Analysis and Control of Contagion Processes on Networks

by

Kimon Drakopoulos

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author .. Signature redacted .....
Department of Electrical Engineering and Computer Science
May 20, 2016

Certified by..... Signature redacted .....
Asuman Ozdaglar
Professor
Thesis Supervisor

Certified by..... Signature redacted .....
John N. Tsitsiklis
Professor
Thesis Supervisor

Accepted by..... Signature redacted .....
Leslie A. Kolodziejski
Professor
Chair, Department Committee on Graduate Students

Analysis and Control of Contagion Processes on Networks

by

Kimon Drakopoulos

Submitted to the Department of Electrical Engineering and Computer Science
on May 20, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Computer Science and Engineering

Abstract

We consider the propagation of a contagion process ("epidemic") on a network and study the problem of dynamically allocating a fixed curing budget to the nodes of the graph, at each time instant. We provide a dynamic policy for the rapid containment of a contagion process modeled as an SIS epidemic on a bounded degree undirected graph with n nodes. We show that if the budget r of curing resources available at each time is $\Omega(W)$, where W is the CutWidth of the graph, and also of order $\Omega(\log n)$, then the expected time until the extinction of the epidemic is of order $O(n/r)$, which is within a constant factor from optimal, as well as sublinear in the number of nodes. Furthermore, if the CutWidth increases sublinearly with n, a sublinear expected time to extinction is possible with only a sublinearly increasing budget r.

In contrast, we provide a lower bound on the expected time to extinction under any such dynamic allocation policy, for bounded degree graphs, in terms of a combinatorial quantity that we call the resistance of the set of initially infected nodes, the available budget, and the number of nodes n. Specifically, we consider the case of bounded degree graphs, with the resistance growing linearly in n. We show that if the curing budget is less than a certain multiple of the resistance, then the expected time to extinction grows exponentially with n. As a corollary, if all nodes are initially infected and the CutWidth of the graph grows linearly in n, while the curing budget is less than a certain multiple of the CutWidth, then the expected time to extinction grows exponentially in n.

The combination of these two results establishes a fairly sharp phase transition on the expected time to extinction (sublinear versus exponential) based on the relation between the CutWidth and the curing budget.

Finally, in the empirical part of the thesis, we analyze data on the evolution and propagation of influenza across the United States and discover that compartmental epidemic models enriched with environment dependent terms have fair prediction accuracy, and that the effect of inter-state traveling is negligible compared to the effect of intra-state contacts.

Thesis Supervisor: Asuman Ozdaglar
Title: Professor

Thesis Supervisor: John N. Tsitsiklis
Title: Professor

To my parents, Silia and Kostas

Acknowledgments

In the clearing stands a boxer, and a fighter by his trade,
And he carries the reminders of every glove that laid him down
And cut him till he cried out, in his anger and his shame,
"I am leaving, I am leaving." But the fighter still remains

Paul Simon

On the first year of my PhD, I met with an alumnus from the Laboratory for Information and Decision Systems (LIDS) who argued that his years at MIT had been the best years of his life. At that point, that statement sounded absurd. Seven years after this conversation, I can now relate to his experience. Being a person who seeks adventure, I admit that the past years have been the most challenging, life-changing and exciting adventure so far. There are several people who made this adventure more didactic, more enlightening and definitely more enjoyable.

First and foremost, I would like to express my gratitude to my phenomenal advisors, Asu Ozdaglar and John Tsitsiklis as well as my committee member Daron Acemoglu. Their guidance throughout my graduate studies made me a better researcher but most importantly a better person. I will always be indebted to them for their support, their ideas and admittedly their patience.

Asu, I do not have enough words to describe how much I owe you. Thank you for nurturing me since my first day at MIT, thank you for exposing me to exciting problems and research directions, thank you for teaching me how to look at the big picture of things, thank you for your enthusiasm, the bright outlook and the "pep talks" in your office that was always open regardless of your busy schedule and countless responsibilities.

John, you have been one of the most influential people in my life. Thank you for introducing me to the magical powers of simplicity, thank you teaching me how to structure my thoughts and my approach to problems. I should even thank you (and I never thought I would say this) for your comments, edits and the suspense of sending you drafts of my work. Interacting with you has been invaluable both for my

academic and my personal maturation.

Daron, there is no need to praise your academic ingenuity or your incredible mentorship, as at this point these are widely recognized. I should thank you, though, for finding the time for me to discuss research problems and life decisions.

During my graduate studies at MIT, I was privileged to serve as a TA for John Tsitsiklis, Patrick Jaillet, Asu Ozdaglar, Polina Goland, Itai Ashlagi and as an instructor for the Leaders for Global Operations. I would like to thank both the instructors and the students for helping me mature as a lecturer and for introducing me to one of the most rewarding aspects of academic life, teaching.

Moreover, I would like to extend my gratitude to the LIDS staff, and in particular Roxana Hernandez, Jennifer Donovan, Lynne Dell, Petra Aliberti and Brian Jones for their help with various administrative tasks as well as for letting me in my office on an almost daily basis. Furthermore, I would like to thank the funding sources of this thesis: Draper Laboratories and the Army Research Office's Multidisciplinary University Research Initiative .

I was extremely fortunate to be part of the stimulating and energetic LIDS community and in particular the Optimization and Network Game Theory, and Systems, Networks and Decisions groups. I would like to thank all members of LIDS and in particular Kostas Bimpikis, Ozan Candogan, Alireza Salehi, Mihalis Markakis, Ali Makhdoumi, Elie Adam, Spyros Zoumpoulis, Paul Njoroge, Ermin Wei, Mert Gürbüzbalaban for being incredible collaborators, friends and coffee-drinking partners.

The past several years would not have been as enjoyable or even bearable if it was not for the company of my dear friends, Ampelale, the Greeks of Boston, and in particular Georgios Angelopoulos with whom we started and ended the PhD journey, my dear friend and roommate Dimitris Chatzigeorgiou, and Yola Katsargyri who was a supporting companion during the emotionally challenging and stressful last PhD years.

Concluding, throughout my life and my PhD studies in particular, I have been spoiled by the support, kindness and generosity of two extraordinary individuals. Mom and Dad, this thesis is dedicated to you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem and Motivation

With infectious diseases frequently dominating news headlines, public health and pharmaceutical industry professionals, policy makers, and infectious disease researchers increasingly need to understand their transmission dynamics, make better predictions, and design effective intervention policies.

Clearly, such contagion processes (processes spreading over contact networks) do not only apply in the context of infectious diseases but also in the context of propagation of information [2], viral marketing [37], spread of computer viruses [23], diffusion of innovations [54] or financial contagion [1].

The theoretical part of this thesis (cf. Chapters 2-6) is concerned with efficient dynamic intervention for the control of such contagion processes, under limited curing resources. Our main motivation comes from infectious disease epidemics, although without aiming at a faithful representation of the details of real-world situations.

A relevant example is the recent outbreak of the Ebola virus which causes an acute and serious illness [67]. Ebola was associated with a high fatality rate in the rural forest communities of Guinea in December 2013 which spiraled into an epidemic that ravaged West Africa and evoke fear around the globe [5]. The virus spreads through human-to-human transmission via direct contact (through broken skin or mucous membranes) with the blood, secretions, organs or other bodily fluids

of infected people, and with surfaces and materials (e.g., bedding, clothing) contaminated with these fluids [67]. However, supplies of experimental medicines, e.g., the prototype drug ZMapp, are limited and "will not be sufficient for several months to come," as stated in [69]. In view of the limited availability of treatment for the virus, [50] poses the following question: "Ebola Drug Could Save a Few Lives. But Whose?".

There are many other examples of communicable diseases including measles, influenza, and tuberculosis. The mechanism of transmission of infections is now known for most diseases and generally they can be split in two categories: (i) diseases transmitted by viral agents, such as influenza, measles, rubella (German measles), and chicken pox which confer immunity against reinfection and (ii) diseases transmitted by bacteria, such as tuberculosis, meningitis, and gonorrhea which do not confer immunity against reinfection.

The wide applicability and major significance of contagion processes has led to extensive work on modeling their evolution and understanding the resulting dynamics [32, 14]. Many models have been proposed in the field of mathematical epidemiology to describe and study infectious diseases [3]. The main characteristic of these models is the presence of an underlying *contact network*. Depending on the context, the network may represent contacts between individuals [31], influence among them ([36], [4]), or influence among different blogs in the blogspace ([2], [37], [26]). Given the network, there are two main approaches to modeling epidemics [1].

(i) **SIR** type models where agents can be in one of three states: *susceptible* to infection, *infected* or *removed* from the system (recovered). These models are used to describe situations where reinfection is not possible.

(ii) **SIS** type models where agents can be in one of two states: *susceptible* to (re)infection or *infected*. These models are used to describe situations where reinfection is possible.

For each agent, transitions between these states depend on the details of the model and may happen deterministically or stochastically. The analysis of these

[1]Many extensions of these models have been proposed in the literature, including more states for agents, such as exposed but asymptomatic, quarantined etc.

models in the literature has been pursued through three different routes of increasing mathematical difficulty and modeling granularity:

(i) *Approximations using a single differential equation:* This approach has been used in the literature to study simplistic deterministic versions of SIR and SIS models. This is the most traditional approach and takes a macroscopic view of the system by focusing on aggregate metrics of infections. The paper [28] provides an excellent survey of this approach.

(ii) *Approximations using systems of differential equations:* This approach has been used in the literature to approximate stochastic versions of SIR and SIS models. Such approaches focus more on the details of the infection state of the system thus providing greater modeling flexibility. See [63] and references therein for a concrete study on the application and accuracy of such *Mean Field Approximations.*

(iii) *Stochastic Analysis of Exact Dynamics:* Results for the general case of stochastic infections and recoveries on arbitrary underlying networks are scarce, mostly due to the complex structure of the problem. See [21, 40] and references therein for main results. The theoretical part of this thesis (cf. Chapters 2-6) provides results using these exact models on arbitrary graphs.

The different models for studying evolution and propagation of epidemics described above have been widely used in the literature due to their tractability and/or insightful interpretation. However, limited work [46] has been done to understand the effectiveness of these models in describing real phenomena.

On the other hand, forecasting of epidemics is an extremely active research area and the approaches that have been developed to predict the spread fall into two categories: time-series modeling [61, 53, 12] and non parametric forecasting ([65] and references therein).

In the empirical part of this thesis (cf. Chapter 7), we use real data on the propagation of influenza related infections in the United States in order to evaluate

these more traditional models by testing their predictive accuracy. Clearly, the scope and purpose of this study is not to improve on the benchmark for epidemic prediction. Instead, we seek to understand whether the traditional epidemiological models are rich enough to fairly describe real contagion phenomena such as influenza propagation and we investigate the effect of inter-state traveling.

1.2 Related literature

Several approaches to the problem of optimal intervention have been proposed, in which the curing rate allocation is static (open-loop) ([13, 27, 11, 52]), and the proposed methods were either heuristic or based on mean-field approximations of the evolution process; see [45] for a survey.

In this thesis, we study the dynamic *control* of contagion processes (from now on called *epidemics*) under limited curing resources. Specifically, we study *dynamic* allocation policies that use information on the underlying structure of contacts and on the *infection state* of individuals, and we evaluate performance in terms of the expected time until the epidemic becomes extinct.

Specifically, our work involves an extension of the canonical SIS epidemic model: the epidemic spreads on the underlying network from an initial set of *infected* nodes to *healthy* nodes and at the same time, infected nodes can be cured. Healthy nodes get infected at a constant and common *infection rate* (that we assume equal to 1) by each of their infected neighbors. In contrast to the standard SIS model, which assumes a common curing rate for all infected nodes at all times, we assume instead a node and time-specific *curing rate.* A *curing policy,* to be applied by a central controller, is a choice, at each time instant, of the curing rates at each node, taking into account the history of the epidemic and the network structure, subject to a budget constraint on the sum of the curing rates applied at each time. We denote the available budget by r. The resulting process is a controlled finite Markov chain with a unique absorbing state: the state where all nodes are healthy. We say that the epidemic becomes *extinct* when that absorbing state is reached. Under mild assumptions (positive total budget

on infected nodes) on the curing budget and given any set of initially infected nodes, the epidemic becomes extinct in a random but finite amount of time. The goal of the network planner is to minimize the expected extinction time subject to the budget constraint.

Several approaches to studying this problem have been proposed in the literature, and below we describe the main contributions.

Static and node-independent policies

Traditionally, the literature has focused on the *uncontrolled* version of the contact process where the *curing rate* is equal to a constant, i.e., the case where all infected nodes receive the same, constant amount of curing. The latter can be considered as a special case of a dynamic curing policy. Several papers and books, such as [40], [49], [20] focus on analyzing the behavior of the expected extinction time for special cases of graphs such as line graphs, star graphs and lattices. The seminal paper [22], on the other hand, obtains strong results on the effect of network topology on the behavior of the expected extinction time, assuming that all nodes are initially infected:

(a) if $r > n\rho(A)$, where $\rho(A)$ is the spectral radius of the graph Laplacian, then the expected extinction time is $O(\log n)$.

(b) If $r < n\eta$, where η is the isoperimetric constant of the graph, then the expected extinction time is exponential in the number of nodes.

More intuitively, all these papers identify two regimes: depending on the parameters and the underlying graph properties, extinction can be *fast*, in which case expected extinction time scales *sublinearly* with the number of nodes, or *slow*, in which case the expected extinction time scales *exponentially* with the number of nodes.

Static and node-specific

As a first approach to node-specific (but still static) policies, the authors of [9] let the curing rates be proportional to the degree of each node and independent of the current state of the network, which may however result in having curing resources wasted

	fast extinction	slow extinction
$\rho_v(t) = \rho$	$r > n\rho(A)$	$r < n\eta$
$\rho_v(t) = d_v$	$r > Cn$	-
any policy	not applicable	expander graphs and $r < C'n$

Table 1.1: Existing Results: Conditions for fast and slow extinction under different curing policies, assuming all nodes initially infected.

on healthy nodes. For bounded degree graphs, the policy in [9] achieves sublinear expected time to extinction (small), but requires a curing budget that is proportional to the number of nodes (large). More precisely, under this more sophisticated control policy, for which $\rho_v(t) = d_v$, where d_v denotes the degree of node v the authors obtain significant improvement in the performance:

(a) if $r > Cn$, where C is an appropriately chosen constant, then the expected extinction time is $O(\log n)$.

(b) If $r < C'n$, where C' is an appropriately chosen constant **and the graph is an expander**, then, for any curing policy, the expected extinction time is exponential in the number of nodes.

Intuitively, this more sophisticated policy achieves fast extinction using total curing resources that scale *linearly* with the number of nodes. Moreover, they argue that if the underlying graph is an expander, then the curing resources required to achieve fast extinction scale linearly with the number of nodes.

The main question addressed in this thesis is whether better performance is achievable by applying *dynamic* and node-specific curing policies. Specifically, we identify conditions and the corresponding dynamic curing policies under which both extinction time and curing budget is small (sublinear).

1.3 Simple examples

In this section we discuss two examples where intuitive dynamic curing policies perform better than the existing policies in terms of required curing budget for fast

Figure 1-1: An efficient dynamic curing policy for the line graph would allocate all curing resources to the right-most infected node.

extinction. Specifically, we will go over two examples: the line graph and the two dimensional square grid.

1.3.1 Example I: the line graph

As discussed in the previous subsection both the constant rate curing as well as the degree based curing (which are almost identical in this case) require total curing budget that scales linearly in the number of nodes to achieve fast extinction.

In contrast, consider a policy which allocates all curing resources to the right-most infected node. In this case the number of infected nodes increases at a rate equal to one (since there is only one edge connecting infected and healthy nodes) and decreases at a rate that is equal to r, the curing budget. Therefore, as long as the curing budget is larger than 1 the expected extinction time can be made linear (compared to exponential for the constant rate curing). Moreover, as long as the curing budget is larger than $n/\log n + 1$ (compared to linear for the degree based curing) the expected extinction time can be made $\log n$.

1.3.2 Example II: the two dimensional grid

As discussed in the previous subsection both the constant rate curing as well as the degree based curing (which are almost identical in this case) require total curing budget that scales linearly in the number of nodes to achieve fast extinction.

In contrast, consider a policy which allocates all curing resources to the first (in the lexicographic order) infected node and assume that the budget is equal to $10\sqrt{n} + n/\log(n)$. Consider a situation where the set of infected nodes is a rectangle, similar to the one depicted in Figure 1-2. In this case the number of infected nodes increases at a rate equal to $\sqrt{n}$ (since there are only $\sqrt{n}$ edges connecting infected and

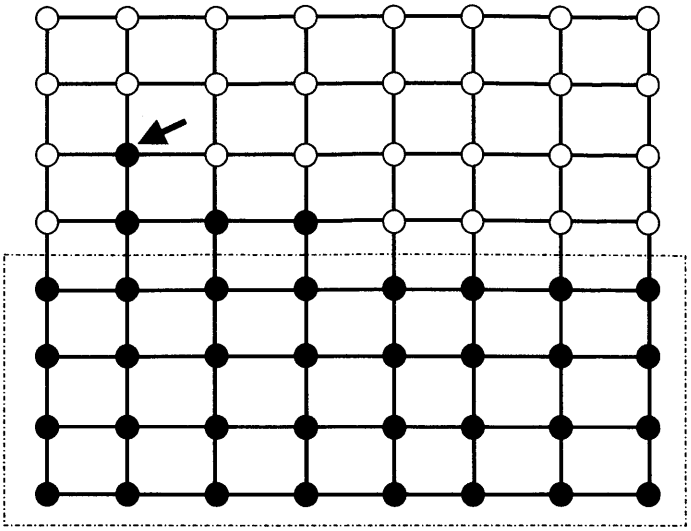


Figure 1-2: An efficient dynamic curing policy for the grid graph would allocate all curing resources to the top-right infected node.

healthy nodes) and decreases at a rate equal to r, the curing budget. If there were no infections, the time until the height of the rectangle decreases by one is equal to $\sqrt{n}/r$. Assuming that there are some infections that are changing the shape of the rectangle (such as the nodes outside the rectangle in Figure 1-2), then all the curing budget is allocated to these nodes. If there is a small number of such infections the number of infected nodes still increases at a rate roughly equal to $\sqrt{n}$. Therefore, since the budget is chosen sufficiently high (equal to $10\sqrt{n} + n/\log(n)$), the process returns with high probability to the rectangle-shape and the time to decrease the height of the rectangle by one is indeed (roughly) equal to a multiple of $\sqrt{n}/r$. Therefore the total extinction time is (roughly) equal to $\sqrt{n} \cdot \sqrt{n}/r$ and hence is $O(\log n)$.

1.4 Contributions of this thesis

The main results of this thesis can be categorized in the following two categories:

I. Theoretical contributions

(i) In Chapter 2 we introduce novel graph theoretic quantities that capture the "hardness to cure" for a given subset of nodes.

(ii) In Chapter 3 we propose a dynamic policy which achieves order-optimal performance (expected extinction time) when the curing budget is sufficiently higher than the CutWidth of the underlying graph. Our results have originally appeared in [15].

(iii) We use this result to show that for bounded degree graphs with small CutWidth (sublinear in the size of the graph), efficient performance (sublinear extinction time) can be achieved *economically*, i.e., by properly allocating a sublinear curing budget, hence demonstrating the increased effectiveness of dynamic policies.

(iv) In Chapters 4 and 5 we establish a converse result for graphs with large CutWidth, namely, for graphs whose CutWidth scales linearly in the number of nodes. In particular, we show that if $r \leq c_r W$, where $c_r > 0$ is an absolute constant (depending only on the degree bound and on c_γ), then, for some initial states, the expected time to extinction is at least exponential, under any curing policy. Our results have originally appeared in [16] and [17].

(v) Using these results we draw an important qualitative distinction between networks in which (i) the spread of the epidemic is hard to stop with the given curing budget, so that the expected time to extinction grows exponentially with the number of nodes, and (ii) the curing resources are adequate, so that the expected time to extinction grows slowly (sublinearly) with the number of nodes.

II. Empirical contributions

(i) We enrich traditional epidemiological models with environment-dependent parameters. Specifically, we include an unknown dependence on absolute humidity to improve existing models and allow for better predictions.

(ii) We develop a recurrent neural network approach to estimate these models. The estimated models have fair predictive accuracy although they are extremely dependent on absolute humidity.

(iii) We use our estimates to evaluate the effect of interstate traveling and discover that the latter is negligible compared to intra-state contacts.

1.5 Structure of the thesis

This thesis is organized as follows. In Chapter 2 we introduce our model and the main problem under consideration. Moreover, we present several combinatorial graph theoretic results that will be widely used throughout the thesis. In Chapter 3 we describe and analyze our dynamic curing policy and obtain performance guarantees. In Chapter 4 we provide a lower bound on the performance of all dynamic policies for graphs with very large CutWidth while in Chapter 5 we obtain a similar result but in the general case of linear CutWidth. In Chapter 6 we summarize our theoretical findings and pose an open problem for future research. Finally, in Chapter 7 we describe the empirical part of the thesis and present our findings.

Chapter 2

Model and Graph Theoretic Preliminaries

In this chapter we introduce our model as well as several important concepts and quantities that will be widely used in the rest of this thesis.

2.1 Controlled Contact Process

We consider a network, represented by an undirected graph $G = (V, E)$, where V denotes the set of nodes and E denotes the set of edges. We use n to denote the number of nodes. Two nodes $u, v \in V$ are *neighbors* if $(u, v) \in E$. We restrict to graphs for which the node degrees are upper bounded by Δ, which we take to be a given constant throughout the thesis.

We let $I_0 \subseteq V$ be a set of intially infected nodes, and assume that the infection spreads according to a controlled contact (or SIS) process, where the rate at which infected nodes get cured is determined by a network controller. Specifically, each node can be in one of two states: *infected* or *healthy*. The controlled contact process is a right-continuous, continuous-time, controlled Markov process $\{I_t\}_{t \geq 0}$ on the state space $\{0, 1\}^V$, where I_t stands for the set of infected nodes at time t. We refer to I_t as the *infection process*. We will sometimes use I_{t-} as a short-hand for the value $\lim_{s \uparrow t} I_s$ just before time t.

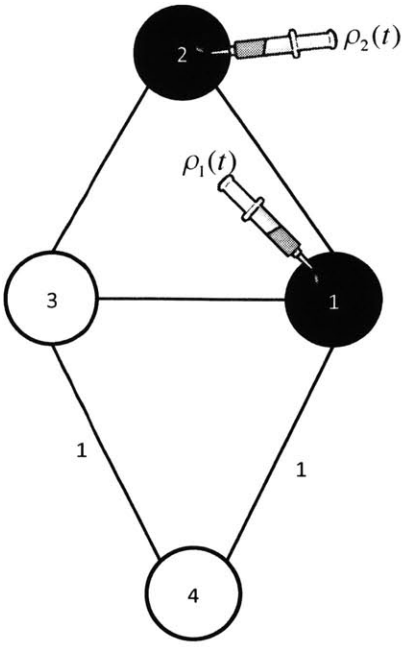


Figure 2-1: The controlled contact process: each node can be either infected (black) or healthy (white). Infected nodes infect their healthy neighbors according to a Poisson process with rate 1. Infected nodes get cured according to a Poisson process with rate $\rho_v(t)$ that is determined by the network controller.

At any point in time, state transitions at each node occur independently, according to the following rates. (These rates essentially define the generator matrix of the continuous-time Markov process under consideration.)

a) The process is initialized at the given initial state I_0.

b) If a node v is healthy, i.e., if $v \notin I_t$, the transition rate associated with a change of the state of that node to being infected is equal to a positive infection rate β times the number of infected neighbors of v, that is,

$$\beta \cdot \big|\{(u,v) \in E : u \in I_t\}\big|,$$

where we use $|\cdot|$ to denote the cardinality of a set. Any transition of this type will be referred to as an *infection*. By rescaling time, we can and will assume throughout the thesis that $\beta = 1$.

c) If a node v is infected, i.e., if $v \in I_t$, the transition rate associated with a change of the state of that node to being healthy is equal to a curing rate $\rho_v(t)$ that is determined by the network controller, as a function of the current and past states of the process. We are assuming here that the network controller has access to the entire history of the process. Any transition of this type will be referred to as a *recovery.*

2.2 Main Problem

So far we discussed the dynamics of infection and curing events. In this section we discuss the problem that the network controller is facing. Specifically, we impose a *budget constraint* of the form

$$\sum_{v \in V} \rho_v(t) \leq r, \tag{2.1}$$

for each time instant t, reflecting the fact that curing is costly.

A *curing policy* is a mapping which at any time t maps the past history of the process to a curing vector $\rho(t) = \{\rho_v(t)\}_{v \in V}$ that satisfies (2.1).

We define the *time to extinction* as the first time when the process first reaches the absorbing state where all nodes are healthy:

$$\tau = \min\{t \geq 0 : I_t = \emptyset\}.$$

In this thesis, we focus on the *expected time to extinction* (the expected value of τ), as the performance measure of interest. At a high level, the network planer is interested in solving the following optimization problem with respect to all *curing policies* $\rho(t)$.

$$\begin{aligned} \underset{\rho(\cdot)}{\text{minimize}} \quad & \mathbb{E}_{I_0}[\tau] \\ \text{subject to} \quad & \sum_{v \in V} \rho_v(t) \leq r, \text{ for all } t. \end{aligned}$$

Without loss of generality, we can and will restrict to policies that at any point in time allocate the entire budget to a single infected node, if one exists. We can do

this because it is not hard to show that there exist optimal policies (i.e., policies that minimize the expected time to extinction) with this property.[1] Under this restriction, the empty set (all nodes being healthy) is a unique absorbing state, and therefore the time to extinction is finite, with probability 1.

Finally, the aforementioned optimization problem is a Dynamic Program with a state space that scales exponentially in the number of nodes. Specifically, since each node at every time instant can be either infected or healthy and because of the Markovian nature of the dynamics, the state space of the problem is $\{0,1\}^n$. Hence, the resulting problem is inherently combinatorial and obtaining the optimal solution is hard.

Instead, in this thesis, we focus on

(a) Obtaining an order-optimal policy (Chapter 3).

(b) Understanding the fundamental limits of this problem with respect to the structure of the underlying graph (Chapters 4 and 5)

2.3 Discussion on the Model

The model described above is an extension of the canonical SIS model. Several of the modeling assumptions that are made both in this work as well as the prior literature are noteworthy and are discussed in this section.

(i) **Re-infections:** The canonical SIS model assumes that nodes after recovering from the infection are susceptible to re-infection. This assumption, although realistic in some situations (as explained in Section 1.1) is not natural in many other applications, such as most infectious diseases where agents develop immunity after recovery.

[1] A formal proof of this statement (which we only outline) goes as follows. We write down the Bellman equation for the problem of minimizing the expected time to extinction and observe that the right-hand side of Bellman's equation is linear in $\rho(t)$. We then recall that $\rho(t)$ is constrained to lie in a certain simplex, and conclude that we can restrict, without loss of optimality, to the vertices of that simplex. Any such vertex corresponds to allocating the entire budget to a single infected node.

(ii) **Intervention:** In this thesis, we assume that the network planner intervenes to the evolution through (stochastically) curing a subset of the infected nodes. In practice, several other intervention actions can be considered such as removing nodes from the network, quarantining subset of the network or reducing the contact rates on a subset of edges of the graph. Our work focuses on curing, mostly due to tractability but extending this work to other intervention actions is an extremely interesting and important research direction.

(iii) **Curing budget:** In this work, we assume that the budget constraint takes the form of a constant amount of curing resources R *at each time instant.* With this assumption, we aim to model situations where due to production or logistical constraints, the network planner has access to a specific and limited amount of resources per time unit (day, week etc.). In many situations, budget constraints take different forms, such as a total curing budget available at the beginning to be allocated over time, or a time-varying capacity over time to be determined by the network-planner according in a static or dynamic manner.

(iv) **Objective function:** In this thesis, the network planner seeks to minimize the expected time extinction time. This objective, although natural for applications were the goal is to return the system to normal operation (such as financial networks) may seem unnatural for other applications (such as infectious diseases) where the total number of infections would be the main concern. We chose to work with this specific objective function due to tractability as well as the ability to compare and benchmark our results against the existing literature.

2.4 Graph Theoretic Preliminaries

In the remainder of this chapter, after giving some elementary definitions and notation, we introduce and examine a deterministic version of the problem under consideration. Variants of such deterministic problems have been studied in the literature [34, 47] and involve the concept of the *CutWidth* of a graph. Loosely speaking, the

CutWidth is the maximum cut encountered during the *deterministic* extinction of an epidemic on a graph, starting from all nodes infected, in the absence of any reinfections of nodes that have become healthy, and under the best possible sequence with which nodes are cured. (A formal definition will be given shortly.)

We also introduce and study two natural extensions of the concept of the CutWidth, for the more general case where only a subset of the nodes is initially infected; we refer to them as the *resistance* and the *impedance* of the subset. These objects turn out to contain important information about the evolution of an epidemic, starting from the corresponding subset, and will serve as a low-dimensional summary of the state of an infection process.

2.4.1 Notation and Terminology.

For convenience, we use the term **bag** to refer to a "subset of V." For any bags A, B, we define

$$A \setminus B = \{v \in A : v \notin B\},$$

which is the set of nodes that belong in A but not in B, and

$$A \triangle B = (A \setminus B) \cup (B \setminus A),$$

which is the set of nodes at which A and B differ. Finally, for any node v, we write

$$A + v = A \cup \{v\}, \qquad A - v = A \setminus \{v\}.$$

We next define the concept of a crusade from A to B as a sequence of bags that starts at A and ends at B, with the restriction that at each step of this sequence, arbitrarily many nodes may be added to the previous bag, but at most one can be removed. The formal definition follows.

Definition 1. *For any two bags A and B, an $(A$-$B)$-**crusade** ω is a sequence $(\omega_0, \omega_1, \ldots, \omega_k)$ of bags, of length $k+1$, with the following properties:*

(i) $\omega_0 = A$,

(ii) $\omega_k = B$, *and*

(iii) $|\omega_i \setminus \omega_{i+1}| \leq 1$, *for* $i = 0, \ldots, k-1$.

