathbf{e}}-\mathbf{e}|}{|\mathbf{e}|} \times 100\%$. Within 10 days, The relative error on the initial condition is reduced to below 50% in 60% of the cases. Even if the prior distribution falls far away from the ground-truth value, the estimation error is always reduced to below 95% with the first few observations. Moreover, after 33 periods, the estimated initial conditions converge to the ground-truth value within 1% error in over 90% realizations, which shows agreement with our earlier conjecture, as suggested by Figure 4.3b.
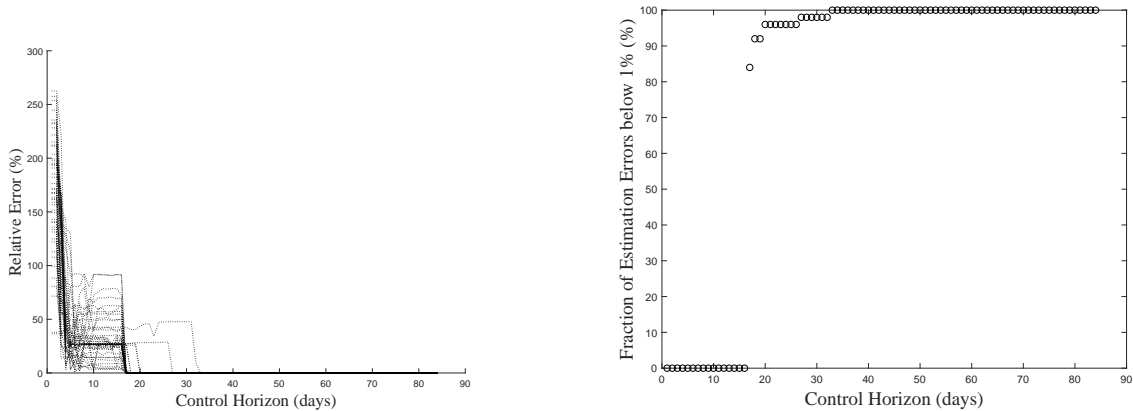
(a) Estimation error of the latent model with accumulating observations.

(b) Fraction of simulations with errors reduced to below 1%.

Figure 4.3: Convergence results of the latent model.

As a result, the average objective is reduced to $13,634 across the 50 simulations, with average infections of 4,054 incidents, and a total death of 262 incidents. The objective is reduced by

46.0% compared to the case without the control actions, with 70.5% less infections and 77.2% less mortality counts. Therefore, the disease propagation is significantly mitigated with the applied control actions. Moreover, the average cost due to the impacts caused by the control actions are $217 and $6,741, respectively, which compose a significant portion of the total cost.

The quarantine resource allocation to each group in each decision epoch is presented in Figure 4.4. The usage of the quarantine resource reduces over decision epochs (except for epoch 1, which is likely due to biased perception of the system state), and it stops screening the population in the last two epochs. The reasons are two-folds: i) screening the population at an earlier stage is more effective, because high-degree susceptible individuals would be infected sooner than the others and then infect more of the others, and timely quarantine can better preserve these high-degree individuals than delayed quarantine; ii) the infected population have grown to a significant amount in the later epochs such that quarantine the infected leads to a remarkably higher cost. Moreover, most of the quarantine efforts are spent on the students group. This might look counter-intuitive at the first glance, as the faculties have a high average degree. It is because not only that students require a lower screening cost per capita (i.e. $c_{q,s}/N_s < c_{q,f}/N_f$), but also that disease transmission among students are more frequent (high average degree) and faster (high unit infection rate).
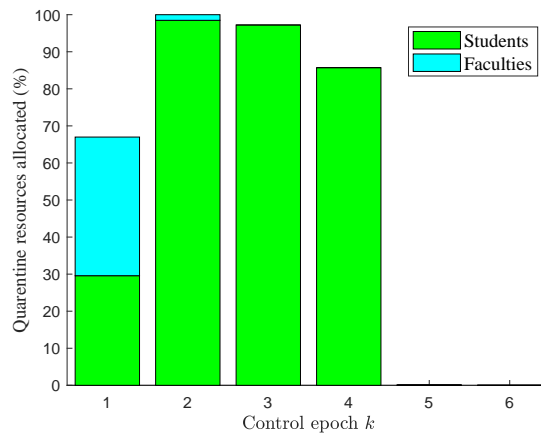
Figure 4.4: Total quarantine resources allocated to each group over control epochs, averaged over 50 realizations.

Moreover, the health advisories are usually posted at a relatively early stage of the control horizon. For example, in 21 of the simulations, the advisories to the students are posted at the beginning of the first epoch, i.e., $\tilde{t}_s = 1$ day, with an average intensity of $\tilde{\gamma}_s = 0.5012$; in another 18 simulations, $\tilde{t}_s = 15$, and $\tilde{\gamma}_s = 0.4067$, and in the rest of the simulations the advisory is posted to students at the third epoch with $\tilde{\gamma}_s = 0.3965$. These results are intuitive: the later the advisory is posted, the higher intensity it requires to achieve the same effect in mitigating the disease propagation, and the same phenomenon is observed on the faculty group as well. Furthermore, for the

faculty group, the advisory is almost always posted no earlier than that to the students, i.e. $\tilde{t}_s \leq \tilde{t}_f$ in 49 simulations, and it is strictly later in 20 of these simulations. The average advisory posting times are at day 12.2 and at day 18.6, with mean intensity of 0.4442 and 0.4536 to students and to faculties, respectively. The slight delay of advisories to faculties is interesting but also intuitive: as the effect of advisories would be diminishing over time, a delayed advisory renews the effect of the waned control and might eventually outperform than posting all advisories simultaneously. Finally, for the same reason that quarantine resources are mostly allocated to students, the health advisories are posted to them prior to that to faculties as well.

4.4.2 Sensitivity analysis

We perform several expriments to test the sensitivity of the results, validate the robustness of the framework, and to reveal operational insights on these two control strategies. We change the diminishing rate of the epidemic warning η_i, the objective weighing parameters α_m, the resource budget for quarantine $B_{q,t}$, the connectivity of the network λ_{ij}, as well as the disease transmission and death rates r_{ij}'s and $u_{i,d}$'s, respectively, and present the objective and operation decisions in these cases. We denote the benchmark case in the previous subsection as Case 0, and compare the results of the different cases under various settings. The results are averaged over 50 realizations in each case, and are summarized in Table.4.1.

Table 4.1: Sensitivity analysis results of optimal epidemic control with quarantine and health advisories.

Cases	Advisory decisions				Quarantine resource		Costs ($\times 10^3$\$)				
	Posting time t_i (day)		Posting intensity $\tilde{\gamma}_i$		allocation (%)		Infection	Death toll	Quarantine	Advisory	Total
	Students	Faculties	Students	Faculties	Students	Faculties	cost	cost	impact	impact	
0. (benchmark)	12.2	18.6	0.44	0.45	51.8	6.5	4.05	2.62	0.22	6.74	13.63
1. ($\eta_i = 0.02$)	4.9	49.2	0.48	0.71	28.9	7.5	2.08	1.44	0.21	10.81	14.54
2. ($\eta_i = 0.06$)	9.1	13.0	0.31	0.25	44.7	10.6	5.02	3.00	0.31	6.27	14.60
3. (α_3, α_4 reduces by 50%)	4.6	7.7	0.22	0.22	69.8	10.8	2.21	1.33	0.16	5.50	9.21
4. (α_3, α_4 increases by 50%)	6.3	25.9	0.59	0.33	33.0	16.3	5.89	3.92	0.78	7.60	18.18
5. ($\lambda_{s,s} = 3.0$)	-	33.5	-	0.86	27.0	2.5	0.44	0.46	0.03	0.08	1.00
6. ($\lambda_{s,s} = 12.0$)	13.0	22.5	0.09	0.35	56.7	7.2	12.92	7.83	1.95	13.41	36.11
7. ($B_{q,t} = 0.04$)	5.8	6.0	0.46	0.24	66.0	4.7	4.21	2.61	0.22	6.76	13.79
8. ($B_{q,t} = 0.12$)	6.0	6.0	0.41	0.28	61.7	4.6	3.29	2.02	0.47	7.61	13.39
9. (r_{ij} reduces by 50%)	-	-	-	-	52.8	4.7	0.75	0.69	0.10	0.00	1.54
10. (r_{ij} increases by 50%)	11.9	22.6	0.12	0.26	61.3	15.6	9.93	6.04	2.11	13.15	31.23
11. ($u_{i,d}$ reduces by 50%)	14.4	25.9	0.48	0.39	50.6	14.4	4.90	1.56	0.30	5.93	12.69
12. ($u_{i,d}$ increases by 50%)	5.2	13.9	0.39	0.31	47.0	6.0	2.96	2.65	0.19	7.70	13.50

The outcome of all these cases fall within our expectations, and the decisions yields some interesting insights on the model. Comparing the benchmark case with the Cases 1 and 2, it is observed that the diminishing rate of the health advisory effect η_i has a complex impact on the objective. While a lower diminishing rate (Case 1) prevents more infections and further mitigates the disease propagation, the detrimental impact caused by the control measure is also non-negligible. There-

fore, the objective could be higher with either a lower or a higher diminishing rate η_i. Moreover, the lower diminishing rate the longer time interval between the two advisory posts, which is as well expected as the effect of the early post lasts longer.

Cases 3 and 4 evaluate the impacts of the objective weights. The result turns out to be quite sensitive to these parameters. With a lower weight of the control approach penalties, the decision-maker can adopt stronger controls: in Case 3, the advisories are posted earlier than the benchmark case at a higher intensity level; in Case 4, the advisory posting decisions are more complex, but the cost of advisory impact indicates that these advisories are less intense than the benchmark case ($6.74 \times 150\% > 7.60$). Moreover, the quarantine resource usage is monotonically decreasing with the weight factors α_3.

When the average degree among the students is changed by ± 4.5, respectively, the control decisions and the results are completely different. In Case 5, both the impacts of disease propagation and controls have been significantly mitigated because of the reduced network connectivity; the result is on the contrary in Case 6, where not only the impacts of disease propagation is drastically increased, but in order to preserve the population from an even worse situation, stronger control actions are adopted which leads to higher penalties as well. Moreover, the impact of quarantine resource amount is negligible on both the control decisions as well as the objective, which is likely due to that the resource level is still too low to cause significant changes.

Furthermore, the disease transmission rate has a significant impact on the results as shown by Cases 9 and 10, similar to that of the connectivity among students, which is as expected. However, Cases 11 and 12 suggests that the effect of the death rate is more complex. For one thing higher death rate leads to higher mortalities, and on the other hand it also mitigates the disease propagation process such that the infected number is reduced. As a result of these mutually canceling effects, the total objective is not significantly changed.

As a final observation, we find that the first health advisory is always posted to the students group at a relatively early stage of the epidemic, while that to the faculties could be delayed and the delay depends on the diminishing rate of the effect of the first advisory. Moreover, the students are always allocated with more quarantine resources in these experiments.

CHAPTER 5 – DYNAMIC VACCINATION GAME

5.1 Introduction

While the human society are freed from the fear of many infectious diseases in the modern world, largely thanks to the invention of vaccines, people start to cast vaccines aside as an option rather than a necessity (The College of Physicians of Philadelphia, 2018). A handful of reports show that due to various reasons, many people may choose to avoid vaccination during an epidemic event (Sun, 2018; Thomas, 2018; Calandrillo, 2003; Myers and Goodwin, 2011; Chor et al., 2009; Tozzi et al., 2009). Therefore, "vaccine-phobia" is not merely a subculture among a certain group of people, but rather a serious issue that demands attention in the public health industry. During an epidemic outbreak, each individual makes its own vaccination decision to minimize its expected disutility from both vaccine-phobia and the risk of infection. Such a problem is known as a vaccination game, as people's vaccination decisions not only affect their own disutilities but those of all others through probabilistic disease transmissions. It is important to study the complex behavior of individuals in a vaccination game so as to understand the development and consequence of an epidemic outbreak, and in turn to reveal insights on epidemic controls.

This chapter investigates the autonomous vaccination behavior of a heterogeneous mixing population. To stay focused, we consider population heterogeneity in terms of a person's number of contacts, while every individual shares the same level of vaccine-phobia. The rest of this chapter is organized as follows. Section 5.2 presents the developed SIRVA (susceptible-infected-removed-vaccinated-activated) dynamics model and the analytical results on the final epidemic size; some properties of the vaccination game equilibrium are also presented, based on which a heuristic solution approach to solve for the equilibrium solution is proposed. Next, Section 5.3 discusses several different vaccination schemes, including delayed and non-differential vaccinations. Section 5.4.2 presents the case study and the sensitivity analysis, and discusses the insights revealed by the experiments.

5.2 Methodology

5.2.1 Vaccination dynamics

We consider a dynamic vaccination game in a large population with heterogeneous degree distributions, where the natural disease propagation follows an SIR process. In order to accommodate the vaccination process, we further define two compartments based on individuals' health status: vaccinated (V) and activated (A). A susceptible node becomes vaccinated once being injected with a vaccine and thus obtaining immunity. When a vaccinated node receives an infectious contact, instead of becoming infected (like a susceptible node), it activates its immunity and thus become activated. The transition among the different health states are summarized below:

- A susceptible node becomes infected upon receiving an infectious contact from one of its neighbors;
- An infected individual deceases or recovers by itself (with immunity) following the natural removal rate;
- A susceptible node becomes vaccinated at a certain rate (based on its own decision process);
- A vaccinated node becomes activated upon receiving an infectious contact from one of its neighbors.[a]

We now use $S(t), I(t), R(t), V(t), A(t)$ to denote the fraction of susceptible, infected, removed, vaccinated, and activated individuals at time t, respectively.[b] The size of population is fixed, such that the equivalent of (2.7) now becomes:

$$S(t) + I(t) + R(t) + V(t) + A(t) = 1, \forall t \geq 0. \tag{5.1}$$

Furthermore, we differentiate nodes by their degrees; denote X_k as the fraction of degree-k nodes in compartment $X \in \{S, I, R, V, A\}$. Naturally, we have

$$\sum_{k=1}^{K} X_k(t) = X(t), \forall t \geq 0, X \in \{S, I, R, V, A\}, \tag{5.2a}$$

$$\sum_{X \in \{S,I,R,V,A\}} X_k(t) = p_k, \forall t \geq 0, k = 1, \ldots, K. \tag{5.2b}$$

where $K \geq 0$ is the maximum possible degree of a node.

In the vaccination game, we consider the Nash game; i.e., each individual tries to minimize

---

[a]We assume perfect vaccination without vaccine failures, such that vaccine uptake guarantees immunity. This assumption can be easily relaxed in the dynamics model, but in order to stay focused, we will not incorporate the risk of vaccine failures in this dissertation.

[b]It might appear cumbersome to differentiate the activated from the vaccinated as they both represent the same immunized status. However, it is quite necessary to do so, because the model described by (2.3) - (2.7) focuses on the status of edges (i.e., whether or not they have passed infectious contacts), and hence the status of the nodes can only be derived from the edge status.

its own expected penalty by deciding its vaccination rate, given others' vaccination decisions. We use $v_k(t)$ to denote the average vaccination rate of degree-k nodes at time t. The expected penalty includes the psychological penalty from vaccine-phobia (if vaccinated), and the penalty of eventually getting infected (if not vaccinated).

