國立交通大學

資訊科學與工程研究所

博士論文

網路導向式計算流行病學：
整合疾病動態與人類社會網路的多層次
傳染病模型架構

Network-based Computational Epidemiology:
A Multilayer Framework Integrating Social Networks with Epidemic Dynamics

研 究 生：蔡宇軒

指導教授：孫春在　教授

中 華 民 國 一 百 年 一 月

網路導向式計算流行病學：
整合疾病動態與人類社會網路的多層次傳染病模型架構
Network-based Computational Epidemiology:
A Multilayer Framework Integrating Social Networks with Epidemic Dynamics

研 究 生：蔡宇軒　　Student：Yu-Shiuan Tsai

指導教授：孫春在　　Advisor：Chuen-Tsai Sun

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

博 士 論 文

A Dissertation

Submitted to

Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Computer Science

January 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年一月

網路導向式計算流行病學：
整合疾病動態與人類社會網路的多層次傳染病模型架構

學生：蔡宇軒　　指導教授：孫春在博士

國立交通大學資訊科學與工程研究所博士班

摘　　要

網路導向式計算流行病學利用電腦與理論或真實網路拓樸結構研究人類疾病動態和社會趨勢。本論文的主旨在於探討網路導向式計算流行病的重要性、研究現況、優勢與建模過程，並詳述三項原創研究。首先，第一項研究以理論探討無尺度網路下個體資源和疾病傳播成本對於疾病傳播關鍵門檻值的影響，並於流行病模型的基礎上提出解析方程式來解釋關鍵門檻值在無尺度網路下的存在性。該研究指出個體資源和疾病傳播成本的控管對於在無尺度網路下疫情擴散防治的可行性。其次，第二項研究提出整合真實社會網路、個體觀點、國家觀點的多層流行病學架構－多層流病動態模擬器（MEDSim），並以該架構模擬 2009 年 A 型 H1N1 流感疫情在台灣爆發的情形，測試該架構對於不同爆發地點和傳染阻絕方案的靈活性，希望藉此釐清複雜的個體接觸行為對於疾病傳播動態的影響。最後，在第三項研究中分析網路導向式計算流行病學的潛在優勢，並針對網路導向式計算流行病學初學者給予建立網路導向式流行病模型的方法。該研究的目標在於

協助擁有較低電腦技能者建立流行病學模型、決定合適的模擬參數與建立操作流程。本論文期望透過上述三項研究，利用電腦模擬來分析多層次的個體互動行為，進而協助傳染阻絕政策的制定。

Network-based Computational Epidemiology:

A Multilayer Framework Integrating Social Networks with Epidemic Dynamics

Student：Yu-Shiuan Tsai　　Advisor：Dr. Chuen-Tsai Sun

Institute of Computer Science and Engineering
National Chiao Tung University

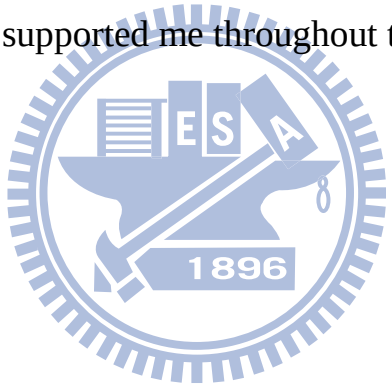
ABSTRACT

Network-based computational epidemiologists use computers and either theoretical or actual network topologies to study the transmission dynamics of human diseases and social trends. In this dissertation I discuss the importance, current status, advantages, and modeling procedures of network-based computational epidemiology, specifically presenting three original studies in detail. The first study is an investigation of how resources and transmission costs influence diffusion dynamics and tipping points in scale-free networks. An epidemic model based on an analytic equation is proposed to explain the existence of epidemic critical thresholds in scale-free networks. Study results suggest the possibility of controlling the spread of epidemics in scale-free networks by manipulating resources and costs associated with an infection event. In the second study, a proposal for a multilayer epidemiological framework that integrates realistic social networks, called the Multilayer Epidemic Dynamics Simulator (MEDSim), is

described from individual and national perspectives. Model flexibility and generalizability are tested using outbreak locations and intervention scenarios for the 2009 A/H1N1 influenza epidemic in Taiwan. The results coincide with the dynamic processes of epidemics under different intervention scenarios, thus clarifying the effects of complex contact structures on disease transmission dynamics. In the third study, the potential benefits of epidemic simulations and instructions for building network-based epidemic models by novices learning network-based computational epidemiology approaches is investigated. The goal is to help individuals with less advanced computing skills build epidemiological models, determine appropriate simulation parameters, and construct operational procedures. It is my hope that the studies presented in this dissertation can assist in efforts by public health organizations to correctly implement intervention strategies by using simulations to analyze multilayer interactions.

Acknowledgements

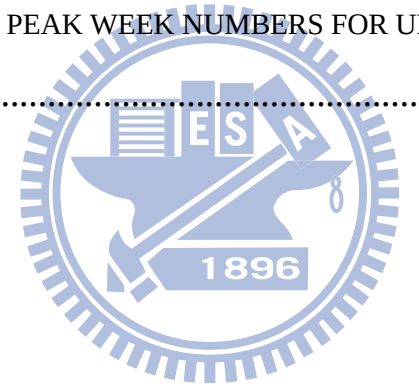
I am greatly thankful to my advisor, Professor **Chuen-Tsai Sun**, whose encouragement, suggestions and support from the proposal to the concluding level enabled me to develop an understanding of the dissertation. I am deeply indebted to my admirable friends, Dr. **Chung-Yuan Huang** and Dr. **Tzai-Hung Wen**, for their continuous support and intellectual suggestions. I would like to thank my family, and my lovely girlfriend, **Pin-Han Wang**. For their love and encouragement, I can express only an inadequate acknowledgement of my appreciation. I dedicate this dissertation to them. I also dedicate this dissertation to my many friends who have supported me throughout the process.