We use the notation $\Omega(A)$ *to refer to the set of all (A-*$\emptyset$*)-crusades, i.e., crusades that start with a bag* A *and eventually end up with the empty set.*

Property (iii) states that at each step of a crusade, arbitrarily many nodes can be added to, but *at most one* node can be removed from the current bag. Note that the definition of a crusade allows for *non-monotone* changes, since a bag at any step can be a subset, a superset, or not comparable to the preceding bag.

We also consider a special case of crusades, the **monotone crusades** for which only removal of nodes is allowed at each step, as defined below.

Definition 2. *For any two bags* A *and* B, $A, B \subseteq V$, *an* $(A \downarrow B)$***-monotone crusade*** ω *is an* $(A - B)$*-crusade* $(\omega_0, \omega_1, \ldots, \omega_k)$ *with the additional property:*

$$\omega_i \supseteq \omega_{i+1},$$

for $i \in \{0, \ldots, k-1\}$. *We denote by* $\Omega(A \downarrow B)$ *the set of all* $(A \downarrow B)$*-crusades.*

2.4.2 Cuts, CutWidth, and Resistance.

The number of edges connecting a bag A with its complement will be called the cut of the bag. Its importance lies in that it is equal to the total rate at which new infections occur, when the set of currently infected nodes is A.

Definition 3. *For any bag* A, *its* ***cut****,* $c(A)$, *is defined as the cardinality of the set of edges*

$$\{(u, v) : u \in A,\ v \in A^c\}.$$

In Lemma 1 below, we record, without proof, some elementary properties of cuts.

Lemma 1. *For any two bags* A *and* B, *we have*

(i) $\left|c(A) - c(B)\right| \leq \Delta \cdot \left|A \triangle B\right|$.

(ii) If $A \subseteq B$, *and* $v \in A$, *then*

$$c(A - v) - c(A) \leq c(B - v) - c(B).$$

Note that Lemma 1(ii) states the well-known submodularity property of the function c(·) ([19]), and thus of the infection rate.

We now define the width of a crusade as the maximum cut that it encounters.

Definition 4. *The* ***width*** $z(\omega)$ *of an (A-B)-crusade* $\omega = (\omega_0, \ldots, \omega_k)$ *is defined by*

$$z(\omega) = \max_{1 \leq i \leq k} \{c(\omega_i)\}.$$

Note that in the above definition, the maximization starts at the first step of the crusade, i.e., we exclude ω_0 from consideration. The reason is the important Monotonicity property in Lemma 2(i), in the next subsection, which would otherwise fail to hold.

We next define the *resistance* of a bag A as the minimum crusade width, over all $(A\text{-}\emptyset)$-crusades. Intuitively, this is the maximum cut encountered after the first step, during a crusade that "cures" all nodes in A in an "optimal" manner.

Definition 5. *The* ***resistance*** $\gamma(A)$ *of a bag* A *is defined by*

$$\gamma(A) = \min_{\omega \in \Omega(A)} z(\omega).$$

We finally define the *impedance* a bag A as the minimum crusade width, over all $(A \downarrow \emptyset)$-crusades. Intuitively, this is the maximum cut encountered *including* the first step, during a monotone crusade that "cures" all nodes in A in an "optimal" manner.

Definition 6. *The* ***impedance*** $\delta(A)$ *of a bag* A *is defined by*

$$\delta(A) \doteq \min_{\omega \in \Omega(A \downarrow \emptyset)} \max\{z(\omega), c(A)\}. \tag{2.2}$$

We say that a (monotone) crusade $(A \downarrow B)$-crusade $\omega = (\omega_0, \ldots, \omega_{k-1})$ is *optimal* if it attains the minimum in Eq. (2.2).

The *CutWidth* W of the graph is the impedance of the set of all nodes V, i.e.,

$$W = \delta(V).$$

In other words, the problem of finding the CutWidth of a graph is the problem of deterministically curing one node at a time, starting from all nodes infected, so that the maximum cut (or the width) encountered during the curing process is minimized.

Note that traditionally, the CutWidth of a graph is defined in terms of monotone crusades, but [8] and [34] prove that even if general crusades are considered, the minimum width is the same, as the following theorem illustrates.

Theorem 1 ([8, 34]). *For any graph* $G = (V, E)$,

$$\delta(V) = \gamma(V)$$

We close this section by observing that the resistance of a bag A satisfies the Bellman equation

$$\gamma(A) = \min_{|A \backslash B| \leq 1} \left\{ \max\{\mathrm{c}(B), \gamma(B)\} \right\}, \tag{2.3}$$

while the impedance of a bag satisfies the Bellman equation

$$\delta(A) = \max\left\{\mathrm{c}(A),\ \min\{\delta(B) : B \subseteq A,\ |A \backslash B| = 1\}\right\}. \tag{2.4}$$

Note that along an optimal crusade, we have $\delta(\omega_{i+1}) \leq \delta(\omega_i)$, for $i = 0, 1, \ldots, k-1$.

2.4.3 Properties of the resistance.

This section develops some properties of the resistance. Lemma 2(i) states that if A and B are two bags with $A \subseteq B$, then $\gamma(A) \leq \gamma(B)$. Intuitively, this is because one can construct a crusade from A to $\emptyset$ as follows: The crusade starts from A,

then continues to the first bag encountered by a B-optimal crusade ω^B, and then follows ω^B. The constructed crusade and ω^B are the same except for the respective initial bags. By the definition of the resistance, the initial bag does not affect the maximization and thus the width of the new crusade is equal to $\gamma(B)$. An optimal crusade from A can do no worse.

Lemma 2(ii) states that if two bags A and B differ by only m nodes, then the corresponding resistances are at most $m\Delta$ apart. Intuitively, this is because if $m = 1$ and $A \triangle B = \{v\}$, one can attach node v to the optimal crusade for the smaller of the two bags, thus obtaining a crusade that starts at the larger bag and encounters a maximum cut which is at most Δ different from the original. The result for general m is obtained by moving from A to B by adding or removing one node at a time.

Lemma 2. *Let A and B be two bags.*

(i) [Monotonicity] If $A \subseteq B$, then $\gamma(A) \leq \gamma(B)$.

(ii) [Smoothness] We have that $\left|\gamma(A) - \gamma(B)\right| \leq \Delta \cdot \left|A \triangle B\right|$.

Proof. Recall that $\Omega(A)$ stands for the set of all (A-$\emptyset$)-crusades. Let also Ω^A be the set of all such crusades that achieve the minimum in the definition of the resistance, i.e.,

$$\Omega^A = \{\omega \in \Omega(A) : z(\omega) = \gamma(A)\}.$$

(i) Suppose that $A \subseteq B$. Let $\omega^B = (\omega_0^B, \ldots, \omega_k^B) \in \Omega^B$. Consider the sequence $\hat{\omega} = (\hat{\omega}_0, \ldots, \hat{\omega}_k)$ of bags with $\hat{\omega}_0 = A$, and $\hat{\omega}_i = \omega_i^B$, for $i = 1, \ldots, k$. We claim that $\hat{\omega}$ is a crusade $\hat{\omega} \in \Omega(A)$. Indeed,

(a) $\hat{\omega}_0 = A$;

(b) $\hat{\omega}_k = \omega_k^B = \emptyset$;

(c) $|\hat{\omega}_0 \setminus \hat{\omega}_1| = |A \setminus \hat{\omega}_1| \leq |B \setminus \omega_1^B| = |\omega_0^B \setminus \omega_1^B| \leq 1$, where the first inequality follows from $A \subseteq B$ and $\hat{\omega}_1 = \omega_1^B$. Moreover, for $i = 0, \ldots, k-1$, we have $|\hat{\omega}_i \setminus \hat{\omega}_{i+1}| = |\omega_i^B \setminus \omega_{i+1}^B| \leq 1$.

Clearly,

$$z(\hat{\omega}) = \max_{1 \le i \le k} \{c(\hat{\omega}_i)\} = \max_{1 \le i \le k} \{c(\omega_i^B)\} = \gamma(B).$$

Using the definition of $\gamma(A)$, and the fact that $\hat{\omega} \in \Omega(A)$, we conclude that

$$\gamma(A) = \min_{\omega \in \Omega(A)} z(\omega) \le z(\hat{\omega}) = \gamma(B).$$

(ii) If $|A \triangle B| = m$, we can go from bag A to bag B in a sequence of m steps, where at each step, we add or remove a single node. It thus suffices to show that each one of these steps can change the resistance by at most Δ. Accordingly, we only need to conside the case where $B = A + v$, for some $v \notin A$.

Let $\omega^A = (\omega_0^A, \ldots, \omega_k^A) \in \Omega^A$. Consider the sequence $\hat{\omega} = (\hat{\omega}_0, \ldots, \hat{\omega}_{k+1})$ of bags with $\hat{\omega}_i = \omega_i^A + v$, for $i = 0, \ldots, k$, and $\hat{\omega}_{k+1} = \emptyset$. Clearly, $\hat{\omega}$ is a crusade in $\Omega(B)$ and, therefore,

$$\gamma(B) \le z(\hat{\omega}) = \max_{1 \le i \le k} \{c(\omega_i^A + v)\} \le \max_{1 \le i \le k} \{c(\omega_i^A)\} + \Delta = \gamma(A) + \Delta,$$

where the second inequality follows because the addition of one node can change the cut by at most Δ (Lemma 1(i)).

□

An immediate corollary of Lemma 2(i) is that for any bag A, we have $\gamma(A) \le W$.

2.4.4 Relating cuts to the resistance.

This section explores a connection between cuts and resistances at the times that the resistance is reduced. It shows that, whenever the resistance is high and gets reduced, the total infection rate is also high. This observation will play a central role in the proof of our main results.

Lemma 3. *Let A be a bag and suppose that $\gamma(A - v) < \gamma(A)$, for some $v \in A$. Then,*

$$c(A - v) \ge \gamma(A).$$

Proof. Let $B = A - v$. Since $|A \setminus B| = 1$, Eq. (2.3) implies that

$$\gamma(A) \leq \max\{\mathrm{c}(B), \gamma(B)\}. \tag{2.5}$$

Having assumed that $\gamma(B) < \gamma(A)$, Eq. (2.5) implies that $\gamma(A) \leq \mathrm{c}(B)$. □

We call a bag for which $\gamma(A - v) < \gamma(A)$ for some $v \in A$, an *improvement* bag and denote by $\mathcal{C}$ the set of all improvement bags, i.e.,

$$\mathcal{C} = \{A \subseteq V : \exists v \in A, \gamma(A - v) < \gamma(A)\}. \tag{2.6}$$

2.4.5 Properties of the impedance

In this subsection we discuss two important properties of the impedance of a bag. First, it follows from the definition that the impedance of a bag A is at least $\mathrm{c}(A)$, which in general may be much larger than the CutWidth. This is a concern because the stochastic nature of the infections can always bring the process to a bag with high impedance, and therefore high subsequent infection rates. The next lemma provides an upper bound on the impedance of a bag A in terms of the CutWidth W of the graph and the cut of A. Its proof is given in the Appendix.

Lemma 4. *For any bag A, we have*

(i) $\delta(A) \geq c(A)$,

(ii) $\delta(A) \leq W + c(A)$.

Proof. (i) Follows from Definition 6

(ii) Consider a monotone crusade $\omega \in \mathcal{C}(V \downarrow \emptyset)$ whose width is equal to the CutWidth W. This crusade starts with V and removes nodes one at a time, until the empty set is obtained. Let $v_1, v_2, \ldots, v_n$ be the nodes in V, arranged in the order in which they are removed.

Let us now fix a bag A. We construct a monotone crusade $\hat{\omega} \in \mathcal{C}(A \downarrow \emptyset)$ as follows. We start with A and remove its nodes one at a time, according to the order

prescribed by ω. For example, if $n = 4$, and $A = \{v_2, v_4\}$, the monotone crusade that starts from A first removes node v_2 and then removes node v_4.

At any intermediate step during the crusade $\hat{\omega}$, the current bag is of the form $A \cap \{v_k, \ldots, v_n\}$, for some k. It only remains to show that the cut of this bag is upper bounded by $\mathrm{c}(A) + W$. Let $R = \{v_1, \ldots, v_{k-1}\}$. Note that

$$\mathrm{c}(R) \leq W,$$

because of the definition of the width and the assumption that the width of ω is W. Note also that the current bag is simply $A \cap R^c$.

For any two sets S_1 and S_2, let $e(S_1, S_2)$ be the number of edges that join them. We have that

$$\begin{aligned}
\mathrm{c}(A \cap R^c) &= e\big(A \cap R^c, (A \cap R^c)^c\big) \\
&= e(A \cap R^c, A^c \cup R) \\
&\leq e(A \cap R^c, A^c) + e(A \cap R^c, R) \\
&\leq e(A, A^c) + e(R^c, R) \\
&= \mathrm{c}(A) + \mathrm{c}(R) \\
&\leq \mathrm{c}(A) + W.
\end{aligned}$$

We conclude that the cut associated with any intermediate bag in the crusade $\hat{\omega}$ is upper bounded by $\mathrm{c}(A) + W$. It follows that the width of $\hat{\omega}$, and therefore $\delta(A)$ as well, is also upper bounded by that same quantity. □

2.4.6 Resistance and Impedance

In the preceding subsections, we defined two different concepts for a subset of nodes A, the *resistance* $\gamma(A)$ and the *impedance* $\delta(A)$. The definitions of these two concepts are related but their behavior can differ significantly.

Intuitively, the impedance of a bag is useful for the *curing* problem. Specifically, when designing a dynamic curing policy, the network planner may decide to allocate

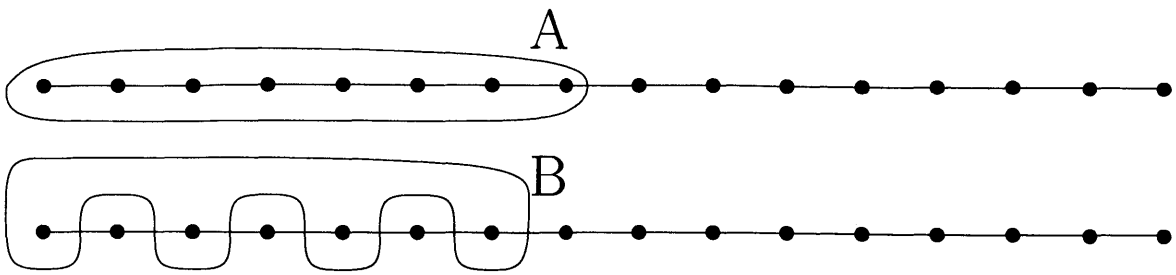


Figure 2-2: A line graph with n nodes (n is a multiple of 4). Bag A consists of $n/2$ nodes. Bag B consists of the odd numbered $n/4$ nodes of bag A. The impedance and the resistance of bag A coincide and are equal to 1. However, the impedance of bag B is equal to $n/2 - 1$ while the resistance of bag B is equal to 1.

curing resources to any node of the bag. The definition of impedance involves monotone crusades and hence provides a recipe for the order of these curing decisions. On the other hand, the resistance of a bag involves non-monotone crusades, and hence allows for new "infections" during the curing process. Allowing for non-monotonicities during the process, resistance is useful when studying the evolution of the process without restricting the network planner to a specific dynamic curing policy. In other words, the resistance of a bag is a crucial concept when studying the behavior of the contagion process under arbitrary curing policies and hence, when exploring lower bounds on the performance of the *optimal* dynamic curing policy.

Hence, in order to obtain meaningful upper and lower bounds on the performance of dynamic curing policies, we should be able to relate these two central concepts, impedance and resistance. The rest of this section explores the connection between the two, starting with two examples.

Example 1 For the case of V, impedance and resistance coincide, as Theorem 1 suggests.

Example 2 Consider a line graph with n nodes, where n is a multiple of 4. Its CutWidth is easily seen to be equal to 1: if all nodes are initially infected, we can cure them one at a time, starting from the left; the cuts encountered along the way are all equal to 1. Consider a bag with $n/2$ nodes such as bag A of Figure 2-2. The impedance of A is equal to 1 since an optimal monotone crusade consists of curing

nodes one at a time starting from the right. Similarly, the resistance of A is equal to one as any crusade cannot achieve width less than one. In contrast, consider a subset B of bag A which only consists of odd numbered nodes. Note that the cut of bag B is equal to $n/2-1$ and hence by Lemma 4 (ii), its impendance is at least equal to $n/2-1$. A monotone crusade that achieves this lower bound consists of curing nodes one at a time starting from the right. Hence the impedance of B is equal to $n/2-1$. On the other hand, the resistance of B is equal to 1 since an optimal crusade would first infect all even numbered nodes of bag A and then proceed by curing one node at a time starting from the right, hence achieving width of 1. The implications of this discrepancy are significant for the curing problem. The difficulty of curing the two bags is comparable (especially given that one is a subset of the other), but a curing policy that is "monotone" would face higher drifts in the case of bag B. Example 2 seems to suggest that for bags with high cuts, there is a potential discrepancy between the impedance and the resistance of a bag, and that the resistance is the more relevant one.

In the rest of this section we show that for bags with small cuts, the distance between impedance and resistance is also small. We start by showing the existence of an optimal crusade with several desirable properties. Specifically, we argue that for any bag A, there exists an optimal crusade that adds nodes (and perhaps removes one) only at the first step of the crusade and from then on, the crusade is monotone (properties (i)-(ii)) .

Lemma 5. *For any bag A there exists an optimal crusade $\hat{\omega} = (\hat{\omega}_0, \hat{\omega}_1, \ldots, \hat{\omega}_k) \in \Omega^A$ with the following properties:*

(i) For $i \in \{0, \ldots, k-1\}$, $\hat{\omega}_i \neq \hat{\omega}_{i+1}$

(ii) For $i \in \{1, \ldots, k-1\}$, $\hat{\omega}_{i+1} \subset \hat{\omega}_i$.

Proof. We assign to every $(A-\emptyset)$-crusade $\omega \in \Omega^A$ a value

$$P(\omega) = \left(\sum_{i=0}^{|\omega|-1} (\mathrm{c}(\omega_i) + 1), \sum_{i=0}^{|\omega|-1} |\omega_i| \right).$$

Let $\hat{\omega} \in \text{argmin}_{\omega \in \Omega^A} P(\omega)$, where the minimum is taken with respect to the lexicographic ordering.

(i) We first prove that for all $i \in \{1, \ldots, k-1\}$,

$$\hat{\omega}_i \neq \hat{\omega}_{i+1}. \tag{2.7}$$

For the purposes of contradiction, assume that for some $q \in \{1, \ldots, k-1\}$, $\hat{\omega}_q = \hat{\omega}_{q+1}$, and construct a crusade $\tilde{\omega} = (\tilde{\omega}_0, \ldots, \tilde{\omega}_{k-1})$ by setting $\tilde{\omega}_i = \hat{\omega}_i$ for all $i \leq q$, and $\tilde{\omega}_i = \hat{\omega}_{i+1}$ for $i = q+1, \ldots, k-1$.

Clearly, $\tilde{\omega}$= $(\tilde{\omega}_0, \ldots, \tilde{\omega}_{k-1})$ is a crusade, i.e., $\tilde{\omega} \in \Omega(A - \emptyset)$. Moreover, $\tilde{\omega} \in \Omega^A$, because $\max_{1 \leq i \leq k-1} \mathrm{c}(\tilde{\omega}_i) = z(\hat{\omega}) = \gamma(A)$. But $\sum_{i=0}^{k-1}(\mathrm{c}(\tilde{\omega}_i) + 1) < \sum_{i=0}^{k}(\mathrm{c}(\hat{\omega}_i) + 1)$, which implies that $P(\tilde{\omega}) < P(\hat{\omega})$, and contradicts the minimality of $\hat{\omega}$.

(ii) The idea of the proof of this property is borrowed from [8], and is based on the submodularity of $\mathrm{c}(\cdot)$. We first argue that for all $i \in \{1, \ldots, k-1\}$,

$$\mathrm{c}(\hat{\omega}_{i+1} \cup \hat{\omega}_i) \geq \mathrm{c}(\hat{\omega}_i). \tag{2.8}$$

For the purposes of contradiction, assume that there exists some $q \in \{1, \ldots, k-1\}$ such that

$$\mathrm{c}(\hat{\omega}_{q+1} \cup \hat{\omega}_q) < \mathrm{c}(\hat{\omega}_q), \tag{2.9}$$

and construct the sequence of bags $\tilde{\omega} = (\tilde{\omega}_0, \ldots, \tilde{\omega}_k)$, by setting $\tilde{\omega}_i = \hat{\omega}_i$ for all $i \neq q$ and $\tilde{\omega}_q = \hat{\omega}_{q+1} \cup \hat{\omega}_q$.

We first claim that $\tilde{\omega}$ is a crusade, i.e., $\tilde{\omega} \in \Omega(A - \emptyset)$. Indeed, since $\hat{\omega}$ is a crusade, we get $|\hat{\omega}_q \setminus \hat{\omega}_{q+1}| \leq 1$ and $|\hat{\omega}_{q-1} \setminus \hat{\omega}_q| \leq 1$. Therefore,

$$|\tilde{\omega}_{q-1} \setminus \tilde{\omega}_q| = |\hat{\omega}_{q-1} \setminus (\hat{\omega}_{q+1} \cup \hat{\omega}_q)| \leq |\hat{\omega}_{q-1} - \hat{\omega}_q| \leq 1,$$

where the first equality follows from the construction of $\tilde{\omega}$ and the second inequality from $\hat{\omega}_{q+1} \cup \hat{\omega}_q$. Furthermore,

$$|\tilde{\omega}_q \setminus \tilde{\omega}_{q+1}| = |(\hat{\omega}_{q+1} \cup \hat{\omega}_q) \setminus \hat{\omega}_{q+1}| \leq |\hat{\omega}_q - \hat{\omega}_{q+1}| \leq 1,$$

where the the first equality follows from the construction of $\tilde{\omega}$ and the second inequality from $\hat{\omega}_{q+1} \cup \hat{\omega}_q \supset \hat{\omega}_{q+1}$.

Moreover, we claim that $\tilde{\omega} \in \Omega^A$. Indeed

$$\begin{aligned}\max_{1\leq i\leq k} \mathrm{c}(\tilde{\omega}_i) &= \max\{c(\tilde{\omega}_q), \max_{1\leq i\leq k, i\neq q} \mathrm{c}(\hat{\omega}_i)\} \\ &\leq \max_{1\leq i\leq k} \mathrm{c}(\hat{\omega}_i) = \gamma(A),\end{aligned}$$

where the inequality follows from (2.9).

On the other hand, it follows from (2.9) that $\sum_{i=0}^{k}(\mathrm{c}(\tilde{\omega}_i)+1) < \sum_{i=0}^{k}(\mathrm{c}(\hat{\omega}_i)+1)$ and thus $P(\tilde{\omega}) < P(\hat{\omega})$, which contradicts the minimality of $\hat{\omega}$. We have therefore established (2.8).

Using the submodularity of the cut as well as Eq. (2.8), we have that for all $i \in \{1, \ldots, k-1\}$,

$$\mathrm{c}(\hat{\omega}_{i+1} \cap \hat{\omega}_i) \leq \mathrm{c}(\hat{\omega}_{i+1}). \tag{2.10}$$

We now prove that $|\hat{\omega}_{i+1} \cap \hat{\omega}_i| \geq |\hat{\omega}_{i+1}|$ for all $i \in \{1, \ldots, k-1\}$.

For the purposes of contradiction, assume that there exists some $q \in \{1, \ldots, k-1\}$ such that

$$|\hat{\omega}_{q+1} \cap \hat{\omega}_q| < |\hat{\omega}_{q+1}|. \tag{2.11}$$

Construct the sequence $\tilde{\omega}_i = \hat{\omega}_i$ for all $i \neq q+1$ and $\tilde{\omega}_{q+1} = \hat{\omega}_{q+1} \cap \hat{\omega}_q$.

We first claim that $\tilde{\omega}$ is a crusade, i.e. $\tilde{\omega} \in \Omega(A - \emptyset)$. Indeed, since $\hat{\omega}$ is a crusade we get $|\hat{\omega}_q \setminus \hat{\omega}_{q+1}| \leq 1$ and $|\hat{\omega}_{q+1} \setminus \hat{\omega}_{q+2}| \leq 1$. Therefore,

$$|\tilde{\omega}_q \setminus \tilde{\omega}_{q+1}| = |\hat{\omega}_q \setminus (\hat{\omega}_{q+1} \cap \hat{\omega}_q)| = |\hat{\omega}_q - \hat{\omega}_{q+1}| \leq 1,$$

where the the first equality follows from the construction of $\tilde{\omega}$ and the second inequality from $\hat{\omega}_{q+1} \cap \hat{\omega}_q \subset \hat{\omega}_{q+1}$. Furthermore,

$$|\tilde{\omega}_{q+1} \setminus \tilde{\omega}_{q+2}| = |(\hat{\omega}_{q+1} \cap \hat{\omega}_q) \setminus \hat{\omega}_{q+2}| \leq |\hat{\omega}_{q+1} - \hat{\omega}_{q+2}| \leq 1,$$

where the the first equality follows from the construction of $\tilde{\omega}$ and the second inequal-

ity from $\hat{\omega}_{q+1} \cap \hat{\omega}_q \subset \hat{\omega}_{q+1}$. Moreover, we claim that $\tilde{\omega} \in \Omega^A$. Indeed

$$\begin{aligned}\max_{1\le i\le k} \mathrm{c}(\tilde{\omega}_i) &= \max\{c(\tilde{\omega}_q), \max_{1\le i\le k, i\neq q+1} \mathrm{c}(\hat{\omega}_{q+1})\} \\ &\le \max_{1\le i\le k} \mathrm{c}(\hat{\omega}_i) = \gamma(A),\end{aligned}$$

where the inequality follows from (2.10).

On the other hand, it follows from (2.10) that $\sum_{i=0}^{k}(\mathrm{c}(\tilde{\omega}_i)+1) \le \sum_{i=0}^{k}(\mathrm{c}(\hat{\omega}_i)+1)$ and from (2.11) that $\sum_{i=0}^{k}|\omega_i| < \sum_{i=0}^{k}|\tilde{\omega}_i|$. Therefore, $P(\tilde{\omega}) < P(\hat{\omega})$, which contradicts the minimality of $\hat{\omega}$.

Therefore we established that $|\hat{\omega}_{i+1} \cap \hat{\omega}_i| \ge |\hat{\omega}_{i+1}|$ for all $i \in \{i, \ldots, k-1\}$. The latter implies that for all $i \in \{1, \ldots, k-1\}$, $\hat{\omega}_{i+1} \subseteq \hat{\omega}_i$. Using part (i) of the lemma, it follows that that for $i \in \{1, \ldots, k-1\}$, $\hat{\omega}_{i+1} \subset \hat{\omega}_i$. □

The next corollary summarizes our findings regarding the relationship between the resistance and the impedance of a bag.

Corollary 1. *For any bag A*

$$\delta(A) \le \gamma(A) + c(A) + \Delta$$

Proof. Consider any bag A and construct an optimal crusade $\hat{\omega} = (\hat{\omega}_0, \hat{\omega}_1, \ldots, \hat{\omega}_k)$ with the properties of Lemma 5. We consider two cases.

If $\hat{\omega}_1 \subseteq A$, then by property (ii), $\hat{\omega}$ is a monotone crusade an hence

$$\gamma(A) = \delta(A).$$

Otherwise, there exists $B \subseteq A$, with $|A \setminus B| \le 1$, such that $B \subset \hat{\omega}_1$. For all $i \in \{1, \ldots, k-1\}$, let $v_i = \hat{\omega}_i \setminus \hat{\omega}_{i+1}$ be the sequence of nodes that are removed at each stage of the crusade. Then for all $i \in \{1, \ldots, k-1\}$,

$$\mathrm{c}(\hat{\omega}_1 - \{v_1, \ldots, v_i\}) \le \gamma(A),$$

by the optimality of $\hat{\omega}$. Consider the crusade $\tilde{\omega} \in \Omega(B \downarrow \emptyset)$ for which $\tilde{\omega}_0 = B$ and $\tilde{\omega}_i = B \setminus \{v_1, \ldots, v_i\}$ for $i \in \{1, \ldots, k-1\}$. Note that by the assumption $\hat{\omega}_1 \supset B$ we get that for all $j \in \{1, \ldots, k-1\}$, $B \setminus \{v_1, \ldots, v_j\} \subseteq \hat{\omega}_1 - \{v_1, \ldots, v_j\}$. Hence, using (ii) of Proposition 1, we obtain

$$\begin{aligned} &\mathrm{c}(B \setminus \{v_1, \ldots, v_{j+1}\}) - \mathrm{c}(B \setminus \{v_1, \ldots, v_j\}) \\ &\quad \leq \mathrm{c}(\hat{\omega}_1 \setminus \{v_1, \ldots, v_{j+1}\}) - \mathrm{c}(\hat{\omega}_1 \setminus \{v_1, \ldots, v_j\}). \end{aligned}$$

Adding all the inequalities that correspond to $j \leq i$, we obtain that for all $i \in \{1, \ldots, k-1\}$,

$$\mathrm{c}(B \setminus \{v_1, \ldots, v_i\}) - \mathrm{c}(B) \leq \mathrm{c}(\hat{\omega}_1 \setminus \{v_1, \ldots, v_i\}) - \mathrm{c}(\hat{\omega}_1).$$

Therefore,

$$\begin{aligned} \mathrm{c}(\tilde{\omega}_i) &\leq \mathrm{c}(\hat{\omega}_1 \setminus \{v_1, \ldots, v_{j+1}\}) + (\mathrm{c}(B) - \mathrm{c}(\hat{\omega}_1)) \\ &< \mathrm{c}(\hat{\omega}_1 \setminus \{v_1, \ldots, v_{j+1}\}) + \mathrm{c}(B) = c(\hat{\omega}_i) + c(B), \end{aligned}$$

and hence

$$\delta(A) \leq \mathrm{c}(\tilde{\omega}_i) \leq c(\hat{\omega}_i) + \mathrm{c}(B) \leq \gamma(A) + c(B),$$

where the first inequality follows from the definition of the impedance and from the fact that $\tilde{\omega}$ is a monotone path, while the last inequality follows from the optimality of $\hat{\omega}$. Finally, note that by Lemma 1, $\mathrm{c}(B) \leq \mathrm{c}(A) + \Delta$, hence the result. $\square$

Chapter 3

An Efficient and Optimal Curing Policy

In this chapter we develop one of the main contributions of this thesis, an (order) optimal dynamic curing policy. Specifically we show that when the budget r of curing resources available at each time is $\Omega(W)$, where W is the CutWidth of the graph, and also of order $\Omega(\log n)$, then the expected extinction time of the epidemic is of order $O(n/r)$, which is within a constant factor from optimal, as well as sublinear in the number of nodes. Consequently, if the CutWidth increases only sublinearly with n, a sublinear expected time to extinction is possible with a sublinearly increasing budget r.