For notational brevity, in the rest of this section we omit the time argument t for all time dependent variables such as θ, ϕ, v_k, S_k, I_k, R_k, V_k and A_k unless otherwise specified. Now we study the status of a degree-k node who has not received any infectious contacts from its neighbors. The fraction of such nodes is given by $p_k\theta^k$, and these nodes can either be susceptible or vaccinated, which yields the following equation:

$$S_k + V_k = p_k\theta^k, \forall t \geq 0, k = 1, \ldots, K. \tag{5.3}$$

The dynamics of S_k contain two parts: the vaccinations and the infections of the degree-k population, respectively. The former is simply $-v_kS_k$. Moreover, the test node being vaccinated or not does not affect the infectious contacts it receives, and thus the node being vaccinated or susceptible is random (as long as there is no infectious contact yet). Therefore, the latter is the rate of reduction among those who have not received infectious contacts, $d\left(p_k\theta^k\right)/dt$, times the fraction of susceptible nodes, $S_k/p_k\theta^k$. Furthermore, the dynamics of θ still follow (2.3) by definition. As such, we have

$$\dot{S}_k = -v_kS_k + \frac{d\left(p_k\theta^k\right)}{dt}\frac{S_k}{p_k\theta^k} = \left(-v_k - kr\frac{\phi}{\theta}\right)S_k, \forall t \geq 0, k = 1, \ldots, K. \tag{5.4}$$

The dynamics of ϕ, i.e. (2.4), should be slightly modified. Recall its dynamics include three parts: (i) transmission of disease along edges to the test node, (ii) removal of the infected neighbors, and iii) new infections of the susceptible neighbors. Vaccination of the susceptible does not affect the already infected nodes nor occurrence of infectious contacts, thus the first two terms remain unchanged. The third term is essentially the rate at which a susceptible neighbor of the test node becoming infected, i.e., infectious contacts happening to the neighbor while the neighbor is not vaccinated. The probability that a random neighbor of the test node is a degree-k node who has not received any infectious contacts from others is $kp_k\theta^{k-1}/g'(1)$. The rate at which infectious contacts happen to this neighbor is then $d\left(-kp_k\theta^{k-1}/g'(1)\right)/dt$. Using the same argument that leads to (5.4), the rate of the neighbor being infected is $-\left(S_k/p_k\theta^k\right) \cdot d\left(kp_k\theta^{k-1}/g'(1)\right)/dt$. Therefore, (2.4) becomes

$$\dot{\phi} = \left(-r - u + \frac{r}{g'(1)\theta^2}\sum_{k=1}^{K}S_kk(k-1)\right)\phi.\forall t \geq 0. \tag{5.5}$$

Finally, a degree-k vaccinated node becomes activated at the same rate as a degree-k susceptible

node becomes infected. Therefore,

$$\dot{A}_k = kr\frac{\phi}{\theta}V_k, \forall t \geq 0, k = 1, \ldots, K, \tag{5.6}$$

To this point, we have formulated the complete system dynamics with (2.3), (2.5), and (5.1)-(5.6). One can evaluate the impacts of the population's vaccination decisions by iteratively computing the dynamic equations over time from the beginning of the epidemic to the end. This is computationally burdensome for a large population and cannot easily reveal insights on the vaccination game. In light of this, we further assume that people's vaccination strategies v_k follow a special form, such that the outcome of the epidemics can be derived analytically in closed form. This will be the focus of the remainder of this paper.

5.2.2 Final epidemic size

In reality, people often determine their vaccination rate based on the current propagation rate of the disease. Therefore, we consider a vaccination strategy in which the vaccination rate is proportional to the fraction of infectious contacts that is going to happen. From now on, we use an explicit argument t for all the time-dependent variables so as to distinguish them from time-invariant parameters. Then the vaccination strategy is presented as follows:

$$v_k = \mu_k r\frac{\phi}{\theta}, \forall t \geq 0, k = 1, \ldots, K, \tag{5.7}$$

where $\mu_k \geq 0$ is a constant parameter determined by the degree-k population. It reflects the likelihood of the population taking vaccines, and we call it *vaccine adoption level*.[c] The higher vaccine adoption level μ_k among the population, the lower infection risk but a greater psychological suffering (due to the vaccine-phobia).

With (5.7), we have the following equation:

$$S_k = p_k\theta^{k+\mu_k}, \forall t \geq 0. \tag{5.8}$$

Readers can easily verify this by taking derivative over (5.8) and comparing it with (5.4).

Then we use the same approach as in Miller (2011) to obtain the final epidemic size. First, we denote $\theta(\infty)$ as θ_∞, and the total immunized fraction as M, i.e., $M = V(\infty) + A(\infty) = $

---

[c]Here μ_k represents an averaged vaccine adoption level to capture the overall vaccination rate of the degree-k population. Yet in fact each individual could make its own decision and their vaccine adoption levels may differ. However, as we will show later, individual decisions should be equal under equilibrium for nodes with the same degree. Thus it is safe to use an aggregated rate in the following derivations.

$\sum_k \int_0^\infty v_k(\tau) S_k(\tau) d\tau$. With (5.7) and (5.8), we have

$$M = \sum_{k=1}^{K} \frac{\mu_k p_k}{k + \mu_k} \left(1 - \theta_\infty^{k+\mu_k} \right).$$

(5.9)

By the end of the epidemic, the degree-k uninfected population is the summation of susceptible, vaccinated, and activated nodes:

$$S_k(\infty) + V_k(\infty) + A_k(\infty) = \frac{p_k k \theta_\infty^{k+\mu_k} + \mu_k p_k}{k + \mu_k}, \forall k = 1, \ldots, K.$$

(5.10)

Moreover, we let π be the probability that a random neighbor of the test node has never been infected by the end of the epidemic, given that the test node does not transmit the disease. This could happen under two mutually exclusive events: (i) no infectious contact has ever happened to this neighbor, and (ii) this neighbor has already been vaccinated when an infectious contact arrives, i.e. this neighbor is activated. The first probability is given by $g'(\theta_\infty)/g'(1)$. The second can be computed via integrating $\sum_{k=2}^{K} \frac{V_k}{p_k \theta^k} d \left(-\frac{p_k k \theta^{k-1}}{g'(1)} \right) /dt$ from $t = 0$ to ∞. The result is presented as follows:

$$\pi = \frac{1}{g'(1)} \left(g'(\theta_\infty) + \sum_{k=2}^{K} p_k k \left(\frac{\mu_k}{k + \mu_k - 1} - \theta_\infty^{k-1} + \frac{k-1}{k + \mu_k - 1} \theta_\infty^{k+\mu_k-1} \right) \right).$$

(5.11)

We further denote the *transmissibility* of disease as $T = \frac{r}{r+u}$. It gives the probability that, conditional on the fact that one end of a random edge is infected, an infectious contact happens along this edge before that infected end is removed (recall, that both events follow Poisson processes). The probability that a random edge of the test node has never transmitted an infectious contact is then $1 - T(1 - \pi)$, which, from (5.11), yields the following nonlinear equation:

$$\theta_\infty = 1 - T + \frac{T}{g'(1)} \left(g'(\theta_\infty) + \sum_{k=2}^{K} p_k k \left(\frac{\mu_k}{k + \mu_k - 1} - \theta_\infty^{k-1} + \frac{k-1}{k + \mu_k - 1} \theta_\infty^{k+\mu_k-1} \right) \right).$$

(5.12)

The solution can be easily found by searching from 1 to 0 with a small step size. Then, the final uninfected population, from (5.10), as well as the final epidemic size, $R(\infty) = 1 - (S(\infty) + V(\infty) + A(\infty))$, can be computed accordingly. Moreover, the fraction of nodes in other compartments in

the final state can be computed as follows:

$$A_k(\infty) = p_k \left(\frac{\mu_k}{k + \mu_k} - \theta_\infty^k + \frac{k}{k + \mu_k} \theta_\infty^{k+\mu_k} \right), \forall k = 1, \ldots, K, \qquad (5.13)$$

$$I_k(\infty) = 0, \forall k = 1, \ldots, K, \qquad (5.14)$$

$$R_k(\infty) = p_k \left(1 - \frac{k\theta_\infty^{k+\mu_k} + \mu_k}{k + \mu_k} \right), \forall k = 1, \ldots, K. \qquad (5.15)$$

5.2.3 Vaccination game

We are now ready to formulate the vaccination game that determines μ_k for each degree-k subpopulation. We denote $\boldsymbol{\mu} = \{\mu_k\}_{\forall k}$ as the overall vaccine adoption level across the population. Since the vaccine guarantees perfect immunity, the probability of a degree-k individual i eventually getting infected is the probability that it is infected conditional on it is not vaccinated (NV) times the probability that it is not vaccinated. Given its own vaccine adoption level, denoted μ_k^{i+}, and all others' decisions, denoted $\boldsymbol{\mu}^{i-} = \{\mu_{k'}\}_{\forall k' \neq k} \bigcup \{\mu_k^{i-}\}$, the infection probability of i is[d]:

$$Pr \left\{ \text{Infected} \, \big| \mu_k^{i+}, \boldsymbol{\mu}^{i-}, k \right\}$$
$$= Pr \left\{ \text{Infected} \, \big| \text{NV}, \mu_k^{i+}, \boldsymbol{\mu}^{i-}, k \right\} Pr \left\{ \text{NV} \big| \mu_k^{i+}, \boldsymbol{\mu}^{i-}, k \right\}$$
$$= Pr \left\{ \text{Infected} \, \big| \text{NV}, \boldsymbol{\mu}^{i-}, k \right\} Pr \left\{ \text{NV} \big| \mu_k^{i+}, \boldsymbol{\mu}^{i-}, k \right\}, \forall k = 1, \ldots, K.$$

We assume $f(\cdot)$ is a function that captures the psychological penalty caused by vaccine-phobia. Given people's vaccine-phobia level, the penalty function f is identical for every individual, strictly increasing w.r.t. μ_k, and satisfying $f(0) = 0$. We assume that the total expected penalty (or disutility) for a degree-k node i is a linear combination of its own psychological penalty and infection risk, as follows:

$$E_k^i \left[\mu_k^{i+}, \boldsymbol{\mu}^{i-} \right] = \alpha_1 f(\mu_k^{i+}) + \alpha_2 Pr \left\{ \text{Infected} \, \big| \mu_k^{i+}, \boldsymbol{\mu}^{i-}, k \right\}, \forall k = 1, \ldots, K, \qquad (5.16)$$

where $\alpha_1, \alpha_2 \geq 0$ are weights.

In a large population, the vaccination decision of one individual does not significantly affect the overall vaccination coverage, i.e., $\boldsymbol{\mu}^{i-}$ is approximately the same for all i's, and hence approximated by $\boldsymbol{\mu}$. The expected penalty of i can be written as:

$$E_k^i \left[\mu_k^{i+}, \boldsymbol{\mu}^{i-} \right]$$
$$\approx \alpha_1 f(\mu_k^{i+}) + \alpha_2 Pr \left\{ \text{Infected} \, \big| \text{NV}, \boldsymbol{\mu}, k \right\} Pr \left\{ \text{NV} \big| \mu_k^{i+}, \boldsymbol{\mu}, k \right\}, \forall k = 1, \ldots, K. \qquad (5.17)$$

---

[d]Again, we adopt the perfect vaccination assumption here, such that only the unvaccinated individuals could be infected. The infection probability should be derived differently if there is non-negligible risk of vaccine failures.

Since all not-vaccinated degree-k nodes are probabilistically identical, the probability for one of them to get infected shall be simply the following ratio:

$$Pr\left\{\text{Infected} \mid \text{NV}, \boldsymbol{\mu}, k\right\} = \frac{R_k(\infty, \boldsymbol{\mu})}{R_k(\infty, \boldsymbol{\mu}) + S_k(\infty, \boldsymbol{\mu})}, \forall k = 1, \ldots, K. \tag{5.18}$$

Moreover, we need to find out the probability of i not getting vaccinated, $Pr\left\{\text{NV} \mid \mu_k^{i+}, \boldsymbol{\mu}, k\right\}$. We first denote the probability of node i being in status X at time t as $X_k^i(t)$, $\forall X \in \{S, I, R, V, A\}$. Conditional on that i is susceptible at time t, the vaccination rate is v_k^{i+}, and the infection rate is the same with all other degree-k susceptible nodes. Therefore, the dynamics of S_k^{i+} follow a similar form as (5.4):

$$\dot{S}_k^i = \left(-v_k^{i+} - kr\frac{\phi}{\theta}\right) S_k^i, \forall t \geq 0, k = 1, \ldots, K. \tag{5.19}$$

We easily obtain $S_k^i(t) = \theta^{k+\mu_k^{i+}}$ as a solution. The probability that i eventually being vaccinated or activated is $\int_0^\infty v_k^{i+} S_k^i dt$, which leads to the following equation:

$$Pr\left\{\text{NV} \mid \mu_k^{i+}, \boldsymbol{\mu}, k\right\} = \frac{k + \mu_k^{i+}\left(\theta_\infty\left(\boldsymbol{\mu}\right)\right)^{k+\mu_k^{i+}}}{k + \mu_k^{i+}}, \forall \mu_k^{i+}, \boldsymbol{\mu} \geq 0, k = 1, \ldots, K. \tag{5.20}$$

(5.20) is a strictly decreasing and strongly convex function with respect to μ_k^{i+} for any $\theta_\infty \in (0, 1)$, which can be verified from its first and second order derivatives. If $f(\mu_k^{i+})$ is convex with respect to μ_k^{i+}, (5.17) is also strongly convex with respect to μ_k^{i+} for any $\alpha_1, \alpha_2 > 0$. In this case, there exists a unique μ_k^{i+*} that minimizes the expected penalty for i, and it is the *best response* of individual i given all others' decisions:

$$B_i\left(\boldsymbol{\mu}^{i-}\right) = \mu_k^{i+*} = \underset{\mu_k^{i+}}{\arg\min} E_k^i\left[\mu_k^{i+}, \boldsymbol{\mu}\right], \forall k = 1, \ldots, K. \tag{5.21}$$

These results naturally lead to the following proposition, which highlights the Nash Equilibrium among the subpopulation with the same degree.

Proposition 4. *For any fixed average vaccination rates of all other subpopulations, $\mu_{k'} r\phi/\theta, \forall k' \neq k$, and $\forall \alpha_1, \alpha_2 > 0$, if the penalty function for vaccination $f(x)$ is convex, the vaccination rates among all degree-k nodes should be equal at Nash Equilibrium; i.e., $B_i\left(\boldsymbol{\mu}^{i-}\right) = B_j\left(\boldsymbol{\mu}^{j-}\right)$ if i and j both have degree $k, \forall k = 1, \ldots, K$.*

Proposition 4 states that the optimal vaccination rate of nodes with the same degree should be equal under equilibrium. This is as well intuitive because nodes with the same degree are probabilistically identical. The following proposition further speaks to the existence and uniqueness of such a solution.

Proposition 5. *(Sufficient condition) Suppose $f(x)$ is a convex function and Proposition 4 holds. The solution to following equations:*

$$\alpha_1 \frac{\partial f(\mu_k^*)}{\partial \mu_k} + \alpha_2 \frac{R_k(\infty, \boldsymbol{\mu}^*)}{R_k(\infty, \boldsymbol{\mu}^*) + S_k(\infty, \boldsymbol{\mu}^*)} \frac{\partial}{\partial \mu_k} \left(\frac{k + \mu_k^* \theta_\infty^{k+\mu_k^*}}{k + \mu_k^*} \right) = 0, k = 1, \ldots, K, \quad (5.22)$$

$$\boldsymbol{\mu}^* \geq 0, \quad (5.23)$$

denoted $\boldsymbol{\mu}^* = \{\mu_k^*\}_{\forall k}$, exists and characterizes the vaccination decisions among the entire population under Nash Equilibrium. Moreover, if $f'(0) = 0$ and f is strongly convex, there always exists a unique $\boldsymbol{\mu}^*$ that satisfies (5.22) and (5.23).

Proof. (5.22) are simply the first order conditions of an individual's objective function, which come naturally with the equilibrium definition and Proposition 4. To show the existence and uniqueness of the equilibrium, we first observe that for any $\boldsymbol{\mu} \geq 0$, and any fixed $\tilde{\theta}_\infty \in [\theta_\infty(\mathbf{0}), 1)$, where $\theta_\infty(\mathbf{0})$ is the final state of θ under $\boldsymbol{\mu} = \mathbf{0}$, there always exists a unique solution $\tilde{\mu}_k$ to

$$\alpha_1 \frac{\partial f(\tilde{\mu}_k)}{\partial \mu_k} + \alpha_2 \frac{k - k\tilde{\theta}_\infty^{k+\mu_k}}{k + \mu_k \tilde{\theta}_\infty^{k+\mu_k}} \frac{\partial}{\partial \mu_k} \left(\frac{k + \tilde{\mu}_k \tilde{\theta}_\infty^{k+\tilde{\mu}_k}}{k + \tilde{\mu}_k} \right) = 0, k = 1, \ldots, K. \quad (5.24)$$

This is because while $f'(\mu_k)$ is an strictly increasing function from 0 (at $\mu_k = 0$), the second term is also strictly increasing with a negative starting value at $\mu_k = 0$, and it approaches 0 when $\mu_k \to \infty$. Therefore, there always exists a unique solution to (5.24), and we can write it as a function of $\tilde{\theta}_\infty$, namely $\tilde{\mu}_k\left(\tilde{\theta}_\infty\right), k = 1, \ldots, K$.