Contents

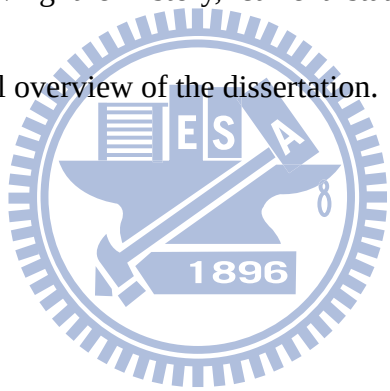
List of Tables

List of Figures



Chapter 1. Introduction

Network-based computational epidemiologists use computers and either theoretical or realistic network topologies to study the reasons, conditions, and transmission dynamics of human diseases and social trends. In this chapter I will summarize several network-based computational epidemiological issues, and introduce some of the details of both computational and network-based computational epidemiology. After reviewing the history, current status, and importance of the subject, I will give a general overview of the dissertation.



1.1. Computational and Network-based Computational Epidemiology

Epidemiologists study the distribution of individuals who are healthy or infected during a contagious disease outbreak, as well as conditions and factors supporting the spread of a disease (Lloyd & May, 2001). The two most common research approaches are observation and experimentation. In the first, epidemic diseases are analyzed using empirical data collected from clinical cases, epidemic monitoring surveys, and other investigation tools to determine contagious patterns or disease properties. In the second, subjects are randomly divided into two groups, members of one group are treated with the experimental variable, and a comparison of the two groups determines the positive, negative, or null effects of the variable.

Computational epidemiologists construct mathematical models and use computing techniques to obtain epidemic results. Researchers validate results by comparing them with observable empirical data, or use their results to explain experimental variable characteristics. The most commonly used mathematical tool for computational epidemiology studies is differential equations, in which individuals in a population are divided into finite states representing different health statuses. Relations among states are determined by differential equations using mathematical

symbols. A major advantage of a differential equation system is that the number of each state can be easily computed; a disadvantage is that they are weak in terms of describing social properties such as social distance.

To compensate for the lack of social properties, epidemiologists are incorporating social networks into their mathematical models, reflecting the idea that there is no distance between individuals with the same health status—that is, there are no restrictions on any two individuals being in contact with each other. Social network structures consist of nodes (objects) and links (social relations). For example, in a friendship network, nodes represent individuals and links represent whether or not two nodes are friends. Due to its ability to represent social relations, network-based computational epidemiology has grown in popularity. However, since individual characteristics and their corresponding integrated mathematical models are so complex, powerful computers are required to solve equation systems. As the size and detail of a social network increases, so does the need for increased computation time and power.

1.2. Computational Epidemiology History

The first approaches used in computational epidemiology were based on differential equation systems. One of the first contagious epidemiology models, proposed by Kermack and McKendrick (1927), is known as the compartmental SIR model. In this model, all individuals in a population are classified as *Susceptible* (vulnerable to infection but not yet infected), *Infected* (and capable of infecting others), or *Removed* (recovered, dead, or otherwise not posing any further threat). Differential equations mark the progress of each state. In the past 80 years, numerous compartmental models have been created and improved for research purposes (for examples, see Bailey, 1950; Bartlett, 1956; Diekmann, Heesterbeek & Metz, 1990; Hyman & Stanley, 1988; Rollett, 1945). Major progress was made in the 1990s, with the addition of other states and model revisions to emphasize cyclic characteristics (Ahmed & Agiza, 1998; Anderson & May, 1991; Wang, 2006).

Describing epidemic dynamics using compartmental models based on differential equation systems is an easy method for representing the time dimension, but such approaches lack a spatial dimension. Individuals in the same compartment are modeled as one group, implying that any two group members are directly connected—a flawed concept, since it ignores a long list of potential real-world

differences among individuals. To overcome this flaw, Von Neumann (1966) introduced his cellular automata model (which considers spatial differences and the movement of individuals) to epidemic propagation research (Fuentes & Kuperman, 1999; Sirakoulis, Karafyllidis & Thanailakis, 2000). Other researchers focused on integrating compartmental and cellular automata models to support epidemiological models (Liu & Jin, 2005; Mikler, Venkatachalam & Abbas, 2005; White, del Rey & Sánchez, 2007).

In addition to using the cellular automata model to consider spatial effects, social network models are increasingly being used by mathematical epidemiologists. Watts and Strogatz (1998) have proposed the concept of a "small-world" phenomenon to explain why any two individuals in the world can be contacted via a small number of connecting individuals. Barabasi and Albert (1999) then proposed a "scale-free network" algorithm to explain the phenomenon of "the rich becoming richer." Unlike theoretical random networks (Erdos & Renyi, 1960), social networks are much closer to the real world situation, and hence can be used to depict individual contacts in network-based epidemic model studies (Dezső & Barabási, 2002; Grais, Hugh Ellis & Glass, 2003; Meyers, Newman, Martin & Schrag, 2003; Newman, 2002; Pourbohloul et al., 2005; Parham & Ferguson, 2006; Handcock & Jones, 2006). Other researchers have focused on the influences of social network structure on compartmental models

(Barthelemy, Barrat, Pastor-Satorras & Vespignani, 2004, 2005; Draief, 2006; Pastor-Satorras & Vespignani, 2001b; Shirley & Rushton, 2005; Silva, Ferreira & Martins, 2007; Wang, 2002; Yang et al., 2007; Zhou, Yan & Wang, 2005;).