3.1 Description of the CURE policy

In this section, we present our curing policy and we study the resulting expected time to extinction, starting from an arbitrary initial set of infected modes. Loosely speaking, the policy, at any time, tries to follow a certain desirable (monotone) crusade, called a target path, by allocating all of the curing resources to a single node, namely, the node that should be removed in order to obtain the next bag along the target path. On the other hand, this ideal scenario may be interrupted by infections, at which point the policy shifts its attention to newly infected nodes, and attempts

to return to a bag on the target path. It turns out that under certain assumptions, this is successful with high probability and does not take too much time. However, with small probability, the process veers far off from the target path; in that case the policy "restarts" in a manner that we will make precise in the sequel.

Waiting period. A typical attempt starts at some bag A, with a waiting period. (If this is the first attempt, then $A = I_0$. Otherwise, A is the bag at the end of the preceding attempt.) During the waiting period, all curing rates $\rho_v(t)$ are kept at zero.[1] The waiting period ends at the first subsequent time that[2]

$$\mathrm{c}(I_t) \leq r/8.$$

Let B be the bag I_t right at the end of the waiting period, and let $\omega^B = (\omega_0^B, \ldots, \omega_{|B|}^B)$ the corresponding optimal crusade, which we refer to as the *target path.*

Segments. Each segment of an attempt starts either at the end of the waiting period or at the end of a preceding segment of the same attempt. In all cases, the segment starts with a bag on the target path. For the first segment, this is guaranteed by the definition of the target path. For subsequent segments, it will be guaranteed by our specifications of what happens at the end of the preceding segment. Let $v_1, \ldots, v_m$ be the nodes in the bag at the beginning of a segment, arranged in the order according to which they are to be removed along the target path. For example, the bag at the beginning of the segment is $\omega_0^B = \{v_1, \ldots, v_m\}$, the next bag is $\omega_1^B = \{v_2, \ldots, v_m\}$, etc. The node v_1 is called the *target node*; the goal of the segment is to cure the target node and reach the bag $C = \{v_2, \ldots, v_m\}$. For all t during the segment, we define $D_t = I_t \setminus C$; this is the set of infected nodes that do not belong to the next bag on the target path. At the beginning of the segment, $I_t = C \cup \{v\}$ and therefore $D_t = \{v_1\}$.

[1]During the waiting period the curing budget is wasted and not allocated to any of the nodes. Note that the cut of I_t during the waiting phase could be linear in the number of nodes, while we focus on the regime where the available budget is sublinear. Therefore, regardless of the allocation, during the waiting period the process would have an upward drift. For this reason, allocating budget to a subset of nodes in this period would not have a significant effect on the performance.

[2]Note that the waiting period is guaranteed to terminate in finite time, with probability 1. This is because if it were infinite, then healthy nodes would keep getting infected until eventually $I_t = V$. But $\mathrm{c}(V) = 0$, which means that at some point the condition $\mathrm{c}(I_t) \leq r/8$ would be satisfied and the waiting period would be finite, a contradiction.

During the segment, the entire curing budget is allocated to an arbitrarily chosen node from D_t. Note that $\rho_v(t) = 0$ for $v \in C$ during the segment and therefore, we always have $I_t \supseteq C$.

The segment ends when either:

(i) all nodes have been cured, i.e., $I_t = \emptyset$; in this case, the attempt is considered successful and the process is over.

(ii) $I_t = C$ and $C \neq \emptyset$ in which case the target node is cured, the process is on the target path, and we are ready to start the next segment. In this case, we say that we have a *short* segment.

(ii) $|D_t| \geq r/8\Delta$, in which case we say that the segment was *long*, and that the attempt has failed. In this case, the attempt has no more segments, and a new attempt will be initiated, starting with a waiting period.

3.2 Performance Analysis

We now proceed to establish an upper bound on the expected time to extinction, under the assumption that $r \geq 4W$, for any set of initially infected nodes. If the process always stayed on the target path, that is, if we had no infections, the expected time to extinction would be the time until all nodes (at most n of them) were cured. Given that nodes are cured at a rate of r, the expected time to extinction would have been $O(n/r)$. On the other hand, infections do delay the curing process, by increasing $|D_t|$ during segments, and we need to show that these do not have a major impact.

There are two kinds of segments to consider, *short* ones, at the end of which $|D_t| = 0$, and *long* ones, at the end of which $|D_t| \geq r/8\Delta$. During a segment, the size of D_t (the "distance" from the target path) is at most $r/8\Delta$. Using also an upper bound on the size of the cut along the target path, we can show that the infection rate throughout a segment is smaller than the curing rate. For this reason, during a segment, the process $|D_t|$ has a downward drift. As a consequence, using a standard argument, the expected duration of a segment is small and there is high probability

that the segment ends with $|D_t| = 0$, so that the segment is short and we continue with the next segment. As a result, the expected duration of an attempt behaves similar to the case of no infections and is also of order $O(n/r)$. Finally, by studying the number of failed attempts until a successful one, we can establish an upper bound for the overall policy. A formal version of this argument is the content of the rest of this section.

3.2.1 Segment analysis

Let us focus on a particular segment, and let $M_t = |D_t|$. The process M_t evolves on the finite set $\{0, 1, \ldots, r/8\Delta\}$. (For simplicity, and without loss of generality, we assume that $r/8\Delta$ is an integer.) Recall that C was defined as the bag on the target path that we were trying to reach at the end of the segment. The difference D_t at the time that the segment starts consists of exactly one node: the target node. Thus, the process M_t is initialized at 1, at the beginning of the segment. The process M_t is stopped as soon one of the two boundary points, 0 or $r/8\Delta$, is reached. At each time before the process is stopped, there is a rate equal to r of downward transitions. Furthermore, there is a rate $\mathrm{c}(I_t)$ of upward transitions, corresponding to new infections.

Lemma 6. *The rate $c(I_t)$ of upward transitions during a segment satisfies $c(I_t) \leq r/2$.*

Proof. The definition $D_t = I_t \setminus C$ implies that $I_t \subseteq C \cup D_t$. Consequently,

$$\begin{aligned} \mathrm{c}(I_t) &\leq \mathrm{c}(C) + \mathrm{c}(D_t) \leq \mathrm{c}(C) + \Delta \cdot |D_t| \\ &= \mathrm{c}(C) + \Delta \cdot M_t \leq \mathrm{c}(C) + \frac{r}{8}. \end{aligned} \tag{3.1}$$

We have used here Proposition 1, in the first and second inequality, together with the fact $M_t \leq r/8\Delta$.

On the other hand, C is on the target path associated with B, the bag obtained at the end of the waiting period. As remarked at the end of Section 2.4.2, the impedance does not increase along an optimal crusade, and therefore, $\delta(C) \leq \delta(B)$. Using also

Lemma 4, we have

$$c(C) \leq \delta(C) \leq \delta(B) \leq W + c(B).$$

Recall now that a waiting period ends with a bag whose cut is at most $r/8$. Therefore, $c(B) \leq r/8$. It follows that $c(C) \leq W + r/8$. Using this fact, together with the assumption $r \geq 4W$ and Eq. (3.1), we obtain

$$c(I_t) \leq c(C) + \frac{r}{8} \leq \left(W + \frac{r}{8}\right) + \frac{r}{8} \leq \frac{r}{4} + \frac{r}{8} + \frac{r}{8} = \frac{r}{2}.$$

□

We now establish the properties of the segments that we have claimed earlier; namely, that segments are short, with high probability, and do not last too long.

Lemma 7. *a) The probability that the segment is long is at most*

$$p = \frac{1}{2^{r/8\Delta} - 1}.$$

b) The expected length of a segment is upper bounded by $2/r$.

Proof. a) Using Lemma 6, the process M_t is stochastically dominated by a process N_t on the same space $\{0, 1, \ldots, r/8\Delta\}$, which is initialized to be equal to the value of M_t at the beginning of the segment (which is 1), has a rate r of downward transitions, a rate $r/2$ of upward transitions, and stops at the first time that it reaches one of the two boundary values. Note that the ratio of the downward to the upward drift is equal to 2. The probability, denoted by p, that the process N_t will first reach the upper boundary is a well-studied quantity and is given by the expression in part (a) of the lemma. The proof is standard and can be found in Section 2.1 of [38] (for a non-martingale based proof) or Section 2.3 of [60] (for a martingale based proof). Since M_t is stochastically dominated by N_t, the probability that M_t will first reach the upper boundary is no larger.

b) For simplicity, let us suppose that the segment starts at time $t = 0$. We define the process

$$H_t = M_t + \frac{r}{2}t$$

and the stopped version, $\hat{H}_t$ which stops at the time T that the segment ends. It is straightforward to verify that $\hat{H}_t$ is a supermartingale, because the upward drift of the process is $\beta c(I_t) \leq r/2$ and the downward drift is r, so that the total downward drift at least $r/2$. Furthermore, $\hat{H}_0 = H_0 = M_0 = 1$. Using Doob's optional stopping theorem we obtain

$$1 = \mathbb{E}[M_0] = \mathbb{E}[\hat{H}_0] \geq \mathbb{E}[\hat{H}_T] + \frac{r}{2} \cdot \mathbb{E}[T] \geq \frac{r}{2} \cdot \mathbb{E}[T],$$

from which we conclude that

$$\mathbb{E}[T] \leq \frac{2}{r}.$$

□

Note that if $r \geq \alpha \log n$, where α is a sufficiently large constant, then p can be made smaller that $1/n^2$, so that np tends to zero. We will be using this observation later on. We will now bound the length of a waiting period.

Lemma 8. *The expected length of a waiting period is bounded above by* $8n/r$.

Proof. A waiting period involves at most n infections. The waiting period ends as soon as $c(I_t) \leq r/8$. Therefore, during the waiting period, infections happen at a rate of at least $r/8$. In particular, during the waiting period, the expected time between consecutive infections is at most $8/r$. For a maximum of n infections, the expected time is upper bounded by $8n/r$. □

We can now combine the various bounds we have derived so far in order to bound the expected time to extinction under our policy.

Theorem 2. *Suppose that* $r \geq 4W$ *and that* r *is large enough so that* $np < 1$, *where* p *is as defined in Lemma 7. For any initial bag, the expected time to extinction under*

the CURE policy is upper bounded by

$$\frac{1}{1-np} \cdot \frac{10n}{r}.$$

Proof. We start by upper bounding the expected duration of an attempt. The expected length of the waiting period of an attempt is upper bounded by $8n/r$, by Lemma 8.

The number of segments during an attempt is at most n since each segment is associated with one target node and there can be at most n different target nodes. By Lemma 7, the expected length of a segment is at most $2/r$.

Putting everything together, the expected duration of an attempt is at most $(8n/r) + (2n/r) = 10n/r$.

Each attempt involves n segments. During each segment, there is probability at most p that the segment is long and that the attempt fails. Therefore, the overall probability that an attempt will fail is at most np (here we used the union bound). We note that his upper bound (np) on the failure probability holds regardless of the initial bag at the beginning of an attempt. It follows that the attempt is stochastically dominated by a geometric random variable with parameter $1-np$. For this reason, the expected number of attempts is at most $1/(1-np)$, and the desired result follows. □

3.3 Corollaries and near-optimality of the CURE policy

Theorem 2 has a number of interesting consequences, which we collect in the corollary that follows. We argue that if all nodes are initially infected, then the expected time to extinction under any policy is at least n/r. Furthermore, in a certain regime of parameters, our policy achieves $O(n/r)$ expected time to extinction and is therefore optimal within a multiplicative constant. Finally, if the CutWidth increases sublinearly with the number of nodes, then the expected time to extinction can be made sublinear in n, using only a sublinear budget.

Corollary 2. *a) For any graph with n nodes and with all nodes initially infected, the expected time to extinction is at least n/r, under any policy.*

b) Suppose that the budget r satisfies

$$r \geq 4W, \qquad r \geq 16\Delta \log_2 n.$$

Then, for large enough n, and for any initial set of infected nodes, the expected time to extinction under the CURE policy is at most $26n/r$, which is sublinear in n and within a multiplicative factor from optimal.

c) Suppose that the maximum degree is bounded, i.e., Δ is $O(1)$. If the CutWidth increases sublinearly with n, then it is possible to have sublinear time to extinction with a sublinear budget.

Proof. a) Since nodes are cured at a rate of at most r, and there are n nodes to be cured, the expected time to extinction must be at least n/r, even in the absence of infections.

b) When $r \geq 16 \cdot \log_2 n \cdot \Delta$, we have $r/8\Delta \geq 2\log_2 n$, and $2^{r/8\Delta} \geq n^2$. Thus, the probability p in Lemma 7 is of order $O(1/n^2)$, and np is of order $O(1/n)$. In particular, for large enough n, the factor $1/(1-np)$ is less than 2. By Theorem 2, the expected time to extinction is at most $20n/r$. This is sublinear in n, because r tends to infinity. Order optimality follows from part (a).

c) Suppose that the budget r satisfies the conditions in part (b), together with the condition

$$r = \Omega(n/\log n).$$

Then, it follows from part (b) that the expected time to extinction under the CURE policy is of order $O(\log n)$. If W increases sublinearly with n, we can satisfy the conditions in parts (b) and (c) while keeping r sublinear in n, and still achieve sublinear, e.g., $O(\log n)$ expected time to extinction.

□

We continue with some examples. For a line graph with n nodes, the CutWidth is equal to 1 and $\Delta = 2$. Therefore, by part (b) of Corollary 2 we can guarantee an approximately optimal expected time to extinction, of order $O(n/r)$, as long as $r \geq 16 \cdot \log_2 n \cdot \Delta = 32 \log_2 n$. We note, however, that for this example, our analysis is not tight, and the requirement $r \geq 32 \log_2 n$ is stronger than necessary.

For a square grid-graph with n nodes, the Cut-Width is approximately $\sqrt{n}$ and $\Delta = 4$. In this case, the requirement $r \geq 4W \approx 4\sqrt{n}$ is the dominant one, and suffices to guarantee an approximately optimal expected time to extinction, of order $O(n/r)$.

In both of these examples, we can of course let r be much larger than the minimum required, which was $O(\log n)$ and $O(\sqrt{n})$, respectively, in order to obtain a smaller expected time to extinction, e.g., the $O(\log n)$ expected time to extinction in part (c) of the corollary.

3.4 Performance of the CURE Policy under arbitrary initial infections

The results of Section 3.3 are stated in terms of n and W which are global characteristics of the network and do not take into account the possibility of a favorable set of initially infected nodes. In this section we obtain performance guarantees for our policy as a function of $|A|$ and $\delta(A)$, where A is the bag of initially infected nodes. Our goal is to explore conditions under which the CURE policy is (order) optimal, i.e., achieves expected extinction time of order $O(|A|/r)$.

Note that if $\mathrm{c}(A) > r/8$, a waiting phase is initiated. By the end of the waiting phase a superset of A (potentially the whole graph) is infected and thus the performance of the CURE policy cannot be related to the properties of A. For this reason, we focus on the case where $\mathrm{c}(A) < r/8$. Section 3.3 illustrates that when the budget is larger than $4W$ then, the CURE policy is (order) optimal. In this section we are interested in the case where the impedance of the initial bag, $\delta(A)$, is smaller than the CutWidth of the graph, i.e., $\delta(A) < W$. Under such conditions, we expect to

require less curing budget in order to attain (order) optimal extinction time; the main theorem of this section confirms this fact.

First we establish some properties of the *first* attempt of the CURE policy,when $r \geq \max\{4\delta(A), 8\mathrm{c}(A)\}$. Note the similarity between the latter condition and that of Corollary 2(a)

Lemma 9. *Suppose that the set of initially infected nodes is A, and that*

$$r \geq \max\{4\delta(A), 8c(A)\}.$$

Let τ_S denote the duration of a segment and let $\mathcal{S}$ denote the event that the segment is short. Moreover, we write $p_l = \mathbb{P}(\mathcal{S}^c)$. Then, for the first attempt the following properties hold:

a) The probability p_l that a segment is long is at most

$$p = \frac{1}{2^{r/8\Delta} - 1}.$$

b) The expected length of a segment is upper bounded by $2/r$, i.e., $\mathbb{E}[\tau_s] \leq 2/r$.

c) The conditional expectation of a segment, given that it is short, $\mathbb{E}[\tau_s \mid S]$, is upper bounded by $2/(r(1-p))$.

Proof. a,b) Note that since $\mathrm{c}(A) \leq r/8$ there is no waiting phase and the target path of the first attempt is the crusade associated with $\delta(A)$. Given this observation, the proofs are identical to Lemma 7 after replacing W by $\delta(A)$ in all arguments.

c) We have,

$$\begin{aligned}\mathbb{E}[\tau_s] &= \mathbb{E}[\tau_s \mid \mathcal{S}](1-p_l) + \mathbb{E}[\tau_s \mid \mathcal{S}^c]p_l \\ &\geq \mathbb{E}[\tau_s \mid \mathcal{S}](1-p_l) \geq \mathbb{E}[\tau_s \mid \mathcal{S}](1-p).\end{aligned}$$

Solving for $\mathbb{E}[\tau_s \mid \mathcal{S}]$ and using part (b) the result follows.

□

We now combine the bounds we derived in order to bound the expected time to extinction under our policy.

Lemma 10. *Suppose that the set of initially infected nodes is A with*

$$r \geq \max\{4\delta(A), 8c(A)\}.$$

Moreover, suppose that r is large enough so that $|A|p < 1$ and let $\mathcal{E}$ denote the event that the first attempt is successful. Then

$$\mathbb{E}[\tau \mid \mathcal{E}] \leq |A| \frac{2}{(1-p)r}.$$

Proof. First, the conditional expectation is well defined since $\mathbb{P}(\mathcal{E}) \geq 1 - |A|p > 0$ by the assumptions of the lemma. Conditioned on the success of the first attempt, the number of segments is $|A|$ and the result follows from Lemma 9c. □

Lemma 10 is mainly relevant in the regime where $|A|$ grows to infinity with

$$r \geq \max\{4\delta(A), 16\Delta \log_2 |A|, 8c(A)\}. \tag{3.2}$$

In this regime, the budget is sufficiently high for the first attempt to be successful with high probability. Thus, the performance indicated by Lemma 10 is achieved conditioned on an event which occurs with high probability, as the following theorem states.

Theorem 3. *Suppose that the budget satisfies Eq. (3.2) and that the set of initially infected nodes is A, whose size $|A|$ grows to infinity. Let $\mathcal{E}$ be the event that the first attempt is successful. Then, $\mathbb{P}(\mathcal{E}) = 1 - o(1)$, $\mathbb{E}[\tau \mid \mathcal{E}]$ is of order $O(|A|/r)$, and thus our policy is (order) optimal with high probability.*

Proof. Following similar reasoning as in Corollary 2,under the condition (3.2), the probability p in Lemma 9 is of order $O(1/|A|^2)$. This implies that

$$\lim_{|A|\to\infty} \mathbb{P}(\mathcal{E}) \geq \lim_{|A|\to\infty} (1 - |A|p) = 1.$$

Moreover, for large enough $|A|$, $1-p$ is larger than $1/2$ and thus, by Lemma 10 the expected time to extinction, conditioned on $\mathcal{E}$ is at most $4|A|/r$ and thus $O(|A|/r)$. $\square$

Note that Theorem 3 establishes (order) optimality with high probability, which is weaker than (order) optimality in Corollary 2. This is due to the fact that the lower budget requirements ($r \geq \max\{4\delta(A), 16\Delta \log_2 |A|, 8c(A)\}$ vs. $r \geq \max\{4W, 16 \log_2 n \cdot \Delta\}$) come at a cost: if we have a long segment and a failed attempt (which is a small probability event) the process can potentially be uncontrollable and the extinction time from then on large.

3.5 Simulation Results

In this section, we evaluate the practical performance of our policy, against the previously proposed degree based policy as well as the dynamic version of the latter. Specifically, for each amount of total curing budget, r we evaluate the following policies:

1. **Static Degree:** The curing rate at each node v at each time instant t is equal to
$$\rho_v(t) = \frac{d_v}{\sum_{u \in V} d_u} r.$$
2. **Dynamic Degree:** The curing rate at each node v at each time instant t is equal to
$$\rho_v(t) = \frac{d_v X_v(t)}{\sum_{u \in V} d_u X_u(t)} r,$$
where $X_v(t) = 1$ ($X_v(t) = 0$) if node v is infected (healthy) at time t.
3. **CURE Policy:** this policy is described in Section 3.1.

We simulate the performance of these policies in two graphs. We calculate the expectation of the extinction time by averaging over 100 samples, for each value of the curing budget r.

Star graph: The first is a star graph with 51 nodes. In this case we are able to find the optimal monotone crusade for any subset of nodes and hence implement the

CURE policy as described above. Specifically, the optimal monotone crusade for a bag that contains only leaves consists of removing leaves one at a time at any order. The optimal monotone crusade for a bag that contains less than 25 leaves and the center consists of first removing the center and then the leaves one by one. Finally, the optimal monotone crusade for a bag that contains more than 25 leaves and the center consists of first removing leaves one at a time until 25 leaves remain in the bag, then remove the center and finally the rest of the leaves one by one.

Mesh graph: Our second experiment uses a 5×10 mesh graph. In this case, due to the hardness of finding the optimal monotone crusade for a given subset of infections, we use the following heuristic for finding a monotone crusade: at each step of the proposed crusade we remove the largest (in the lexicographic order) infected node. Note that our heuristic for constructing a monotone crusade is order optimal.

Figures 3-1 and 3-2 contain the results of our simulations. From these results we conclude that dynamic curing policies greatly outperform static policies. Moreover, when curing resources are sparse, i.e. curing budget is small, then there is a significant difference in the performance of the CURE and the dynamic degree policy. In contrast, when the curing budget is large, both dynamic policies yield comparable results.

We also observe that in the mesh graph, the CURE policy significantly outperforms the dynamic degree policy for a large range of curing budgets. This behavior is in agreement with our theoretical findings. Specifically, the CutWidth of the star graph is $O(n)$ (in particular it is equal to $\lfloor (n-1)/2 \rfloor$) and therefore the required curing budget for the CURE policy to be efficient (achieve small expected extinction time) is also $O(n)$. On the other hand, when the budget is $O(n)$ the dynamic degree policy also performs significantly well, hence the policies are comparable. On the other hand, the mesh graph has CutWidth that is equal to $\sqrt{n}$ (in particular, for a $m \times n$ mesh graph it is equal to $\min\{m, n\} + 1$). Therefore, the CURE policy performs optimally when the Curing budget is $\Omega(\sqrt{n})$ while the dynamic degree policy performs well when the total budget is $\Omega(n)$, hence the significant difference in the performance.

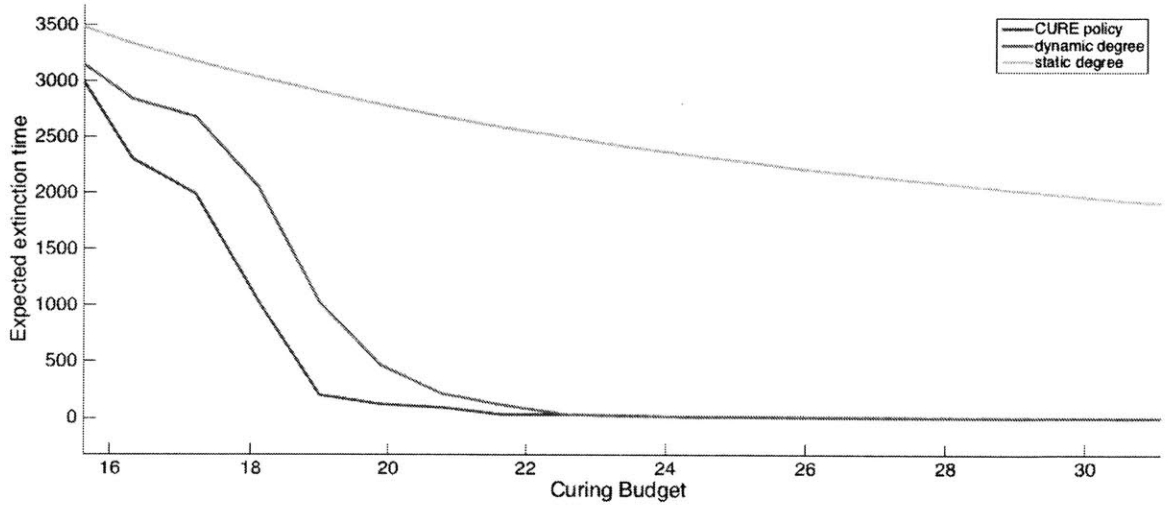


Figure 3-1: Performance of three policies on a star graph with 51 nodes.

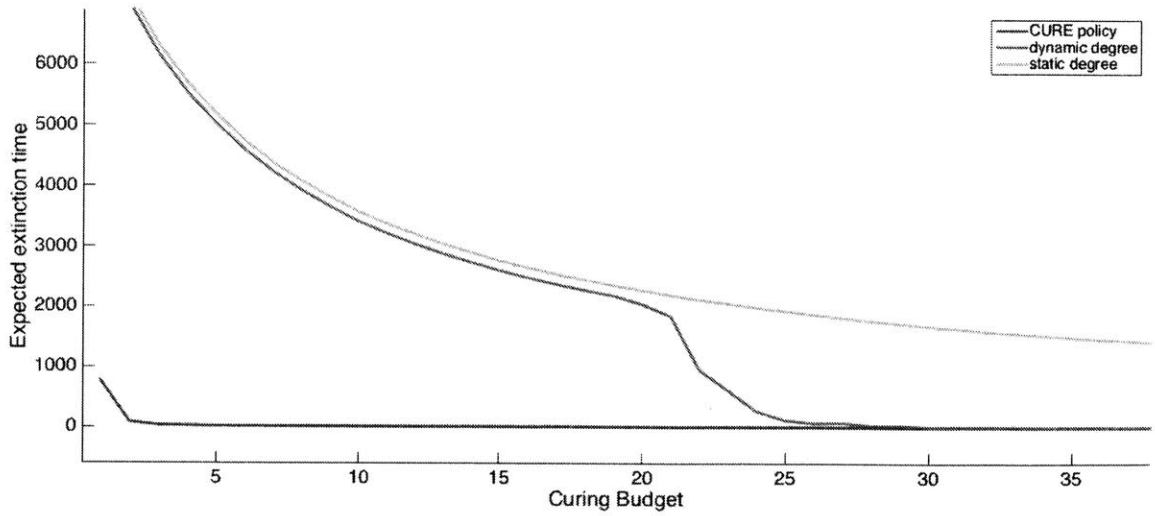


Figure 3-2: Performance of three policies on a 5×10 mesh graph.

3.6 Discussion and Conclusions

We have presented a dynamic curing policy which achieves sublinear expected time to extinction, using a sublinear curing budget when the CutWidth of the underlying graph is sublinear in the number of nodes. This policy applies to any subset of initially infected nodes and the resulting expected time to extinction is order-optimal when the available budget is sufficiently large.

The analysis of the extinction time under our policy is based on a drift analysis of the epidemic process. The upward drift is equal to the cut of the set of infected nodes $c(I_t)$ and the downward drift is proportional to the curing budget r. While the process is on the target path, $c(I_t)$, and therefore the upward drift, can be bounded from above by the impedance of the starting bag. On the other hand, when the process deviates from the target path this is no longer the case. For this reason we invoke the maximum degree Δ of the graph in order to bound the change of the cut during each such deviation. Note that none of our results (except for Corollary 1c) requires bounded degree. The maximum degree appears in the minimum budget requirement but is not required to be bounded. Furthermore, our results indicate that under our policy, the process has low probability of deviating significantly from the target path and therefore only the locally maximum degree is relevant to the analysis, and not the global maximum. In other words, as long as the infection does not reach high degree nodes we should have results similar to those for the bounded degree case. However, the performance analysis for this case is expected to be significantly harder and the statement of the results more complicated.

Our policy allocates all the available budget to one node at every time instant. This is permitted by our formulation but in practice each infected agent can only be offered a bounded amount of curing resources. Our policy, cannot be directly generalized to account for such a constraint but the insights of our solution can be directly adapted to such a scenario.

A drawback of the CURE policy is computational complexity because calculating the impedance of a bag or finding a target path is computationally hard. Like

many other interesting graph problems, CutWidth is NP-complete [24], even if we restrict to planar graphs or graphs with maximum degree three [42] but in general fixed parameter linear [62]. Several approximation algorithms have been developed for computing the CutWidth of a graph. Specifically, there is a polynomial time $O(\log^2 n)$-approximation algorithm for general graphs [35], and a polynomial time constant factor approximation algorithm for dense graphs [56]. We leave it as an interesting future direction to develop such algorithms for computing the impedance of a bag. Finally, we have argued in this paper that the CURE policy is efficient in the sense of attaining near-optimal, $O(n/r)$ expected time to extinction, in a certain parameter regime. It is an interesting problem to look for approximately optimal policies over a wider set of regimes.