Moreover, by taking partial derivatives on (5.24) over $\tilde{\theta}_\infty$, we see that $\tilde{\mu}_k$ should be continuous and strictly decreasing with $\tilde{\theta}_\infty, \forall k = 1, \ldots, K$. Furthremore, $\theta_\infty(\boldsymbol{\mu})$ is continuous and monotonically increasing with respect to $\mu_k, \forall k = 1, \ldots, K$. Therefore, there must exists a unique $\tilde{\theta}_\infty \in [\theta_\infty(\mathbf{0}), 1)$, such that $\theta_\infty(\boldsymbol{\mu}^*) = \tilde{\theta}_\infty$, where $\boldsymbol{\mu}^* = \left\{\tilde{\mu}_k\left(\tilde{\theta}_\infty\right)\right\}_{\forall k}$. $\square$

Proposition 5 characterizes the vaccination game equilibrium condition using a simple first order condition, and proposes a sufficient condition for the existence and uniqueness of the equilibrium. The requirements that $f'(0) = 0$ and f be strongly convex are essentially stating that the marginal psychological penalty of vaccination increases with the vaccine adoption level, which could be easily satisfied in many real-world cases.[e]

It is not possible to directly obtain an analytical solution to the equilibrium, however, because θ_∞ lacks an explicit form in terms of $\boldsymbol{\mu}^*$, and requires solving the non-linear equation (5.12). In

---

[e]The convexity of the penalty function is analogous to an individual's risk aversion behavior in general, which is represented by a concave utility function, or equivalently a convex disutility function.

the next subsection, we utilize Proposition 5 to design a heuristic algorithm that solves for the equilibrium point.

5.2.4 Solution algorithm

The basic idea of the heuristic solution algorithm is to iteratively find the optimal vaccine adoption level for a constant value of θ_∞, and then update θ_∞ with the new vaccination rate. Given θ_∞ in each iteration, the problems are convex so the optimal vaccination rates are found via a bisection search; then θ_∞ is updated by solving (5.12).

Algorithm 2 Vaccination game equilibrium solver

Step 2.1 Set $n = 0$, $\boldsymbol{\mu}^{(n)} = \mathbf{0}$. Solve for $\theta_\infty^{(n)} = \theta_\infty\left(\boldsymbol{\mu}^{(n)}\right)$ and $R_k\left(\infty, \boldsymbol{\mu}^{(n)}\right)$ with (5.12) and (5.15), respectively, and solve for $\hat{\mu}_k$ such that $\alpha_1 f\left(\hat{\mu}_k\right) = \alpha_2 R_k\left(\infty, \boldsymbol{\mu}^{(n)}\right)$, $\forall k$;

Step 2.2 Use 0 and $\hat{\mu}_k$ as the lower and upper bounds, respectively, and perform bisection search to find the solution $\mu_k^{(n+1)}$ to the following equations:

$$
\alpha_1 \frac{\partial f\left(\mu_k^{(n+1)}\right)}{\partial \mu_k} + \alpha_2 \frac{k - k\left(\theta_\infty^{(n)}\right)^{k+\mu_k^{(n)}}}{k + \mu_k^{(n)}\left(\theta_\infty^{(n)}\right)^{k+\mu_k^{(n)}}} \frac{\partial}{\partial \mu_k}\left(\frac{k + \mu_k^{(n+1)}\left(\theta_\infty^{(n)}\right)^{k+\mu_k^{(n+1)}}}{k + \mu_k^{(n+1)}}\right) = 0,
$$
$$
k = 1, \ldots, K. \tag{5.25}
$$

Step 2.3 Solve for $\theta_\infty^{(n+1)}$ using (5.12) with $\boldsymbol{\mu}^{(n+1)}$;

Step 2.4 If any of the following conditions are satisfied, terminate and return $\boldsymbol{\mu}^{(n+1)}$; otherwise, let $n \leftarrow n + 1$ and go to Step 1. The termination conditions include:

- Maximum iteration number is reached;
- $\left|\boldsymbol{\mu}^{(n+1)} - \boldsymbol{\mu}^{(n)}\right| \leq \epsilon$, where ϵ is a user-defined convergence threshold; and
- $\left|\theta_\infty^{(n+1)} - \theta_\infty^{(n)}\right| \leq \epsilon$.

In each iteration, we need to solve K nonlinear equations as suggested by (5.25), which might appear formidable when K is large. Luckily, existing research show that real-life contact networks are more suitably described by degree distributions without a heavy tail (Kossinets and Watts, 2006; Srinivasan et al., 2006), which means the probability of k exceeding a finite large number can be neglected. Moreover, bisection search solves each one of the one-dimensional nonlinear equations fairly efficiently.

Although theoretical convergence of the above algorithm is not guaranteed, it is not difficult to see that the algorithm terminates when the equilibrium point is reached. In fact, good convergence performances are observed through intensive experiments in many realistic Nash game contexts. In addition, we also found that the convergence and/or stability of the algorithm is rather insensitive to the step size.

5.3   Other Vaccination Schemes

In the dynamic vaccination game, we have implicitly assumed that people vaccinate themselves during the course of the epidemic outbreak (i.e., called *delayed vaccination*), and people make different decisions based on degree heterogeneity (i.e., called *heterogeneous vaccination*). Note, however, that in the real world, people may also choose vaccination prior to the onset of the outbreak (which will be called *early vaccination*), and/or enforce a vaccination rate indiscriminatively (e.g. vaccination programs enforced by a public health agency, which will be called *homogeneous vaccination*). While it is expected that early vaccination should outperform its delayed counterpart, we aim to quantitively measure the impacts caused by the delay and in so doing provide useful insights for decision makers.

To evaluate the impacts of these alternative vaccination options, we compare the following three schemes (I) early and homogeneous vaccination; i.e., prior to the disease outbreak, a fraction of $M = V(\infty) + A(\infty)$ population are randomly vaccinated regardless of their degrees; (II) early and heterogeneous vaccination; i.e., prior to the disease outbreak, a $(A_k(\infty) + V_k(\infty))/p_k$ fraction among the degree-k population are vaccinated, $\forall k = 1, \ldots, K$; and (III) delayed and homogeneous vaccination; i.e., during the epidemic propagation process, the population are vaccinated homogeneously following the same total rate as the vaccination game, i.e., there is a fraction of $\sum_k v_k(t) S_k(t) dt$ susceptible individuals vaccinated in the time interval $[t, t+dt)$, $\forall t \geq 0$. In each of these schemes, the total number of individuals eventually receiving vaccination is set to be the same as that of the dynamic vaccination game in the previous section. However, the disease propagation process and the final epidemic size may be different.

We first follow Newman (2002) to compute the final epidemic size of Scheme I. The vaccination scheme is equivalent to randomly removing a fraction of M nodes from the network before the disease starts to propagate. After removing these nodes, not only the total number of nodes, but also the degree distribution in the remaining network changes. To find the new degree distribution in the remaining network, we note that randomly removing M fraction of nodes from the network is equivalent to removing M fraction of the stubs. The probability of a degree-k node having m stubs connected to a remaining stub is given by a binomial distribution with probability $\binom{k}{m}(1-M)^m M^{k-m}$. Then the degree generating function of the remaining nodes, denoted as

$g_1(x)$, is:

$$
\begin{aligned}
g_1(x) &= \sum_m x^m \sum_k p_k \binom{k}{m} (1-M)^m M^{k-m} \\
&= \sum_k p_k \sum_m \binom{k}{m} (x(1-M))^m M^{k-m} \\
&= \sum_k p_k \left(x(1-M) + M \right)^k \\
&= g \left(x(1-M) + M \right).
\end{aligned}
\tag{5.26}
$$

The same result is also derived in Buldyrev et al. (2010).

When the infectious disease starts to propagate in the network, there are two possibilities: (i) there are only local outbreaks such that the epidemic size does not scale with the population; and (ii) there is a pandemic outbreak such that a single disease seed leads to a giant connected subnetwork of infected nodes. Newman (2002) presents the critical condition on the disease transmissibility to determine whether a pandemic outbreak could happen; i.e., a pandemic outbreak may occur only when $T > T_c$, where

$$
T_c = \frac{g_1'(1)}{g_1''(1)} = \frac{g'(1)}{(1-M)\, g''(1)}.
\tag{5.27}
$$

If there is no pandemic outbreak, the average number of infected nodes, is given by:

$$
s_0 = 1 + \frac{T g_1'(1)}{1 - T g_1''(1)/g_1'(1)} = 1 + \frac{T\,(1-M)\, g'(1)}{1 - T\,(1-M)\, g''(1)/g'(1)}.
\tag{5.28}
$$

Otherwise, if the pandemic outbreak occurs, the fraction of the giant infected subnetwork out of the unvaccinated population is given by $1 - g_1(1 + (y-1)T)$, where y is the solution to the following self-consistency relation:

$$
y = \frac{1}{g_1'(1)} g_1'(1 + (y-1)T).
\tag{5.29}
$$

Then the following equation yields the pandemic size in terms of fraction among the original population:

$$
s_1 = (1-M)\left(1 - g_1\left(1 + (y-1)T\right)\right).
\tag{5.30}
$$

Moreover, the average outbreak size among the nodes not belong to be giant subnetwork is still given by (5.28).

For Scheme II, the population is heterogeneously vaccinated. Although there exists a similar analytical approach in finding the final epidemic size, e.g. as in Huang et al. (2011), it requires the vaccination probability of nodes to follow a certain special form. Therefore, in this paper, we

will simply use stochastic simulations to compute the system evolution and the final epidemic size. Finally, we note that Scheme III is a special case of dynamic vaccination. Hence, the outcome of Scheme III can be easily computed using the developed SIRVA dynamics model.

5.4   Numerical Results

In this section, we first validate the proposed SIRVA dynamic model (Eqs. (2.3), (2.5), and (5.1) - (5.6)) as well as the derived final epidemic size (Eqs. (5.11) and (5.12)), by comparing these analytical formulas with agent-based simulation outcomes. Then we present the results of the vaccination game and evaluate the impacts of delayed and homogeneous vaccinations. Sensitivity analysis is then conducted to reveal interesting insights.

The benchmark case involves a population with a Poisson degree distribution and a hypothetical epidemic event with moderate disease transmissibility. Poisson distribution is chosen because it does not possess heavy-tailed properties and has been widely recognized in the existing literature (Newman et al., 2001; Barthélemy et al., 2005; Volz, 2008). The average degree $g'(1) = 7$, with a cutoff $K = 22$ (such that $P(k \geq K) \leq 10^{-5}$). The disease parameters are as follows: the infection rate $r = 0.01$, the removal rate $u = 0.01$, and consequently the disease transmissibility $T = 0.5$. Moreover, the initial condition at time $t = 0$ is set as follows: $\theta(0) = 1.0$, and $\phi(0) = 0.001$.

5.4.1   Model validation

In the agent-based stochastic simulations, configuration model random networks are built following Molloy and Reed (1995), where open edges are generated and assigned to the nodes following the degree distribution, and then randomly paired up with other open edges. The number of nodes (i.e., the population size) is set to be 10^4; in every simulation, initial disease seeds are randomly located at 10 of these nodes.

We perform a total of 50 simulations, each with a new realization of network topology and initial disease seeds. In each discretized time step (e.g., a day), disease transmission, node removal, and node vaccination, and node activation are randomly simulated with corresponding probabilities. The vaccination rate follows (5.7), with $\mu_k = 0.4k, \forall k$. In the simulations, since θ and ϕ are virtual quantities which cannot be measured directly, we use the following approximation of ϕ/θ to compute v_k:

$$r\frac{\phi}{\theta} \approx \frac{\Delta I}{\sum_k k S_k},$$

(5.31)

where ΔI is the fraction of recent infections.

The simulation results can be compared with those from the system dynamics equations, as
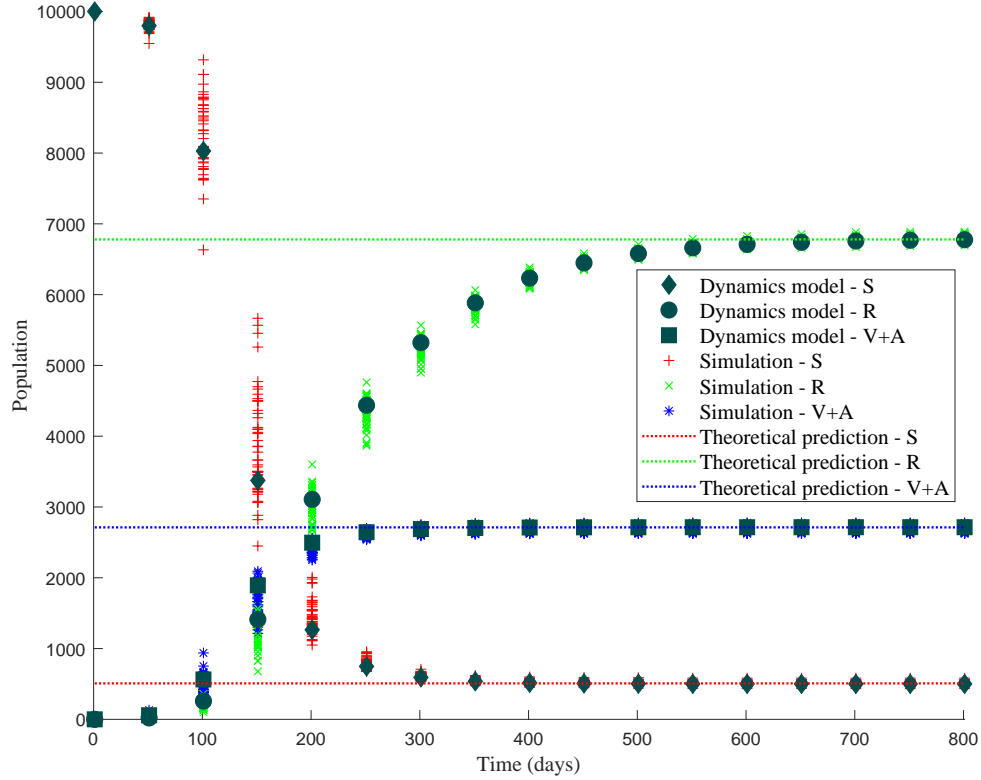
Figure 5.1: Compare simulation with system dynamic equations and theoretical final states.

well as the analytical formulas (5.12) and (5.11). The colored dots in Figure 5.1 show results from the 50 stochastic simulations, sampled every 50 days. While these simulated samples show some variations in the early stages of the simulation, the final state of the epidemic converges fairly well to the same value across all simulations: the total fraction of immunized population is 0.2687, that of the remaining susceptible is 0.0510, and that of the removed is 0.6803. The results from system dynamics equation coincide well with the simulation results: the final epidemic size $R(\infty)$ computed from (2.3), (2.5), and (5.1)-(5.6) is 0.6786, which is only 0.3% different from that of the simulations, and the fraction of remaining susceptible is 0.0501. Considering the stochasticity of the simulations, the error shows excellent agreement between the SIRVA dynamic model and the simulations. In fact, we believe the relative difference will further reduce if the population size increases. Furthermore, Eqs. (5.12) and (5.11) can be used directly to compute $\theta_\infty = 0.6738$ and the final epidemic size equals 0.6780. The relative error between the theoretical final epidemic size and that of the system equations is negligible regardless of initial condition of ϕ. These results give us reasonable confidence to use the established systems dynamics models and the derived analytical results to study vaccination games.

5.4.2 Vaccination game results

We proceed to present the results of a hypothetical vaccination game with the same population and disease. For convenience, we set the weight factors $\alpha_1 = 10^{-4}$ and $\alpha_2 = 1$. The penalty for vaccination is set to be a polynomial function in the form of $f(x) = x^b$ where $b = 2$. As such, $f(x)$ is a monotonically increasing function, and it satisfies the equilibrium uniqueness requirements in Proposition 5, thus the vaccination game has a unique equilibrium solution. The heuristic algorithm terminates after 11 iterations when ϵ reaches 1×10^{-4}.
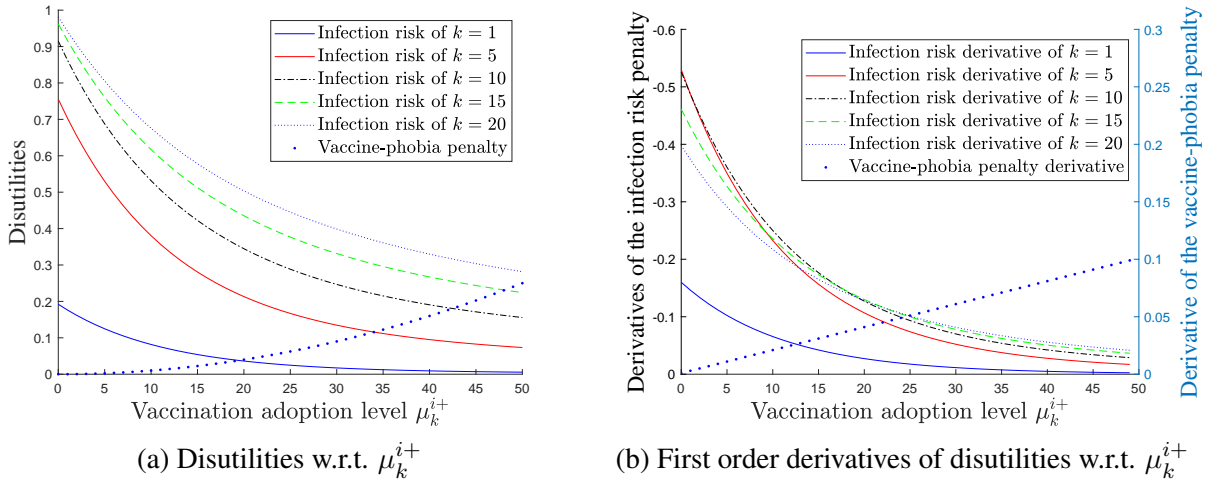

Figure 5.2: Validation of the vaccination game results

Figure 5.2 presents an individual's disutilities due to vaccine-phobia and infection risk, as well as the first order derivatives of these disutilities at the equilibrium. As expected, the infection risk penalties are strictly decreasing functions of μ_k^{i+} for all k's, and that of vaccine-phobia is a strongly convex and increasing function; see Figure 5.2a. Moreover, Figure 5.2b shows the first order derivatives of these disutilities. For each k, the intersection of the vaccination derivative curve and that of the infection is exactly the solution to the equilibrium condition (5.22).