Agent-based differential equation system approaches emphasize heterogeneity and interactions among individuals. In these approaches, individuals are represented as agents whose interactions can be modeled in the form of rules (Boguñá & Pastor-Satorras, 2002; Huang, Sun, Hsieh & Lin, 2004). The advantage of such an approach is that it supports simulations of the movement of individuals, which in turn supports an understanding of epidemic contagion routes. Using this kind of approach, Barrett et al. (2005) constructed a society of 1.6 million agents to simulate the daily behaviors of individuals in Portland, Oregon, and Epstein (2009) studied the 2009 influenza A (H1N1) epidemic by constructing a model containing 6.5 billion agents to simulate international human contact and daily movement. Unlike compartmental models that focus on the behaviors of whole populations, agent-based models focus on individual behaviors.

A geometric structure has recently been integrated into epidemic models. Due to the limitations of standard cellular automata, in this study geographical cellular automata are used to simulate an environment (Liu, Xia, Yeh, Qiang & Jia, 2007; Zhou, Sun & Xie, 1999). An actual geographic area can be defined as cells to study epidemic

dynamics in social and geometric transformations (Flache & Hegselmann, 2001; Menard, 2008). Other kinds of cellular automata have been tied to epidemic contagious behaviors via network-based compartmental models (Zhong, Huang & Song, 2009). To visualize the dynamics of a regional epidemic, at least two research teams have integrated a geometric information system (GIS) into a mathematical epidemiological model (Xu, Zhang & Mendes, 2007; Zhou, 2008).

1.3. Current Status of Computational Epidemiology

Epidemiologists are currently emphasizing temporal and spatial depictions of infectious disease occurrences and pathogenic mechanisms. Regarding the temporal aspect, researchers have focused on understanding spreading trends and dynamic changes in infectious diseases. The most common approach is to construct analytically systematic epidemiological models with differential equations, and then derive stable solutions (Feng, Huang & Castillo-Chavez, 2005; Inaba, 2007; Langlais & Naulin, 2003; Li & Jin, 2005; Shim, Feng, Martcheva & Castillo-Chavez, 2006; Supriatna, Soewono & Van Gils, 2008; Wang & Zhao, 2005). Populations can be broken down into infection stages such as *Susceptible*, *Latent*, *Infectious* and *Recovered*, and changes in subpopulations over time can be modeled using system dynamic differential equations (Feng et al., 2005; Inaba, 2007; Langlais & Naulin, 2003; Li & Jin, 2005; Shim et al., 2006; Supriatna et al., 2008; Wang & Zhao, 2005). Using suitable parameter values (e.g., transmission rate, recovery rate), infectious dynamics and transmission thresholds that become endemic above and vanish below those values can be derived to acquire analytic solutions from equations (Huang, Tsai & Sun, 2009; Huang, Tsai, Sun, Hsieh & Cheng, 2010; Pastor-Satorras & Vespignani, 2001, 2002;

Tsai, Sun & Huang, 2008). According to the different transmission capabilities of epidemic diseases, basic reproduction numbers can be derived to estimate how many individuals will be infected from the first infected individual (Hethcote, 2000; Keeling & Grenfell, 2000).

Regarding the spatial aspect, researchers have focused on understanding the distribution of infected individuals (which can be determined from medical case reports) to help in monitoring and immunization efforts. Because of the advantages of computer technology, GIS data on absolute distance and the properties of geographical regions are now commonly applied in research (Rae, 2009; Wylie, Shah & Jolly, 2007). Many researchers are also integrating GIS into epidemic disease monitoring and prevention efforts (Edwards & Clarke, 2009; Jeger, Pautasso, Holdenrieder & Shaw, 2007; Mao & Bian, 2010; Thakur & Sharma, 2009). By analyzing medical cases and collecting data on environmental factors, geographic spatial distribution information can be determined, and epidemic pathogenic mechanisms can be analyzed. For example, using spatial clustering analysis, it is possible to analyze abnormal clusters that exceed an expected number of infected cases, thus supporting efforts to understand the extent of disease clustering relative to increases in disease vectors (Kan et al., 2008). Kan et al. have used this approach to explain the smaller number of cases of dengue hemorrhagic fever in Taiwan compared to Southeast Asian countries. However, it is important to

take advantage of both temporal and spatial aspects when analyzing infectious disease propagation, therefore many researchers are trying to integrate both temporal and spatial factors into their epidemiological models (Barrett, Eubank & Marathe, 2006; González, Hidalgo & Barabási, 2008; Yang, Atkinson & Ettema, 2008).

1.4. Trends in Social Network Integration in Computational Epidemiology

The past decade has witnessed significant advancements in social network research, ever since Watts and Strogatz (1998) first described small-world networks characterized by highly clustered connections and short paths between node pairs. Their work represents a fundamental change in our knowledge of human relationships, which has influenced research avenues in a wide range of disciplines such as epidemiology. (Diosan & Dumitrescu, 2007; Montoya & Solé, 2002; Vázquez, Flammini, Maritan & Vespignani, 2003).