Chapter 4

A lower bound for graphs with very large CutWidth: a special case

In this chapter, we provide a lower bound on the optimal expected extinction time as a function of the available budget, the epidemic parameters, the maximum degree, and the CutWidth of the graph, for graphs with large CutWidth (close to the largest possible). Under a budget which is sublinear in the number of nodes, our lower bound scales exponentially with the size of the graph.

This chapter focuses on the case of graphs whose CutWidth is close to the largest possible, while the next chapter provides a similar result for all graphs whose CutWidth scales linearly in the number of nodes. The reason for including the current chapter and the special result is twofold. First, it conveys much of the intuition of the general proof and secondly it relies on a combinatorial result (cf. Lemma 12) of independent interest.

If the graph G is complete, all policies that always allocate the entire curing budget to infected nodes are essentially equivalent, in the sense that the dynamics of $|I_t|$, the number of infected nodes, are identical under all such policies. Furthermore, I_t evolves as a birth-death Markov chain which is easy to analyze, and it is not hard to show that the expected time to extinction increases exponentially with n. On the other hand, for more general graphs with large CutWidth but bounded degree the analysis is more challenging: an analysis using a one-dimensional birth-death chain

or a simple Lyapunov function does not seem possible.

A related, and conceptually simple, way of deriving lower bounds for more general graphs would be to try to show that the process must make consistent progress through configurations (subsets I of V) where the total curing rate is significantly lower than the total infection rate. Such progress must then be a low-probability event, implying an exponential lower bound on the time to extinction. Unfortunately, it is not clear whether this line of argument, based only on the instantaneous infection rates, suffices.

The main technical contribution in this chapter is to show that in the regime examined (large CutWidth), the above outlined simple approach to deriving lower bounds is successful. Based on some nontrivial combinatorial properties of the CutWidth, and the related concept of the *resistance* of a set of nodes, we will show that there is a sizeable part of the configuration space in which $|I_t|$ has a strong upward drift.

Before we proceed, let us specify the meaning of large CutWidth. The largest possible value of a cut, for graphs with maximum degree Δ, is $n\Delta/2$, and therefore the CutWidth is also upper bounded by $n\Delta/2$. For notational convenience, we define

$$E = \frac{2}{\Delta}\left(\frac{(n+2)\Delta}{2} - W\right), \tag{4.1}$$

and observe that $E \geq 2/\Delta$. Note that "small" values of E indicate that the CutWidth is not too far from the largest possible value, $n\Delta/2$. In Section 4.1.1 we relate E to cuts and show that when E is small, then bags with large resistance, as defined below, also have a large cut.

We first explore several graph theoretic (combinatorial) properties of the resistance of a bag and the corresponding optimal crusades and then use these results to obtain our first lower bound.

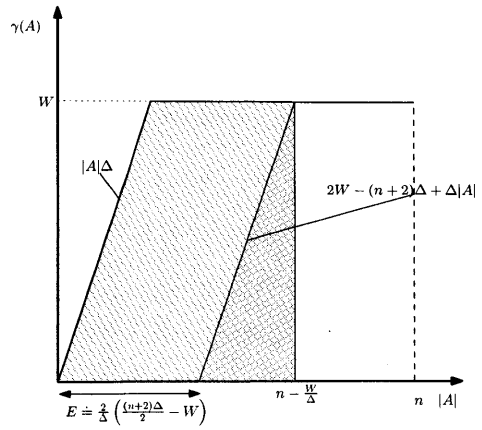


Figure 4-1: Admissible region for the pair $(\gamma(A), |A|)$. If $\gamma(A) < W$, Lemma 11 implies that $(\gamma(A), |A|)$ belongs to the parallelogram shown in the figure. On the other hand, there is no restriction on the size $|A|$ of bags with $\gamma(A) = W$, and so the admissible region also includes the horizontal line segment at the top of the figure.

4.1 More properties of optimal crusades and implications

This section can be viewed by the reader as a continuation of Section 2. We start by exploring the connection between the size of a bag and its resistance. We first obtain a bound on $\gamma(A)$ by considering a crusade which removes all nodes of A, one at a time, in an arbitrary order. We then obtain a related bound on W by constructing a crusade in $\Omega(V - \emptyset)$ that removes the nodes of the complement of A, one at a time, and then uses an A-optimal crusade ω^A. These two observations imply certain constraints (an "admissible region") for the pair $(\gamma(A), |A|)$ on the two dimensional plane, which are illustrated in Figure 4-1. Finally, using the properties of the function $\gamma(\cdot)$ that have been established so far, we obtain a refinement of the admissible region, which is again illustrated in Figure 4-1.

Lemma 11. *Consider a graph with $W \geq \Delta$ and a bag A. Let E be as defined in Eq.* (4.1),

(i) $\gamma(A) \leq |A|\Delta$;

(ii) If $\gamma(A) < W$, *then* $W \leq (n - |A|)\Delta$;

(iii) If $\gamma(A) < W$, *then* $\gamma(A) \geq \Delta(|A| - E)$.

Proof. (i) Consider some enumeration $(a_1, a_2, \ldots, a_{|A|})$ of the nodes of A. We construct a crusade $\hat{\omega} \in \Omega(A-\emptyset)$ by letting $\hat{\omega}_0 = A$, and $\omega_i = \omega_{i-1} \backslash \{a_i\}$ for $i = 1, \ldots, |A|$. By Proposition 1(i), the maximum cut encountered by $\hat{\omega}$ is bounded by $|A|\Delta$. Therefore,

$$\gamma(A) \leq z(\hat{\omega}) \leq |A|\Delta.$$

(ii) Consider some enumeration $(a_1^c, a_2^c, \ldots, a_{n-|A|}^c)$ of the nodes of A^c, the complement of A. Let $\omega^A = (\omega_0^A, \ldots, \omega_k^A) \in \Omega^A$. We construct a crusade $\omega \in \Omega(V - \emptyset)$ by letting $\omega_0 = V$, $\omega_i = \omega_{i-1} \setminus \{a_i^c\}$ for $i = 1, \ldots, n - |A|$, and $\omega_i = \omega_{i-n+|A|}^A$ for $i = n - |A| + 1, \ldots, k + n - |A|$. Then,

$$\begin{aligned} W = \gamma(V) \leq z(\omega) &= \max\{\gamma(A), \max_{i \in \{1,\ldots,n-|A|\}} c(\omega_i)\} \\ &\leq \max\{\gamma(A), (n - |A|)\Delta\}. \end{aligned} \tag{4.2}$$

The first equality above follows from Theorem 1; the second equality follows from the construction of ω; the last inequality follows from Proposition 1(i). Using the assumption $\gamma(A) < W$, Eq. (4.2) implies that

$$W \leq (n - |A|)\Delta.$$

(iii) Consider some bag A for which $\gamma(A) < W$. From (ii),

$$|A| \leq n - W/\Delta. \tag{4.3}$$

Let $C \subseteq V \setminus A$ be some nonempty bag with $|C| = n - \lfloor W/\Delta \rfloor - |A| + 1$. Note that Eq. (4.3) implies that

$$|C| = n - \lfloor W/\Delta \rfloor - |A| + 1 \geq n - W/\Delta - |A| + 1 \geq 1,$$

and that the assumption $W \geq \Delta$ implies that

$$|C| = n - \lfloor W/\Delta \rfloor - |A| + 1 \leq n - |A|.$$

This shows the existence of a bag with the desired properties exists.

We define $F = A \cup C$. Note that

$$\begin{aligned} |F| &= |A| + |C| \geq n - \lfloor W/\Delta \rfloor + 1 \\ &> n - \lfloor W/\Delta \rfloor \geq n - W/\Delta, \end{aligned}$$

and thus

$$W > (n - |F|)\Delta.$$

Then, part (ii) of the Lemma implies that $\gamma(F) = W$. The resistance of F satisfies

$$W = \gamma(F) = \gamma(A \cup C) \leq \gamma(A) + |C|\Delta,$$

where the inequality follows from applying Lemma 2(ii) $|C|$ times. Therefore,

$$\begin{aligned} \gamma(A) &\geq W - |C|\Delta \\ &= W - (n - \lfloor W/\Delta \rfloor - |A| + 1)\Delta \\ &\geq W - (n - W/\Delta + 1 - |A| + 1)\Delta \\ &\geq 2W - (n+2)\Delta + \Delta|A| \\ &= \Delta(|A| - E), \end{aligned}$$

which concludes the proof. □

4.1.1 Characterization of optimal crusades and some implications

In this section we prove that when E is small, i.e., when the CutWidth is close to the largest possible value, then bags with large resistance also have large cuts.

Lemma 12. *Suppose that $W \geq \Delta$ and that the bag A satisfies $0 < \gamma(A) < W$. Then,*

$$c(A) \geq \gamma(A) - 2(E+2)\Delta.$$

The rest of the section is devoted to proving this property. We start with a characterization of optimal crusades for a given bag A. Specifically, Lemma 5 of Section 2.4.6 states that for any bag A, there exists an optimal crusade which: (i) can add nodes, and potentially remove one node at the first step; (ii) cannot add nodes (i.e., is monotone) after the first step (parts (i)-(ii)). Moreover, we argue that except for trivial cases, an improvement bag must be encountered before the end of the crusade (part (vi)). These properties allow us to make a connection between resistance and cuts.

Lemma 13. *For any nonempty bag A with $\gamma(A) > 0$, there exists a crusade $\hat{\omega} = (\hat{\omega}_0, \hat{\omega}_1, \ldots, \hat{\omega}_k) \in \Omega^A$ with the following properties:*

(i) *For $i \in \{1, \ldots, k\}$, $\hat{\omega}_i \neq \hat{\omega}_{i-1}$*

(ii) *For $i \in \{2, \ldots, k\}$, $\hat{\omega}_i \subset \hat{\omega}_{i-1}$.*

(iii) *For $i \in \{0, \ldots, k\}$, $\gamma(\hat{\omega}_i) \leq \gamma(A)$.*

(iv) *$\gamma(\hat{\omega}_1) \geq \gamma(A) - \Delta$.*

(v) *$c(A) \geq c(\hat{\omega}_1) - \Delta(E+2)$.*

(vi) *Let $l = \min\{i \geq 0 : \hat{\omega}_i \in \mathcal{C}\}$. Then, $l < \infty$.*

Proof. We assign to every $(A - \emptyset)$-crusade $\omega \in \Omega^A$ a value

$$P(\omega) = \left(\sum_{i=0}^{|\omega|-1} (\mathrm{c}(\omega_i) + 1), \sum_{i=0}^{|\omega|-1} |\omega_i| \right).$$

Let $\hat{\omega} \in \operatorname{argmin}_{\omega \in \Omega^A} P(\omega)$, where the minimum is taken with respect to the lexicographic ordering.

(i)-(ii) These properties have been established in Lemma 5 of Section 2.4.6.

(iii) We prove the result by induction. First, observe that $\gamma(\hat{\omega}_0) = \gamma(A)$. Assume that $\gamma(A) \geq \gamma(\hat{\omega}_i)$. Moreover, by (2.3), for all $i \in \{1, \ldots, k-1\}$, $\gamma(\hat{\omega}_i) = \max\{\gamma(\hat{\omega}_{i+1}), c(\hat{\omega}_{i+1})\} \geq \gamma(\hat{\omega}_{i+1})$. Therefore, $\gamma(\hat{\omega}_{i+1}) \leq \gamma(A)$.

(iv) We consider two cases. Assume that $\hat{\omega}_1 \supset A$. Then $\gamma(\hat{\omega}_1) \geq \gamma(A) \geq \gamma(A) - \Delta$. Otherwise, by the definition of a crusade we get $|A \setminus \hat{\omega}_1| \leq 1$. Therefore, we can write $\hat{\omega}_1 = A \cup D - v$, for some set D (disjoint from A) and some $v \in A$. Using Lemma 2(ii), and then Lemma 2(i), we obtain

$$\gamma(\hat{\omega}_1) \geq \gamma(A \cup D) - \Delta \geq \gamma(A) - \Delta.$$

(v) From (iii) we obtain $\gamma(\hat{\omega}_1) \leq \gamma(A)$. Therefore, using Lemma 11(iii), we conclude that

$$|\hat{\omega}_1| \leq \frac{\gamma(A)}{\Delta} + E. \tag{4.4}$$

Moreover, by Lemma 11(i), we get

$$|A| \geq \frac{\gamma(A)}{\Delta}. \tag{4.5}$$

We consider two cases. Assume that $\omega_1 \supset A$. Then,

$$|\hat{\omega}_1 \triangle A| = |\hat{\omega}_1| - |A|.$$

Otherwise, by the definition of a crusade we get $|A \setminus \hat{\omega}_1| \leq 1$. Therefore, we can write $\hat{\omega}_1 = A \cup D - v$, where D is disjoint from A and $v \in A$. Thus,

$$|\hat{\omega}_1 \triangle A| = |D| + 1 = |\hat{\omega}_1| - |A| + 2.$$

Therefore, in both cases,

$$|\hat{\omega}_1 \triangle A| \leq |\hat{\omega}_1| - |A| + 2.$$

We then use Eqs. (4.4) and (4.5) to obtain

$$|\hat{\omega}_1 \triangle A| \leq E + 2.$$

The result follows by applying Proposition 1(iv).

(vi) Note that $\gamma(\hat{\omega}_k) = \gamma(\emptyset) = 0$. Note also that any single-element set B satisfies $\gamma(B) = 0$. Suppose that $\gamma(\hat{\omega}_1) > 0$. Then, there exists some $i \in \{1, \ldots, k-1\}$ such that $\gamma(\hat{\omega}_{i+1}) < \gamma(\hat{\omega}_i)$, and $\hat{\omega}_i$ is an improvement bag, as defined in Section 2.4.4.

Suppose now that $\gamma(\hat{\omega}_1) = 0$. If $\hat{\omega}_1 = \emptyset$, then $\gamma(A) = 0$, which contradicts the assumption $\gamma(A) > 0$. If $\hat{\omega}_1$ is nonempty, then we must have $\mathrm{c}(\hat{\omega}_2) = 0$, so that $\hat{\omega}_2$ is empty and $\hat{\omega}_1$ is a singleton. Since $\gamma(A) > 0$, the set A is not a singleton. Since at most one element can be removed in going from A to $\hat{\omega}_1$, it follows that A consists of two elements and that a single element was removed from A. In that case, $A = \hat{\omega}_0$ is an improvement bag.

In both cases, we see that there exists some i for which $\hat{\omega}_i$ is an improvement bag and therefore l is well-defined and finite. □

Proof. (of Lemma 12) Consider a crusade $\hat{\omega}_A$ with the properties in Lemma 5, and let $l \geq 0$ be such that $B = \hat{\omega}_l^A$ is the fist improvement bag encountered.

From Lemma 5(vi), l is well-defined and finite. We consider three cases:

(i) $l = 0$: If A is itself an improvement bag, then from Lemma 3, $\mathrm{c}(A) \geq \gamma(A) - \Delta$.

(ii) $l = 1$: In this case, $\hat{\omega}_1$ is an improvement bag. From Lemma 3, $\mathrm{c}(\hat{\omega}_1) \geq \gamma(\hat{\omega}_1) - \Delta$. Then, from Lemma 5(iv), we obtain

$$\mathrm{c}(\hat{\omega}_1) \geq \gamma(A) - 2\Delta.$$

Moreover, from Lemma 5(v), we get

$$\mathrm{c}(A) \geq \mathrm{c}(\hat{\omega}_1) - (E+2)\Delta \geq \gamma(A) - (E+4)\Delta.$$

(iii) $l > 2$: In this case, by property (ii) in Lemma 13(ii), it folows that $B \subset \hat{\omega}_1$ and

$$|B \triangle \hat{\omega}_1| = |\hat{\omega}_1| - |B|.$$

Moreover, since B is the first improvement bag that is encountered, $\gamma(B) = \gamma(\hat{\omega}_1)$. We use Lemma 11(i) to obtain

$$|B| \geq \gamma(B)/\Delta = \gamma(\hat{\omega}_1)/\Delta,$$

and Lemma 11(iii) to obtain

$$|\hat{\omega}_1| \leq \gamma(\hat{\omega}_1)/\Delta + E.$$

Combining the above,

$$|B \triangle \hat{\omega}_1| = |\hat{\omega}_1| - |B| \leq E,$$

from which we conclude that

$$\mathrm{c}(\hat{\omega}_1) \geq \mathrm{c}(B) - E\Delta \geq \gamma(\hat{\omega}_1) - (E+1)\Delta.$$

where the first inequality follows from Proposition 1(iv) and the second from the fact that B is an improvement bag and $\gamma(B) = \gamma(\omega_1)$. Therefore, from Lemma 13(v), we obtain

$$\mathrm{c}(A) \geq \gamma(\omega_1) - (2E+3)\Delta.$$

Finally, using Lemma 5(iv), we conclude that

$$\mathrm{c}(A) \geq \gamma(A) - 2(E+2)\Delta,$$

which completes the proof of Lemma 12.

The combinatorial properties of the resistance derived in this section will be used next to obtain a lower bound on the expected extinction time, in the regime where

$\gamma(I_0) \gg r$.

4.2 Exponential Lower Bound

In this section we state and prove the main result of this chapter. Specifically, we use Lemma 12 to argue that the process must traverse a region in which the number $|I_t|$ of infected nodes has an upward drift, which in turn leads to the desired lower bound.

We first recall a basic result on the standard continuous time random walk on the integers. Specifically, let Z_t denote the state of a Markov process with the following dynamics:

$$\begin{aligned} Z_t &: i \to i+1, \text{ with rate } \mu, \\ Z_t &: i \to i-1, \text{ with rate } \lambda. \end{aligned} \tag{4.6}$$

Fix some integers M and L with $0 < M < L$, and let $\mathbb{P}_M$ be the probabilty measure that describes the process when initialized at $Z_0 = M$. Let

$$\tau_L = \inf\{t : Z_t = 0 \text{ or } Z_t = L\}$$

denote the first time that the process Z_t visits state 0 or L, which is a *stopping time.* Moreover, let

$$\hat{\tau} = \inf\{t : Z_t = 0\}$$

denote the first time that Z_t hits 0.

The following result is standard; see, e.g., Section 2.1 of [38] or Section 2.3 of [60].

Lemma 14. *Consider the process Z_t and the stopping times τ_L and $\hat{\tau}$. Then,*

$$\mathbb{P}_M(Z_{\tau_L} = L) = \frac{1-(\lambda/\mu)^M}{1-(\lambda/\mu)^L}, \tag{4.7}$$

Consider now a related Markov process Y_t, whose transition rates are as follows:

$$\begin{aligned} Y_t &: i \to i+1, \text{ with rate } \mu, \\ Y_t &: i \to i-1, \text{ with rate } \lambda, \end{aligned} \tag{4.8}$$

for $i \in \{1, \ldots, L-1\}$ while

$$Y_t : i \to i-1, \text{ with rate } \lambda,$$

for $i = L$ and

$$Y_t : i \to i+1, \text{ with rate } \mu,$$

for $i = 0$.

We are looking for a lower bound on the expected time that it takes for the process Y_t to hit 0 for the first time, assuming that it starts at $L-1$. Let p be the probability that Y_t hits level L before hitting 0 starting from state $L-1$, which is given by Lemma 14, with $M = L-1$. We consider the case where $\lambda < \mu$, so that $p > 1/2$. Each time that the process is at state $L-1$, the process regenerates, and we have a new trial, which succeeds in hitting state 0 before state L, with the same probability $1-p$. Let N be the number of trials and note that its expected value is $1/(1-p)$. In between trials, there needs to be a transition from state L to state $L-1$, whose expected time is $1/\lambda$. Thus, the total expected time elapsed until state 0 is hit for the first time is $\mathbb{E}[N-1]/\lambda = p/(1-p)\lambda$. Using Lemma 14 and some straightforward algebra, we obtain that this expected time is at least as large as

$$\frac{1}{2}\left(\left(\frac{\mu}{\lambda}\right)^{L-1} - 1\right)\frac{1}{\lambda}. \tag{4.9}$$

Theorem 4. *Suppose that* $\gamma(I_0) \geq \Delta(9E+12)+3r$. *Then,*

$$\mathbb{E}_{I_0}[\tau] \geq \frac{1}{2r}\left(\left(\frac{\gamma(I_0)-(9E+12)\Delta}{3r}\right)^{\frac{\gamma(I_0)}{3\Delta}-1}-1\right).$$

Proof. We define a process V_t which is coupled with the process I_t as follows.

$$V_t = \begin{cases} |I_t|, & \text{if } |I_t| \leq \left\lfloor \frac{2\gamma(I_0)}{3\Delta} \right\rfloor, \\ \left\lfloor \frac{2\gamma(I_0)}{3\Delta} \right\rfloor, & \text{if } |I_t| > \left\lfloor \frac{2\gamma(I_0)}{3\Delta} \right\rfloor. \end{cases} \tag{4.10}$$

The dynamics of V_t are as follows. If $i < \lfloor 2\gamma(I_0)/3\Delta) \rfloor$, then

$$V_t : i \to i+1, \text{ with rate } c(I_t),$$
$$V_t : i \to i-1, \text{ with rate } r.$$

Furthermore, if $i = \lfloor 2\gamma(I_0)/(3\Delta)) \rfloor$, then

$$V_t : \ i \to i-1, \text{ with rate } r(I_t),$$

where $r(I_t) \leq r$.

Consider the stopping time

$$\begin{aligned} \tau^* &= \inf\left\{t \geq 0 : |I_t| \leq \left\lfloor \frac{\gamma(I_0)}{3\Delta} \right\rfloor\right\} \\ &= \inf\left\{t \geq 0 : V_t \leq \left\lfloor \frac{\gamma(I_0)}{3\Delta} \right\rfloor\right\}. \end{aligned}$$

For every sample path, $\tau^* \leq \tau$. Therefore,

$$\mathbb{E}_{I_0}[\tau] \geq \mathbb{E}_{I_0}[\tau^*].$$

Suppose now that $|I_t|$ satisfies

$$\frac{\gamma(I_0)}{3\Delta} \leq |I_t| \leq \frac{2\gamma(I_0)}{3\Delta}.$$

Then, by parts (i) and (iii) of Lemma 11, we obtain that

$$\frac{\gamma(I_0)}{3} - E\Delta \leq \gamma(I_t) \leq \frac{2\gamma(I_0)}{3}.$$

Furthermore, Lemma 12 implies that

$$c(I_t) \geq \gamma(I_t) - 2(E+2)\Delta \geq \frac{\gamma(I_0)}{3} - (3E+4)\Delta.$$

It follows that the process V_t stochastically dominates a process Y_t, described above, with parameters $\lambda = r$, $\mu = \gamma(I_0)/3 - (3E+4)\Delta$, and $L = \gamma(I_0)/3\Delta$. Therefore, using Eq. (4.9) in the Appendix,

$$\mathbb{E}_{I_0}[\tau] \geq \frac{1}{2r}\left(\left(\frac{\gamma(I_0) - (9E+12)\Delta}{3r}\right)^{\frac{\gamma(I_0)}{3\Delta}-1} - 1\right).$$

□

Note that when $3r < \gamma(I_0) - 9E\Delta - 12\Delta$, the optimal expected extinction time scales exponentially in the resistance of the set of the initially infected nodes. In Chapter 3 we focused on the case where $I_0 = V$ (the worst case) and proved that if the CutWidth of the graph is a sublinear function of the number of nodes, and if $r = o(n)$, then, the expected extinction time is $o(n)$. In contrast, the following result considers the case where W scales linearly in the number of nodes and provides an exponential lower bound on the expected extinction time. Specifically, using Theorem 1 to replace $\gamma(V)$ by W, and using also the definition of E, we can write our lower bound as

$$\mathbb{E}_V[\tau] \geq \frac{1}{2r}\left(\left(\frac{19W - 9n\Delta - 30\Delta}{3r}\right)^{W/3\Delta-1} - 1\right),$$

and we obtain the corollary that follows.

Corollary 3. *Fix a constant $C > 1$ and consider those graphs[1] for which*

[1]Regarding the existence of such graphs, consider as an example the Ramanujan Δ-regular graphs. Their CutWidth is at least equal to $n(\Delta - 2\sqrt{\Delta-1})/2$, since their edge-expansion is at least equal to $\Delta - 2\sqrt{\Delta-1}$

$$W \geq \frac{9C}{19} n\Delta.$$

Moreover, assume that $r < C'n$, *where* $C' < C/3$. *Then,*

$$\mathbb{E}_V[\tau] = \Omega(2^n).$$

Chapter 5

A lower bound for graphs with linear CutWidth: the general case

In the previous chapter, we showed that for graphs with large CutWidth, when the curing budget scales sublinearly, the expected extinction time is exponential for all dynamic curing policies. In this chapter we prove the general result by not restricting our analysis to graphs with large CutWidth. Instead, we consider graphs whose CutWidth scales linearly in the number of nodes (but not necessarily close to $n\Delta/2$).

5.1 The main result and the core of its proof

In this section we state our main result and provide the key elements of its proof in the form of two lemmas. Loosely speaking, the result states that if the resistance of the initial bag scales linearly with the number n of nodes, and the budget scales only as a small constant multiple of n, then the expected time to extinction is exponentially large.

Theorem 5. *Consider a graph with n nodes and a set I_0 of initially infected nodes, and suppose that for some constant c_γ,*

$$\gamma(I_0) \geq c_\gamma n.$$

Suppose, furthermore that all node degrees are bounded above by Δ. Then, there exist positive constants c_r and c, which only depend on c_γ and Δ, such that if

$$r \leq c_r n,$$

then

$$\mathbb{P}\Big(\tau \geq ce^{cn}\Big) \geq \frac{1}{2},$$

under any policy, and for all large enough n. In particular,

$$\mathbb{E}[\tau] \geq \frac{1}{2} ce^{cn}.$$

Remark: An immediate corollary of Theorem 5 is obtained by letting $I_0 = V$, so that $\gamma(I_0)$ coincides with the CutWidth W: if the CutWidth scales linearly in n, and the curing budget is less than a certain multiple of the CutWidth, then the expected time to extinction grows exponentially in n. As a further corollary, if the curing budget grows sublinearly with n, fast extinction is possible only if the CutWidth grows sublinearly in n. This is a converse to the results of Chapter 3 , which establish that if the CutWidth grows sublinearly in n, then fast extinction is possible with a sublinear budget.

The proof of our main result involves the following line of argument.

a) In the first, deterministic, part of the proof (Lemma 15), we show that for graphs with large CutWidth, the time interval until the extinction of the epidemic must contain a substantially long subinterval during which the expected total infection rate is significantly larger than the budget, yet the realized ratio of infections to recoveries is relatively small, and in particular, fairly different than the ratio of the corresponding expected rates.

b) In the second, stochastic, part of the proof (Lemma 16), we argue that for a given time interval to have the properties in a), a "large deviations" event, with exponentially small (in n) probability, must occur. This is used to conclude

that, with significant probability, it will take an exponentially long amount of time until an interval with the properties in a) emerges.

Proof of Theorem 5. We start the proof by fixing a graph with n nodes, and the initial set I_0 of infected nodes. For convenience, from now on, we will use the short-hand notation γ instead of $\gamma(I_0)$. We assume that c_γ and Δ have been fixed, and that $\gamma \geq c_\gamma n$. Note that for sufficiently large n, γ will be much larger than Δ, so that we can use freely inequalities such as $\Delta < \gamma/4$, or $\gamma/4\Delta > 1$. In order to keep notation simple and avoid the use of ceilings and floors, we will also assume from now on that $\gamma/4\Delta$ is an integer. The proof for the general case, is essentially the same.

The first part of the proof corresponds to the following lemma.

Lemma 15. *Consider a sample path for which $\tau < \infty$. For that sample path, there exist times t' and t'', with $0 \leq t' \leq t'' \leq \tau$, such that:*

(i) $c(I_t) \geq \gamma/4$, for all $t \in [t', t'']$;

(ii) we have $b = (\gamma/4\Delta) - 1$ recoveries during the interval $[t', t'']$;

(iii) we have no more than $n + b$ infections during the interval $[t', t'']$.

The times t' and t'' in the preceding lemma are random variables (they depend on the sample path). However, they are not necessarily stopping times of the underlying stochastic process.

Note that it suffices to prove the existence of a time interval $[t', t'']$ with just properties (i) and (ii). This is because there are only n nodes in the graph. If we have b recoveries during a time interval, the number of infections cannot exceed $n+b$, and property (iii) follows automatically.

For the stochastic part of the proof, let us introduce some notation: for any $c > 0$, we define B_c to be the event that there exist times t', t'', with the properties in Lemma 15, together with the additional property $t'' \leq ce^{cn}$.

Lemma 16. *Having fixed c_γ and Δ, there exist small enough positive constants c_r and c such that if $r \leq c_r n$, then*

$$\mathbb{P}(B_c) \leq \frac{1}{2},$$

for all large enough n.

Lemmas 15 and 16 immediately imply Theorem 5. To see this, Lemma 15 implies that t'' is well defined for any sample path. For any sample path that satisfies $\tau \leq ce^{cn}$, we must also have $t'' \leq ce^{cn}$. Thus, the event $\{\tau \leq ce^{cn}\}$ is a subset of the event B_c. Using Lemma 16, we conclude that $\mathbb{P}(\tau \leq ce^{cn}) \leq \mathbb{P}(B_c) \leq 1/2$, as long as c_r and c are suitably chosen.

5.2 Proof of Lemma 15

Lemma 4 is the central — and least obvious — part of the proof. Before continuing with a formal argument, we provide a high-level informal overview, intended to enhance comprehension. The overall plan is to argue that $\gamma(I_t)$, whose initial value is γ, must eventually (at some time T) drop to $\gamma/2$, and that while the value $\gamma/2$ is approached, there must be a sufficiently long interval during which $\mathrm{c}(I_t)$ is at least $\gamma/4$. Indeed, if $\mathrm{c}(I_t) \geq \gamma/4$ for all times in $[0, T]$ (this is Case 1 below), the cut remains relatively large (and larger than the budget), which implies that the process is moving in a direction opposite to its drift; in particular, the probability of this happening is small.