The equilibrium solutions are presented in Figure 5.3. Figure 5.3a shows that the equilibrium vaccination adoption level increases with the node degree; meanwhile, the total disutilities are strictly increasing with k. Figure 5.3b shows the final state compositions of nodes with different degrees. Interestingly, even though the equilibrium vaccination rate strictly increases with k, the probability for a node to eventually get vaccinated first increases but then declines. This is because, the vaccination rate v_k and infection rate are both proportional to $r\phi/\theta$, whereas the multiplicative factor equals μ_k for vaccination and k for infection. Thus the infection rate grows proportionally to k, yet the vaccination rate grows sub-linearly as shown in Figure 5.3a. Consequently, the ratio of vaccination rate to infection rate declines as the node degree increases. Implementing the vaccination decisions at the equilibrium in the systems dynamics model (2.3), (2.5), and (5.1)-(5.6), the

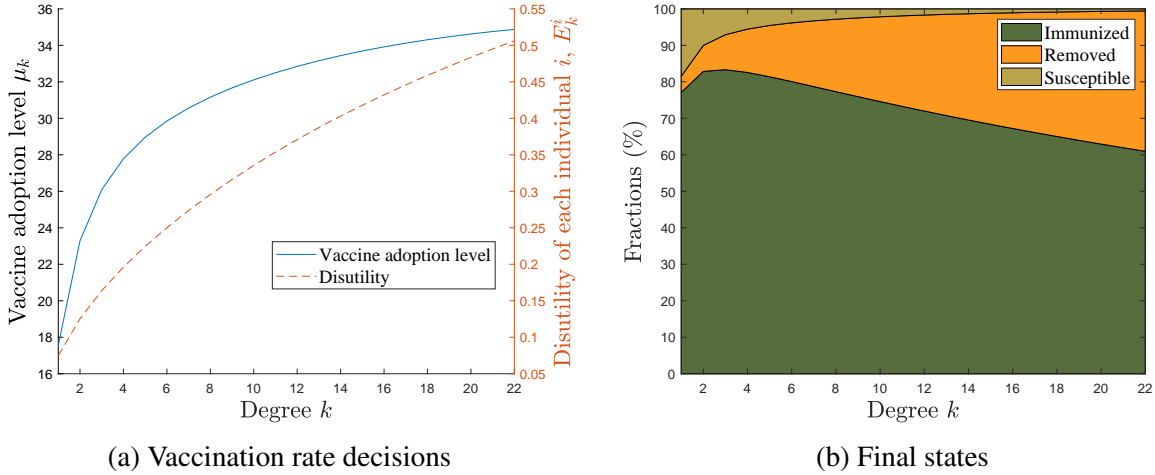| (a) Vaccination rate decisions	(b) Final states

Figure 5.3: Vaccination game equilibrium results

final epidemic size is 0.1770, with $\theta_\infty = 0.9124$, and a total fraction of 0.7866 of the population are eventually vaccinated. Compared to the results in Sec. 5.4.1 where a relatively low vaccinate rate is adopted, we find that the disease propagation has been significantly mitigated as a result of the vaccination game.

5.4.3 Delayed and homogeneous vaccinations

Following Scheme I and considering vaccinating the same fraction 0.7866 of the population (as in the vaccination game) indiscriminatingly before the epidemic outbreak, we find that $T_c = 4.69 > T$ from (5.27). There will only be several endemic infections, and the average outbreak size is 1.8 from (5.28). This means that each disease seed is expected to infect only 0.8 additional susceptible individual during the entire outbreak. Since there are 10 initial disease seeds, and in the large population their neighbors are highly unlikely to overlap, the total infected fraction is expected to be approximately 0.0018. Scheme II requires that the population be vaccinated heterogeneously based on their degrees prior to the epidemic outbreak. The mean infection fraction over 50 random simulations is 0.0051.

The comparison between early and delayed vaccinations is consistent with our expectation. In Schemes I and II, the infection fractions are 99.0% and 97.1% smaller than that of the vaccination game case, respectively. This shows that, with preventive vaccination ahead of time, the pandemic outbreak is prevented from happening; in contrast, delaying the vaccination allows devastating outbreak to take place and infect a large fraction of the population. This is intuitive – the higher-degree individuals are more easily infected and then pose larger threats to others once infected, thus vaccinating these individuals before the outbreak can greatly counteract the disease

propagation. Furthermore, if we compare Schemes I and II, we find that non-discriminative vaccination is helpful if done prior to epidemic outbreak. Scheme I outperforms Scheme II because the high-degree nodes have higher vaccination probabilities in Scheme I, as implied by Figure 5.3b.

Scheme III, with both delayed and homogeneous vaccination, yields a final epidemic size of 0.1805. Comparing Scheme III and the dynamic vaccination game from Sec.5.4.2, we observe that homogeneous vaccination is slightly less effective than heterogeneous vaccination if they are "delayed.âĂİ This is in sharp contrast to the comparison between Schemes I and II, mainly because the vaccination rate (instead of vaccination probability) of the high degree nodes are higher in the vaccination game as suggested by Figure 5.3a, which helps protect the other susceptible individuals.

These results are interesting and insightful. The coverage and timing of vaccination for the high-degree nodes almost dictate the outcome of a disease outbreak. As discussed in Section 5.4.2, in a vaccination game people with more social connections tend to take vaccines sooner. Therefore, it would be more effective to encourage voluntary vaccine uptake than blindly vaccinate the whole population once the epidemic outbreak has started. This finding would be particularly helpful when vaccines are scarce during the course of an epidemic outbreak. However, despite their high vaccine adoption level, the high-degree nodes are likely to be quickly infected before they can be vaccinated, as we have also seen in Section 5.4.2. This facilitates disease propagation and leads to a much wider spread of disease. Therefore, it is important to vaccinate as many high-degree nodes as possible (prior to the outbreak) to provide them with timely protection. In this case, non-discriminative vaccination is not a bad option âĂŞ compared to the voluntary (but delayed) vaccination scheme, the high-degree nodes now have a better chance of getting vaccinated in time. Targeting individuals with many social connections would greatly cut the possibility of disease propagation in both early and delayed vaccination schemes. Nevertheless, we shall also notice that such a measure is often not practical and may raise social/equity concerns.

5.4.4 Sensitivity analysis

We conduct additional experiments to investigate the outcome of vaccination games under different disease-related, network-related, and penalty function related parameters. In all these tested cases, the equilibrium solution was found in no more than 15 iterations. For convenience, we call the vaccination game in Section 5.4.2 as Scenario 0.

We first change the degree distribution of the population, by considering the following cases: Scenarios (i) and (ii) change the average degree and the degree cutoff of the Poisson distribution, to $g'(1) = 4$ and $K = 16$, and $g'(1) = 10$ and $K = 27$, respectively; Scenario (iii) changes the degree distribution to a uniform distribution with the same average degree $g'(1) = 7$; then Scenarios (iv)

and (v) consider uniform distributions with the same average degrees as Scenarios (i) and (ii), respectively. The vaccination rate decisions of all these scenarios are presented in Figure 5.4.
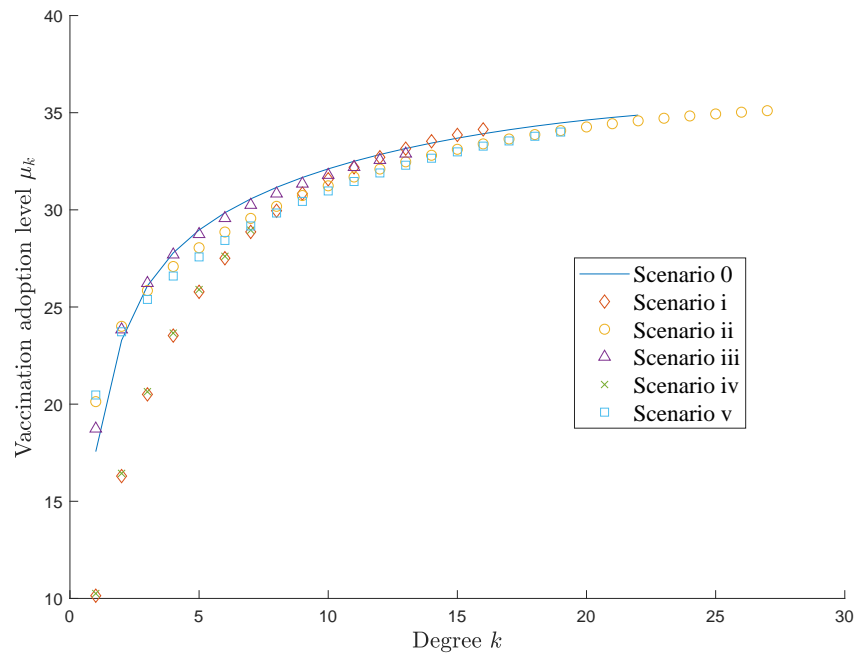

Figure 5.4: Compare vaccination decisions under different degree distributions.

Interestingly, if the degree distributions, no matter Poisson or uniform, share the same average degree, they would lead to similar vaccination decisions for individuals of the same degrees. As the figure shows, Scenarios in the followings pairs, 0 and (iii), (i) and (iv), and (ii) and (v), all show exceptional agreement with each other. This suggests that the vaccine adoption level may not be affected by the shape of the degree distribution as long as they have the same average degree. However, it is worth mentioning that same vaccination decisions do not necessarily lead to the same epidemic outbreak outcome, including people's disutilities and final epidemic sizes, as the degree distributions still affects the disease propagation process. For instance, the final epidemic sizes in these scenarios are 0.1086, 0.2368, 0.1755, 0.1079, and 0.2287, respectively.

Moreover, the relationship between the average degree in the population and the equilibrium vaccination decisions is somewhat more complex. Figure 5.4 shows that for low degree nodes (e.g., $k < 5$), higher average degree leads to higher vaccination rate; however, it is the opposite for the high degree nodes (e.g, $k \geq 15$). This phenomenon is probably due to two contradicting mechanisms. First, it is straightforward that high network connectivity increases the infection risk and thus encourages higher vaccination rate. However, with higher network connectivity, high degree nodes will be infected much faster such that in order to even slightly reduce their infection risk, a remarkably high vaccination adoption level is required. This induces a high psychological

66

suffering due to vaccine-phobia. Consequently, they would simply give up vaccination. A similar phenomenon can be observed in a later example.

Next, we examine the impact of the disease-related parameters r and u. Note that in all the equations used to solve for the equilibrium, these two parameters appear together in the form of the disease transmissibility T. Therefore, we only need to tune the value of T. Since in Scenario 0 T takes a moderate value of 0.5, we consider three additional cases: Scenario (vi) and (vii) where $T = 0.3$ and 0.7, respectively; and an extreme case Scenario (viii) where $T = 1$ (i.e., the removal rate $u = 0$).

The results are presented in Figure 5.5. Again, the influence of disease transmissibility on people's vaccination decisions appears to be highly complex. The impact seems greater on low-degree nodes but becomes negligible for the high-degree nodes. Furthermore, the final epidemic sizes of these three cases are 0.1611, 0.1873, and 0.1931, respectively, suggesting that disease transmissibility does not have a significant impact on the outcome of the epidemic in a vaccination game. This is probably because the overall vaccination coverage is already sufficiently high.
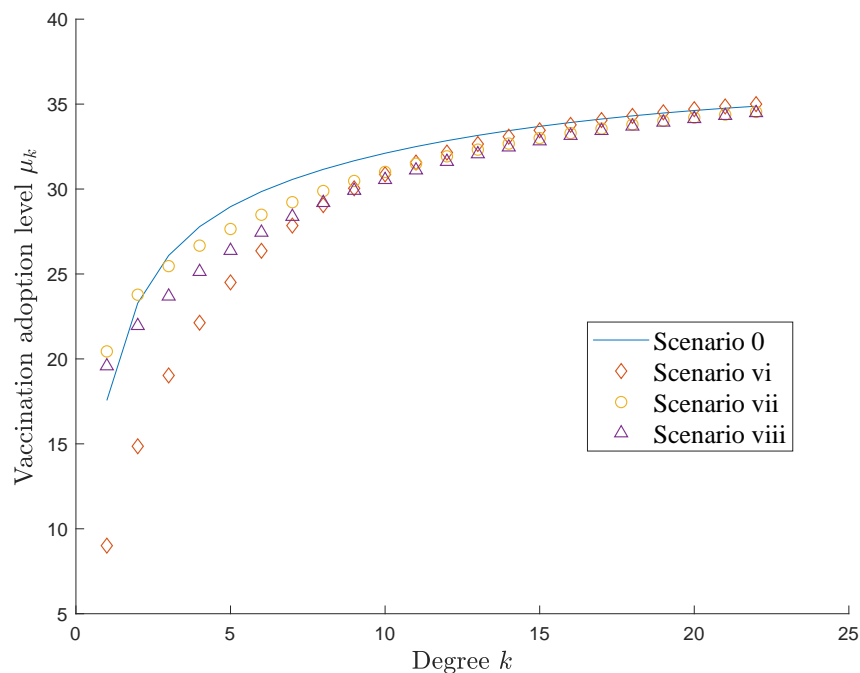
Figure 5.5: Compare vaccination decisions with different disease transmissibilities.

Finally, we investigate the impact of the weight factors and the penalty function parameter. This is to examine the influence of different levels of vaccine-phobia compared to various infection penalties. Different combinations of the vaccination penalty parameters α_1 and b are tested, with α_2 fixed to be 1, and the consequential infection and vaccination fraction, i.e. $R(\infty)$ and

$V\left(\infty\right)+A\left(\infty\right)$ are presented, respectively. The results are summarized in Table 5.1. These results show that people's vaccination decisions and the final outcome of the epidemic is largely dependent on the weighing of infection risk over vaccination, and the vaccination penalty function. It suggests that vaccine-phobia has a huge impact on population vaccination decisions, and therefore the outcome of the epidemic outbreak.

Table 5.1: Sensitivity analysis on the penalty function parameters.

| Cases	$\alpha_1 = 1 \times 10^{-5}, b = 1.5$	$\alpha_1 = 1 \times 10^{-4}, b = 1.5$	$\alpha_1 = 1 \times 10^{-3}, b = 1.5$
Infection fraction	0.0263	0.0736	0.2009
Vaccination fraction	0.9201	0.8805	0.7596
Cases	$\alpha_1 = 1 \times 10^{-5}, b = 2.0$	$\alpha_1 = 1 \times 10^{-4}, b = 2.0$	$\alpha_1 = 1 \times 10^{-3}, b = 2.0$
Infection fraction	0.0790	0.1767	0.3646
Vaccination fraction	0.8792	0.7851	0.6024
Cases	$\alpha_1 = 1 \times 10^{-5}, b = 2.5$	$\alpha_1 = 1 \times 10^{-4}, b = 2.5$	$\alpha_1 = 1 \times 10^{-3}, b = 2.5$
Infection fraction	0.1594	0.2950	0.4958
Vaccination fraction	0.8033	0.6711	0.4742

Moreover, a fairly counter-intuitive phenomenon is observed when the weighing of vaccination penalty is relatively high. Taking the case where $\alpha_1 = 1 \times 10^{-3}$ and $b = 2.5$ for instance, we demonstrate the vaccine adoption level of the population in Figure 5.6a. It shows that high degree nodes (e.g., $k \geq 10$) tend to give up on vaccinating themselves, as their vaccine adoption level is even lower than those with fewer degrees. Figure 5.6b demonstrates that at the equilibrium, the marginal benefit in reducing the infection risk intersects with the marginal vaccine-phobia penalty at a low vaccine adoption level for the high degree nodes, suggesting it is difficult for the high degree nodes to further reduce their infection risk without inducing a much higher vaccination penalty. This implies that when the vaccine-phobia level is high, or equivalently the penalty of infection is low, those who can be easily infected would rather take the infection risk than imposing a high psychological suffering on themselves. In this case, non-differential vaccination enforced by the public health agency may outperform the result of the vaccination game.