Complex networks can be used to model real-world complexity. A complex network is a structure containing numerous nodes and edges. Nodes can represent objects such as individuals, locations, organisms, or World Wide Web pages. Depending on node type, edges can represent relationships such as human friendships, food chains for non-human organisms, or links between web pages. Several network indexes have been developed to measure relationships (Boccaletti, Latora, Moreno, Chavez & Hwang, 2006). For example, degree of clustering has been used to determine why our friend's friend is often also our friend, degree of separation has been used to measure how small the world is in terms of weak links, and connectivity distribution

has been used to explain the existence of super nodes (Huang, Tsai & Sun, 2010). Such topological characteristics have also been used as epidemiological indexes to measure the spreading speed of an epidemic disease (Edmunds, O'Callaghan & Nokes, 1997; Estrada & Hatano, 2008; Hwang, Kim, Ramanathan & Zhang, 2008).

Infectious diseases spread through individual contact, and many epidemiologists are using social networks to model individual contact behavior. Social networks, one type of complex network that is also considered a social structure model, emphasize individual heterogeneity, individual interaction, and network topological structure (Boguñá & Pastor-Satorras, 2002; Huang et al., 2004). They are often used to model populations, with nodes representing individuals and links representing contacts. Social network topological structures have been used in many epidemic studies over the past decade. Based on human epidemic disease or computer virus features, different social network structures have been proposed to analyze epidemic spreading dynamics and transmission rate thresholds (see, for example, Huang, Sun, Hsieh, Chen & Lin, 2005; Langlais & Naulin, 2003; May & Lloyd, 2001; Pastor-Satorras & Vespignani, 2001b). In addition, traffic networks such as daily commuting routes have been used to analyze the spread of diseases via human transportation networks (Barrett et al., 2005, 2006). Social network studies comparing the efficiencies of various public health policies have been conducted by Huang et al. (2004), Huang, Sun & Lin (2005), and Pastor-Satorras

and Vespignani (2001b, 2002).

New epidemiological models integrate spatial and social network factors. The most commonly used approach adds various network topologies (e.g., small-world network, scale-free network) to determine different epidemic spatial distributions (Huang et al., 2004; Pastor-Satorras & Vespignani, 2001b, 2002). After building social network models, parameters such as initial infected agent, and epidemic attributes such as transmission and recovery rates, are manipulated to calculate disease propagation within the defined network (Huang et al., 2004; Wang & Ruan, 2004). According to epidemic properties, different simulation scenarios (e.g., network topologies, contact patterns, agent attributes such as age or gender) can be studied using simulations in order to develop effective public health policies. For example, HIV research entails looking at how heterosexual sexual contact, homosexual sexual contact, or illegal drug use affects virus transmission and propagation in a social network (Morris, 1997; Sumodhee, Hsieh, Sun, Huang & Chen, 2005).

1.5. Advantages of Network-based Computational Epidemiology

Understanding the spreading dynamics of infectious diseases and the spatial distribution of infected individuals is the primary concern of agencies involved in infectious disease control and prevention (Hethcote, 2000; Moore & Newman, 2000; Pastor-Satorras & Vespignani, 2002). Efforts to understand social network associations among geographical characteristics such as coordinates, population size, and census data represent a current trend in computation epidemiology. The advantages of understanding these associations are as follows:

1. Epidemic disease properties such as the transmission capability of a virus and recovery days among individuals are connected to geographical location (Barrett et al., 2005; Larsen, Axhausen & Urry, 2006). For example, the transmission capability of influenza in urban areas is greater than in rural areas because of population density differences, therefore when setting epidemic parameters, transmission rate should be higher in urban areas. Network integration into compartmental models can be used to represent individual heterogeneity. Associating social networks with geographical characteristics has the advantage of accurately describing the topology of individual social relations in the real world (Barrett et al., 2005; Davis, Yoo & Baker, 2003).

2. Public transportation systems such as aircraft, subways, commuter trains, and buses support the spreading of a virus (Colizza, Barrat, Barthélemy & Vespignani, 2006; Grais et al., 2003; Kaza, Xu, Marshall & Chen, 2009). Modern public transportation systems make it easy to move between distant locations, and pathogens can be carried long distance within a matter of hours or days. In 2009, the swine-origin H1N1 virus emerged in Mexico and rapidly spread throughout South America, Europe, and Asia within a few weeks; by mid-November of that year, 6,770 deaths were reported in 206 geographic locations (Smith et al., 2009). This underscores the importance of considering such factors as the location of public transportation systems in epidemiological studies.

3. During a contagious disease outbreak, medical officials and/or public health experts must consider balances among many factors when determining how to best use medical resources and enact prevention policies (Riley et al., 2003; Molinari et al., 2007). In addition, differences in resources and population densities among administrative and geographical divisions must be considered (Sypsa, Pavlopoulou & Hatzakis, 2009; Wylie et al., 2007). From the perspective of medical system utility, a suitable mix of intervention policies is required to efficiently control a disease outbreak according to limitations of medical resources (Tsai & Huang, 2010; Huang et al., 2010). In addition, time of maximum number of infected individuals in each

division must be considered when planning the timing of interventions across administrative divisions.