Recall now that the cut is approximately equal to the resistance at those times that the resistance drops. Thus, $\mathrm{c}(I_T)$ is approximately equal to $\gamma/2$. If $\mathrm{c}(I_t)$ drops below $\gamma/4$ before time T (this is Case 2 below), there must exist an interval $[T', T]$ during which $\mathrm{c}(I_t) \geq \gamma/4$, and during which the cut increases from $\gamma/4$ to $\gamma/2$. We want to argue that such an increase must be accompanied by a large number of recoveries (which will consist a low-probability event). The difficulty is that cut increases may be caused by either recoveries or infections. In order to isolate the effects of recoveries, we look at a "bottleneck process" Θ_t that starts the same as I_t at time T', and which keeps track of the recoveries in I_t, while ignoring the infections. Similar to I_t, there will be a time at which the resistance of Θ_t will drop to $\gamma/2$ (this is due to the fact that $\Theta_t \subset I_t$, and monotonicity), and at that time, $\mathrm{c}(\Theta_t)$ will be roughly equal to $\gamma/2$. Thus, $\mathrm{c}(\Theta_t)$ also increases from $\gamma/4$ to $\gamma/2$. However, because Θ_t only changes

whenever the process I_t has a recovery, it follows that there must be $O(\gamma)$ recoveries in the process Θ_t and, therefore, for the process I_t as well (Lemma 17).

We can now start with the formal proof. Let us fix a particular sample path for which $\tau < \infty$. Let T be the first time that $\gamma(I_t)$ drops to a value of $\gamma/2$ or less:

$$T = \inf\{t \geq 0 : \gamma(I_t) \leq \gamma/2\}.$$

Given that $\gamma(I_\tau) = \gamma(\emptyset) = 0$, it is clear that such a time T exists and satisfies $0 \leq T \leq \tau$.

We distinguish between two cases:

Case 1: Suppose that throughout the interval $[0, T]$, we also have $\mathrm{c}(I_t) \geq \gamma/4$. Because of the monotonicity property of $\gamma(\cdot)$ (Lemma 2(i)), $\gamma(I_t)$ decreases only when the set I_t decreases, that is, only when there is a recovery. Furthermore, using the smoothness property in Lemma 2(ii), each time that there is a recovery, $\gamma(I_t)$ can drop by at most Δ. Therefore, the number of recoveries during the time interval $[0, T]$ is at least

$$\frac{\gamma(I_0) - \gamma(I_T)}{\Delta} \geq \frac{\gamma - \gamma/2}{\Delta} = \frac{\gamma}{2\Delta}.$$

We can then find some $\hat{T} \leq T$ such that during the time interval $[0, \hat{T}]$, we have exactly $\gamma/4\Delta - 1$ recoveries, and properties (i)-(ii) in the statement of Lemma 15 are satisfied by letting $t' = 0$ and $t'' = \hat{T}$.

Case 2: Suppose now that there exists some $t \in [0, T]$, with $\mathrm{c}(I_t) < \gamma/4$, which is the more difficult case. Note that just before time T, we have $\gamma(I_{T-}) > \gamma/2$. Furthermore, $\gamma(I_T) \leq \gamma/2$. With our continuous-time Markov chain model, only one event (infection or recovery) can happen at any time. Since $\gamma(I_T) < \gamma(I_{T-})$, and since $\gamma(\cdot)$ is monotonic, it follows that we had a recovery and, therefore, $I_T = I_{T-} - v$, for some node v. Lemma 3 applies, with $A = I_{T-}$ and $A - v = I_T$, and we obtain

$$\mathrm{c}(I_T) \geq \gamma(I_{T-}) > \frac{\gamma}{2}.$$

We now define

$$T' = \sup\{t \leq T : \mathrm{c}(I_t) < \gamma/4\},$$

so that $\mathrm{c}(I_{T'-}) < \gamma/4$. Furthermore, since $\mathrm{c}(I_t)$ can change by at most Δ at each transition (Lemma 1), we must actually have

$$\mathrm{c}(I_{T'}) < \frac{\gamma}{4} + \Delta, \tag{5.1}$$

which implies that $T' \neq T$ and $0 \leq T' < T$.

We will show that the interval $[t', t'']$, with $t' = T'$ and $t'' = T$ has properties (i)-(ii) in the statement of Lemma 15. Indeed, the definition of T' implies that

$$\mathrm{c}(I_t) \geq \frac{\gamma}{4}, \qquad \forall\, t \in [T', T],$$

which is property (i). The proof of Lemma 15 is completed by showing property (ii), namely, that the increase in $\mathrm{c}(I_t)$, from a value smaller than $(\gamma/4) + \Delta$ (at time T'), to a value above $\gamma/2$ (at time T) together with a drop of the resistance from a value above $\gamma/2$ (at time T') to a value below $\gamma/2$ (at time T), must be accompanied by at least $(\gamma/4\Delta) - 1$ recoveries. This is the content of the next lemma.

Lemma 17. *The number of recoveries during the time interval* $[T', T]$ *is at least*

$$\frac{\gamma}{4\Delta} - 1.$$

Lemma 17 is a rather simple statement, but we are not aware of a simple proof or of a transparent intuitive explanation. Our proof relies on an auxiliary process, the *bottleneck process*, coupled with I_t, which is introduced and analyzed in the next section.

5.3 Proof of Lemma 17

The first step in proving Lemma 17 is the construction of a process which is coupled

with the infection process. Observe that a sample path of the infection process defines a crusade in which, at each step, a single node is added to or removed from the current bag. To any such crusade, we associate a *bottleneck sequence*, which is a sequence of bags consisting of subsets of the bags in the original crusade, with several important properties. Consider a crusade $\omega = (A_0, A_1, \ldots, A_l)$ in which $|A_i \triangle A_{i-1}| = 1$, for $i = 1, \ldots, l$. In particular, we always have $A_i \subset A_{i-1}$ or $A_i \supset A_{i-1}$. We associate with ω a related sequence of bags $(\Theta_0, \ldots, \Theta_l)$, by letting

$$\Theta_i = \bigcap_{k=0}^{i} A_k, \qquad i = 0, \ldots, l. \tag{5.2}$$

It is clear from our construction that Θ_i is always a subset of A_i, and that $\Theta_i \supseteq \Theta_{i-1}$. We have the following interpretation: Θ_0 starts the same as A_0. Whenever a node is removed from a bag in the original sequence, the same is done in the bottleneck sequence, as long as this is possible. On the other hand, whenever a node is added to a bag in the original sequence, nothing is done in the bottleneck sequence.

Lemma 18. *Consider a sequence $(A_0, A_1, \ldots, A_l)$ of bags such that $|A_i \triangle A_{i-1}| = 1$, for $i = 1, \ldots, l$, and the associated bottleneck sequence $(\Theta_0, \ldots, \Theta_l)$. The following hold:*

(i) $\Theta_i \subseteq A_i$.

(ii) If $c(\Theta_i) > c(\Theta_{i-1})$, *then* $A_i \subset A_{i-1}$.

(iii) $c(\Theta_i) - c(\Theta_{i-1}) \leq \Delta$.

Proof. Proof:

(i) Follows directly from the definition.

(ii) Suppose that $c(\Theta_i) > c(\Theta_{i-1})$. Then, $\Theta_i \neq \Theta_{i-1}$. From the definition of the bottleneck sequence, we see that it if $A_i \supset A_{i-1}$, then $\Theta_i = \Theta_{i-1}$. Therefore, we must have that $A_i \subset A_{i-1}$.

(iii) If $A_i \supset A_{i-1}$, then, $\Theta_i = \Theta_{i-1}$, and $c(\Theta_i) - c(\Theta_{i-1}) = 0$. On the other hand, if $A_i \subset A_{i-1}$, and using the assumption $|A_i \triangle A_{i-1}| = 1$, we write $A_i = A_{i-1} - v$ for

some $v \in A_{i-1}$, and from Eq. (5.2) we obtain $\Theta_i = \Theta_{i-1} - v$. The result then follows from Lemma 1(i). $\square$

We now complete the proof of Lemma 17. Let $A_0, \ldots, A_l$ be the sequence of bags that arise during the evolution of I_t, between times T' and T. In particular, $A_0 = I_{T'}$ and $A_l = I_T$. Let $\Theta_0, \ldots, \Theta_l$ be the corresponding bottleneck sequence, so that $\Theta_0 = A_0 = I_{T'}$. Using property (i) in Lemma 18, we have $\Theta_i \subseteq A_i$, for all i. Using the nonotonicity of $\gamma(\cdot)$, we obtain $\gamma(\Theta_i) \leq \gamma(A_i)$, for all i. In particular,

$$\gamma(\Theta_l) \leq \gamma(A_l) = \gamma(I_T) \leq \frac{\gamma}{2} < \gamma(I_{T'}) = \gamma(\Theta_0).$$

(The second and third inequalities follow from the definition of T and the fact $T' < T$, respectively.) This implies that there exists some $i \in \{1, \ldots, l\}$ for which

$$\gamma(\Theta_i) \leq \frac{\gamma}{2} < \gamma(\Theta_{i-1}).$$

We apply Lemma 3 and obtain that $\mathrm{c}(\Theta_i) \geq \gamma(\Theta_{i-1}) > \gamma/2$. Thus, the bottleneck sequence starts with $\mathrm{c}(\Theta_0) = \mathrm{c}(I_{T'}) < (\gamma/4) + \Delta$ (cf. Eq. (5.1)) and eventually its cut rises to a value above $\gamma/2$. From part (ii) of Lemma 18, $\mathrm{c}(\Theta_i)$ can increase only when there is a recovery. From part (iii) of Lemma 18, $\mathrm{c}(\Theta_i)$ can increase by at most Δ at each recovery. Thus, in order to obtain an increase from $(\gamma/4) + \Delta$ to $\gamma/2$, we must have had at least $(\gamma/4\Delta) - 1$ recoveries in the process I_t between times T' and T.

A schematic summary of the two cases introduced in Section 5.2 is provided in Figure 5-1.

5.4 Proof of Lemma 16

Lemma 16 is a fairly routine "large deviations" result. It is useful to provide some intuition by considering the special case in which the times t' and t'' are fixed (not random), and $\mathrm{c}(I_t) = \gamma/4$ throughout the interval $[t', t'']$ (as opposed to $\mathrm{c}(I_t) \geq \gamma/4$). In this case, we have a Poisson process (recoveries) with rate r and an independent

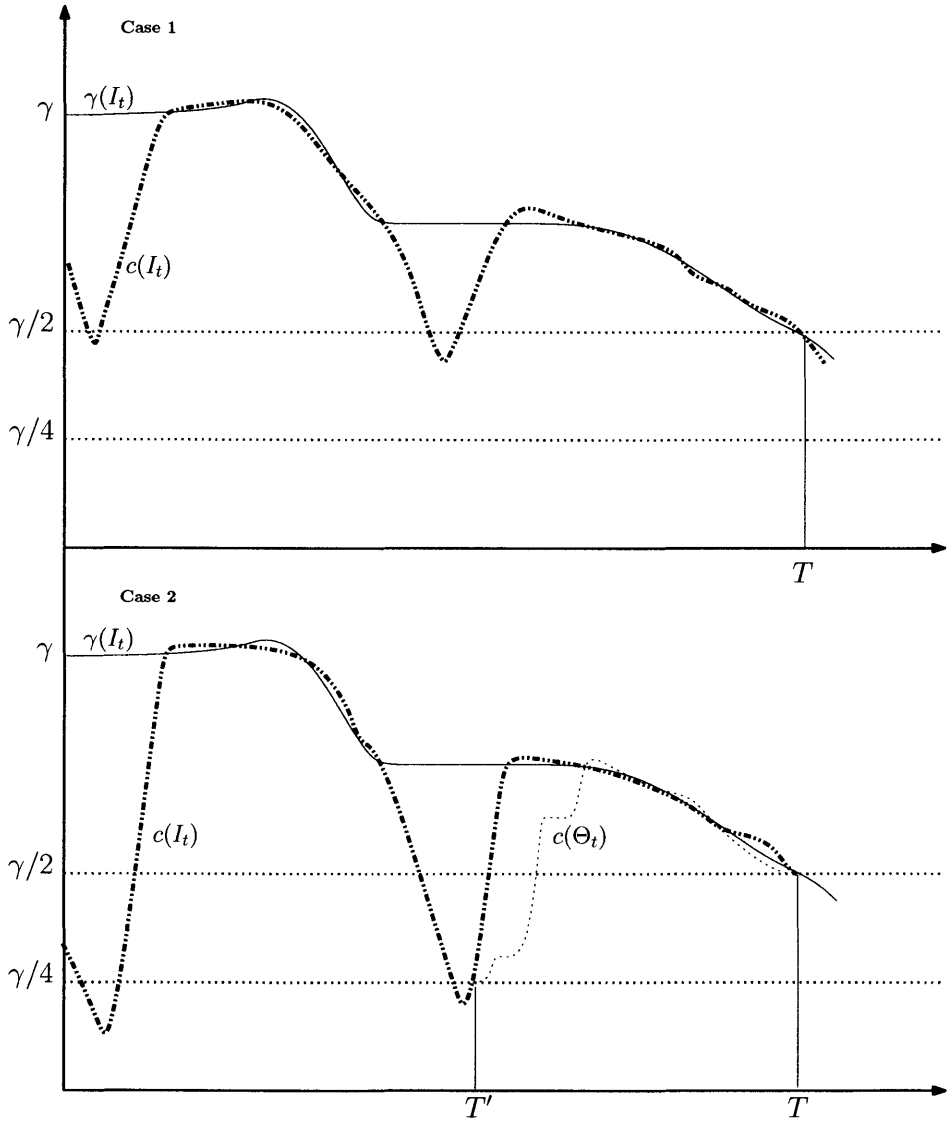


Figure 5-1: **Case 1:** In the first case, $c(I_t)$ remains at least $\gamma/4$ throughout the interval $[0, T]$. Moreover, since the resistance drops from γ to $\gamma/2$, at least $\gamma/2$ recoveries must occur. **Case 2:** In the second case, $c(I_t)$ drops below $\gamma/4$. The last time that it does so (time T'), the resistance is above $\gamma/2$ and needs to drop to a value below $\gamma/2$. Therefore, $c(I_t)$ needs to grow above (roughly) $\gamma/2$. In principle, this increase may happen through infections and not only through recoveries. This is why we define the auxiliary process Θ_i, whose cut also needs to increase to $\gamma/2$ but can only increase through recoveries, implying that at least (roughly) $\gamma/4\Delta$ recoveries occur.

Poisson process (infections) with rate $\gamma/4$; their ratio is $4r/\gamma$. For properties (i) and (ii) in Lemma 15 to hold, the empirical ratio of observed recoveries to infections must be at least

$$\frac{b}{n+b} = \frac{(\gamma/4\Delta)-1}{n+(\gamma/4\Delta)-1},$$

where b is as defined in Lemma 15. When r is small compared to $\gamma/4$, which is the case if we choose c_r small enough, we have an empirical ratio of recoveries to infections which is above the theoretical ratio by a constant factor. Large deviations theory implies that this event has exponentially small probability. We then argue that within the time horizon of interest, $[0, ce^{cn}]$, there are only $O(ne^{cn})$ intervals that need to be considered. By choosing c small enough and using the union bound, the overall probability that there exist t' and t'' with the desired properties can be made small.

The proof for the general case runs along the same lines but involves a coupling argument to show that when $\mathrm{c}(I_t)$ can exceed $\gamma/4\Delta$, then the event of interest (relatively few infections or, equivalently, too many recoveries) is even less likely to occur.

5.4.1 Decomposing the event of interest

Let c be a small enough constant — how small it needs to be will be seen at the end of the proof. Let $t^* = ce^{cn}$, which is the time horizon of interest in Theorem 5. Recall our definition of the event B_c in Section 5.1: event B_c occurs if and only if there exists a time interval $[t', t'']$ with $t'' \leq ce^{cn} = t^*$, with exactly $b = (\gamma/4\Delta) - 1$ recoveries, with at most $n + b$ infections, and during which $\mathrm{c}(I_t) \geq \gamma/4$.

Our first step is to show that only a finite number of intervals $[t', t'']$ need to be considered. The recovery process behaves as a Poisson process with rate r, as long as the absorbing state has not been entered. To simplify the presentation, let us redefine the process, so that recoveries take place forever, according to a Poisson process. Any recovery that occurs after the extinction time τ is "dummy" and has no effect on the process $\{I_t\}$.

For $i \geq 1$, let t_i, be the time of the ith recovery (actual or dummy). We consider

the time interval $[t_i, t_{i+b-1}]$, which is the interval until $b-1$ new recoveries are observed, after the time t_i of the ith recovery.

For $i \geq 1$, we define B_i as the event that throughout the interval $[t_i, t_{i+b-1}]$ we have $\mathrm{c}(I_t) \geq \gamma/4$ and at most $n + b$ infections.

Lemma 19. $B_c \subseteq \bigcup_{i=1}^{\infty} B_i$.

Proof. Proof: Consider a sample path that belongs to B_c, so that there exists an interval $[t', t'']$ with the properties in the definition of B_c. In particular, there exists some $i \geq 1$ such that the interval $[t', t'']$ contains the times $t_i, \ldots, t_{i+b-1}$, i.e.,

$$t' \leq t_i \leq t_{i+b-1} \leq t'';$$

furthermore, $\mathrm{c}(I_t) \geq \gamma/4$ during that interval, and we have at most $n + b$ infections. But in that case, the interval $[t_i, t_{i+b-1}]$ has all of the properties that are required for event B_i to hold. □

Let K be the total number of recoveries (real or dummy) during the time interval $[0, t^*]$. Using Lemma 19 and the union bound, we obtain

$$\mathbb{P}(B_c) \leq \sum_{i=1}^{4rt^*} \mathbb{P}(B_i) + \mathbb{P}(K > 4rt^*) \leq \sum_{i=1}^{4rt^*} \mathbb{P}(B_i) + \frac{1}{4}, \tag{5.3}$$

where the last inequality is obtained from the fact that K is a (Poisson) random variable with mean rt^*, and the Markov inequality.

It remains to bound the sum of the $\mathbb{P}(B_i)$. Since t^* grows exponentially with n, we are looking for an exponentially small upper bound on each B_i. This is the subject of the next subsection.

5.4.2 Bounding $\mathbb{P}(B_i)$

The main obstacle in characterizing $\mathbb{P}(B_i)$ is that the infection process has a time-varying rate. We will handle this issue through a coupling with a Poisson process that has a constant rate.

For $t \geq t_i$, let $M_i(t)$ be the number of infections during the interval $[t_i, t]$, Let also

$$C_i(t) = \big\{\mathrm{c}(I_t) \geq \gamma/4,\ \forall t \in [t_i, t]\big\},$$

which is the event that $\mathrm{c}(I_t)$ remains "large" during the interval $[t_i, t]$. Then, the event B_i can be expressed as

$$B_i = \big\{M_i(t_{i+b-1}) \leq n + b\big\} \cap C_i(t_{i+b-1}).$$

For the remainder of the proof, we assume that c_r is chosen (based only on c_γ and Δ, as in the statement of the theorem) so that

$$c_r < \frac{c_\gamma^2}{40\Delta} \tag{5.4}$$

By rearranging terms, it is then seen that we can fix a constant $\bar{t}$ that again depends only on c_γ and Δ, which satisfies

$$c_r \bar{t} < \frac{c_\gamma}{5\Delta} \qquad \text{and} \qquad \frac{c_\gamma}{4}\bar{t} > 2 \tag{5.5}$$

For some interpretation and an outline of the rest of the argument, $\bar{t}$ is chosen so that, with high probability, the interval $[t_i, t_i + \bar{t}]$ has fewer than $b - 1$ recoveries, but more than $n + b$ infections if the cut remains "large." As will be seen, this property of $\bar{t}$ implies that, with high probability, the event B_i does not occur.

We define the event $\overline{B}_i$ by

$$\overline{B}_i = \big\{t_{i+b-1} < t_i + \bar{t}\big\} \cup \Big(\big\{M_i(t_i + \bar{t}) \leq n + b\big\} \cap \mathrm{c}(t_i + \bar{t})\Big).$$

We will now show that $B_i \subseteq \overline{B}_i$. Consider a sample path in B_i. If that sample path also satisfies $t_{i+b-1} < t_i + \bar{t}$, then it is also an element of $\overline{B}_i$. Suppose now that the sample path satisfies $t_{i+b-1} \geq t_i + \bar{t}$. Using the monotonicity of the counting process $M_i(\cdot)$, we obtain $M_i(t_i + \bar{t}) \leq M_i(t_{i+b-1}) \leq n + b$, where the last inequality holds because the sample path belongs to B_i. Furthermore, since the sample path belongs

to B_i, it must belong to $C_i(t_{i+b-1})$, which implies that it must also belong to $C_i(t_i+\bar{t})$. Thus, the sample path belongs to $\{M_i(t_i+\bar{t}) \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})$, and is therefore an element of $\overline{B}_i$. This concludes the proof that $B_i \subseteq \overline{B}_i$. It then follows, using the union bound, that

$$\mathbb{P}(B_i) \leq \mathbb{P}(\overline{B}_i) \leq \mathbb{P}(t_{i+b-1} < t_i+\bar{t}) + \mathbb{P}\Big(\{M_i(t_i+\bar{t}) \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})\Big). \quad (5.6)$$

Our next step is to derive an upper bound for each of the two terms on the right-hand side of Eq. (5.6), in terms of the Poisson distribution. For the first term, this is simple. The event $\{t_{i+b-1} < t_i+\bar{t}\}$ is the event that starting from time t_i, at least $b-1$ recoveries occur within $\bar{t}$ time units. Since the recovery process is Poisson with rate r, we have

$$\mathbb{P}(t_{i+b-1} < t_i+\bar{t}) = \mathbb{P}\big(R > b-1\big), \quad (5.7)$$

where R is a Poisson random variable with mean $r\bar{t}$.

To study the second term, we use $\mathbb{1}_C$ to denote the indicator function of the event $C_i(t_i+\bar{t})$. For those sample paths that belong to $C_i(t_i+\bar{t})$, and during the interval $[t_i, t_i+\bar{t}]$, the counting process $M_i(\cdot)$ maintains a rate that is larger than or equal to $\gamma/4$. Thus, on that time interval, $M_i(\cdot)$ can be coupled with a Poisson process $\overline{M}(\cdot)$ with rate equal to $\gamma/4$, in a way that guarantees that

$$M_i(t_i+\bar{t})\mathbb{1}_C \geq \overline{M}_i(t_i+\bar{t})\mathbb{1}_C,$$

for every sample path. Using this dominance relation, we obtain

$$\begin{aligned}
\mathbb{P}\Big(\{M_i(t_i+\bar{t}) \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})\Big) &= \mathbb{P}\Big(\{M_i(t_i+\bar{t})\mathbb{1}_C \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})\Big) \\
&\leq \mathbb{P}\Big(\{\overline{M}_i(t_i+\bar{t})\mathbb{1}_C \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})\Big) \\
&= \mathbb{P}\Big(\{\overline{M}_i(t_i+\bar{t}) \leq n+b\} \cap \mathrm{c}(t_i+\bar{t})\Big) \\
&\leq \mathbb{P}\big(\overline{M}_i(t_i+\bar{t}) \leq n+b\big) \\
&= \mathbb{P}(\overline{M} \leq n+b), \quad (5.8)
\end{aligned}$$

where $\overline{M}$ is a Poisson random variable with mean $\gamma\bar{t}/4$.

We are now ready to apply large deviations results for Poisson random variables. Note that a Poisson random variable with mean λn can be viewed as a sum of n independent Poisson random variables with mean λ, and therefore, by the Chernoff bound, the probability of deviating from the mean by a constant factor falls exponentially with n. We record this fact in the lemma that follows, which just asserts the fact that we have a positive large deviations exponent.

Lemma 20. *There exists a function $\epsilon(\lambda, \lambda')$, defined for positive λ and λ', and which is positive whenever $\lambda \neq \lambda'$, with the following properties.*

(i) Let X be a Poisson random variable with mean bounded above by λn. If $\lambda' > \lambda$, then

$$\mathbb{P}(X \geq \lambda' n) \leq e^{-\epsilon(\lambda,\lambda')n}, \qquad \forall\ n.$$

(ii) Let X be a Poisson random variable with mean bounded below by λn. If $\lambda' < \lambda$, then

$$\mathbb{P}(X \leq \lambda' n) \leq e^{-\epsilon(\lambda,\lambda')n}, \qquad \forall\ n.$$

The random variable R in Eq. (5.7) is Poisson with mean $r\bar{t} \leq c_r\bar{t}n$. Note that, for large enough n, we have $b-1 = (\gamma/4\Delta) - 2 \geq (\gamma/5\Delta) \geq (c_\gamma/5\Delta)n$, where the last inequality follows from the fact that $\gamma \geq c_\gamma n$. . We apply Lemma 20(i), with $\lambda = c_r\bar{t}$ and $\lambda' = c_\gamma/5\Delta$:

$$\mathbb{P}\big(R > b-1\big) \leq \mathbb{P}\big(R > (c_\gamma/5\Delta)n\big) \leq e^{-\epsilon_1 n}.$$

Because of our assumptions on c_r and $\bar{t}$ (cf. Eq. (5.5)), we have $\lambda' > \lambda'$, and ϵ_1 is a positive number determined by c_r, c_γ, and Δ.

Similarly, the random variable $\overline{M}$ in Eq. (5.8) is Poisson with mean $\gamma\bar{t}/4 \geq (c_\gamma\bar{t}/4)n$. For any graph, γ is bounded above by $n\Delta$, and this implies that $b = (\gamma/4\Delta) - 1 \leq n$. We apply Lemma 20(ii), with $\lambda = c_\gamma\bar{t}/4$ and $\lambda' = 2$:

$$\mathbb{P}\big(\overline{M} \leq n + b\big) \leq \mathbb{P}\big(\overline{M} \leq 2n\big) \leq e^{-\epsilon_2 n}.$$

Because of our assumptions on c_γ and $\bar{t}$ (cf. Eq. (5.5)), we have $\lambda' < \lambda$, and ϵ_2 is a positive number determined by c_γ.

We have therefore established that each of the two terms on the right-hand side of Eq. (5.6) is bounded above by an exponentially decaying term. By letting $\epsilon = \min\{\epsilon_1, \epsilon_2\} > 0$, we obtain that

$$\mathbb{P}(B_i) \leq 2e^{-\epsilon n}. \tag{5.9}$$

5.4.3 Completing the proof of Lemma 16.

For the given c_γ and Δ, we choose a suitably small c_r as in Eq. (5.4). This allows us to set $\bar{t}$ as in Eq. (5.5), leading to a positive ϵ in Eq. (5.9). We then use Eq. (5.9) to bound the terms $\mathbb{P}(B_i)$ in the inequality (5.3), and also make use of the facts that $t^* = ce^{cn}$ and $4rt^* \leq 4c_r nce^{cn}$, to obtain

$$\mathbb{P}(B_c) \leq 4c_r nce^{cn} 2e^{-\epsilon n} + \frac{1}{4} \leq \frac{1}{2},$$

provided that c is small enough (it just needs to be chosen a little smaller than ϵ) and n is large enough. This concludes the proof of Lemma 16.

5.5 Conclusions

We have considered the control of an epidemic (contagion process) given a limited curing budget, and provided an exponential lower bound on the expected time to extinction, for bounded degree graphs under any dynamic curing policy. For the interesting (and least favorable) case where all nodes are initially infected, our assumption was that the CutWidth of the graph scales linearly with the number of nodes, and that the curing budget is bounded above by a small enough multiple of the number of nodes.

It remains an open problem to develop lower bounds for more general bounded-degree graphs, whose CutWidth scales sublinearly with the number of nodes.

Chapter 6

Summary of Theoretical Results and Open Questions

In the preceding chapters, we have considered the control of an epidemic (contagion process) given a limited curing budget. In Chapter 3 we proposed an order optimal policy, while in Chapters 4 and 5 we provided lower bounds on the expected time to extinction for any dynamic curing policy.

In this chapter we contrast these results and describe an open problem that remains unanswered.

6.1 Case I: All nodes initially infected

We start our discussion by considering the case of all nodes being initially infected. A standard coupling argument reveals that this is the worst case scenario for the optimal dynamic curing policy with respect to the achieved expected extinction time.

With regards to our discussion in Section 1.4 and in order to clearly illustrate the contributions of this thesis, we define a network to be **resilient** if and only if there exists a policy that achieves sublinear extinction time with sublinear budget, assuming that all nodes are initially infected.

Chapters 3, 4 and 5 provide necessary and sufficient conditions for the resilience of a bounded degree network, as the following theorem illustrates.

Theorem 6. *A network is resilient if and only if its CutWidth is sublinear.*

Proof. Sufficiency follows from Corollary 2 (d). Necessity follows from Theorem 5. □

We now focus on the set of graphs whose CutWidth scales linearly in the number of nodes (large CutWidth). Our results if combined show that for such graphs, the ratio of the curing resources to the CutWidth is the key factor that distinguishes between slow and fast extinction, as shown in Figure 6-1.

Figure 6-1: Slow vs. fast extinction for graphs with large CutWidth. The ratio of the curing resources to the CutWidth is the key factor that distinguishes between slow and fast extinction.

Our results do not provide a sharp conclusion for the case of graphs whose CutWidth scales sublinearly in the number of nodes (small CutWidth) as Figure 6-2 illustrates. More specifically, Corollary 2(c) specifies conditions under which fast extinction is achievable by the CuRe policy. On the other hand, Theorem 5 applies only to the case of graphs with large CutWidth and it remains an open problem to develop similar lower bounds for such more general bounded-degree graphs. In some cases, this is easy. For example, for a square mesh with n nodes, the CutWidth is of order $O(\sqrt{n})$. Using the fact that any subset of the mesh with (roughly) $n/2$ nodes has a cut of size $\Omega(\sqrt{n})$, one can show that a curing budget that scales at least as fast as $\sqrt{n}$ is necessary for fast extinction. The same argument applies whenever we deal with families of graphs that satisfy suitable isoperimetric inequalities.