In another angle, we see that in a vaccination game, people's vaccination decisions are mostly affected by their relative perception of infection risk over vaccine-phobia, instead of the population connectivity or the disease transmissibility. Therefore, if a public health agency wishes to mitigate the propagation of an infectious disease by encouraging voluntary vaccine uptake among the population, it is important to alleviate people's fear towards vaccines and to properly inform them about the seriousness of the disease infection. Reducing population connectivity and disease transmissibility (e.g., using other disease control approaches such as quarantine) would not greatly affect people's vaccination behaviors; however, these measures can mitigate the propagation of the disease and should be considered during the course of disease spread.
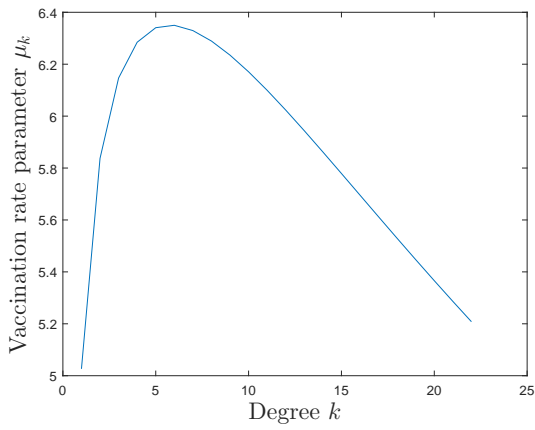
(a) Vaccination rate decisions
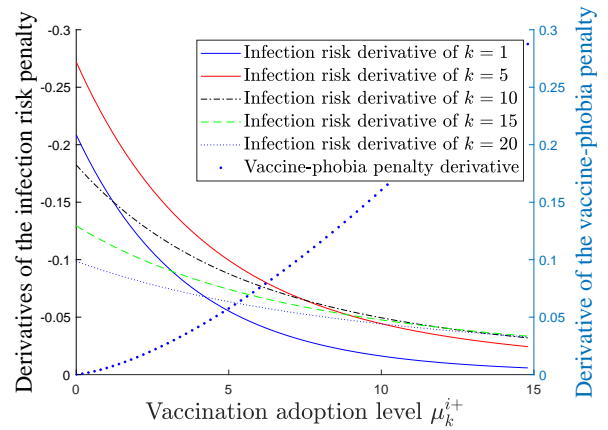
(b) First order derivatives of disutilities w.r.t. μ_k^{i+}

Figure 5.6: Vaccination decisions and equilibrium conditions of the high vaccination penalty scenario

CHAPTER 6 – OPTIMAL VACCINE ALLOCATION IN A POPULATION WITH SPATIAL HETEROGENEITY

6.1 Introduction

In the previous chapter, we investigate a dynamic vaccination game where each individual in a population makes its own vaccination decisions during a disease outbreak. The detrimental impacts of delayed vaccination is revealed in the numerical experiments. If there are insufficient vaccine resources at the beginning of an epidemic event (which happens quite often in reality), such delay will take place. Therefore, another important question for public health agencies is, given limited vaccines, where to allocate the vaccine resources to minimize the expected impact of an upcoming epidemic outbreak.

In a large geographical region, a population's distribution and their contact network structure may possess high spatial heterogeneity. For instance, a population in an area with higher density might interact more frequently among themselves than those with lower densities do; two populations living in nearby areas might also have a higher connectivity with each other than with those in farther regions. Moreover, the probability that the outbreak starts in a certain area also varies greatly depending on the spatial information. For example, port cities and metropolises have a significantly greater chance of being the source of disease outbreak than inland and rural areas. Therefore, for infectious disease transmitted via direct contacts, it is important to account for such spatial heterogeneities when considering optimal vaccine allocation.

To address these challenges, this chapter investigates the optimal vaccine allocation that minimizes the expected impact of a disease outbreak with consideration of population spatial heterogeneity. Following the previous chapters of this dissertation, we consider heterogeneous mixing populations using random networks. Moreover, spatial heterogeneity is accounted for in the following three aspects: i) population distribution, ii) subpopulation connectivity, and iii) disease source distribution.

The rest of this chapter is organized as follows. Section 6.2 develops an approach to evaluate the effectiveness of an allocation decision by computing the expected disease outbreak size. Section 6.3 formulates two optimization problems to solve for the minimum required resource and

optimal allocation, and proposes solution approaches to solve for the two optimization problems, respectively. Next, Section 6.4 performs numerical experiments on a hypothetical case study to validate the modeling frameworks and solution approaches.

6.2 Vaccine Allocation and Expected Outbreak Sizes

Consider a planar area Ω where a population of size N lives within, and the entire area is partitioned into m subregions. The population is distributed heterogeneously across Ω, and each subregion contains a fraction of λ_i population such that $\sum_{i=1}^{m} \lambda_i = 1$. The contact network between the population in subregions i and j is described by the probability generating function of the degree distribution:

$$g_{ij}(x_{ij}) = \sum_{k=0}^{\infty} p_{ij}(k)x_{ij}^k, \forall i, j = 1, \ldots, m. \tag{6.1}$$

where $p_{ij}(k)$ represents the probability that an arbitrary node in subregion i has k edges connected to the nodes in subregion j, and x_{ij} is a dummy variable. Degree correlations between different pairs of subregions are igonored. Naturally, the following degree conservation law should always hold:

$$\lambda_i g'_{ij}(1) = \lambda_j g'_{ji}(1), \forall i, j = 1, \ldots, m. \tag{6.2}$$

In this spatially-heterogeneous population contact network, the decision-maker needs to allocate vaccines to different subregions to minimize the expected outbreak size. It is the expected number of infections caused by a single initial infection, including the seed node. The outbreak size depends on the location of the seed, because the network connectivity is spatially heterogeneous. We denote the conditional expected outbreak size given the initial seed lives in subregion i by s_i. The agency may have an anticipation over the seed distribution, and we let w_i be the probability that the epidemic starts in subregion i, such that $\sum_{i=1}^{m} w_i = 1$. The objective of the agency is to minimize the expected outbreak size incorporating such prior distribution:

$$\min_{\boldsymbol{\delta}} \sum_{i=1}^{m} w_i s_i, \tag{6.3}$$

where δ_i is the fraction of vaccinated population in subregion i $\forall i = 1, \ldots, m$, and $\boldsymbol{\delta}$ is the collection of all δ_i's. Furthermore, since there is a limit amount of total vaccines, denoted Q, we have the following budget constraint:

$$N\boldsymbol{\lambda}^T \boldsymbol{\delta} \leq Q. \tag{6.4}$$

We assumed that vaccines directly remove the nodes when they are applied to the population,

71

and the edges connected to the vaccinated nodes are removed as well. Since the connections are random, for an unvaccinated subregion-i node with a degree of k to subregion j, the probability that r out of the k edges are connected to unvaccinated subregion j nodes is given by Binomial distribution $\binom{k}{r}(1-\delta_j)^r\delta_j^{k-r}$. Therefore, the degree distribution from subregion i to j of the unvaccinated subregion i nodes is given by:

$$g_{ij,1}(x_{ij}) = \sum_{k=0}^{\infty}\sum_{r=0}^{k}\binom{k}{r}p_{ij}(k)(1-\delta_j)^r\delta_j^{k-r}x_{ij}^r = g_{ij}\left((1-\delta_j)x_{ij}+\delta_j\right), \forall i,j = 1,\ldots,m.$$

(6.5)

When all of the nodes in subregion j are removed, $g_{ij,1}(x_{ij}) = 1$, which means the resulting degree distribution $p_{ij,1}(k) = 0, \forall k \geq 1$. After vaccination, degree conservation law still holds, such that:

$$(1-\delta_i)\lambda_i g_{ij,1}'(1) = (1-\delta_j)\lambda_j g_{ji,1}'(1), \forall i,j = 1,\ldots,m.$$

(6.6)

The proof is straightforward using (6.5).

Moreover, we inherit the disease transmissibility, denoted T, from previous chapters. By sampling the edges in the remaining network after vaccination with probability T, we can construct a (possibly fragmented) network containing only the selected edges. We choose a random unvaccinated node as the initial seed, then the expected size of the component containing this node in this network represents the expected outbreak size. The mathematical derivation for the degree distribution of such a network is highly similar to (6.5), as the number of selected edges of a node also follows a Binomial distribution, given by $\binom{k}{r}T^r(1-T)^{k-r}$. We let $g_{ij,2}$ be the corresponding degree distribution:

$$g_{ij,2}(x_{ij}) = g_{ij,1}(T(x_{ij}-1)+1) = g_{ij}(T(x_{ij}-1)(1-\delta_j)+1), \forall i,j = 1,\ldots,m.$$

(6.7)

As such, the actual disease transmission network can be properly described by $g_{ij,2}$, with which the expected outbreak size can be found, and the existence of a giant connected component can be examined. To begin with, we first select a random edge that connects subregion i and j, and follow it to the j end, the probability that excluding this edge, the j end has k edges to subregion j', namely the *excess degree*, is denoted by $p_{ijj'}(k)$. Its distribution is given by the generating function $g_{ijj'}(x_{jj'})$:

$$g_{ijj'}(x_{jj'}) = \begin{cases} \frac{g_{ji,2}'(x_{ji})}{g_{ji,2}'(1)}, & \text{if } j' = i, \\ g_{jj',2}(x_{jj'}), & \text{o.w.} \end{cases}$$

(6.8)

Then, let $H_{ij}(x)$ be the generating function for the distribution of the size of components that

are reached by choosing a random (i, j) edge and following it to the j end. The following self-consistency equation must be satisfied:

$$
\begin{aligned}
H_{ij}(x) =& x\left[\prod_{j'=1}^{m} p_{ijj'}(0) + \sum_{j''=1}^{m} H_{jj''}(x)p_{ijj''}(1)\prod_{j'\neq j''} p_{ijj'}(0) + \dots\right] \\
=& x\prod_{j'=1}^{m} g_{ijj'}\left(H_{jj'}(x)\right), \forall i, j = 1, \dots, m.
\end{aligned}
\tag{6.9}
$$

To illustate, Figure 6.1 demonstrates a simple example with 2 subregions. A square means a component found by following an edge to one of its end, and a circle represents a node. Following a $(2, 1)$ edge to the subregion-1 end, the node may have k_1 and k_2 edges to subregion 1 and 2, respectively, $\forall k_1, k_2 \geq 0$. The probability that the end node is vaccinated is given by δ_i; and the probability distributions of k_1 and k_2, i.e., the excess degrees, are given by g_{211} and g_{212}, respectively.
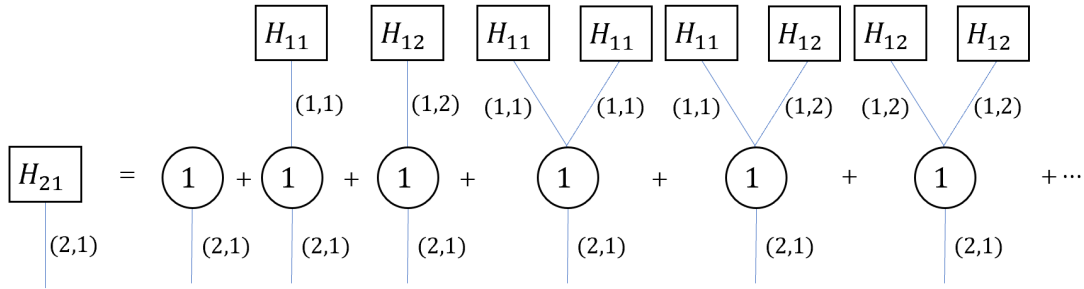
Figure 6.1: An explanatory example for component size distribution with 2 subregions.

(6.9) gives the probability distribution of component sizes found by selecting an edge and following it to one end. Starting at a randomly chosen unvaccinated node in subregion i, the generating function for the component size distribution is given as follows:

$$
H_i(x) = x\prod_{j=1}^{m} g_{ij,2}\left(H_{ij}(x)\right), \forall i = 1, \dots, m.
\tag{6.10}
$$

Finally, the average component size obtained by selecting a random node in subregion i can be

computed as follows:

$$s_i = (1 - \delta_i) H_i'(1)$$

$$= (1 - \delta_i) \left[1 + \sum_{j=1}^m g_{ij,2}'(1) H_{ij}'(1) \right]$$

$$= (1 - \delta_i) \left[1 + \sum_{j=1}^m T(1 - \delta_j) g_{ij}'(1) H_{ij}'(1) \right], \forall i = 1, \ldots, m. \tag{6.11}$$

where $H_{ij}'(1)$ is the solution to the following system of equations, obtained from (6.8) and (6.9). For notational brevity, we denote $H_{ij}'(1)$ as y_{ij}:

$$y_{ij} = 1 + \sum_{j'=1}^m g_{ijj'}'(1) y_{jj'}$$

$$= 1 + \frac{g_{ji,2}''(1)}{g_{ji,2}'(1)} y_{ji} + \sum_{j' \neq j} g_{jj',2}'(1) y_{jj'}$$

$$= 1 + T(1 - \delta_i) y_{ji} \frac{g_{ji}''(1)}{g_{ji}'(1)} + \sum_{j' \neq i} T(1 - \delta_{j'}) g_{jj'}'(1) y_{jj'}, \tag{6.12}$$

$$\forall i, j = 1, \ldots, m.$$

To differentiate these two types of component sizes in the following discussion, we call them node and edge component sizes, respectively. In the case where there is only one subregion, the result is consistent with Newman (2002). Therefore, for a given combination of vaccination decisions $(\mathbf{q}, \boldsymbol{\delta})$, we can compute the expected outbreak size. For convenience, we call this linear system with respect to (w.r.t.) y_{ij}'s given $\boldsymbol{\delta}$ as $LP(\boldsymbol{\delta})$. We observe that the solution $\mathbf{y}$ to $LP(\boldsymbol{\delta})$ is non-increasing and continuous for $\delta_i \in [0, 1], \forall i = 1, \ldots, m$. In particular, when $\boldsymbol{\delta} = \mathbf{1}$, $y_{ij} = 1$ is the minimum value, $\forall i, j = 1, \ldots, m$.

As a final remark, an implicit assumption of (6.9), which can also be observed from Figure 6.1, is that the connected components contain no loops. When the population size is large and the component sizes are small, the probability that a loop exists among the connected components is negligible since the connections are random. Therefore, the proposed approach to calculate the expected outbreak size is valid with negligible errors. However, if a connected component takes a non-negligible portion of the population, this assumption can be easily violated. In a simplest 1-subregion case, the solution to $LP(\delta)$ is given by $y = \frac{1}{1 - T(1-\delta)\frac{g''(1)}{g'(1)}}$. Suppose $T > g'(1)/g''(1)$, when δ approaches $\hat{\delta} = 1 - \frac{g'(1)}{Tg''(1)}$ from right, y diverges to ∞. Indeed, $\hat{\delta}$ is the critical threshold where a giant component starts to form, namely an epidemic outbreak occurs. In this case, (6.9) and

all the results obtained thereupon may not hold. Therefore, these equations start to lose accuracy when the magnitude of the connected component size grows comparable to that of the population.

6.3 Optimal Allocation and Solution Approaches

6.3.1 Preventing an epidemic and minimizing local infections

We first formualte the optimal vaccine allocation problem into a mathematical program, and discuss a few analytical properties of the problem. It can be written as follow.

$$\mathbf{P1} \quad \min_{\boldsymbol{\delta},\mathbf{y},\mathbf{s}} z = \mathbf{w}^T \mathbf{s} \tag{6.13}$$

$$\text{s.t. } s_i = (1 - \delta_i)\left[1 + \sum_{j=1}^{m} T\left(1 - \delta_j\right) g'_{ij}(1) y_{ij}\right], \forall i = 1, \ldots, m, \tag{6.14}$$

$$1 \leq \mathbf{y} \leq \Lambda, \tag{6.15}$$

$$0 \leq \boldsymbol{\delta} \leq 1, \tag{6.16}$$

and (6.4), (6.12).

The objective (6.13) is to minimize the expected outbreak size with consideration of the disease source distribution $\mathbf{w}$. Constraints (6.14) are directly from the previous derivation. Since y_{ij} represents an expected edge component size, it can only be a non-negative number, and from (6.12) we can derive its lower bound as 1. Moreover, Constraints (6.15) impose a numerical upper bound Λ for y_{ij}'s, where $\Lambda > 1$ is an insignificant number compared to the size of the population, i.e., $\Lambda \ll N$. This numerical treatment is to satisfy the premise that the probability that a loop exists among the connected components can be neglected. Furthermore, Constraints (6.16) regulate that vaccinated fraction in each subregion can only be between 0 and 1. Note that in this problem, $\mathbf{y}$ and $\mathbf{s}$ are both auxiliary variables, which can be computed correspondingly once $\boldsymbol{\delta}$ is determined. Therefore, the objective z can also be expressed as a function of $\boldsymbol{\delta}$.