4. GIS is a suitable tool for graphically representing epidemics. By using visualization tools, large bodies of complex data can be analyzed spatially. Based on experience with newly emerging viruses such as SARS, avian influenza (H5N1), and swine-adapted influenza (H1N1), public health officials must deal with the potential of one such virus becoming pandemic (Fraser et al., 2009; Kuiken, Rimmelzwaan, Van Amerongen & Osterhaus, 2003; Tomlinson & Cockram, 2003). However, traditional epidemic models cannot adequately work with geographic information due to limitations associated with equation size, therefore geographic network-based computation epidemiology with GIS has value in terms of studying virus spreading trends.

1.6. Dissertation Overview

The rest of this dissertation is organized as follows: in Chapter 2 I will present a brief overview of a preliminary study involving network-based computational epidemiology, especially a network topology proposal from an original study conducted by Huang, Tsai and Sun (2010).

In Chapter 3 I will present details from an original research project conducted by Huang et al. (2010), Tsai & Huang (2010), and Tsai, Sun & Huang (2010) that used network-based computational epidemiology with a theoretically complex network topology. Based on considerations of resource limitations and transmission costs, I will propose an epidemic model that uses analytic equations to identify critical epidemic thresholds in scale-free networks.

In Chapter 4 I will discuss the details of an original research project by Tsai et al. (2010) to integrate realistic social networks with standard epidemiological models, and then describe a multilayer epidemiological framework—Multilayer Epidemic Dynamics Simulator, or MEDSim—from national and individual perspectives. The framework was used to compute outbreak locations and intervention scenarios for the 2009 A/H1N1 influenza epidemic as a means of testing model flexibility and generalizability.

In Chapter 5 I will present details of an original research project on the potential benefits of epidemic simulations, and describe the building of a network-based epidemic model for epidemiology students with little computing experience who are interested in studying computational epidemiology and public health education (Hsieh, Huang, Sun & Tsai, 2009; Huang, Tsai & Wen, 2010a, 2010b). In Chapter 6 I will summarize my conclusions and give suggestions for future research.

Chapter 2. Preliminaries

In this chapter, I will first introduce the most commonly used epidemiological models for network-based computational epidemiological studies, and then briefly explain network-based epidemiology and several social network structures.

2.1. Epidemiological Approaches Overview

The two most commonly used approaches to modeling epidemic spreading dynamics are *population-based* and *network-oriented*. In population-based approaches, hosts that share the same symptoms are modeled or grouped in terms of a limited numbers of *classes* (also known as *compartments*); the main task of researchers is to study and compare their various dynamics (Feng et al., 2005; Inaba, 2007; Langlais & Naulin, 2003; Shim et al., 2006; Supriatna et al., 2008; Wang & Zhao, 2005). Combinations of classes are used to model and analyze population dynamics. For example, the SLIR model puts individuals into one of four infection statuses—*Susceptible*, *Latent*, *Infectious*, or *Recovered* (Li & Jin, 2005)—and differential equations are used to determine transitions between epidemiological phases. Depending on whether removed individuals can become susceptible a second time, diseases can be modeled as SLIR or SLIRS cycles.

Population-based and network-oriented approaches respectively emphasize large-scale population-level and individual-level perspectives. Population-based approaches are suitable for discussing dynamic variation across individuals in the same compartment, but they are weak in terms of modeling individual heterogeneity and addressing human travel networks (Barrett et al., 2005; Huang et al., 2004). Since

individuals are modeled as groups, any two members of the same group are assumed as having a direct connection, which is not true in the real world. Furthermore, individual movement and activity are location-dependent, therefore phenomena cannot be simulated by a population-based approach that assumes a homogeneous population distribution. In contrast, network-oriented approaches may be appropriate for introducing individual heterogeneity, but they are computation-intensive and time-consuming when simulating the behaviors of individuals with multiple attributes in large-scale social environments (Barrett et al., 2005; Epstein, 2009). Many efforts have been made to match individual and population behaviors with heterogeneity and computation requirements when studying epidemic dynamics (Davis et al., 2003; Levin & Durrett, 1996; Sawyer, 2003).

In contrast, network-oriented approaches emphasize individual heterogeneity, interactions among individuals, and network structure (Boguñá & Pastor-Satorras, 2002; Huang et al., 2004). Individuals in a network are represented as nodes, and interactions between them as links. Network nodes can be used to represent the characteristics of individuals, locations, neighborhoods, or cities, and models can incorporate the temporal dynamics of these features. Time frames for links between two nodes can be preferentially defined (Ortiz-Pelaez, Pfeiffer, Soares-Magalhães & Guitian, 2006)—an approach commonly used to represent group structures for

individuals exhibiting interaction or relationship patterns (Barabási & Albert, 1999; Erdos & Renyi, 1960; Newman, 2003; Watts & Strogatz, 1998). Network-oriented approaches are suitable for capturing complex contact patterns among individuals, exploring epidemic dynamics, and assessing the efficacies of public health policies (Pastor-Satorras & Vespignani, 2001b, 2002; Huang et al., 2004, 2005). Lattice networks have been used to determine distance relationships between individuals. In contrast, random networks support features associated with casual contacts among mobile individuals and the low degree of separation commonly observed in social networks (Barrett et al., 2005). These approaches are viewed as reliable for investigating epidemics, with the transmission dynamics of specific network models being manipulated to investigate the spread of emerging infectious diseases (Liu, Lai & Ye, 2003; May & Lloyd, 2001). The topological features of social networks have recently been found to exert considerable influence on the transmission dynamics and critical thresholds of infectious diseases, thus supporting subtle analyses that network-oriented models are incapable of (Draief, Ganesh & Massoulié, 2008; Huang et al., 2005; Langlais & Naulin, 2003; Pastor-Satorras & Vespignani, 2001b).