We conjecture that a similar result is always true: that is, unless the curing budget scales in proportion with the CutWidth, the expected time to extinction will be exponential, if all nodes are initially infected. However, some new tools may have to be developed.

More specifically, the proof of Theorem 5, and in particular Eq. (5.4), shows that the exponential lower bound holds when c_r is smaller than a constant multiple of c_γ^2.

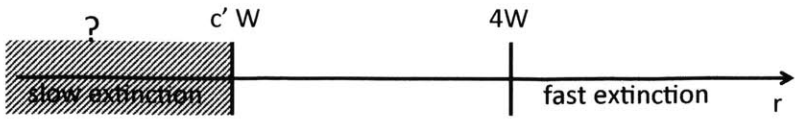


Figure 6-2: Slow vs. fast extinction for graphs with small CutWidth. If the curing budget is larger than the CutWidth of the graph, then fast extinction is achieved. Otherwise, we conjecture that fast extinction is not achievable.

	fast extinction	**slow extinction**
$\rho_v(t) = \rho$	$r > n\rho(A)$	$r < n\eta$
$\rho_v(t) = d_v$	$r > Cn$	-
CuRe	$r > \max\{4\delta(I_0), n/\log(n)\}$	-
any policy	not applicable	~~expander graphs~~ graphs with $W = \Theta(n)$ and $\rho < C''n$

Table 6.1: Existing and New Results: Conditions for fast and slow extinction under different curing policies, assuming all nodes initially infected.

We conjecture that a similar lower bound can be established under the assumption that c_r is smaller than a constant multiple of c_γ. If this is true, the deciding factor will be the ratio between the resistance and the recovery rate in a very concrete sense. However, the proof of this conjecture, if true, will require a much more refined argument.

Table 6.1 summarizes out contributions in the literature for the case of all nodes initially infected.

6.2 Case II: Some nodes initially infected

In this section we consider the more general case where a subset of nodes is initially infected. As we discussed in Chapter 2, there are two notions associated with the curing problem of a subset of nodes A, the resistance $\gamma(A)$ and the impedance $\delta(A)$.

Our results in Chapter 3 are phrased in terms of the impedance of A. Specifically, in Corollary 2 we argue that if the curing budget is larger than $4\delta(A)$ (and larger than $n/\log n$) then sublinear extinction time can be achieved.

In contrast, in Chapters 4 and 5 our results are phrased in terms of the resistance

$\gamma(A)$. For example in Theorem 5, we show that if the resistance $\gamma(A)$ is larger than a linear function of the number of nodes then for sufficiently low curing budget the expected time to extinction is exponential for all dynamic curing policies.

Essentially, if we focus on the regime where $\gamma(I_0) = \Theta(n)$ and consider the relation between the curing budget and the two metrics $\delta(A)$ and $\gamma(A)$ we obtain the following conclusions: (i) if the curing budget is smaller than a constant fraction of the resistance $\gamma(I_0)$, then extinction is slow while (ii) if the curing budget is larger than the impedance $\delta(I_0)$ then the extinction is fast using the CuRe policy, as Figure 6-3 illustrates.

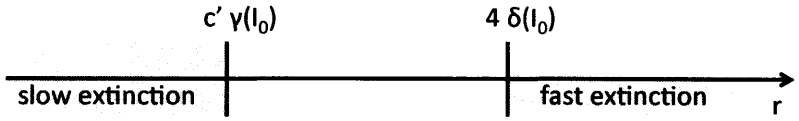


Figure 6-3: If the curing budget is larger than the the impedance of the set of initial infections, then fast extinction is achieved. On the other hand, fast extinction is achievable if the curing budget is smaller than the resistance of the set of initial infections.

Therefore, in order to evaluate the sharpness of this result one should be able to bound the gap between $\gamma(I_0)$ and $\delta(I_0)$. For the case where $I_0 = V$, i.e., the case where all nodes are initially infected (Theorem 1), the gap is zero. For all other cases, Corollary 1 shows that the gap is bounded by $c(I_0)$. In other words, for initial infections with small cut the ratio of the curing resources to the resistance is the key factor that distinguishes between slow and fast extinction. On the other hand, for initial infections with a large cut, our result is not as sharp and remains to be proven whether the impedance or the resistance is the key quantity that determines the speed (slow or fast extinction).

6.3 Conclusions of the theoretical part of the thesis

We have presented a dynamic curing policy with desirable performance characteristics. This policy applies to any subset of initially infected nodes and the resulting expected time to extinction is order-optimal when the available budget is sufficiently large.

Moreover, we provided corresponding lower bounds on the performance of an optimal dynamic curing policy.

Our analysis brings up a number of open problems of both practical and theoretical interest. Specifically, a major drawback of the CuRe policy is its computational complexity, because calculating the impedance of a bag or finding a target path is computationally hard. Therefore, one possible direction is the design of computationally efficient policies with some performance guarantees, perhaps for special cases.

In the same spirit, certain combinatorial optimization tools have been developed for the approximation of the CutWidth of a graph. Such tools can be employed to approximate the impedance of a bag. An interesting direction is the analysis of the performance of the CuRe policy when, instead of optimal crusades, approximately optimal crusades are used.

Furthermore, we have argued in this thesis that the CuRe policy is efficient in the sense of attaining near-optimal, $O(n/r)$ expected time to extinction, in a certain parameter regime. It is an interesting problem to look for approximately optimal policies over a wider set of regimes.

Concluding, our theoretical analysis is based on a set of assumptions that if relaxed give rise to a number of interesting directions for future work:

(i) **Full information:** In this thesis, we assumed that the network planner has full information on the state of each node at every time instant. In reality, though, one is given access to *partial* information on the underlying state (due to corrupted or noisy data, mis-reporting etc.).

(ii) **Budget constraint:** In our analysis we assumed that the network planner allocates a total curing budget R at each time instant t. In more realistic situations though, a network planner would be provided with a total curing budget to be allocated over the course of time, which leads to a more complicated decision problem.

(iii) **Curing and Removal:** In many applications, such as the human epidemics, the network planner can intervene to the evolution of the contagion process

not only through curing nodes but also through removing nodes (quarantine or vaccination). This consideration gives rise to richer models and possibly richer insights.

Chapter 7

Modeling Infectious Diseases: Influenza in the United States

In this chapter, we turn our attention to the modeling and analysis of infectious diseases. Specifically, our goal is to investigate the accuracy of a popular class of models studied in the literature and understand the effect of inter-state traveling on nationwide propagation of influenza-like infections. This latter question is of great importance in the context of epidemic control as it quantifies the effectiveness of state quarantining or regulation of inter-state traveling.

In order to do that, we estimate such models using data[1] on the evolution of influenza-like infections (ILI) in the United States. We then use the estimated models to test their prediction accuracy, and to resolve a hypothesis testing problem aiming to understand whether inter-state traveling is significant in the propagation of ILI.

Starting from the first known result in mathematical epidemiology in 1760 by Daniel Bernoulli [7], mathematical modeling approaches have been the main resort to compare and test theories, as well as to gauge uncertainties in intervention strategies against infectious diseases and *epidemics*.

Modern epidemic models generally assume that the population can be divided into different classes or *compartments* depending on the stage of the disease. The seminal work of Kermack and McKendrick [30] assumes that the population can be divided

[1]Special thanks goes to Vasilis Milias, who helped in the early stages of data-collection.

into two or three compartments depending on the type of the disease:

(i) **Susceptible**: individuals who have no immunity to the infectious agent, so that they might become infected if exposed.

(ii) **Infectious**: individuals who are currently infected and can transmit the infection to susceptible individuals who they contact.

(iii) **Removed** (or **Recovered**): individuals who are immune to the infection, and consequently do not affect the transmission dynamics in any way when they contact other individuals.

Several models have been proposed in the literature to describe the time evolution of the size of each compartment. In this chapter we briefly describe the most popular, networked compartmental models, (Section 7.1) and draw qualitative connections between these models and the model of Section 2 (Section 7.2). In Section 7.3 we present the data that we use to estimate the model and evaluate its predictive accuracy. In Section 7.4 we finalize the structure of the model to be estimated, while in Section 7.5 we present our approach and results. Finally, in Section 7.6 we provide a preliminary attempt to identifying the effect of inter-state traveling on the propagation of ILI.

7.1 Networked Compartmental Models

In this section we present the model that we use in the rest of the chapter. Essentially, the model consists of a system of differential equations describing how the sizes of the compartments change over time. Solutions to these equations will yield, in particular, the size of the each compartment at time t. The numbers of individuals in each compartment must be integers, of course, but if the host population size N is sufficiently large, then they can be treated as continuous variables.

In the simplest version of compartmental models where there is a unique host population, their evolution can be treated through a system of differential equations,

$$\begin{aligned}\dot{S}(t) &= -\beta S(t)I(t),\\ \dot{I}(t) &= \beta S(t)I(t) - mI(t),\\ \dot{R}(t) &= mI(t),\end{aligned} \tag{7.1}$$

where $S(t)$ denotes the number of susceptible individuals, $I(t)$ denotes the size of infected individuals, and $R(t)$ denotes the number of recovered individuals. Here, the transmission (or contact) rate (per capita) is β and the recovery rate is m. Note that the equation for $\dot{R}(t)$ has no effect on the dynamics of $S(t)$ and $I(t)$ (formalizing the fact that removed individuals cannot affect transmission). Intuitively, since the probability that a random contact by an infected individual is made with a susceptible is given by $S(t)/N$, the number of new infections in unit time per infected individual is $(\beta N)(S(t)/N)$, giving a rate of new infections $(\beta N)(S(t)/N)I(t) = \beta S(t)I(t)$.

One of the main and strongest assumptions of the model described above is that the host population is homogeneous, in the sense that contact rate is the same across all individuals or, equivalently, that each infected (or susceptible) individual is randomly and uniformly matched with another individual on the network.

Several models for modeling inhomogeneous populations have been proposed in the literature, by studying SIR-type models on complex networks of interactions (see [44], [6], [57], [48] and references therein). This literature can essentially be divided into *stochastic* models and *deterministic* models. In this chapter we focus on the latter but we briefly review the connections with the former in the next section.

Essentially, one can think of two extremes as follows: Either, a model that considers the whole population of N individuals as in (7.1), or focus on the details of the interactions of each of N individuals, which leads to complex dynamical system (such as the model studied in Chapters 2-5).

Bridging these two approaches, researchers ([39], [33]) have proposed to create $M \ll N$ subpopulations to approximate the dynamics of the entire N-dimensional system. We can then keep track of the state of each subpopulation rather than the

state of each individual in the population, thus providing a more realistic (in terms of homogeneity) approach, but significantly simpler in its description and analysis compared to tracking the state of each individual.

More specifically, let $i \in \{1, ..., M\}$ denote the i-th subpopulation, with N_i individuals, where each individual from the original population with N people is assigned to exactly one subpopulation. In other words, $\sum_{i=1}^{M} N_i = N$. Note that the number of individuals in each subpopulation does not need to be the same. Intuitively, each subpopulation can be thought of as a large city, state, or country.

The main assumption of these *meta-population* models is that each subpopulation i is well-mixed (in the sense that interactions are random within the subpopulation) and that both recovery and transmission rates are homogeneous and equal to m_i and β_{ii} respectively. However, each subpopulation can affect each of the other subpopulations, for example through traveling and "long range" contacts. Hence, there is an infection rate β_{ij} that captures the effect that subpopulation i has on subpopulation j. Note that it is not required that $\beta_{ij} = \beta_{ji}$ nor does it make sense be so. We denote by $I_i(t)$, $S_i(t)$, $R_i(t)$ the number of infected, susceptible and removed individuals of subpopulation i, respectively. The meta-population model for the SIR dynamics consists of the following system of differential equations

$$\begin{aligned}
\dot{S}_i(t) &= -\sum_{j=1}^{M} \beta_{ji} S_i(t) \frac{I_j(t)}{N_j}, \\
\dot{I}_i(t) &= -m_i I_i(t) + \sum_{j=1}^{M} \beta_{ji} S_i(t) \frac{I_j(t)}{N_j}, \\
\dot{R}_i(t) &= m_i I_i(t).
\end{aligned} \tag{7.2}$$

This has now reduced the original N-dimensional system into an M-dimensional one. Essentially, varying the number of subpopulations considered, one can study different levels of granularity. If $M = 1$ we recover the original SIR model of Equation 7.1 while if $M = N$ one can study each individual separately.

7.2 Connections between Compartmental and Stochastic models

In this section we draw qualitative connections between the networked compartmental model of Section 7.1 and the SIS model described in Section 2.1. In particular, the models of Section 7.1 consist of a system of ordinary differential equations in contrast to the model described in Section 2.1 which is a Markov process describing the *stochastic* transitions between infection states of the nodes on a network. Therefore, these models are seemingly unrelated.

More specifically, revisiting the model of Section 2.1 and assuming that (i) infections on the edge (i, j) happen according to a Poisson process with rate β_{ij} (instead of β as described in Section 2.1) and (ii) recoveries happen at a node specific static rate, i.e., that the dynamic curing policy takes the simple form $\rho_i(t) = m_i$, the infection process takes the following simple form:

$$X_i(t) : 0 \to 1 \text{ with rate } \sum_{(j,i)\in E} \beta_{ji} X_j(t),$$
$$X_i(t) : 1 \to 0 \text{ with rate } m_i X_i(t),$$

where $X_i(t) \in \{0, 1\}$ denotes the infection state of node i. We denote by $p_i(t)$ the probability that node i is infected and observe that $p_i(t) = \mathbb{E}[X_i(t)]$. This allows us to write the exact equations for the probability of being infected, for each node i of the SIS model,

$$\dot{p}_i(t) = \frac{d\mathbb{E}[X_i(t)]}{dt} = \mathbb{E}\left[-m_i X_i(t) + (1 - X_i(t)) \sum_{(j,i)\in E} \beta_{ji} X_j(t)\right]. \tag{7.3}$$

Using the linearity of expectation, as well as some algebraic manipulations, we obtain

$$\dot{p}_i(t) = -m_i p_i(t) + \beta \sum_{(j,i)\in E} p_j(t) - \sum_{(j,i)\in E} \beta_{ji} \mathbb{E}[X_i(t) X_j(t)],$$

which does not allow for an explicit solution because the equation for $p_i(t)$ depends on the two-node expectation $\mathbb{E}[X_i(t)X_j(t)]$. Its exact computation requires the knowledge of the joint probability distribution $\mathbb{P}(X_i = 1, X_j = 1)$ for the state of nodes i and j, which is in general impossible. Nevertheless, in order to derive a closed set of N dynamical equations, the *Mean-Field* approximation is commonly enforced which assumes that the states of neighboring nodes are *uncorrelated*, i.e.,

$$\mathbb{E}[X_i(t)X_j(t)] = \mathbb{P}(X_i(t) = 1, X_j(t) = 1) = \mathbb{P}(X_i(t) = 1)\mathbb{P}(X_j(t) = 1) = p_i(t)p_j(t).$$

Using this approximation, we write

$$\dot{p}_i(t) = -m_i p_i(t) + (1 - p_i(t)) \sum_{(j,i)\in E} \beta_{ji} p_j(t). \tag{7.4}$$

This Mean-Field approximation has become extremely popular due to the tractability of the resulting dynamical system, but very few results have been discovered regarding the quality of the approximation. For more information on the derivation, the validity, and the exactness of this approximation the reader may refer to [41] and [64] and references therein.

At this point let us revisit Equation (7.2), and assume that (i) all subpopulations have the same size, $N_i = N_s$, and (ii) all removed nodes are instead susceptible to reinfection, i.e. $N_s = S_i(t) + I_i(t)$. In other words, we slightly change the model to assume that cured individuals instead of developing immunity, become susceptible to reinfection. In that case, (7.2) takes the simpler form

$$\frac{d}{dt}\frac{I_i(t)}{N_s} = -m_i \frac{I_i(t)}{N_s} + \left(1 - \frac{I_i(t)}{N_s}\right) \sum_{j=1}^{M} \beta_{ji} \frac{I_j(t)}{N_j}. \tag{7.5}$$

Note that Equations (7.4) and (7.5) are exactly the same if we set $p_i(t) = I_i(t)/N_s$. In other words, the compartmental networked model coincides with the evolution of the probability of infection in the stochastic SIS networked model (after making the mean field assumption). This observation has a fairly natural interpretation: instead

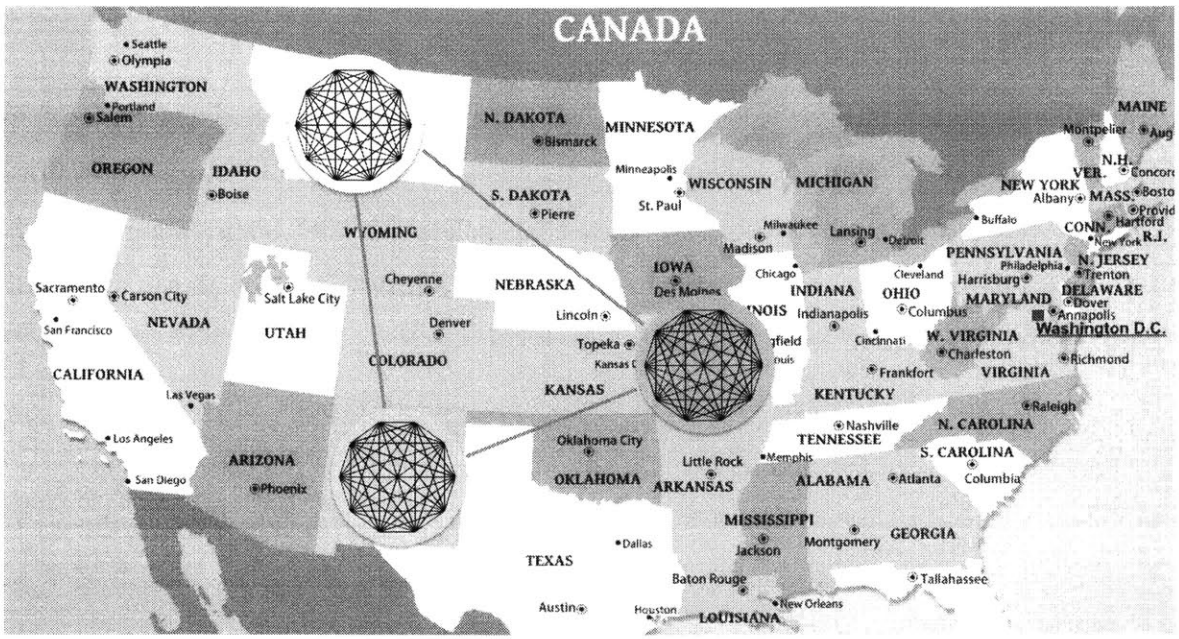


Figure 7-1: The metapopulation model for the United States

of considering a node as an individual in the stochastic SIS model, we may interpret it as a well mixed subpopulation with connections to other nodes/subpopulations. Under this new interpretation, the probability that a specific node is infected at time t coincides with the fraction of infected individuals in that subpopulation.

For the rest of this chapter we will be working with a variant of the meta-population model as described in Equation (7.2), and use it to understand and study the evolution of influenza type infections in the United States as shown in Figure 7-1.

7.3 Data

In this section we present the data that we use to estimate and test a meta-population model. Our data consists of weekly samples of influenza related infections, percentage of vaccinated population, absolute humididy for all states over 8 seasons as well as a static estimate of inter-state travel intensity for each pair of states in the country.

Influenza

Influenza, commonly known as "the flu", is an infectious disease caused by an influenza virus [68]. Symptoms can be mild to severe. The most common symptoms include high fever, runny nose, sore throat, muscle pains, headache, coughing, and fatigue. These symptoms typically begin two days after exposure to the virus and usually last less than a week.

Three types of influenza viruses affect people, called Type A, Type B, and Type C. Usually, the virus is spread through the air from coughs or sneezes [68] and hence is believed to occur mostly over relatively short distances. It can also be spread by touching surfaces contaminated by the virus and then touching the mouth or eyes. A person may be infectious to others both before and during the time they are sick. The infection may be confirmed by testing the throat, sputum, or nose for the virus. A number of rapid tests are available; however, people may still have the infection if the results are negative.

Yearly vaccinations against influenza are recommended by the World Health Organization for those at high risk. The vaccine is usually effective against three or four types of influenza viruses and is usually well tolerated. A vaccine made for one year may not be useful in the following year, since the virus evolves rapidly.

As Figure 7-2 indicates, influenza spreads around the world in yearly outbreaks, resulting in about three to five million cases of severe illness and about 250,000 to 500,000 deaths per year. In the Northern and Southern parts of the world, outbreaks occur mainly in the winter, while in areas around the equator, outbreaks may occur at any time of the year.

Subpopulations

In the model that we study, the subpopulations of the meta-population model correspond to each of the 50 states of the United States. We assume that each of these states is composed of a well mixed population of individuals. We also assume that the population of each state i remains constant over time as given by [51]. We denote

the size of each subpopulation by N_i.

Infection data

Traditional surveillance systems to monitor influenza spread, including those employed by the U.S. Centers for Disease Control and Prevention (CDC) and the European Influenza Surveillance Scheme (EISS), rely on both virologic and clinical data, including influenza-like illness (ILI) physician visits. CDC publishes national and regional data from these surveillance systems on a weekly basis, typically with a 1-2 week reporting lag.

The data provided from such surveillance systems, although fairly accurate, are not detailed enough to shed light on the structure of models of the form (7.2) for many reasons:

- **Geographic granularity:** Data published from CDC consist of weekly estimates of influenza related hospitalizations per HHS region or Census division. Note that the effective infection rate depends both on the population density (which governs human contacts) and the environmental conditions (which governs probability of transmission as we see in the sections to come). Hence, HHS regions and Census divisions being large areas, have fairly varying environmental conditions and population densities, which make the assumption of fully mixed homogeneous subpopulations very unrealistic.

- **Data granularity:** Data published from traditional surveillance systems take the form of an influenza activity level (scale 1-5) which is not detailed enough to learn a model consisting of a system of ordinary differential equations.

In an attempt to provide faster detection and more detailed reporting, innovative surveillance systems have been created to monitor indirect signals of influenza activity, such as call volume to telephone triage advice lines and over-the-counter drug sales. About 90 million American adults are believed to search online for information about specific diseases or medical problems each year, making web search queries a uniquely valuable source of information about health trends. Previous attempts at using online

activity for influenza surveillance have counted search queries submitted to a Swedish medical website, visitors to certain pages on a U.S. health website, and user clicks on a search keyword advertisement in Canada. A set of Yahoo search queries containing the words "flu" or "influenza" were found to correlate with virologic and mortality surveillance data over multiple years.

Google created a system [25] which builds on these earlier works by utilizing an automated method of discovering influenza-related search queries. By processing hundreds of billions of individual searches from five years of Google web search logs, this system generates comprehensive models for use in influenza surveillance, with regional and state-level estimates of influenza-like illness (ILI) activity in the United States.

In this work, we used the dataset provided by the Google Flu Trends tool, which consists of **weekly** estimates of influenza related infections for all 50 states, over 8 seasons, namely 2008-2015. Note that we *exclude* from the dataset the season 2009-2010, due to the H1N1 influenza virus outbreak which featured non-typical behavior. Moreover, we *exclude* from the dataset the first 10 weeks of each season, as the transition from the previous season to the next one may cause problems both in training and in prediction. A typical snapshot of this data set for a specific state during a specific season is shown in Figure 7-2.

Curing and vaccination

According to the National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention [43], the typical duration of influenza related infection sickness is 5-7 days, and hence since the data set consists of weekly samples, each data point corresponds to new infections. In other words, the relevant difference equation equivalent of our baseline model (7.2) can be written as

$$I_i(t+1) = \sum_{j=1}^{M} \beta_{ji} S_i(t) \frac{I_j(t)}{N_j}$$

where t denotes the week number of the season and ranges from 1 to 42.

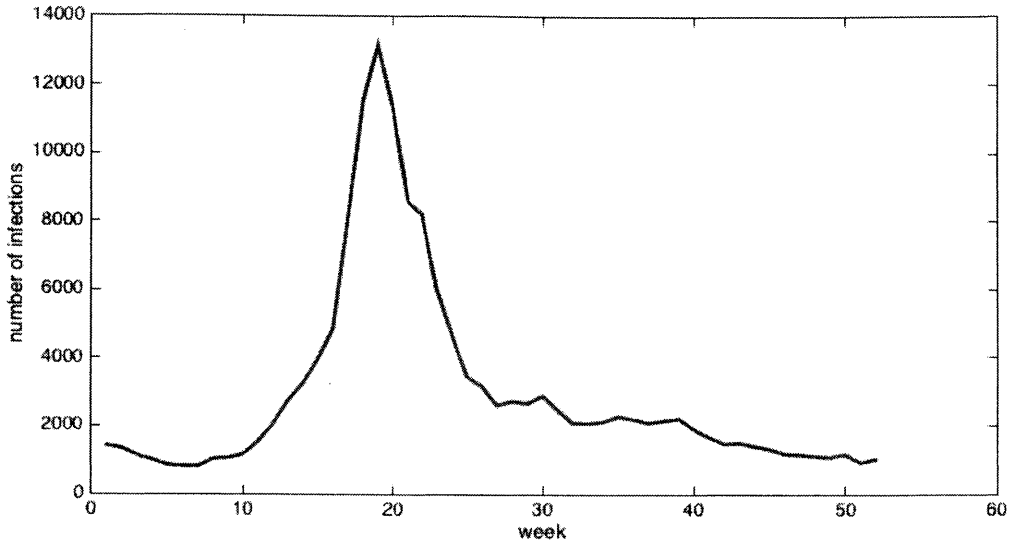


Figure 7-2: Number of influenza related infections in Arizona for the 2008-2009 season. The single peaked shape of the time series is typical in the dataset.

Reinfections from the virus are rare but possible [59] and hence the number of immune (after infection) individuals in week t is proportional to

$$m \sum_{t'=1}^{t-1} I_i(t').$$

Furthermore, immunity may be developed after taking flu vaccines which cause antibodies to develop in the body about two weeks after vaccination. These antibodies provide protection against infection with the viruses that are in the vaccine. The seasonal flu vaccine protects against the influenza viruses that research indicates will be most common during the upcoming season. The evolution of the vaccinated population is crucial in understanding the evolution of the susceptible population in each state. The Center for Disease Control and Prevention (CDC) provides weekly estimates of the vaccinated population for all seasons 2009-2015 through the FluVaxView tool. We denote by $V_i(t)$ the proportion of non vaccinated population in state i on week t and we write

$$S_i(t) = N_i V_i(t) - m \sum_{t'=1}^{t-1} I_i(t').$$

Effect of environmental conditions

Influenza A incidence peaks during winter in temperate regions. The basis for this pronounced seasonality is not understood, nor is it well documented how influenza A transmission principally occurs. Previous studies indicate that several environmental factors may affect seasonality and have analyzed data from laboratory experiments to explore the effects of different parameters on influenza virus transmission and influenza virus survival.

Recent studies (see [58] and references therein) find that absolute humidity constrains both transmission efficiency and, most significantly, influenza virus survival. In the studies presented, 50% of influenza virus transmission variability and 90% of influenza virus survival variability are explained by absolute humidity. In temperate regions, both outdoor and indoor absolute humidity possess a strong seasonal cycle that is smallest during the winter. This seasonal cycle is consistent with a winter-time increase in influenza virus transmission and influenza virus survival and can be used to explain the seasonality of influenza. Thus, differences in absolute humidity provide a single, coherent, physically sound explanation for the observed variability of influenza seasonality.

These observations necessitate the introduction of absolute humidity as a variable in our model. Both influenza virus transmission and influenza virus survival affect the effective contact rate, or equivalently the rate at which infected individuals infect their susceptible contacts. Hence, we modify our infection model to

$$I_i(t+1) = f(H_i(t)) \sum_{j=1}^{M} \beta_{ji} S_i(t) \frac{I_j(t)}{N_j}, \tag{7.6}$$

where $H_i(t)$ corresponds to weekly data on the average absolute humidity at state i during week t as obtained from the National Center for Environmental Information and f is a parametric function to be estimated from the data which quantifies the dependence of the effective contact rate on the absolute humidity. Note that the dataset consists of daily measurements from several locations within the state. We

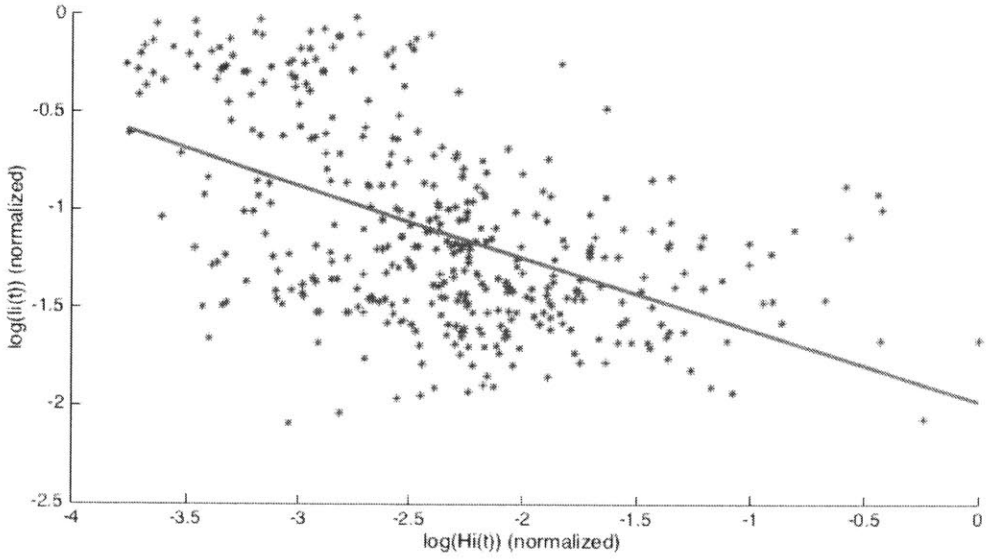


Figure 7-3: Relation between number of infections and absolute humidity of the preceding week in Arizona, for the 2008-2009 season. We plotted the logarithm of each quantity for illustration purposes. The slope for this case is equal to -0.37.

take the average, for every state and every week among all locations and among all days within the week.