Given a vaccine allocation decision $\boldsymbol{\delta}$, if the set of $\mathbf{y}$ defined by (6.12) and (6.15) is empty, such a vaccine allocation decision is insufficient to prevent an epidemic outbreak, and either a better allocation strategy or a larger amount of vaccines is required. Indeed, there are three possible scenarios: i) the disease will not lead to an epidemic outbreak even if there is no vaccination; ii) the epidemic may occur without vaccination, but there are sufficient vaccines to confine the disease within local networks; and iii) the vaccine resource is insufficient to prevent an epidemic outbreak. Although the first case is easy to identify, it is quite difficult to distinguish the latter two scenarios. Unless a feasible vaccine allocation is found that leads to component sizes satisfying

(6.15), we have no proof that the feasible solution set is nonempty. However, even finding an initial feasible solution is quite challenging in this problem due to the non-convexity of the solution set. Therefore, it is important for the agency to know the minimum amount of vaccines required to prevent an epidemic outbreak, which guarantees that the problem is feasible.

To this end, we propose the a mathematical program that solves for the minimum amount of vaccines required. It is presented as follows:

$$\textbf{P2} \quad \min_{\boldsymbol{\delta}, \mathbf{y}} \boldsymbol{\lambda}^T \boldsymbol{\delta} \tag{6.17}$$

$$\text{s.t. } (6.12), (6.15), (6.16).$$

The objective is to minimize total vaccine usage, such that the average edge component sizes are within a limit. We use $\left(\hat{\boldsymbol{\delta}}, \hat{\mathbf{y}}\right)$ to denote the optimal solution to P2, then $\forall Q \geq N\boldsymbol{\lambda}^T \hat{\boldsymbol{\delta}}$, and there exists at least one feasible solution to P1, which can be constructed using $\left(\hat{\boldsymbol{\delta}}, \hat{\mathbf{y}}\right)$.

Therefore, the agency can find the minimum vaccine resources required to prevent an epidemic outbreak by solving P2. Given sufficient resources, P1 finds the optimal vaccine allocation strategy to mitigate local disease propagation and minimize expected infections. In the next section, we discuss the solution algorithms to solve for the optimization problems P1 and P2.

6.3.2 Solution approach to P2

Observing that the difficulty in solving P2 is mainly caused by Constraints (6.12) which contain bilinear terms, we propose a branch and bound algorithm with McCormick relaxation that is guaranteed to find the global optimum. First, we define an *instance* ψ to be a quadruple $(\boldsymbol{\delta}^{\text{UB}}, \boldsymbol{\delta}^{\text{LB}}, \mathbf{y}^{\text{UB}}, \mathbf{y}^{\text{LB}})$, where $\boldsymbol{\delta}^{\text{UB}}, \boldsymbol{\delta}^{\text{LB}}, \mathbf{y}^{\text{UB}}, \mathbf{y}^{\text{LB}}$ are the upper and lower bounds of $\boldsymbol{\delta}$ and those of $\mathbf{y}$, respectively. The natural bounds of these variables are given by Constraints (6.15) and (6.16). Then following McCormick relaxation approach (McCormick, 1976), which tries to replace the nonlinear terms in the contraints with linear approximations, we let $W_{ij} = (1 - \delta_j) y_{ij}$, and con-

sider the following relaxed problem.

$$\textbf{P3}\,(\psi)\quad \min_{\boldsymbol{\delta},\mathbf{y},\mathbf{W}} \boldsymbol{\lambda}^T\boldsymbol{\delta} \tag{6.18}$$

$$\text{s.t. } y_{ij} = 1 + TW_{ji}\frac{g_{ji}''(1)}{g_{ji}'(1)} + \sum_{j'\neq i}Tg_{jj'}'(1)W_{jj'}, \forall i,j = 1,\ldots,m, \tag{6.19}$$

$$W_{ij} \geq \left(1 - \delta_j^{\text{UB}}\right)y_{ij} - y_{ij}^{\text{LB}}\delta_j + \delta_j^{\text{UB}}y_{ij}^{\text{LB}}, \forall i,j = 1,\ldots,m, \tag{6.20}$$

$$W_{ij} \geq \left(1 - \delta_j^{\text{LB}}\right)y_{ij} - y_{ij}^{\text{UB}}\delta_j + \delta_j^{\text{LB}}y_{ij}^{\text{UB}}, \forall i,j = 1,\ldots,m, \tag{6.21}$$

$$W_{ij} \leq \left(1 - \delta_j^{\text{UB}}\right)y_{ij} - y_{ij}^{\text{UB}}\delta_j + \delta_j^{\text{UB}}y_{ij}^{\text{UB}}, \forall i,j = 1,\ldots,m, \tag{6.22}$$

$$W_{ij} \leq \left(1 - \delta_j^{\text{LB}}\right)y_{ij} - y_{ij}^{\text{LB}}\delta_j + \delta_j^{\text{LB}}y_{ij}^{\text{LB}}, \forall i,j = 1,\ldots,m, \tag{6.23}$$

$$\mathbf{y}^{\text{LB}} \leq \mathbf{y} \leq \mathbf{y}^{\text{UB}}, \tag{6.24}$$

$$\boldsymbol{\delta}^{\text{LB}} \leq \boldsymbol{\delta} \leq \boldsymbol{\delta}^{\text{UB}}. \tag{6.25}$$

Upon solving this problem, we are able to obtain a lower bound to the objective of P1 within the solution space defined by (6.12), (6.24), and (6.25). However, using the natural bounds of (6.15) and (6.16) usually over-relaxes the problem and yields a poor lower bound to P1. To address this issue, we adopt the branch and bound method to provide tighter bounds to the problem so as to improve the lower bound to P1. The algorithm is summarized in Algorithm 3.

Given an instance ψ, finding a feasible solution can be done by looking for a minimum stepsize $\theta \in [0,1]$ such that the solution $\mathbf{y}$ to $LP\left(\theta\boldsymbol{\delta}^{\text{LB}} + (1-\theta)\boldsymbol{\delta}^{\text{UB}}\right)$ satisfies Constraints (6.15). This can be done efficiently via bisection search.

The convergence to the optimal solution of Algorithm 3 is guaranteed, as branch and bound method iteratively partitions the solution space and provides McCormick relaxation with increasingly tighter variable bounds, so that the lower bound could keep approaching the global optimum over iterations (Mitsos et al., 2009). To improve the performance of this algorithm, several treatments are adopted. First, instead of finding a new feasible solution in Step 3.2 in each iteration, it is done with a probability P_n, which is set to be increasing over iterations. This is to reduce the computation time, and it does not hinder the optimal solution to be found eventually. This is because, even if the instance that contains the optimal solution is not used to construct a feasible solution at the current iteration, one of its partitioned instance must still contain the optimal solution and will be added to the instance set H. As the lower bound approaches the optimum over iterations, this instance must be visited again. This process will be repeated until the optimality is attained.

Moreover, before solving for the relaxed problem for newly partitioned instances, Step 3.4 checks the new instance feasibility by evaluating the performance of the largest possible vaccine allocation given the upper bound, and Step 3.5 evaluates the objective of the new instance's lower bound. Only if a new instance has a feasible upper bound and a lower bound better than the best

77

Algorithm 3 P2 solver

Step 3.1 Check if the solution $\mathbf{y}_0$ to $LP(\mathbf{0})$ satisfies Constraints (6.15). If yes, return $\boldsymbol{\delta} = \mathbf{0}$ as the solution and terminate; otherwise, initialize $n \leftarrow 0$, solve $LP(\mathbf{1})$ to get $\tilde{\mathbf{y}}_0$; let $\psi \leftarrow (\mathbf{1}, \mathbf{0}, \mathbf{\Lambda}, \tilde{\mathbf{y}}_0)$, solve P3$(\psi)$ to get the corresponding lower bound, let H $\leftarrow \{\psi\}$; initialize the upper bound z_{UB} at N;

Step 3.2 Select an instance ψ_n from H with the lowest lower bound, denoted as z_{LB}; with probability P_n, construct a feasible solution using ψ_n; if this solution leads to an objective lower than the upper bound z_{UB}, update the best known solution as z_{UB} as well;

Step 3.3 Select a variable δ_i where $i = \mathrm{argmax}_j\, \delta_{\psi_n,j}^{\mathrm{UB}} - \delta_{\psi_n,j}^{\mathrm{LB}}$, and create two new instances ψ_{n_1} amd ψ_{n_2} such that both of them inherit the bounds $\left(\boldsymbol{\delta}_{\psi_n}^{\mathrm{UB}}, \boldsymbol{\delta}_{\psi_n}^{\mathrm{LB}}, \mathbf{y}_{\psi_n}^{\mathrm{UB}}, \mathbf{y}_{\psi_n}^{\mathrm{LB}}\right)$ from ψ_n except for that $\delta_{\psi_{n_1},i}^{\mathrm{LB}} \leftarrow \left(\delta_{\psi_n,i}^{\mathrm{UB}} + \delta_{\psi_n,i}^{\mathrm{LB}}\right)/2$, and $\delta_{\psi_{n_2},i}^{\mathrm{UB}} \leftarrow \left(\delta_{\psi_n,i}^{\mathrm{UB}} + \delta_{\psi_n,i}^{\mathrm{LB}}\right)/2$; remove ψ_n from H;

Step 3.4 If the solution $\tilde{\mathbf{y}}_n$ to $LP\left(\boldsymbol{\delta}_{\psi_n}^{\mathrm{UB}}\right)$ satisfies Constraints (6.15), update $\mathbf{y}_{\psi_n}^{\mathrm{LB}} \leftarrow \tilde{\mathbf{y}}_n$; solve P3$(\psi_{n_2})$ to obtain the lower bound of instance ψ_{n_2}; if this lower bound is no greater than the upper bound z_{UB}, let H $\leftarrow$ H $\cup \{\psi_{n_2}\}$;

Step 3.5 If $N\boldsymbol{\lambda}^T\boldsymbol{\delta}_{\psi_{n_1}}^{\mathrm{LB}}$ is smaller than the upper bound z_{UB}, solve P3(ψ_{n_1}) to obtain the lower bound of instance ψ_{n_1}; if this lower bound is no greater than the upper bound z_{UB}, let H $\leftarrow$ H $\cup \{\psi_{n_1}\}$;

Step 3.6 Find the new lowest lower bound z_{LB} in H, and compute the optimality gap as $\frac{z_{\mathrm{UB}} - z_{\mathrm{LB}}}{z_{\mathrm{B}}} \times 100\%$; if the optimality gap is smaller than a threshold ϵ_1, or the maximum iteration number n_{MAX} is reached, terminated and return the best known solution and the gap; otherwise, let $n \leftarrow n + 1$ and repeat Step 3.2 - Step 3.6.

---

known solution, McCormick relaxation is performed on this instance. This is also to improve the algorithm performance, as feasibility check and objective evaluation take negligible computation time compared to solving the relaxed problem.

With Algorithm 3, we can find the minimum vaccine resource required to eliminate giant components in the population $Q_{\min} = N\boldsymbol{\lambda}^T\hat{\boldsymbol{\delta}}$, as well as a corresponding solution $\left(\hat{\boldsymbol{\delta}}, \hat{\mathbf{y}}\right)$ with known optimality gap. As long as the total vaccine resource is no smaller than $Q_{\min}$, this solution also provides us with a good starting point to solving P1. Without loss of generality, in the following discussions we assume the minimum resource requirement is always satisfied.

6.3.3 Solution approach to P1

Yet with a feasible solution constructed using $\left(\hat{\boldsymbol{\delta}}, \hat{\mathbf{y}}\right)$ knowing that there are sufficient vaccine resources, it is still difficult to solve P1 directly because of the non-convexity of the solution set.

Although a similar approach as the branch and bound in the previous subsection can be devised, (6.14) contain cubic terms and requires an extra step of McCormick relaxation, which makes the convergence speed of the algorithm an even greater issue than in Algorithm 3. To solve P1 efficiently, we propose a modified conditional gradient algorithm as a heuristic solution approach.

It is natural to deem the vaccine allocation variable $\boldsymbol{\delta}$ as the only decision variable in the problem, while $\mathbf{y}$ and $\mathbf{s}$ are intermediate results that solely depend on $\boldsymbol{\delta}$. As such, the number of decision variables is reduced from $\mathcal{O}\left(m^2\right)$ to $\mathcal{O}\left(m\right)$. Moreover, the constraints directly on $\boldsymbol{\delta}$, namely (6.4) and (6.16), are both linear inequalities that can be handled easily. However, doing so also leads to two issues: i) given a vaccine allocation decision $\boldsymbol{\delta}$ that satisfies (6.4) and (6.16), a system of linear equations, namely $LP\left(\boldsymbol{\delta}\right)$, needs to be solved for feasibility check; and ii) the gradient of the objective to $\boldsymbol{\delta}$ is unavailable.

To address these issues and solve the problem, we first adopt an approximate gradient, denoted by $\mathbf{f} \in \mathbb{R}^m$, to reflect the marginal gain when the decision variable $\boldsymbol{\delta}$ is changed by 1 unit. Given a feasible solution $(\boldsymbol{\delta}, \mathbf{y}, \mathbf{s})$, $\mathbf{f}$ is computed as follows:

$$f_i = -w_i \left(1 + \sum_{j=1}^m T\left(1 - \delta_j\right) g'_{ij}\left(1\right) y_{ij}\right) - \sum_{j=1}^m w_j T\left(1 - \delta_j\right) g'_{ji}(1) y_{ji}, \forall i = 1, \ldots, m. \quad (6.26)$$

This approximation is obtained assuming $\mathbf{y}$ does not change with $\boldsymbol{\delta}$. We conjecture that the error induced by such approximation is negligible around the local area of the current solution. Then we can use conditional gradient method such that a vaccine allocation decision $\tilde{\boldsymbol{\delta}}$ satisfying the linear constraints (6.4) and (6.16) can be found. Next, $\tilde{\boldsymbol{\delta}}$ is used in a line search algorithm to update the vaccine allocation decisions and improve the objective. This process is repeated until convergence. The solution approach is presented in Algorithm 4.

In Step 4.1, a convex combination of $\hat{\boldsymbol{\delta}}$ and $\mathbf{1}$ is used as the initial solution of $\boldsymbol{\delta}$ to exploit all the available vaccines. This solution guarantees feasibility and performs reasonably well. Step 4.2 tries to find a search direction that could hopefully improve the objective by solving a linear program, and Step 4.3 performs a line search to update the solution. As such, the non-linear constraints are only encountered in the line search problem, which can be easily handled by one-dimensional enumeration. Unfortunately, it is difficult to prove the convergence of the algorithm. Yet its performance is tested in several numerical experiments, and the quality of the solutions are validated to some extent.

Algorithm 4 P1 solver

Step 4.1 Initialize $n \leftarrow 0$, and construct the initial feasible solution $\left(\boldsymbol{\delta}^{(n)}, \mathbf{y}^{(n)}, \mathbf{s}^{(n)}\right)$ by letting
$$\boldsymbol{\delta}^{(n)} = \frac{N-Q}{N-Q_{\min}}\hat{\boldsymbol{\delta}} + \left(1 - \frac{N-Q}{N-Q_{\min}}\right)\mathbf{1}.$$

Step 4.2 Compute the approximate gradient $\mathbf{f}^{(n)}$ at $\left(\boldsymbol{\delta}^{(n)}, \mathbf{y}^{(n)}, \mathbf{s}^{(n)}\right)$; find a feasible solution $\tilde{\boldsymbol{\delta}}^{(n)}$ by solving the following linear program:

$$\min_{\boldsymbol{\delta}} \mathbf{f}^{(n)^T}\boldsymbol{\delta} \tag{6.27}$$

s.t. (6.4) and (6.16).

Step 4.3 Find the optimal stepsize $t^{(n)}$ to the following problem:

$$\min_{t, \boldsymbol{\delta}, \mathbf{y}, \mathbf{s}} \mathbf{w}^T\mathbf{s} \tag{6.28}$$

$$\text{s.t. } \boldsymbol{\delta} = (1-t)\boldsymbol{\delta}^{(n)} + t\tilde{\boldsymbol{\delta}}^{(n)}, \tag{6.29}$$

$$0 \leq t \leq 1, \tag{6.30}$$

and (6.12), (6.14), (6.15).