2.2. Compartmental Models

In standard epidemiological models, all individuals (nodes) in a population (complex network) can be roughly classified into a limited number of states, including *Susceptible*, *Infected* and *Removed*, as defined in Chapter 1. Epidemiologists use combinations of these states to represent orders of transition between different epidemiological phases, giving names such as “SIR” and “SIS” to their models. The most commonly used model is the SIR (Susceptible $\rightarrow$ *I*nfected $\rightarrow$ *R*ecovery) (Figure 2.1), which can be formulated using ordinary differential equations as follows:

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \alpha I$$
$$\frac{dR}{dt} = \alpha I$$

β, a constant transmission rate, represents the speed at which *Susceptible* individuals become infected, and α is a constant recovery rate used to determine transformation speed from *Infected* to *Recovered*.

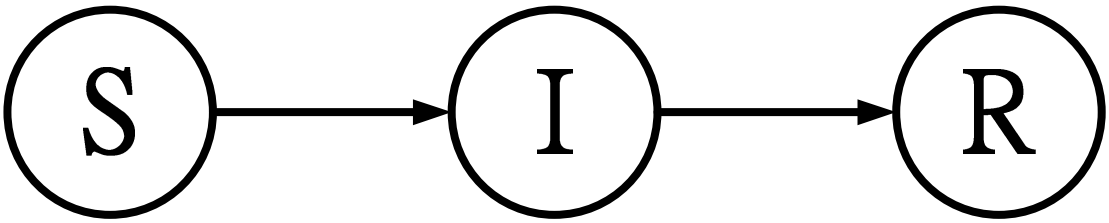


Figure 2.1. Flowchart of the SIR epidemiologic model.

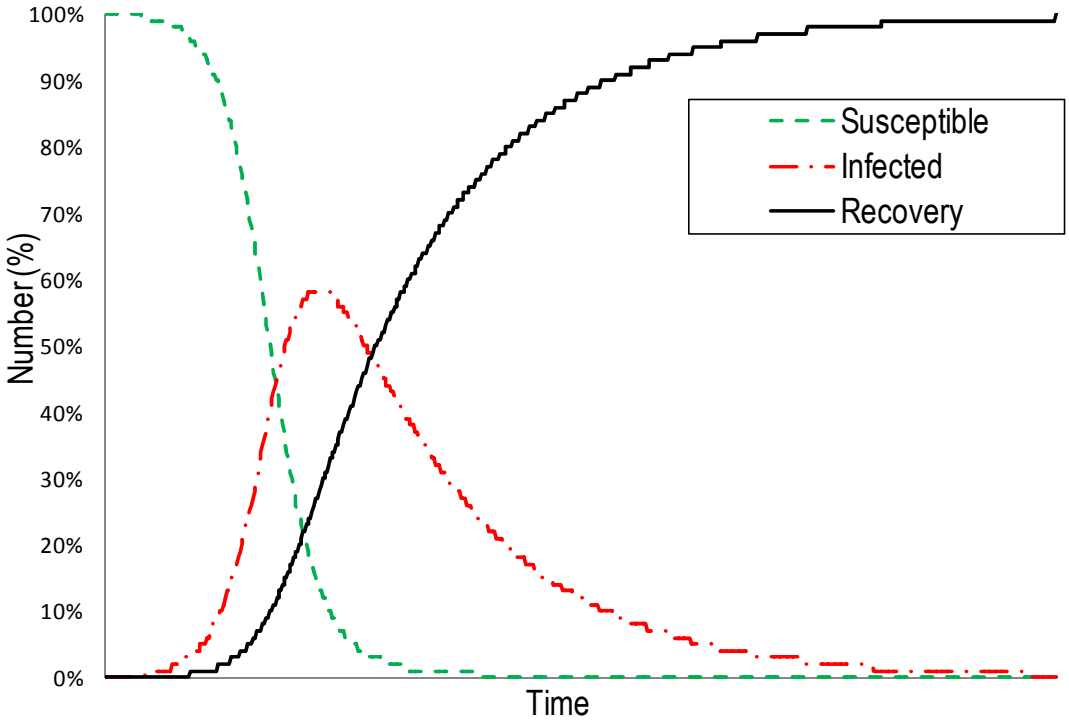


Figure 2.2. The compartment states, *S*, *I*, and *R*, as a function of *t*.