Figure 7-3 shows the strong correlation between number of infections and humidity conditions in the preceding week in the state of Arizona, for the 2008-2009 season.

Inter-state mobility

In our model, we assume a well mixed subpopulation at each state/node, as well as a network of states, corresponding to individuals traveling between states. In other words, in our model (7.4) the weights β_{ji} correspond to the density of individuals traveling from state j to state i.

One approach to determining these parameters β_{ij} would be to "learn" their value from the available data. On the other hand, note that the number of such parameters is in the worst case 2500, since we need two such parameters for each pair of states (to and from intensity). This approach, even imposing sparsity constraints or penalization, would lead to a greatly overdetermined problem of questionable value.

Instead, in our approach we used travel data to estimate the relative travel in-

tensity for each pair of states. Specifically, the National Household Travel Survey (NHTS) is the flagship survey of the U.S. Department of Transportation (DOT) and is conducted periodically to assess the mobility of the American public. The survey gathers trip-related data such as mode of transportation, duration, distance, and purpose, and then links the travel related information to demographic, geographic, and economic data for analysis. Policy makers, individual state DOTs, metropolitan planning organizations, industry professionals, and academic researchers use the data to gauge the extent and patterns of travel, to plan new investments, and for innumerable applications of data on trends in travel for policy and planning applications. The NHTS surveys that we used for our analysis occurred in 1990, 1995, 2001, 2009 and 2014. Each survey contains about 25,000 households representing all 50 U.S. States and the District of Columbia. During the survey period, each household was sent a travel diary and asked to report all travel by household members. For the purposes of our analysis, we identified for each (directed) pair of states the number of individuals traveling from state to state, after normalizing by the population size. Specifically, our calculation for the estimate of β_{ij} is equal to

$$\tilde{\beta}_{ij} = \frac{N_{ij}}{D_i} N_i,$$

where N_{ij} is the total number of individuals in the sample traveling from state i to state j, D_i is the sample size for state i and N_i is the population of state i. Note that $\hat{\beta}_{ij}$ provides a relative measure of travel intensity (or contact rate) between different state pairs. On the other hand, we do not know the absolute magnitude (and effect) of travel intensities as it should appear in the model (and compared to the intra-state contacts) and hence, the effective travel intensity is $\hat{\beta}_{ij} = z\tilde{\beta}_{ij}$ where z is to be estimated using the data. Using this procedure we generate the *network*, as shown in Figure 7-4.

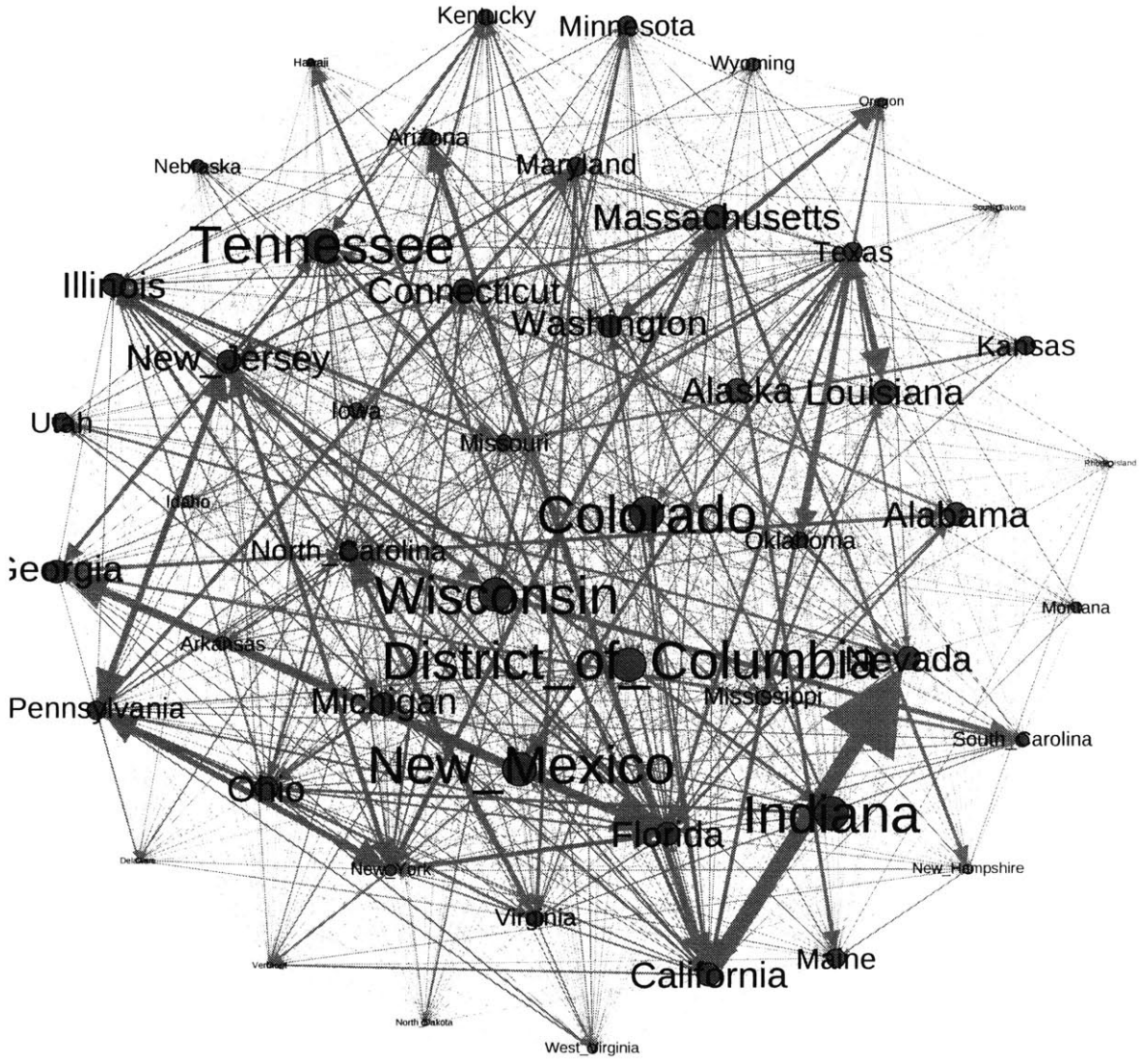


Figure 7-4: Travel intensities (normalized) as calculated using data from the National Household Travel Survey (NHTS).

7.4 Model, Assumptions and Problem

Using the data from various sources as explained above to obtain several parameters in our model, we write

$$I_i(t+1) = f(H_i(t)) \sum_{j=1}^{M} \hat{\beta}_{ji} \left(N_i V_i(t) - m \sum_{t'=1}^{t-1} I_i(t') \right) \frac{I_j(t)}{N_j},$$

where $H_i(t), V_i(t)$ is given by the data. Moreover, we do have estimates for $\hat{\beta}_{ji}$ for all pairs $j \neq i$ (except for their absolute magnitude z) but we do not have estimates for the mixing rates β_{ii}, which correspond to the intra-state human contact rate. Hence, our model can be rewritten as

$$I_i(t+1) = f(H_i(t)) \left(N_i V_i(t) - m \sum_{t'=1}^{t-1} I_i(t') \right) \left(\beta_{ii} I_i(t) + z \sum_{j=1, j\neq i}^{M} \tilde{\beta}_{ji} \frac{I_j(t)}{N_j} \right), \quad (7.7)$$

where z is an unknown parameter which summarizes the effect that infected travelers from other states have on the infections of each state. Equation (7.7) describes the evolution of the number of infected individuals over a single season.

Each season, different strains of flu related viruses cause infections. Each virus (and hence each season) may have different characteristics which change the effective contact rate (dependence on absolute humidity). We model this phenomenon in our analysis by making the function $f(\cdot)$ dependent on the season $s \in \{1, \ldots, 7\}$ as follows:

$$I_i^s(t+1) = f_s(H_i^s(t)) \left(N_i V_i^s(t) - m \sum_{t'=1}^{t-1} I_i^s(t') \right) \left(\beta_{ii} I_i^s(t) + z Q_i^s(t) \right), \quad (7.8)$$

where

$$Q_i^s(t) = \sum_{j=1, j\neq i}^{M} \tilde{\beta}_{ji} \frac{I_j^s(t)}{N_j},$$

denotes the total number of infected individuals traveling into state i during season s.

Several assumptions have been made in developing this model that we summarize here:

(i) The population of each state remains constant over time.

(ii) All infected individuals recover within a week.

(iii) The effect of humidity (i.e., the function f_s) depends only on the season (the strain of the virus) and not on the state. This assumption is crucial for the identifiability of the network effect z.

(iv) The travel intensities are constant over time. In our work, due to lack of more refined travel data we assumed that the travel intensities are constant over time. This assumption is extremely simplistic and the insights could change depending on the quality of the travel data.

Our goal in the rest of this chapter is to estimate this model using the available data, i.e., determine the lower-case parameters m, β_{ii} and z, as well as the functions f_s and evaluate the accuracy of the model by making predictions using these estimates.

7.5 Estimation

In this section we use the data to learn the unknown dependence on absolute humidity and estimate the parameters of the model (7.9). The first step is to decouple the problem of learning the network effect z from learning the function f (in Section 7.5.2 we specify the parametric form of f).

7.5.1 Identifiability of network effect z

Our model (7.8) is linear with respect to the network parameter z. Therefore, assuming that the rest of the unknown parameters and functions are perfectly estimated, we could potentially use linear regression to estimate z. An important question, though, is whether the parameter z can be identified.

For example, consider the case of Colorado in the season 2011-2012, as depicted in Figure 7-5. The time series of infections within the state (I_i^s) is collinear with the time series of traveling infections (Q_i^s) and hence any effect the latter may have on the former cannot be identified by our model.

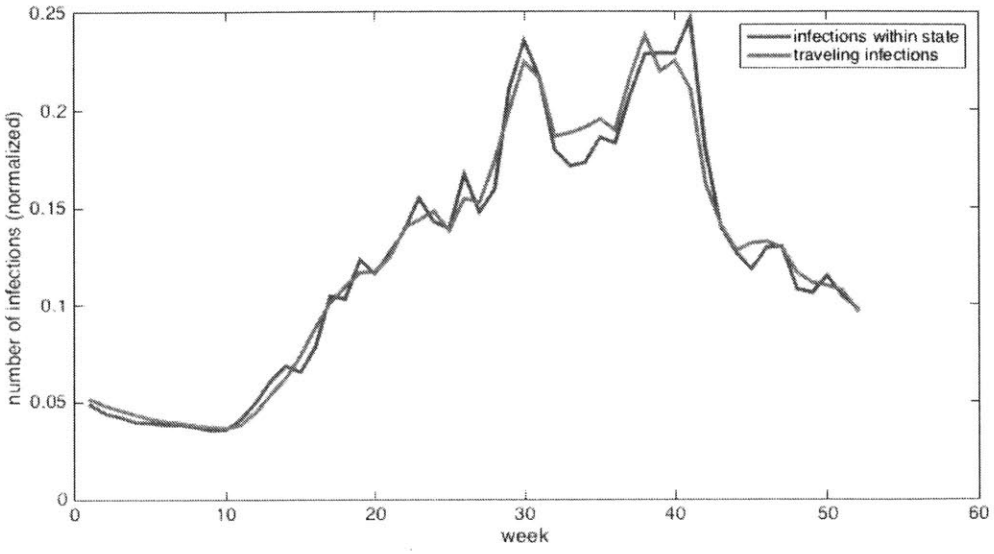


Figure 7-5: Normalized infections within the state of Colorado as well as traveling infections into the state for the season 2011-2012. Identifying network effect z is impossible due to collinearity.

In contrast, consider the case of Kentucky for the same season 2011-2012, as depicted in Figure 7-6. The time series of infections within the state are far from collinear with the time series of traveling infections and hence the effect of the latter on the former may potentially be identified by our model.

In other words, as standard results from linear regression suggest, the parameter z is identifiable only if the covariate vectors $I_i^s(t)$ and $Q_i^s(t)$ are linearly independent. Specifically, let us denote by X the data matrix for each state and each season, i.e.

$$X_i^s = [I_i^s \, Q_i^s],$$

where $I_i^s = \{I_i^s(t)\}_{t=0}^{42}$ and $Q_i^s = \{Q_i^s(t)\}_{t=0}^{42}$ for each specific season. Then the standard error for each estimator of the z variable is proportional to $[(X_i^{s\intercal} X_i^s)^{-1}]_{22}$. Therefore, in order to better understand the orthogonality (or lack thereof) of the in-state infections to the traveling infections, and perform our inference task on states and seasons where identification is possible, we generate for every season and every state,

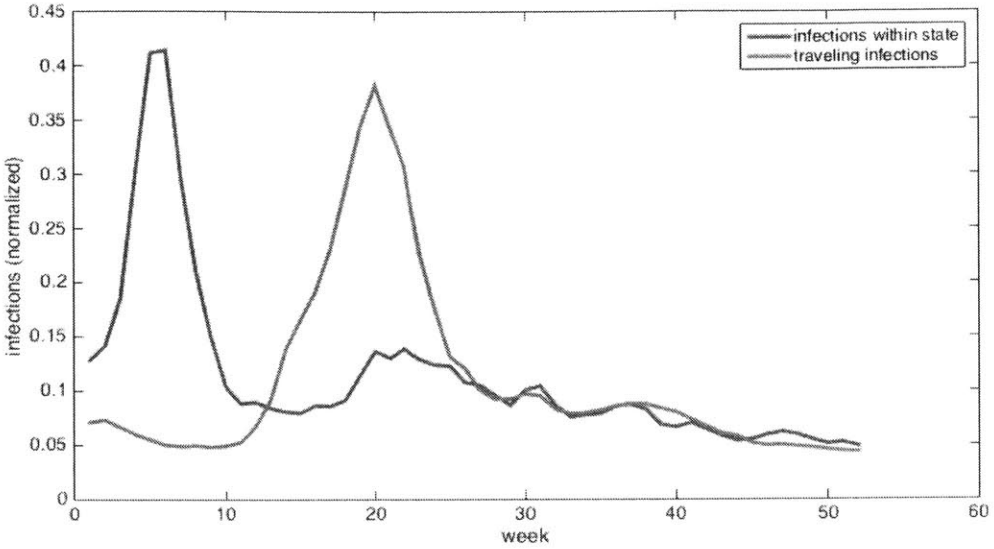


Figure 7-6: Normalized infections within the state of Kentucky as well as traveling infections into the state for the season 2011-2012. Identifying network effect z possible due to orthogonality.

the corresponding entry of the variance matrix:

$$O_i^s = [(X_i^{s\intercal} X_i^s)^{-1}]_{22},$$

for each state $i \in \{1, 50\}$ and each season $s \in \{1, \ldots, 7\}$. Figure 7-7 shows an illustration of these entries. As we can see in the figure, in several state-season pairs the two vectors are orthogonal enough to lead to a small standard error in the coefficient estimation but in many state-season pairs, the two vectors are collinear to the extent that identification is impossible.

In our analysis, we split our dataset into two groups:

(i) **Collinear state-season pairs:** whenever the quantity O_i^s is larger than 5, the two vectors are essentially collinear (and hence the standard error would be inherently high to lead to precise inference). Therefore, in all such season-state pairs, the model can be simplified to

$$I_i^s(t+1) = f_s(H_i(t)) \left(N_i V_i^s(t) - m \sum_{t'=1}^{t-1} I_i^s(t') \right) (\beta_{ii} I_i^s(t)) \,. \tag{7.9}$$

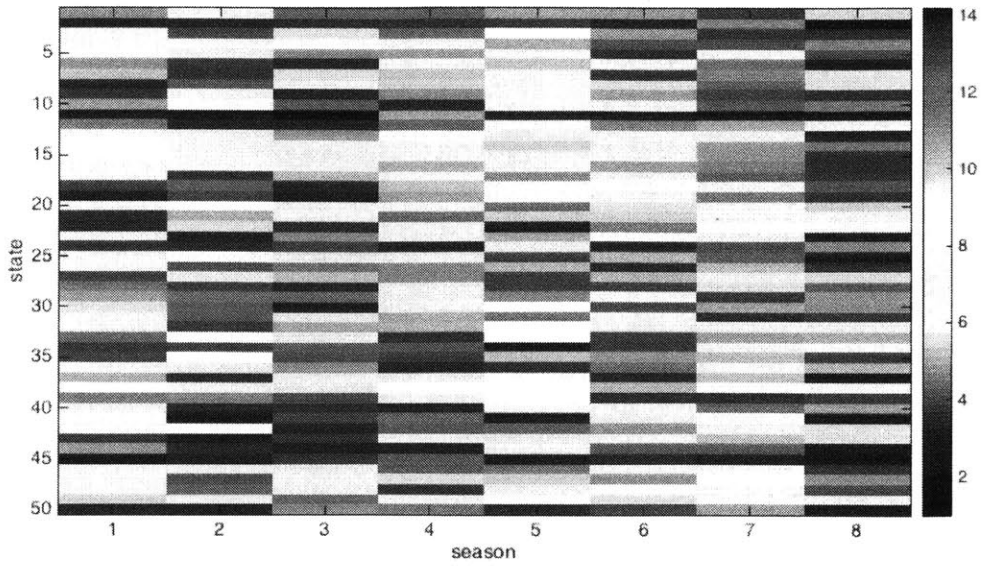


Figure 7-7: Orthogonality between the vectors of in-state infections and traveling infections, for each state $i \in \{1, 50\}$ and each season $s \in \{1, \ldots, 8\}$.

In such state-season pairs (depicted in blue in Figure 7-8, as we will see in the coming sections, we use a neural network approach to estimate $f(\cdot)$ and parameters m, γ_s and β_{ii}.

(ii) **Orthogonal state-season pairs:** We use the estimates obtained from the collinear state-season pairs and perform a linear regression task on the remaining state-season pairs (depicted in yellow in Figure 7-8) in order to identify the network effect z.

7.5.2 Learning dependence on humidity and mixing rates

In this section we explain our approach to estimating the model (7.9) using the available data. As explained in Section 7.5.1, we use a subset of the state-season pairs to estimate model (7.9). We call these season-state pairs the *collinear pairs* and are those which are colored in blue in Figure 7-8. As in most estimation tasks, we use half of these pairs for estimating the model and the rest for testing the performance of the model. We refer to the former as the *training set*, $\mathcal{T}$ and the latter as the

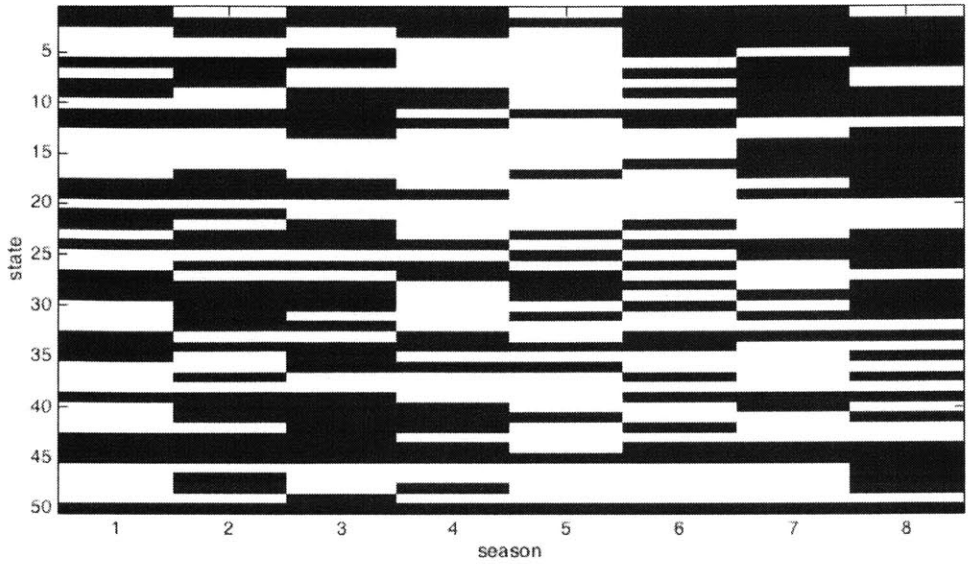


Figure 7-8: In blue: season-state pairs where identification of network effect is impossible. In yellow: season-state pairs where identification of network effect is possible.

validation set, $\mathcal{V}$. We chose the training appropriately so that all states and all seasons are represented at least one in the training set $\mathcal{T}$.

A naive approach to estimation: One Step prediction

For any possible parametric form of $f(\cdot)$, we can use non-linear regression to learn the non linear model (7.9). Specifically, we consider estimates of the form

$$\hat{I}_i^s(t+1) = f(H_i(t)) \left(N_i V_i^s(t) - m \sum_{t'=1}^{t-1} I_i^s(t') \right) \beta_{ii} I_i^s(t), \tag{7.10}$$

where I_i denotes the vector of infections for state i during season s (data), and find the parameters that minimize the sum of squared errors, i.e.,

$$\min_{f_s, m, \{\beta_{ii}\}} \sum_{(i,s)\in\mathcal{T}} \|I_i^s - \hat{I}_i^s\|^2.$$

In the above problem, when we minimize with respect to f_s we minimize with respect to the parameters which appear in a certain parametric form of f_s.

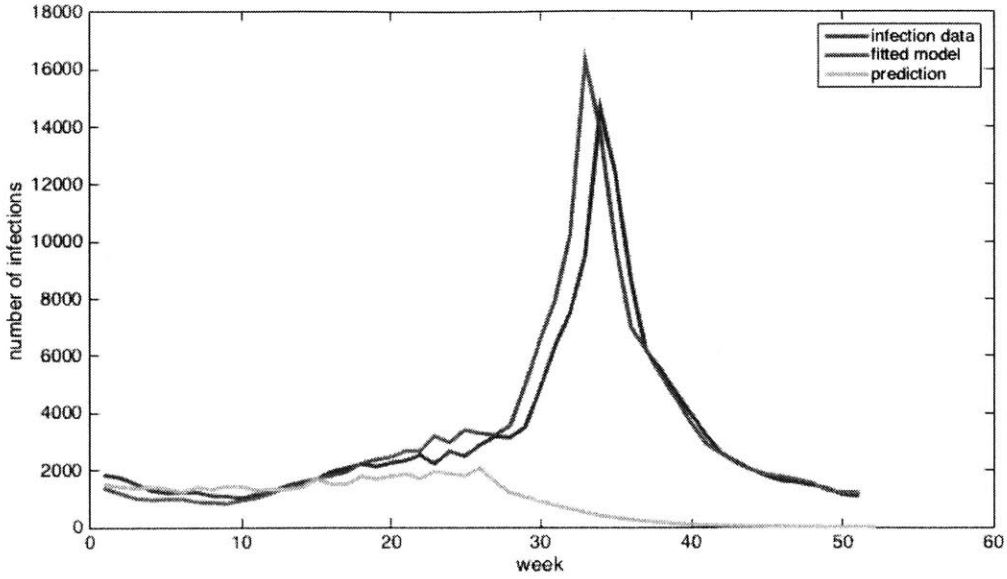
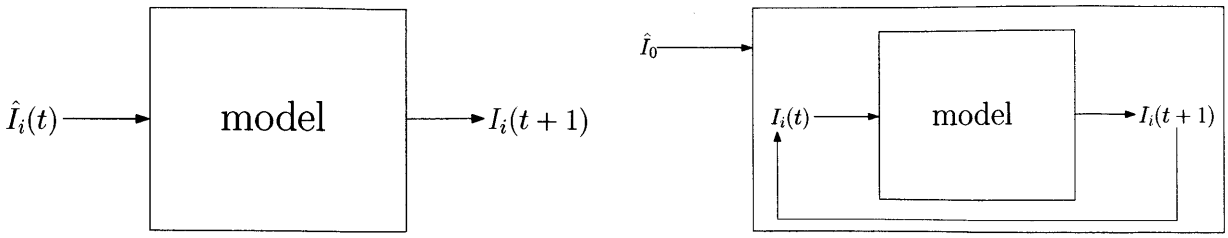


Figure 7-9: Non linear regression: The fitted model is accurate. Prediction using the estimated model is not.

For the purposes of our analysis, we implemented an extensive set of such estimation schemes, trying various parametric forms of $f(\cdot)$ and using the Non Linear Regression package of MATLAB®.

The achieved fit using this approach is excellent with R^2 ranging from 0.908 to 0.974 for different state-season pairs in $\mathcal{T}$. Indeed, as Figure 7-9 shows, 7.10, the fitted model is very close to the data. On the other hand, the *long-term predictive* performance of the estimated model is bad. Specifically, if we use the estimated model and simulate (7.9), the prediction is extremely inaccurate. The mean squared error, ranges from 0.91 to 0.98 for different state-season pairs even within the training set $\mathcal{T}$, as shown in Figure 7-9.

This problem is not uncommon in identification. The non-linear regression task, as outlined above, optimizes with respect to the parameters, in order to achieve the best *one-step* prediction: using the data, $I_i(t)$, match as closely as possible $I_i(t+1)$. On the other hand, since our data consists of weekly samples, the "best" estimation for $I_i(t+1)$ would be very close to $I_i(t)$ and hence, the non linear regression roughly estimates this simple relationship between input $(I_i(t))$ and output $(\hat{I}_i(t+1))$. Hence,



(a) One step prediction: Non linear regression to estimate the model using the relationship between $\hat{I}_i(t)$ (data) and $I_i(t+1)$ (one step estimation).

(b) Long term prediction: Using the initial number of infections ($\hat{I}_i(0)$) we simulate the model for the whole season and minimize the error for the whole season.

Figure 7-10: One step vs Long term prediction

the resulting estimation task is seemingly successful but the long term prediction is not. Therefore, in order to learn the model (7.9) we need to optimize with respect to its long-term prediction accuracy.

Estimation for long term prediction

As discussed above, the simple non-linear regression approach fails to uncover the details of the model (7.10) as the estimated model yields poor predictions. In order to tackle this issue, we explicitly define our estimation objective to incorporate the need for predictive accuracy.

Specifically, for each state-season pair in $\mathcal{T}$, and for each set of parameters, one can simulate the model (7.9) and obtain a vector $\hat{I}_i^s = \{\hat{I}_i^s(t)\}_{t=2}^{42}$, starting from $\hat{I}_i^s(0) = I_i^s(0)$ as given from the data. Our estimation task is to find a set of parameters that minimizes

$$E = \frac{1}{2} \sum_{(i,s)\in\mathcal{T}} \|I_i^s - \hat{I}_i^s\|^2,$$

with the difference that $\hat{I}_i^s$ corresponds to the simulated timeseries as given by model (7.9). Schematically, the difference between the two approaches can be seen in Figure 7-10.

This approach, although fairly natural for our estimation task, involves a non-trivial, non-convex optimization problem. Essentially, the mapping from $\hat{I}_i^s(0)$ to the simulated time series $\hat{I}_i^s$ is extremely complex and highly non-linear. Therefore,

achieving global optimality guarantees for such a problem is hard if not impossible. Instead, as is commonly done in most machine-learning approaches, we obtain a locally optimal solution to our estimation problem using gradient descent.

Note though that the objective, i.e., the squared error between the simulated time series and the data, is fairly complex and hence, calculation of the gradient is a non-trivial task. Specifically,

$$E(f_s, m, \beta_{11}, \ldots, \beta_{ii}) = \frac{1}{2}\sum_{t=2}^{42} E(t)^2,$$

with

$$\begin{aligned} E(t) &= \sum_{(i,s)\in\mathcal{T}} E_{i,s}(t)^2, \\ E_{i,s} &= I_i^s(t) - \hat{I}_i^s(t), \\ I_i^s(0) &= \hat{I}_i^s(0), \end{aligned} \tag{7.11}$$

and

$$\hat{I}_i^s(t+1) = f_s(H_i^s(t))\left(N_i V_i^s(t) - m\sum_{t'=1}^{t-1}\hat{I}_i^s(t')\right)\beta_{ii}\hat{I}_i^s(t). \tag{7.12}$$

In order to perform the update in standard gradient descent, one needs to calculate the gradient of $E(\cdot)$. The problem of efficiently calculating the gradient, in such settings has been encountered in recurrent neural networks and one of the first proposed solutions has been *back propagation through time* [55]. This approach has not been popular due to the need for large computational requirements. Instead, machine learning researchers and practitioners typically use an approach known as *Real-Time Recurrent Learning* [66] that we explain below.

For the purposes of this exposition we assume that the function f_s is determined by only one parameter a_s. Our goal is to estimate the derivatives

$$\Delta w = \frac{\partial E}{\partial w},$$

where $w \in \{a_s, \beta_{ii}, m\}$. We write

$$\Delta w(t) = \frac{\partial E(t)}{\partial w},$$

so that

$$\Delta w = \sum_{t=2}^{42} \Delta w(t). \tag{7.13}$$

Note that

$$\Delta w(t) = \sum_{(i,s)\in\mathcal{T}} E_{i,s}(t) \frac{\partial E_{i,s}(t)}{\partial w} = \sum_{(i,s)\in\mathcal{T}} E_{i,s}(t) \frac{\partial I_i^s(t)}{\partial w}, \tag{7.14}$$

where the last equality follows from (7.11). Therefore, in order to calculate the increment for w, we are looking for an efficient way to calculate the terms $\partial I_i^s(t)/\partial w$ within the sum. To this end, note that differentiating (7.12) with respect to $w \in \{a_s, \beta_{ii}, m\}$ yields a recursion providing $\partial I_i^s(t)/\partial w$ as a function of $\partial I_i^s(0)/\partial w, \ldots, \partial I_i^s(t-1)/\partial w$, that can be computed using the fact that $\partial I_i^s(0)/\partial w_j = 0$. The updates for model (7.9) can be found in Appendix A.