Let $\boldsymbol{\delta}^{(n+1)} \leftarrow \left(1 - t^{(n)}\right)\boldsymbol{\delta}^{(n)} + t^{(n)}\tilde{\boldsymbol{\delta}}$, and construct the corresponding solution $\left(\boldsymbol{\delta}^{(n+1)}, \mathbf{y}^{(n+1)}, \mathbf{s}^{(n+1)}\right)$.

Step 4.4 If any of the following conditions are met, terminate and return the solution $\left(\boldsymbol{\delta}^{(n+1)}, \mathbf{y}^{(n+1)}, \mathbf{s}^{(n+1)}\right)$:

- $t^{(n)} \leq \epsilon_2$;
- $\left|\boldsymbol{\delta}^{(n)} - \tilde{\boldsymbol{\delta}}^{(n)}\right| \leq \epsilon_3$;
- n reaches the maximum iteration number n'_{MAX}.

where ϵ_2 and ϵ_3 are convergence thresholds. Otherwise, let $n \leftarrow n + 1$ and repeat Step 4.2 - Step 4.4 .

6.4 Numerical Experiments

In this section, we demonstrate the capability of the model and the proposed solution algorithms through numerous numerical experiments. We first consider a toy case with 10 subregions to test the optimization framework and the solution approaches, and perform sensitivity analysis to further validate the results and to reveal operational insights.

6.4.1 10-subregion toy case

We consider a hypothetical case where a population of $N = 1 \times 10^6$ are distributed among $m = 10$ subregions in an area. The fraction population in subregion i, $\lambda_i = \frac{i}{55}$. The degree distribution between any pair of subregions is assumed to follow Poisson distribution with mean degree $g'_{ij}(1) = 3.0, \forall i \leq j$, and $g'_{ij}(1)$'s with $i > j$ are computed from (6.2). Therefore, subregions with a smaller index is associated with a smaller population size, but higher average degrees to other subregions, and the average degree across the entire network is 21. The disease transmissibility T is set to be 0.05 such that there will be a significant fraction of population infected if not vaccinated. The probability of the disease starting in any subregion is equal, i.e. $w_i = \frac{1}{m}, \forall i = 1, \ldots, m$. We also set $\Lambda = 30$, such that $N\lambda_i \gg \Lambda, \forall i$. The other parameters in the solution algorithms are given as follows: $\epsilon_1 = 0.05$, $n_{\text{MAX}} = 2 \times 10^5$, $\epsilon_2 = 1 \times 10^{-8}$, $\epsilon_3 = 1 \times 10^{-10}$, and $n'_{\text{MAX}} = 100$.

We first find the minimum required vaccine resources by solving P2 with Algorithm 3. The algorithm is run on a laptop computer with an Intel Core i5-7300U CPU @ 2.60GHz, and the computation is finished within 2 minutes with an optimality gap of 3.8% and an objective of 6.93×10^4 units of vaccines. The corresponding vaccine allocation decision is presented in Figure 6.2. For result validation, we also use a meta-heuristic algorithm known as basin hopping algorithm to solve P2 and compare the results. Python Scipy package built-in function *basinhopping* is used. After 6 minutes and 20 iterations, the basin hopping algorithm yields an objective of 6.72×10^4 units of vaccines, which is 3.0% lower than the objective obtained via Algorithm 3. The corresponding vaccine allocation is presented in Figure 6.2 for comparison as well. The vaccine allocation decisions from both approaches are highly similar with tolerable differences. Due to their higher network connectivities and fewer population, the subregions with smaller indices are vaccinated at a higher rate. Such results are consistent with our expectations.
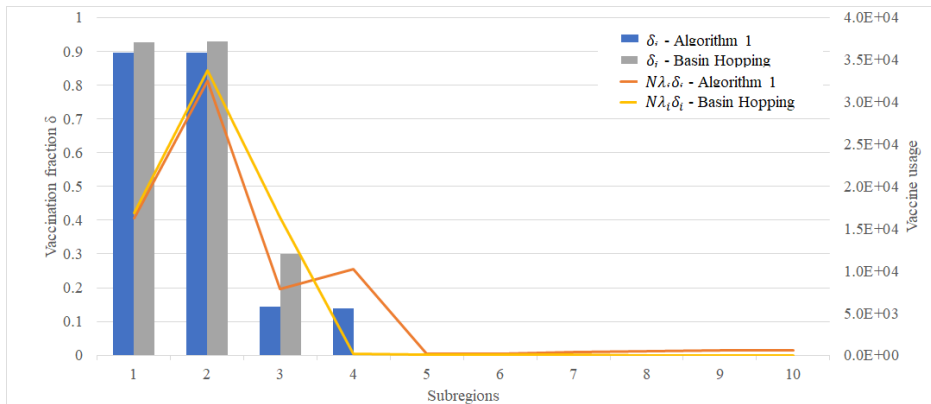

Figure 6.2: Vaccine allocation decisions obtained using Algorithm 1 and basin hopping, respectively.

Next, with the solution to P1 obtained via Algorithm 3, we proceed to solve P1 given a number

of vaccine amount $Q = 1.0 \times 10^5$, which guarantees that the problem is feasible. At the optimal solution, 10% entire population will be vaccinated, and the expected outbreak size will be minimized. Algorithm 4 finishes in 10 seconds on the same computer. The vaccine allocation decision presented in Figure 6.2 yields of objective of 18.5, i.e. approximately 18.5 infections is expected to be caused by a single initial seed. This objective is reduced to 10.0 at the solution given by Algorithm 4, and the resource allocation decision is presented in Table 6.1. It suggests to allocate all the vaccine resources to the subregions with smaller indices. This is because not only these nodes have higher connectivities, but also that as the outbreak probability w_i's are equal $\forall i$ while the population density are different, these individuals have a higher probability to be an initial seed, and thus they are vaccinated with higher priorities.

6.4.2 Sensitivity analysis

We proceed to perform sensitivity analysis based on the toy case. We examine the impact of different resource amount Q and outbreak seed distribution $\mathbf{w}$ on the optimal vaccine allocation, and also investigate the impact of different disease transmissibility T and degree distribution on both the minimum required vaccine problem and the allocation problem. For convenience, we denote the benchmark case of the minimum required vaccine problem Scenario 0, and that of the allocation problem Scenario 0.o. Scenarios 0.i and 0.ii consider different vaccine resources compared to the benchmark case, with $Q = 1.5 \times 10^5$ and $Q = 2.0 \times 10^5$, respectively; 0.iii and 0.iv have different probability distributions of the disease seed location, given as $w_i = \lambda_i, \forall i$, and $w_i = \lambda_i^2 / \sum_j \lambda_j^2, \forall i$, respectively. Scenarios 1 and 2 examines different disease transmissibilities with $T = 0.075$ and $T = 0.025$, respectively; Scenarios 3, 4, and 5 assume different degree distributions among the population, whose details will be given shortly. The minimum required vaccine problem is solved for each of these cases, and the optimal vaccine allocation is solved for the cases o-iv, respectively. For example, the Scenario 1.ii considers the optimal allocation problem with $T = 0.075$ and $Q = 2.0 \times 10^5$. The results are summarized in Table 6.1.

Comparing Scenarios 0, 1, and 2, we find that the results are highly sensitive to the disease transmissibility. With a 50% increase of disease transmissibility T, the minimum require resources is increased by over 360%; meanwhile with a -50% change in T, an epidemic outbreak would not occur even without any vaccination. Consequently, the instances in Scenarios 1.o - 1.iv become infeasible as there are insufficient vaccines to prevent an epidemic outbreak, and the results of Scenarios 2.o - 2.iv show the disease propagation can be very well controlled with the given amount of vaccines. Furthermore, comparing Scenarios o, iii, and iv in both cases 0 and 2, we observe that the higher probability that the epidemic starts in larger population subregions, the larger the expected outbreak size. This is a result of the contradicting effects of low connectivity against high

Table 6.1: Sensitivity analysis results of the minimum vaccine resource and the optimal allocation problems.

Scenarios		1	2	3	4	5	6	7	8	9	10	Computation time (min.)	Gap (%)	Obj. (×10⁴)	P1 Obj. (#)
	Vaccination fraction δ_i (×10⁻¹)												P2		
0. (Benchmark)															
	Min. required vaccines (P2)	8.9	8.9	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	3.8	6.93	19.3
	o) (Benchmark)	10.0	10.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	-	-	7.8
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	10.0	10.0	10.0	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.2	-	-	3.9
	ii) ($Q = 2.0 \times 10^5$)	10.0	10.0	10.0	10.0	2.0	0.0	0.0	0.0	0.0	0.0	0.2	-	-	2.5
	iii) ($w_i = \lambda_i$)	10.0	10.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	-	-	9.4
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	9.0	9.0	1.7	1.7	0.3	0.3	0.4	0.4	0.4	0.4	0.1	-	-	12.0
1. ($T = 0.075$)															
	Min. required vaccines (P2)	9.2	8.4	8.4	8.4	8.4	3.4	0.9	0.9	0.8	0.8	233.8	16.6	32.05	19.2
	o) (Benchmark)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
	ii) ($Q = 2.0 \times 10^5$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
	iii) ($w_i = \lambda_i$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
2. ($T = 0.025$)															
	Min. required vaccines (P2)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
	o) (Benchmark)	10.0	10.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	-	-	1.3
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	10.0	10.0	10.0	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	1.1
	ii) ($Q = 2.0 \times 10^5$)	10.0	10.0	10.0	10.0	2.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	0.9
	iii) ($w_i = \lambda_i$)	10.0	10.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	1.6
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.2	0.1	-	-	1.7
3. (Homogeneous connectivity)															
	Min. required vaccines (P2)	6.5	1.5	0.2	0.1	0.1	0.0	0.0	0.0	0.0	0.0	201.3	5.0	7.94	25.9
	o) (Benchmark)	10.0	10.0	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	13.0
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	10.0	10.0	10.0	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	6.0
	ii) ($Q = 2.0 \times 10^5$)	10.0	10.0	10.0	10.0	2.0	0.0	0.0	0.0	0.0	0.0	0.1	-	-	3.6
	iii) ($w_i = \lambda_i$)	5.6	1.4	0.4	1.4	1.4	0.6	0.6	0.6	0.6	1.4	0.1	-	-	16.4
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.5	0.1	-	-	15.6
4. (Homogeneous connectivity with higher inner-subregion connectivity)															
	Min. required vaccines (P2)	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.6	0.7	3.1	4.5	5.0	8.00	22.4
	o) (Benchmark)	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.8	0.9	3.3	0.1	-	-	14.3
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	0.8	0.8	0.8	0.8	0.8	0.8	0.8	1.3	1.4	3.7	0.1	-	-	7.2
	ii) ($Q = 2.0 \times 10^5$)	1.3	1.3	1.3	1.3	1.4	1.4	1.4	1.8	1.9	4.0	0.1	-	-	4.7
	iii) ($w_i = \lambda_i$)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	2.1	3.3	55.7	-	-	13.2
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	1.6	3.5	0.6	-	-	13.7
5. (Distance-based connectivity)															
	Min. required vaccines (P2)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	2.6	4.3	1.0	11.7	5.0	12.89	15.1
	o) (Benchmark)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
Opt. vaccine allocation (P1)	i) ($Q = 1.5 \times 10^5$)	0.3	0.3	0.3	0.3	0.3	0.4	1.2	2.6	4.0	1.1	0.2	-	-	9.1
	ii) ($Q = 2.0 \times 10^5$)	0.8	0.8	0.8	0.8	0.8	1.2	2.2	3.1	4.0	1.4	0.2	-	-	4.7
	iii) ($w_i = \lambda_i$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible
	iv) ($w_i = \frac{\lambda_i^2}{\sum_j \lambda_j^2}$)	-	-	-	-	-	-	-	-	-	-	-	-	-	Infeasible

disease source probability and large population of the subregions with high indices, which makes it more difficult to effectively allocate the vaccines.

Scenario 3 assumes there is no spatial heterogeneity in population connectivity such that all individuals in the population are probabilistically identical with a total expected degree of 21, and they are connected randomly independent on their subregions, namely $g_{ij} = 21\lambda_j, \forall i, j = 1, \ldots, m$. It requires a larger amount of vaccines to prevent an epidemic outbreak compared to Scenario 0, because there are not subregions with higher connectivity but fewer population which

make targets for vaccination. Therefore, the vaccine allocation decision of P2 in this scenario is more dispersed compared to the benchmark case - in fact any resource allocation is equally effective in this case as long as the total vaccine amount is the same. It also leads to a larger expected outbreak size, as the disease source probabilities are higher in regions with smaller indices. The results of Scenario 2.iii and 2.iv seem to contradict our earlier findings, as the expected outbreak size in 2.iv is slightly lower than that in 2.iii. This is because while the connectivity is homogeneous, in 2.iv subregions that have higher disease source probabilities per capita can be vaccinated with higher priorities, which helps improve the result. Therefore, spatial connectivity and outbreak source distribution together could make the vaccine allocation problem quite complex.

In Scenario 4, the populations are better connected within each subregion, while their connectivities with nodes in other subregions are homogeneous. To be specific, $g'_{ij}(1) = c_1\lambda_j, \forall j \neq i$ and $g'_{ii}(1) = 10.0c_1\lambda_i^{1.5}, \forall i$, where c_1 is a normalization factor such that the average degree of the entire population equals 21.0. As such, the population are better connected within each subregion compared to Scenario 3, and the greater population the better inner-connectivity. In this case, subregions with higher populations have higher total degrees and are thus assigned with more vaccine resources. The minimum required vaccine is lower than Scenario 3, and higher than Scenario 0, which is expected. With more vaccine resources, the expected outbreak size can also be significantly reduced, as suggested by the results of 4.i and 4.ii. In Scenario 4.iii, with a homogeneous disease source probability, the vaccine allocation decision is to focus more on the subregions with higher connectivities. The results of 4.iv, however, appears to be suboptimal, as with higher source probabilities, subregions with large indices (e.g. 8, 9, and 10) ideally should have been allocated with even more resources than they are in Scenario 4.iii, and the objective should be further reduced. This is indeed the case when the vaccine allocation decision δ in Scenario 4.iii is taken as initial solution in Algorithm 4, suggesting that the algorithm converged to a suboptimal solution while solving the problem for Scenario 4.iv.

Finally, Scenario 5 considers distance-based connectivity where population in each subregion are more closely connected with others in index-wise closer subregions. The average degrees are given as: $g'_{ij}(1) = c_2e^{i-j}\lambda_j, \forall i \leq j$, and $g'_{ij}(1)$ where $i > j$ are computed from (6.2), and c_2 is a normalization factor to ensure the overall average degree equals 21. As expected, the required vaccine resources to prevent the epidemic is greater than the previous cases except for Scenario 1. This is straightforward because in this case, subregions with high populations not only have high inner-connectivities, but they are also better connected with each other because their are closer in distance - this is somewhat similar to an assortative mixing population where high degree individuals are linked with higher probabilities. Therefore, it requires large amount of vaccine resources to be used on these subregions to effectively protect the population. Interestingly, subregion 10 is not vaccinated as much as subregions 8 and 9 even though it has the highest inner-

connectivity. This is likely due to that individuals in subregions 8 and 9 have higher total mean degrees than those in 10, and they are the only ones closely connected to 10, it is sufficient to vaccinate a greater fraction of population in these subregions 8 and 9 to preserve the population in 10. Scenarios 5.o, 5.iii, and 5.iv do not have sufficient vaccine resources and are thus infeasible.

Through these experiments, we find that the computation time of both Algorithms 3 and 4 heavily rely on the problem instance. The branch-and-bound-based approach Algorithm 3 can take a tremendous amount of time to converge for some problem instances, which is widely known in the literature. Yet these approaches are still often used for its theoretical guarantee of convergence and being able to provide the optimal gap to validate the solution quality. Algorithm 4 converges with acceptable time in most cases. However, it may converge to some local optimal solutions, which is unfortunately a common drawback of gradient-based approaches in solving non-convex problems. Using different initial feasible solutions as the starting point may help resolve such issues.

CHAPTER 7 – FUTURE RESEARCH

Building upon the modeling framework and preliminary findings from this dissertation, there are several possible future research directions regarding disruption/disease propagations and control in the networked systems.

First, regarding the interdependent infrastructure system, the proposed model can serve as a building block to identify the most critical infrastructure in a given system. Based on this, we can further seek geospatially informed reinforcement or interdiction strategies that will best protect an urban system as a whole or certain vulnerable communities within it. Second, our model assumes that transportation network is rather independent on other infrastructures. Incorporating the interdependencies between transportation network and others will make the model more applicable to situations where transportation network efficiency is influenced by others. Moreover, our modeling framework focuses on the steady state disruption aftermath; the dynamic process of failure propagation, possibly based on Monte-Carlo simulation, may help reveal more insightful and realistic results. In addition, for effective use of this type of model for planning purposes, it will be critical to better understand structural differences in the input data that define the joint infrastructure-community network (e.g., with consideration of communication layer and its interactions with other layers). Finally, we could also use similar approaches for infectious disease modeling and control to investigate complex interdependent infrastructure systems.