When simulating epidemic dynamics in complex networks, epidemiologists usually assume that nodes run stochastically through an SIS cycle, which does not take into account the possibility of an individual's removal due to death or acquired immunization. The SIS model has been widely adopted to study contagious diseases leading to endemic states with a stationary average density of infected individuals. It is worth noting that for many contagious diseases, analyses derived from the SIS model can be readily extended to the SIR and SIRS models (Pastor-Satorras & Vespignani, 2002). During each time step, each susceptible node is subject to a ν probability contagion rate if it is connected to one or more infected nodes. Infected nodes recover at a probability rate δ, and once again become susceptible. An effective spreading rate

λ is defined as $\lambda = \nu/\delta$. Recovery rate δ can be assigned a value of 1, since it only affects the definition of the time scale of contagious disease propagation (Pastor-Satorras & Vespignani, 2003). Pastor-Satorras and Vespignani (2002) define $\rho(t)$ as the density of infected nodes at time step t. When time step t becomes infinitely large, ρ can be represented as a steady-state density of infected nodes. Using these definitions, they applied mean-field theory to a SIS epidemiological model, and used Anderson and May's (1991) *homogeneous mixing hypothesis* according to the topological features of homogeneous networks to obtain (a) a steady-state density ρ of infected nodes during long time periods (Eq. 2.1), and (b) the critical threshold λ_c (Eq. 2.2):

$$\rho = \begin{cases} 0 & \lambda < \lambda_c \\ \dfrac{\lambda - \lambda_c}{\lambda} & \lambda \geq \lambda_c \end{cases} \tag{2.1}$$

$$\lambda_c = \frac{1}{< k >} \tag{2.2}$$

where $\langle k \rangle = \sum_k k p_k$ is the average vertex degree of the network, and p_k the fraction of nodes that have vertex degree k in the network. According to Eqs. 2.1 and 2.2, a positive and nonzero critical threshold λ_c exists in a homogeneous network based on the SIS epidemiological model. The contagion spreads and becomes epidemic if the effective spreading rate exceeds the critical threshold ($\lambda \geq \lambda_c$); otherwise, the contagion dies out. As shown in Figure 2.3, the SIS epidemiological model separates an infected state from a healthy state at critical threshold λ_c. In summary, the primary prediction of

an SIS epidemiological model in a homogeneous network is the presence of a positive critical threshold, proportional to the inverse of the average number of neighbors of each node, below which epidemics die and endemic states are impossible.

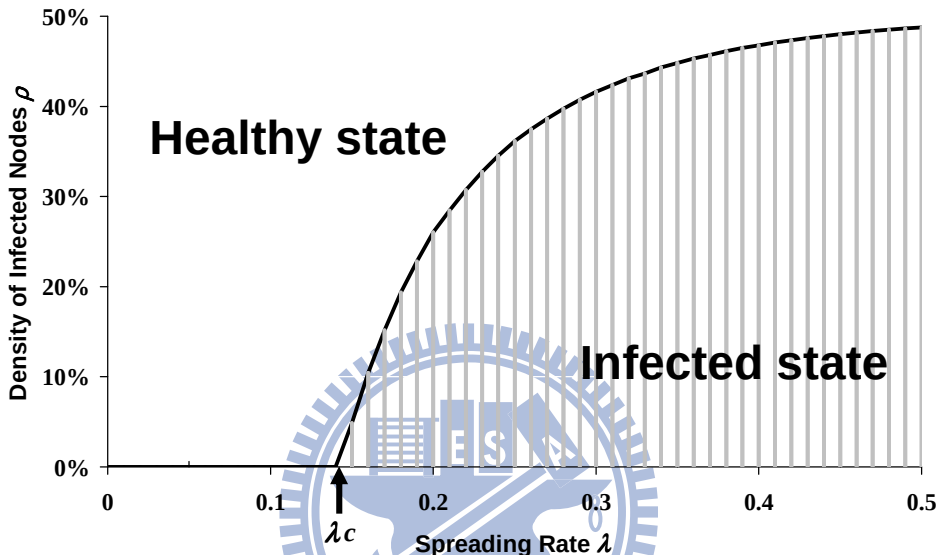


Figure 2.3. Phase transition diagram for epidemic simulations in homogeneous networks.

2.3. Social Network Models and Network-based Epidemiology

Complex networks are commonly used to represent structures for groups of individuals who exhibit interaction or relationship patterns (Barabási & Albert, 1999; Erdos & Renyi, 1960; Newman, 2003; Watts, 2003; Watts & Strogatz, 1998). As shown in Figure 2.4 and Table 2.1, complex networks can be categorized as small world, scale-free, or random according to basic statistical properties such as local clustering, the small world phenomenon, or power-law connectivity distribution. They are popular among researchers who construct computational simulations of virtual societies, contagious diseases, Internet viruses, and the spread of cultural beliefs and influences—all of which are affected by transmission routes.

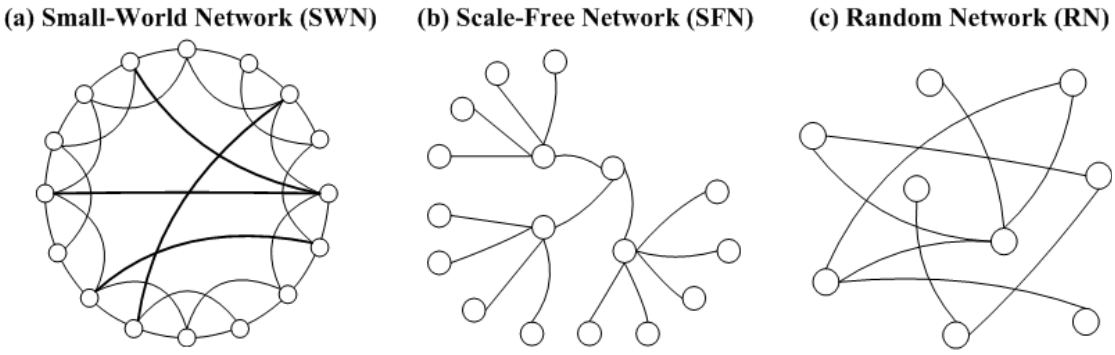


Figure 2.4. Three types of complex networks.