Below, we present the high-level structure of the estimation procedure.

Algorithm 1 Estimation Procedure

1: parameters $\leftarrow$ initialization
2: **while** no convergence **do**
3: generate $\hat{I}_i^s$ for each i, s using (7.9)
4: evaluate $E_{i,s}(t)$ using $\hat{I}_i^s$
5: evaluate $\partial I_i^s(t)/\partial w$, for each $w \in \{a_s, \beta_{ii}, m\}$, using (A.1)-(A.3).
6: Calculate derivatives $\partial E/\partial w$ using (7.13)-(7.14).
7: Perform gradient step using Armijo step size and ADAGRAD.
8: **end while**

Implementation of Real Time Recurrent Learning

Equations (7.14)-(A.3) provide an efficient way of calculating the gradient of the error and in principle provides the necessary ingredients for performing gradient descent in order to calculate the optimal (with respect to mean squared error) parameters. On the other hand, the system that is to be estimated is complex and the objective non-convex and several issues should be taken into account for the implementation.

Teacher forcing: The calculation of $\partial E/\partial w$ for each of the parameters $w \in \{a_s, \beta_{ii}, m\}$ involves calculations using $\hat{I}_i^s(t)$, i.e., the simulated time-series. During the first several iterations of the gradient descent, the simulated time series is far from the data of the training set and hence introduces errors in these calculations which significantly reduce the time to convergence or may lead to instability. Instead, common practice has found that performance is improved when instead of $\hat{I}_i^s(t)$ we use $I_i^s(t)$, i.e., $\partial I_i^s(t)/\partial w$ (line 5 of Algorithm 1) is calculated along the actual trajectory (data) while $E_{i,s}(t)$ is evaluated along the simulated trajectory.

This practice is called *teacher forcing.* In our implementation we tried both approaches but teacher forcing seems to perform better both in terms of convergence and in terms of fit.

System instability: The model under consideration (7.9) is a complex dynamical system and, as such, features complex behavior. Specifically, for a large set of parameter choices, the system becomes unstable and both the calculation of the simulated time series as well as the calculation of the gradient becomes impossible, due to numerical accuracy issues. In the neural network literature, this problem is fairly common. The basic problem is that calculations of large gradients propagated over many stages tend to explode (with much damage to the optimization).

Several approaches ([29] and references therein) have been proposed in the literature to resolve or ease the problem, most of them being problem-specific or extremely convoluted. In our approach, we adopted a simple heuristic, which seems to overcome the instability issue at the expense (perhaps) of the quality of fit. Specifically, in our implementation of the recurrent neural network, whenever the simulated time series exceeds the maximum value in the training set, we replace the former value with the latter. Similarly, whenever the simulated time series is below zero, we replace that value with zero, i.e., we use

$$\hat{I}_i^s(t+1) = \min\left\{0, \max\left\{\max_{t'} I_i^s(t'), \gamma_s f(H_i(t))\left(N_i V_i(t) - m\sum_{t'=1}^{t-1}\hat{I}_i^s(t')\right)\beta_{ii}\hat{I}_i^s(t)\right\}\right\},$$

instead of (7.12) in our implementation.

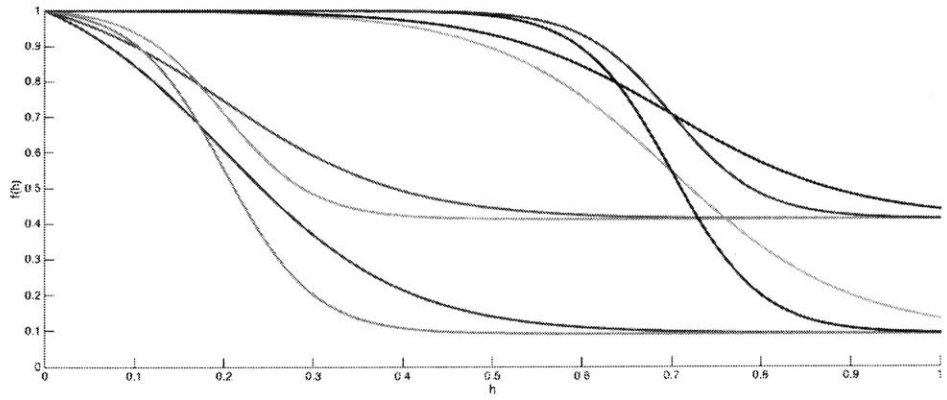


Figure 7-11: Different shapes of logistic functions

Adaptive gradient descent: In our implementation we adaptively chose the stepsize of the gradient descent. Specifically, we use the adaptive gradient algorithm ADAGRAD as described in [18], and chose the increment using backtracking line search (Armijo rule). As is commonly done in most non-convex programs, we allow for the error to increase during the backtracking line search. Specifically, we allow the error to increase by 1% at each step.

Humidity Function: In our analysis we tried several parametric forms for the function $f(\cdot)$, and eventually we chose the logistic function

$$f(h; k, l, m) = \frac{u_s}{1 + e^{-k_s(x-l_s)}} + m_s,$$

which has the flexibility to describe the potentially extreme dependence of the contact rate on the humidity levels. Figure 7-11 shows the shape of f for different values of the parameters, illustrating the modeling flexibility of this parametric form.

Results

In order to provide an easier interpretation, we present our results in terms of the mean absolute normalized error

$$R_{\mathcal{S}} = \frac{\sum_{(i,s)\in\mathcal{S}} \sum_{t=1}^{42} \frac{|I_i^s(t) - \hat{I}_i^s(t)|}{\hat{I}_i^s(t)|}}{42 \cdot |\mathcal{S}|},$$

which measures the absolute error per data point. When $\mathcal{S} = \mathcal{T}$ the corresponding error is the error of the training data, while when $\mathcal{S} = \mathcal{V}$, the error on the validation set ("out of sample" error).

Error on training data Our implementation yields error on the training data equal to $R_{\mathcal{T}} = 0.391$. Figure 7-12 shows an example of the fit for a state-season pair in the *training set.* As expected, the model, being a system of ordinary differential equations is not capable of reproducing the details of the time series. Specifically, the trained model cannot capture the "peak" behavior of weeks 19-25 but is fairly accurate in learning the smoother portions of the time series.

Furthermore, Figure 7-12 provides insight on the effect of the absolute humidity on the evolution of infections. Specifically, low absolute humidity seems to correlate with higher infection rates. Indeed, this effect is captured in the "learned" function $f(\cdot)$, as shown in Fig 7-13.

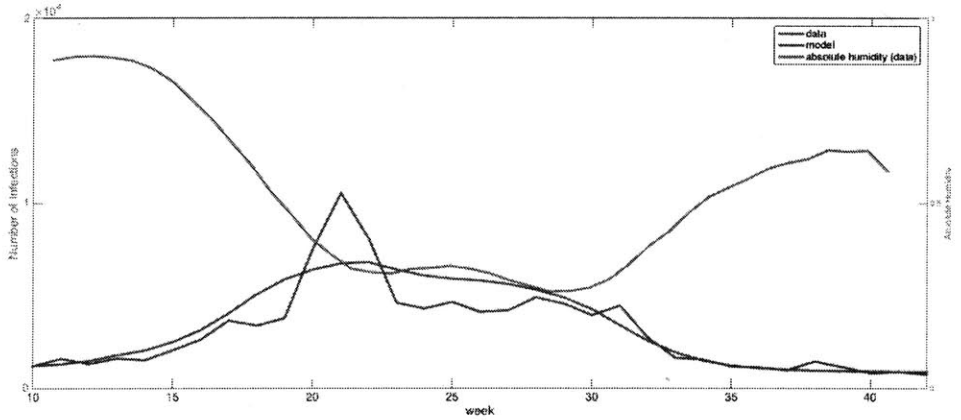


Figure 7-12: Example of fit: simulated vs. real infection time series for a state-season pair in the training set $\mathcal{T}$.

Out of sample error Our implementation yields error $R_{\mathcal{V}} = 0.851$. Figures 7-14 and 7-15 show the predicted time series for two different state-season pairs. In one case the prediction results are extremely good, due to the fact that real infections seem to negatively correlate with the absolute humidity and hence the effect of the latter on the effective contact rate accurately describes the evolution of infections. On the other hand, for other state-season pairs this does not seem to be the case and

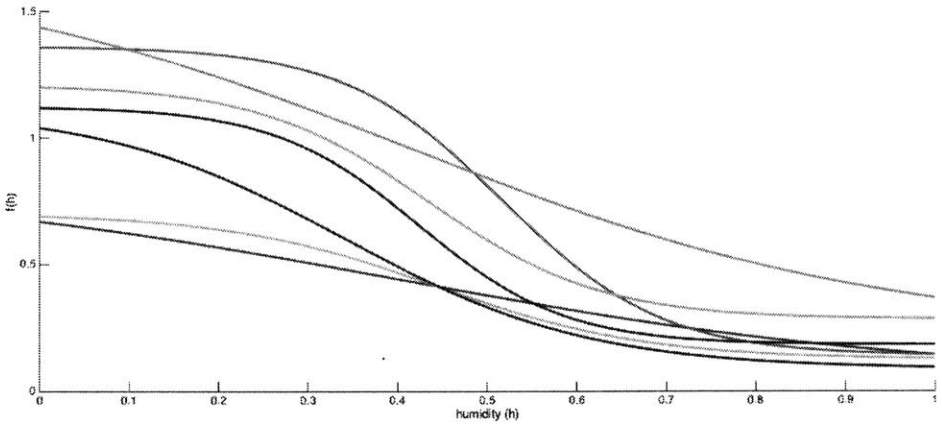


Figure 7-13: Dependence of effective contact rate on absolute humidity for different seasons

in such cases prediction is unsuccesful, leading to large prediction errors.

7.6 A preliminary approach to identifying network effects

Given the estimates of the unknown parameters, denoted as $\hat{\beta}_{ii}$, γ_s, for $s = 1, \ldots 7$ and $i = 1 : 50$, as well as $\hat{m}$ and the estimated function $\hat{f}$ we now proceed to understanding the effect of inter-state traveling. Our approach to this problem is fairly preliminary as we proceed by performing a linear regression on the model (7.9). Specifically we write

$$I_i^s(t+1) = \gamma_s \hat{f}(H_i(t)) \left(N_i V_i(t) - \hat{m} \sum_{t'=1}^{t-1} I_i^s(t') \right) \hat{\beta}_{ii} \left(\alpha \hat{I}_i^s(t) + z Q_i^s(t) \right) + \epsilon_i^s(t), \quad (7.15)$$

and use the Linear Regression package of MATLAB®, to estimate the parameters α and β. The results of the regression are summarized in Table 7.1 and lead to the conclusion that network effects are indeed negligible compared to the effect of intra-state contacts. Specifically, the p-value[2] corresponding to the network parameter z

[2]The p-value corresponds to the test between the null-hypothesis according to which $z = 0$ and the hypothesis that $z > 0$. Note that this p-value is calculated assuming normal and i.i.d. noise which is not necessarily the case in our model.

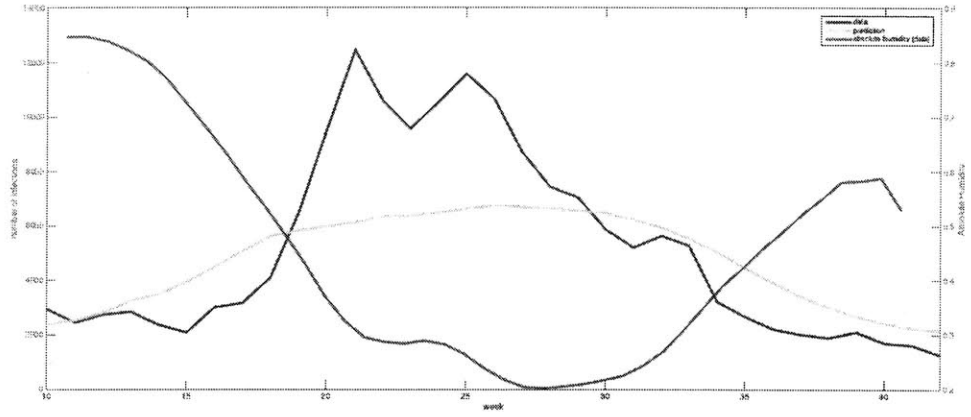


Figure 7-14: Example of prediction: predicted vs. real infection time series for a state-season pair in the validation set $\mathcal{T}$. Prediction error for this case is 0.503, which is the smallest prediction error achieved within the whole validation set.

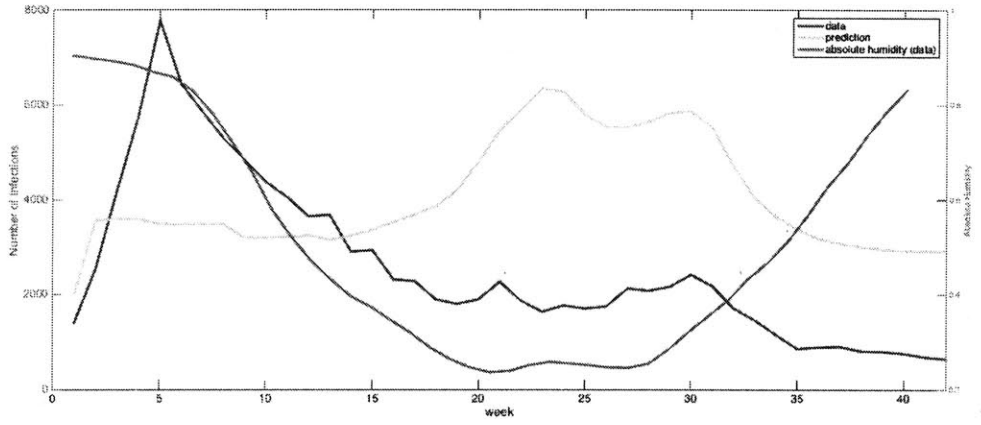


Figure 7-15: Example of prediction: predicted vs. real infection time series for a state-season pair in the validation set $\mathcal{T}$. Prediction error for this case is 1.471, which is one of the largest prediction errors achieved within the whole validation set. For this particular state-season pair, and for those with large prediction error, the absolute humidity does not seem to correlate negatively with the increase of the number of infections. Such pairs are an exception both in the training and the validation set.

	Estimate	SE	pValue
Intercept	0.0002133	$1.7369 \cdot 10^{-05}$	$1.548 \cdot 10^{-34}$
α	0.92924	0.0037379	0
z	0.026663	0.00392	$1.0627 \cdot 10^{-11}$

Table 7.1: Results of linear regression to identify network effects.

is very small (significantly smaller than 0.01 which is the standard benchmark) and hence the corresponding predictor is indeed significant but its value is significantly smaller compared to α, the effect of intra-state contact (note that we normalized the vectors I_i^s and Q_i^s before making the comparison).

This approach for identifying the network effect is overly simplistic as it is neglecting two important effects:

- **temporal correlation:** In this preliminary analysis we have neglected the temporal correlations between the predictors, i.e., the correlations between $I_i^s(t)$, and $I_i^s(t-2), I_i^s(t-3), \ldots$. In other words the noise term in (7.15) is assumed to be *i.i.d.* noise which is not necessarily the case for our purposes.

- **'reflection' problem:** This problem has a long history in econometrics ([10] and references therein) and in our context boils down to the following situation. Assuming that a network effect exists, and assuming that states i and j are neighbors, then $I_i(t-3)$ would affect $I_j(t-2)$ which in turn would affect $I_i(t-1)$. Hence, if this effect is not taken into account, network effects could be overestimated. On the other hand, in our analysis, network effects seem to be negligible and hence the reflection problem will not change our qualitative conclusion.

7.7 Summary and Conclusions

In this chapter, we focused our attention on the modeling and analysis of influenza in the United States. Our goal was twofold. First, to evaluate networked compartmental models based on their predictive power and second identify potential network effects

between different states through traveling. Our approach to tackling these problems can be summarized as follows:

1. **Data collection:** We collected data from various sources describing the evolution of infections, absolute humidity, vaccinated population over the course of eight seasons for all states of the United States. Furthermore, we used Census data as proxies to the travel intensity between each pair of states.

2. **Data pre-processing:** We used the methodology described in Section 7.5.1 to identify state-season pairs for which identification of network effects is possible.

3. **Model estimation:** We used the data corresponding to the state-season pairs for which identification of network effects is not possible in order to learn the parameters of the compartmental models under consideration, using a recurrent neural network approach as explained in Section 7.5.2.

4. **Identification of network effects:** We followed a simplistic approach to identify network effects, using the estimates from the previous step and using the data of the state-season pairs for which identification of network effects is possible.

Our analysis shows that the predictive accuracy of the compartmental models is fair but extremely dependent on the environmental factors. Specifically, our findings indicate that the effective contact rate strongly depends on the absolute humidity and as a result the predictive accuracy of the model boils down to the accuracy of the relation between humidity evolution and the time series of the number of infections (as shown in the extreme example of Figure 7-15).

Several directions can be considered in order to improve the predictive accuracy of the models described in this section, such as adding more compartments (e.g., individuals exposed to infection but not symptomatic, etc.) but if prediction accuracy is the goal, we strongly believe that other model free approaches such as time series prediction ([61, 53, 12]) or non parametric forecasting ([65] and references therein) appear to be more promising.

Regarding the evaluation of network effects, our approach indicates that the effect of traveling infections is minimal compared to intra-state contacts. This result though is obtained through a fairly simplistic approach. Unfortunately, although identification of network effects has been widely studied in the Econometrics literature ([10] and references therein) to the best of our knowledge there is no approach that can be applied to time series data. As argued in this chapter, the latter is a very important problem and one of the many possible extensions of this work.

Finally, the network identification task and our findings (Step 4) are dependent on the estimation step (Step 3). On the other hand neither the fit nor the prediction errors are sufficiently small to allow for high confidence on the estimated models. Hence, an important and interesting direction is the identification of network effects independently of the estimation procedures (model free).

Appendix A

Real-time Recurrent Learning Updates

In this appendix we provide the updates for the model under consideration, i.e. Equation (7.9)

$$\begin{aligned}\frac{\partial I_i^s(t+1)}{\partial m} =& f_s(H_i(t))\left(-\sum_{t'=1}^{t-1} I_i^s(t')\right)\beta_{ii} I_i^s(t) \\ &+ f_s(H_i(t))\left(-m\sum_{t'=1}^{t-1}\frac{\partial I_i^s(t')}{\partial m}\right)\beta_{ii} I_i^s(t) \\ &+ f_s(H_i(t))\left(N_i V_i^s(t) - m\sum_{t'=1}^{t-1} I_i^s(t')\right)\beta_{ii}\frac{\partial I_i^s(t)}{\partial m}.\end{aligned} \tag{A.1}$$

Moreover, note that changes in a_j affect only the values of I_i^j and hence,

$$\frac{\partial I_i^s(t+1)}{\partial a_j} = 0, \text{ for all } j \neq s$$

and

$$\begin{aligned}\frac{\partial I_i^s(t+1)}{\partial a_j} =& f_s'(H_i^s(t))\left(N_iV_i(t) - m\sum_{t'=1}^{t-1} I_i^s(t')\right)\beta_{ii}I_i^s(t) \\ &+ f_s(H_i^s(t))\left(-m\sum_{t'=1}^{t-1}\frac{\partial I_i^s(t')}{\partial a_j}\right)\beta_{ii}I_i^s(t) \\ &+ f_s(H_i^s(t))\left(N_iV_i(t) - m\sum_{t'=1}^{t-1} I_i^s(t')\right)\beta_{ii}\frac{\partial I_i^s(t)}{\partial a_j},\end{aligned} \tag{A.2}$$

where $f_s'(\cdot)$ denotes the derivative of f_s with respect to a. Finally, note that changes in β_{ii} affect only the values of I_i^s and hence,

$$\frac{\partial I_k^s(t+1)}{\partial \beta_{ii}} = 0, \text{ for all } k \neq i$$

and

$$\begin{aligned}\frac{\partial I_i^s(t+1)}{\partial \beta_{ii}} =& f_s(H_i^s(t))\left(N_iV_i(t) - m\sum_{t'=1}^{t-1} I_i^s(t')\right)I_i^s(t) \\ &+ f(H_i^s(t))\left(-m\sum_{t'=1}^{t-1}\frac{\partial I_i^s(t')}{\partial \beta_{ii}}\right)\beta_{ii}I_i^s(t) \\ &+ f(H_i^s(t))\left(N_iV_i(t) - m\sum_{t'=1}^{t-1} I_i^s(t')\right)\beta_{ii}\frac{\partial I_i^s(t)}{\partial \beta_{ii}}.\end{aligned} \tag{A.3}$$

Therefore, the system of equations (7.14) and (A.1)-(A.3) provide a systematic and efficient way to calculate the gradient of E and hence to perform gradient descent.

Bibliography

[1] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, February 2015.

[2] E. Adar and L. A. Adamic. Tracking information epidemics in blogspace. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, WI '05, Washington, DC, USA, 2005. IEEE Computer Society.

[3] L. J. Allen, F. Brauer, P. Van den Driessche, and J. Wu. *Mathematical Epidemiology*. Springer, 2008.

[4] S. Aral and D. Walker. Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341, 2012.

[5] S. Baize, D. Pannetier, L. Oestereich, T. Rieger, L. Koivogui, N. Magassouba, B. Soropogui, M. S. Sow, S. Keïta, H. De Clerck, A. Tiffany, G. Dominguez, M. Loua, A. Traoré, M. Kolié, E. R. Malano, E. Heleze, A. Bocquin, S. Mély, H. Raoul, V. Caro, D. Cadar, M. Gabriel, M. Pahlmann, D. Tappe, J. Schmidt-Chanasit, B. Impouma, A. K. Diallo, P. Formenty, M. Van Herp, and S. Günther. Emergence of zaire ebola virus disease in guinea. *New England Journal of Medicine*, 371(15):1418–1425, 2014. PMID: 24738640.

[6] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, 235(2):275 – 288, 2005.

[7] D. Bernoulli. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Mém Math Phys Acad Roy Sci , Paris*, 1766.

[8] D. Bienstock and P. Seymour. Monotonicity in graph searching. *Journal of Algorithms*, 12(2):239 – 245, 1991.

[9] C. Borgs, J. Chayes, A. Ganesh, and A. Saberi. How to distribute antidote to control epidemics. *Random Structures and Algorithms*, 37(2):204–222, 2010.

[10] Y. Bramoullé, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41 – 55, 2009.

[11] F. R. K. Chung, P. Horn, and A. Tsiatas. Distributing antidote using pagerank vectors. *Internet Mathematics*, 6(2):237–254, 2009.

[12] A. Cliff and P. Haggett. Statistical modelling of measles and influenza outbreaks. *Statistical Methods in Medical Research*, 2(1):43–73, 1993.

[13] R. Cohen, S. Havlin, and D. Ben-Avraham. Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91:247901, 2003.

[14] D. J. Daley and J. M. Gani. *Epidemic modelling : an introduction.* Cambridge University Press New York, 1999.

[15] K. Drakopoulos, A. Ozdaglar, and J. Tsitsiklis. An efficient curing policy for epidemics on graphs. *IEEE Transactions on Network Science and Engineering*, 1(2):67–75, July 2014.

[16] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis. A lower bound on the performance of dynamic curing policies for epidemics on graphs. *In Proceedings of the 54th IEEE Conference on Decision and Control, Osaka, Japan*, 2015.

[17] K. Drakopoulos, A. E. Ozdaglar, and J. N. Tsitsiklis. When is a network epidemic hard to eliminate? *to appear in Mathematics of Operations Research*, 2016.

[18] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12:2121–2159, July 2011.

[19] S. Dughmi. Submodular functions: Extensions, distributions, and algorithms. A survey. *CoRR*, abs/0912.0322, 2009.

[20] R. Durrett and X.-F. Liu. The contact process on a finite set. *Ann. Probab.*, 16(3):1158–1173, 07 1988.

[21] A. Ganesh, L. Massoulié, and D. Towsley. The effect of network topology on the spread of epidemics. In *in IEEE INFOCOM*, pages 1455–1466, 2005.

[22] A. Ganesh, L. Massoulié, and D. Towsley. The effect of network topology on the spread of epidemics. In *IN IEEE INFOCOM*, pages 1455–1466, 2005.

[23] M. Garetto, W. Gong, and D. Towsley. Modeling malware spreading dynamics. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Society*, volume 3. IEEE, 2003.

[24] F. Gavril. Some NP-complete problems on graphs. *11th Conference on Information Sciences and Systems*, 1977.

[25] J. Ginsberg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457:1012–1014, 2009. doi:10.1038/nature07634.

[26] M. Gomez Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, New York, NY, USA, 2010.

[27] E. Gourdin, J. Omic, and P. Van Mieghem. Optimization of network protection against virus spread. In *8th International Workshop on the Design of Reliable Communication Networks (DRCN)*. IEEE, Oct. 2011.

[28] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

[29] Y. B. Ian Goodfellow and A. Courville. Deep learning. Book in preparation for MIT Press, 2016.

[30] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 115(772):700–721, 1927.

[31] L. Kim, M. Abramson, K. Drakopoulos, S. Kolitz, and A. Ozdaglar. Estimating social network structure and propagation dynamics for an infectious disease. In *Social Computing, Behavioral-Cultural Modeling and Prediction*, volume 8393 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014.

[32] R. S. Koff. Infectious diseases of humans: Dynamics and control. *Hepatology*, 15(1):169–169, 1992.

[33] A. Lajmanovich and J. A. Yorke. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences*, 28(3–4):221 – 236, 1976.

[34] A. S. LaPaugh. Recontamination does not help to search a graph. *J. ACM*, 40(2):224–245, Apr. 1993.

[35] T. Leighton and R. Satish. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM*, 46(6):787–832, Nov. 1999.

[36] J. Leskovec, L. A. Adamic, and B. A. Huberman. The dynamics of viral marketing. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, New York, NY, USA, 2006.

[37] J. Leskovec, L. A. Adamic, and B. A. Huberman. The dynamics of viral marketing. *ACM Trans. Web*, 1(1), May 2007.

[38] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2006.

[39] C. Li, R. van de Bovenkamp, and P. Van Mieghem. Susceptible-infected-susceptible model: A comparison of n-intertwined and heterogeneous mean-field approximations. *Phys. Rev. E*, 86:026116, Aug 2012.

[40] T. M. Liggett. *Interacting Particle Systems.* Springer, 1985.

[41] P. V. Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, Feb 2009.

[42] B. Monien and I. Sudborough. Min cut is NP-complete for edge weighted trees . *Theoretical Computer Science*, 58(1–3):209–229, 1988.

[43] National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention. Flu and you.

[44] M. E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66(1):016128, July 2002.

[45] C. Nowzari, V. M. Preciado, and G. J. Pappas. Analysis and Control of Epidemics: A survey of spreading processes on complex networks, May 2015.

[46] O. V. Baroyan, L. A. Rvachev, U. V. Basilevsky, V. V. Ermakov, K. D. Frank, M. A. Rvachev, V. A. Shashkov. Computer modelling of influenza epidemics for the whole country (ussr). *Advances in Applied Probability*, 3(2):224–226, 1971.

[47] T. D. Parsons. The search number of a connected graph. In *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory, and Computing*, 1978.

[48] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, Aug 2015.

[49] R. Pastor-Satorras and A. Vespignani. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14):3200–3203, Apr. 2001.

[50] A. Pollack. Ebola Drug Could Save a Few Lives. But Whose? *New York Times*, August 8th 2014.

[51] Population and H. Unit. United states summary: 2000. *2000 Census of Population and Housing*, April 2004.

[52] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas. Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. *CoRR*, abs/1303.3984, 2013.

[53] P. Quénel and W. Dab. Influenza a and b epidemic criteria based on time-series analysis of health services surveillance data. *European Journal of Epidemiology*, 14(3):275–285.

[54] E. Rogers. *Diffusion of Innovations, 5th Edition.* Simon and Schuster, 2003.

[55] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Neurocomputing: Foundations of research. chapter Learning Representations by Back-propagating Errors, pages 696–699. MIT Press, Cambridge, MA, USA, 1988.

[56] A. Sanjeev, F. Alan, and K. Haim. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. *Mathematical Programming*, 92(1):1–36, 2002.

[57] R. P. Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6):066117+, May 2001.

[58] J. Shaman and M. Kohn. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences*, 106(9):3243–3248, 2009.

[59] T. Sonoguchi, M. Sakoh, N. Kunita, K. Satsuta, H.Noriki, and H. Fukumi. Reinfection with influenza a (h2n2, h3n2, and h1n1) viruses in soldiers and students in japan. *Journal of Infectious diseases*, 153(1):33–40, Jan 1986.

[60] J. Steele. *Stochastic Calculus and Financial Applications*, volume 45. Springer, 2000.

[61] D. F. Stroup, S. B. Thacker, and J. L. Herndon. Application of multiple time series analysis to the estimation of pneumonia and influenza mortality by age 1962–1983. *Statistics in Medicine*, 7(10):1045–1059, 1988.

[62] D. M. Thilikos, M. Serna, and H. L. Bodlaender. Cutwidth I: A Linear Time Fixed Parameter Algorithm. *J. Algorithms*, 56(1):1–24, July 2005.

[63] P. Van Mieghem and R. van de Bovenkamp. Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks. *Phys. Rev. E*, 91:032812, Mar 2015.

[64] P. Van Mieghem and R. van de Bovenkamp. Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks. *Phys. Rev. E*, 91:032812, Mar 2015.

[65] C. Viboud, P.-Y. Boëlle, F. Carrat, A.-J. Valleron, and A. Flahault. Prediction of the spread of influenza epidemics by the method of analogues. *American Journal of Epidemiology*, 158(10):996–1006, 2003.

[66] R. J. Williams and D. Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.*, 1(2):270–280, June 1989.

[67] World Health Organization. Ebola virus disease. *Fact sheet No 103.*

[68] World Health Organization. Influenza (seasonal). *Fact sheet No 211.*

[69] World Health Organization. Statement on the WHO Consultation on potential Ebola therapies and vaccines. September 2014.