On epidemic dynamics and control in a heterogeneous population, there are also a variety of potential research topics. This dissertation mainly considers a fast-spreading disease propagates in a heterogeneous mixing population, and Chapter 4 and Chapter 6 further assume limited resources for epidemic control and consider population heterogeneities in terms of social groups and spatial distribution, respectively. Therefore, when the disease's spreading speed is insignificant as compared to the population birth/death rates, or the individuals change their contact network as a response to the infectious disease, a different dynamics model would be required. Moreover, the configuration model random network largely neglects real-world clustering effects among human contact networks. A different network representation is needed to model the epidemics process in clustered networks, and the optimal control strategies might be different as well .

Furthermore, in Chapter 5, while investigating the population's vaccination behavior under

vaccine-phobia, we made several assumptions: i) the vaccine resources are always abundant, ii) vaccine uptake always guarantees immunity, and iii) the population share similar perceptions toward vaccination. It would thus be very interesting to model a vaccination game where i) people need to compete for limited supply of vaccines, and/or ii) vaccines do not function as expected, or even directly cause morbidity (e.g. Chen and Fu (2019) and Wu et al. (2011)), and/or iii) different vaccine-phobia levels exist among different social groups of population.

On the other hand, vaccine distribution cost/time may largely depend on the distance from distribution centers to destinations, which is however neglected in Chapter 6. It might be interesting to investigate a scenario where there is a limit fund for distribution and/or the time required to deliver the vaccine to some destinations may result in delayed vaccination. Furthermore, in this dissertation we do not cover another potential vaccination strategy, which is to target the individuals who have contacts with the infected ones. This approach is also worth investigating, as targeting the neighbors of the infected individuals can effectively cutoff the disease transmission links and mitigate disease propagation.

Another important issue need to be addressed in the future is real-world data collection. A lot of existing work have been conducted to estimate the disease transmission and removal rates. Moreover, many studies have also tried to understand social networks using survey or social media data. However, these models are not necessarily suitable for modeling infectious disease transmissions (e.g., because of the difference between physical contact and online communications). Therefore, it is imperative to design a proper method to fully understand the population disease transmission contact network structure. Furthermore, population's vaccine-phobia level and the social penalties caused by epidemic control measures also need to be adequately captured.

Finally, it is important to validate these models using real-world observations. If the real-world data is consistent with the model prediction, it suggests the modeling framework could be useful for the decision-makers for application. Otherwise, certain important factors must be been unfortunately neglected in our models, and it is imperative to investigate these omitted factors - this does not damage the value of the existing work, as revealing the key factors in the disease propagation process could also be a potential contribution of this dissertation.

REFERENCES

AGC, 2015. Urban tactical planner URL: http://www.agc.army.mil/Media/FactSheets/FactSheetArticleView/tabid/11913/Article/480931/urban-tactical-planner.aspx. data retrieved from Army Geospatial Center.

Ahmed, S., Mendoza, D., 2014. Ebola hysteria: An epic, epidemic overreaction. CNN 20. URL: www.cnn.com/2014/10/20/health/ebola-overreaction/index.html. accessed: 2018-08-24.

Anderson, R.M., May, R.M., 1992. Infectious diseases of humans: dynamics and control. Oxford university press, Oxford.

Bacaër, N., Guernaoui, S., 2006. The epidemic threshold of vector-borne diseases with seasonality. Journal of Mathematical Biology 53, 421–436.

Bansal, S., Pourbohloul, B., Meyers, L.A., 2006. A comparative analysis of influenza vaccination programs. PLoS Medicine 3, e387.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A., 2005. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. Journal of Theoretical Biology 235, 275–288.

Bauch, C.T., 2005. Imitation dynamics predict vaccinating behaviour. Proceedings of the Royal Society of London B: Biological Sciences 272, 1669–1675.

Bauch, C.T., Earn, D.J., 2004. Vaccination and the theory of games. Proceedings of the National Academy of Sciences of the United States of America 101, 13391–13394.

Bauch, C.T., Galvani, A.P., Earn, D.J., 2003. Group interest versus self-interest in smallpox vaccination policy. Proceedings of the National Academy of Sciences 100, 10564–10567.

Bayram, V., Tansel, B.Ç., Yaman, H., 2015. Compromising system and user interests in shelter location and evacuation planning. Transportation Research Part B: Methodological 72, 146–163.

Bhaskar, V., Lallement, P., 2010. Modeling a supply chain using a network of queues. Applied Mathematical Modelling 34, 2074–2088.

BPR, 1964. Traffic assignment manual for application with a large, high speed computer. Washington : U.S. Dept. of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division.

Buldyrev, S.V., Parshani, R., Paul, G., Stanley, H.E., Havlin, S., 2010. Catastrophic cascade of failures in interdependent networks. Nature 464, 1025.

Calandrillo, S.P., 2003. Vanishing vaccinations: why are so many americans opting out of vaccinating their children. U. Mich. JL Reform 37, 353.

Castilho, C., 2006. Optimal control of an epidemic through educational campaigns. Electronic Journal of Differential Equations (EJDE)[electronic only] 2006, Paper–No.

Chen, X., Fu, F., 2019. Imperfect vaccine and hysteresis. Proceedings of the Royal Society B: Biological Sciences 286, 20182406.

Chor, J.S., Ngai, K.L., Goggins, W.B., Wong, M.C., Wong, S.Y., Lee, N., Leung, T.f., Rainer, T.H., Griffiths, S., Chan, P.K., 2009. Willingness of Hong Kong healthcare workers to accept pre-pandemic influenza vaccination at different WHO alert levels: two questionnaire surveys. Bmj 339, b3391.

Codeço, C.T., 2001. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. BMC Infectious Diseases 1, 1.

Cojocaru, M.G., 2008. Dynamic equilibria of group vaccination strategies in a heterogeneous population. Journal of Global Optimization 40, 51–63.

Di Muro, M., Alvarez-Zuzek, L., Havlin, S., Braunstein, L., 2018. Multiple outbreaks in epidemic spreading with local vaccination and limited vaccines. arXiv preprint arXiv:1805.01564 .

Funk, S., Salathé, M., Jansen, V.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. Journal of the Royal Society Interface , rsif20100142.

Haimes, Y.Y., Horowitz, B.M., Lambert, J.H., Santos, J.R., Lian, C., Crowther, K.G., 2005. Inoperability input-output model for interdependent infrastructure sectors. I: theory and methodology. Journal of Infrastructure Systems 11, 67–79.

Haimes, Y.Y., Jiang, P., 2001. Leontief-based model of risk in complex interconnected infrastructures. Journal of Infrastructure systems 7, 1–12.

Han, D., Sun, M., 2016. An evolutionary vaccination game in the modified activity driven network by considering the closeness. Physica A: Statistical Mechanics and its Applications 443, 49–57.

Hansen, E., Day, T., 2011. Optimal control of epidemics with limited resources. Journal of Mathematical Biology 62, 423–451.

Hawryluck, L., Gold, W.L., Robinson, S., Pogorski, S., Galea, S., Styra, R., 2004. SARS control and psychological effects of quarantine, Toronto, Canada. Emerging Infectious Diseases 10, 1206.

Horner, M.W., Widener, M.J., 2011. The effects of transportation network failure on peopleâĂŹs accessibility to hurricane disaster relief goods: a modeling approach and application to a florida case study. Natural Hazards 59, 1619–1634.

Huang, X., Gao, J., Buldyrev, S.V., Havlin, S., Stanley, H.E., 2011. Robustness of interdependent networks under targeted attack. Physical Review E 83, 065101.

Innes, S., 2015. Doctor says media coverage of ebola âĂŸfanned the hysteriaâĂŹ. Association of Health Care Journalists URL: https://healthjournalism.org/blog/2015/05/doctor-says-media-coverage-of-ebola-fanned-the-hysteria/. accessed: 2018-08-24.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 115, 700–721.

Khan, T., Zaman, G., Chohan, M.I., 2017. The transmission dynamic and optimal control of acute and chronic hepatitis B. Journal of Biological Dynamics 11, 172–189.

Koch, D., Illner, R., Ma, J., 2013. Edge removal in random contact networks and the basic reproduction number. Journal of Mathematical Biology 67, 217–238.

Kossinets, G., Watts, D.J., 2006. Empirical analysis of an evolving social network. Science 311, 88–90.

Lindquist, J., Ma, J., Van den Driessche, P., Willeboordse, F.H., 2011. Effective degree network disease models. Journal of Mathematical Biology 62, 143–164.

Lu, Z., Chi, X., Chen, L., 2002. The effect of constant and pulse vaccination on SIR epidemic model with horizontal and vertical transmission. Mathematical and Computer Modelling 36, 1039–1057.

Ma, J., van den Driessche, P., Willeboordse, F.H., 2013. The importance of contact network topology for the success of vaccination strategies. Journal of Theoretical Biology 325, 12–21.

Madar, N., Kalisky, T., Cohen, R., Ben-avraham, D., Havlin, S., 2004. Immunization and epidemic dynamics in complex networks. The European Physical Journal B 38, 269–276.

McCormick, G.P., 1976. Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems. Mathematical Programming 10, 147–175.

McKee, C., Bohannon, K., 2016. Exploring the reasons behind parental refusal of vaccines. The Journal of Pediatric Pharmacology and Therapeutics 21, 104–109.

Miller, J.C., 2011. A note on a paper by Erik Volz: SIR dynamics in random networks. Journal of Mathematical Biology 62, 349–358.

Miller, J.C., Slim, A.C., Volz, E.M., 2012. Edge-based compartmental modelling for infectious disease spread. Journal of the Royal Society Interface 9, 890–906.

Miller, J.C., Volz, E.M., 2013. Incorporating disease and population structure into models of SIR disease in contact networks. PloS One 8, e69162.

Mitsos, A., Chachuat, B., Barton, P.I., 2009. Mccormick-based relaxations of algorithms. SIAM Journal on Optimization 20, 573–601.

Molloy, M., Reed, B., 1995. A critical point for random graphs with a given degree sequence. Random Structures & Algorithms 6, 161–180.

Motter, A.E., Lai, Y.C., 2002. Cascade-based attacks on complex networks. Physical Review E 66, 065102.

Mukandavire, Z., Garira, W., Tchuenche, J., 2009. Modelling effects of public health educational campaigns on HIV/AIDS transmission dynamics. Applied Mathematical Modelling 33, 2084–2095.

Myers, L.B., Goodwin, R., 2011. Determinants of adults' intention to vaccinate against pandemic swine flu. BMC Public Health 11, 15.

Mylius, S.D., Hagenaars, T.J., Lugnér, A.K., Wallinga, J., 2008. Optimal allocation of pandemic influenza vaccine depends on age, risk and timing. Vaccine 26, 3742–3749.

Newman, M.E., 2002. Spread of epidemic disease on networks. Physical Review E 66, 016128.

Newman, M.E., Strogatz, S.H., Watts, D.J., 2001. Random graphs with arbitrary degree distributions and their applications. Physical Review E 64, 026118.

Nowzari, C., Preciado, V.M., Pappas, G.J., 2016. Analysis and control of epidemics: A survey of spreading processes on complex networks. IEEE Control Systems 36, 26–46.

Odihi, J.O., 1996. Urban droughts and floods in Maiduguri: twin hazards of a variable climate. Berichte des Sonderforschungsbereichs 268, 303–319.

ORNL, 2010. Landscan 2010 high resolution global population data set. URL: http://web.ornl.gov/sci/landscan/. data retrieved from Oak Ridge National Laboratory, UT-Battelle, LLC.

Pang, G., Chen, L., 2007. A delayed SIRS epidemic model with pulse vaccination. Chaos, Solitons & Fractals 34, 1629–1635.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A., 2015. Epidemic processes in complex networks. Reviews of Modern Physics 87, 925.

Pereira, T., Young, L.S., 2015. Control of epidemics on complex networks: Effectiveness of delayed isolation. Physical Review E 92, 022822.

Powell, W.B., 2007. Approximate Dynamic Programming: Solving the curses of dimensionality. volume 703. John Wiley & Sons, Hoboken, NJ.

Reluga, T.C., 2009. An SIS epidemiology game with two subpopulations. Journal of Biological Dynamics 3, 515–531.

Reluga, T.C., Bauch, C.T., Galvani, A.P., 2006. Evolving public perceptions and stability in vaccine uptake. Mathematical Biosciences 204, 185–198.

Reluga, T.C., Galvani, A.P., 2011. A general approach for population games with application to vaccination. Mathematical Biosciences 230, 67–78.

Rinaldi, S.M., Peerenboom, J.P., Kelly, T.K., 2001. Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Systems 21, 11–25.

Sheffi, Y., 1985. Urban transportation networks. volume 6. Prentice-Hall, Englewood Cliffs, NJ.

Smith, P.J., Humiston, S.G., Marcuse, E.K., Zhao, Z., Dorell, C.G., Howes, C., Hibbs, B., 2011. Parental delay or refusal of vaccine doses, childhood vaccination coverage at 24 months of age, and the health belief model. Public Health Reports 126, 135–146.

Smith, R.L., 2016. Proposing a compartmental model for leprosy and parameterizing using regional incidence in Brazil. PLoS Neglected Tropical Diseases 10, e0004925.

Smith, R.L., Gröhn, Y.T., 2015. Use of approximate Bayesian computation to assess and fit models of mycobacterium leprae to predict outcomes of the Brazilian control program. PloS One 10, e0129535.

Srinivasan, V., Motani, M., Ooi, W.T., 2006. Analysis and implications of student contact patterns derived from campus schedules, in: Proceedings of the 12th Annual International Conference on Mobile Computing and Networking, ACM. pp. 86–97.

Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Soldan, V.P., Kochel, T.J., Kitron, U., Elder, J.P., Scott, T.W., 2009. The role of human movement in the transmission of vector-borne pathogens. PLoS Neglected Tropical Diseases 3, e481.

Sun, L.H., 2018. Kids in these u.s. hot spots at higher risk because parents opt out of vaccinations. The Washington Post URL: https://www.washingtonpost.com/news/to-your-health/wp/2018/06/12/kids-in-these-u-s-hotspots-at-higher-risk-because-parents-opt-out-of-vaccinations/?noredirect=on&utm_term=.9dfd0c2ae20b. accessed: 2018-11-26.

The College of Physicians of Philadelphia, 2018. History of anti-vaccination movements URL: https://www.historyofvaccines.org/content/articles/history-anti-vaccination-movements. accessed: 2018-08-04.

Thomas, N., 2018. One-third of US parents plan to skip flu shots for their kids this season. CNN URL: https://www.cnn.com/2018/11/19/health/flu-vaccine-children-study/index.html?no-st=1543269188. accessed: 2018-11-26.

Tien, J.H., Earn, D.J., 2010. Multiple transmission pathways and disease dynamics in a waterborne pathogen model. Bulletin of Mathematical Biology 72, 1506–1533.

Tozzi, A.E., Gesualdo, F., Romano, M., Caione, D., et al., 2009. Parental attitude toward influenza A(H1N1)v vaccination in Italy. Vaccine 27.

Volz, E., 2008. SIR dynamics in random networks with heterogeneous connectivity. Journal of Mathematical Biology 56, 293–310.

Volz, E., Meyers, L.A., 2007. Susceptible–infected–recovered epidemics in dynamic contact networks. Proceedings of the Royal Society of London B: Biological Sciences 274, 2925–2934.

Wang, X., Ouyang, Y., 2013. A continuum approximation approach to competitive facility location design under facility disruption risks. Transportation Research Part B: Methodological 50, 90–103.

Wu, B., Fu, F., Wang, L., 2011. Imperfect vaccine aggravates the long-standing dilemma of voluntary vaccination. PLoS ONE 6, e20577.

Xie, S., Li, X., Ouyang, Y., 2015. Decomposition of general facility disruption correlations via augmentation of virtual supporting stations. Transportation Research Part B: Methodological 80, 64–81.

Xu, D., Xu, X., Xie, Y., Yang, C., 2017. Optimal control of an SIVRS epidemic spreading model with virus variation based on complex networks. Communications in Nonlinear Science and Numerical Simulation 48, 200–210.

Yan, X., Zou, Y., 2008. Optimal and sub-optimal quarantine and isolation control in SARS epidemics. Mathematical and Computer Modelling 47, 235–245.

Zhang, P., Peeta, S., 2011. A generalized modeling framework to analyze interdependencies among infrastructure systems. Transportation Research Part B: Methodological 45, 553–579.

Zhang, Y., 2013. The impact of other-regarding tendencies on the spatial vaccination game. Chaos, Solitons & Fractals 56, 209–215.