Table 2.1. Two Complex Networks Categories

Category	Network Type	Model	Clustering Coefficient	Degree of Separation	Connectivity Distribution
Homogeneous networks	**Small-world**	Watts and Strogatz	high	low	Normal
	Random	Erdös and Renyi	very low	low	Normal
Heterogeneous network	**Scale-free**	Barabási and Albert	very low	low	Power-law

Generating a Watts and Strogatz (1998) small-world network begins with an *n*-dimension ordered network with periodic boundary conditions, in which each node is connected to a z quantity of neighbors, usually $z \geq 2n$ (Figure 2.5a) (Watts & Strogatz, 1998; Newman, 2003). Each link is randomly rewired to a new node with probability p (Figure 2.5b). Under adverse circumstances, this construction method can break the original ordered network into several isolated subgraphs (Figure 2.5d). Newman and Watts (1999) introduced a variation of the original construction method that emphasizes the insertion of long-range shortcuts instead of rewiring links. In their version, two previously unconnected nodes are randomly selected and connected via a newly added link, with users determining the number of links to be added (Figure 2.5c). Newman and Watts' small-world network thus avoids the problem of network breakage, while preserving the positive characteristic of connecting each node in an *n*-dimensional ordered network with $2n$ neighboring nodes. Since both the original and new versions (Newman, 2003) exhibit small world and local clustering properties, they are considered similar to human daily contact networks.

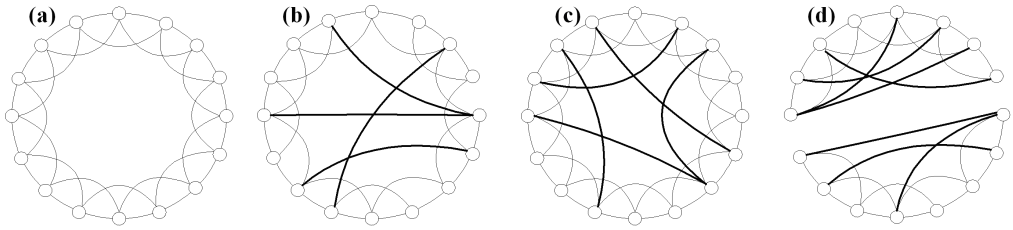


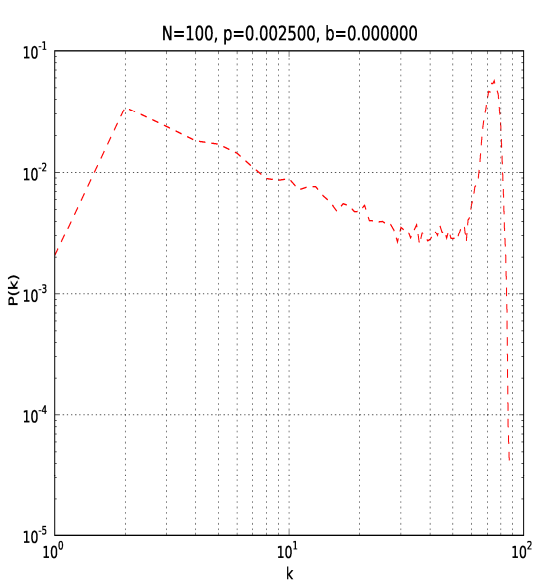
Figure 2.5. (a) One-dimensional ordered network with each node connected to four adjacent nodes. (b) Watts and Strogatz's (Duncan J. Watts & Strogatz, 1998) small world network with four rewired shortcuts. (c) Newman and Watts' (M. E. J. Newman & D. J. Watts, 1999) improved small world network with five additional shortcuts. (d) Example of a broken network in Watts and Strogatz's (Duncan J. Watts & Strogatz, 1998) small world network.

Generating a Barabási and Albert (1999) scale-free network begins with a small number of nodes designated as z_0 (Newman, 2003). During each iteration, a new node is introduced and connected to $z \le z_0$ pre-existing nodes according to a probability based on each node's vertex degree. New nodes are preferentially attached to existing nodes that have large numbers of connections. This type of network exhibits small-world and power-law connectivity distribution properties, implying the existence of a small number of nodes with very large vertex degrees—similar to World Wide Web hyperlinks and human sexual contact webs.

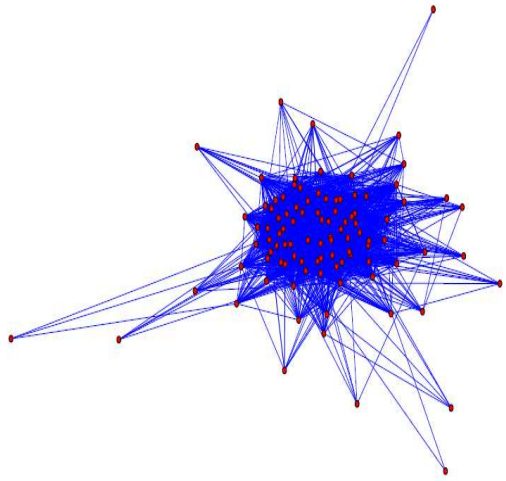
Erdös and Renyi's (1960; Newman, 2003) random networks are generated by adding links between pairs of randomly chosen nodes with certain probabilities. They

are capable of exhibiting small-world properties if sufficient numbers of links are added, but with little or no local clustering—an unusual situation in the real world.

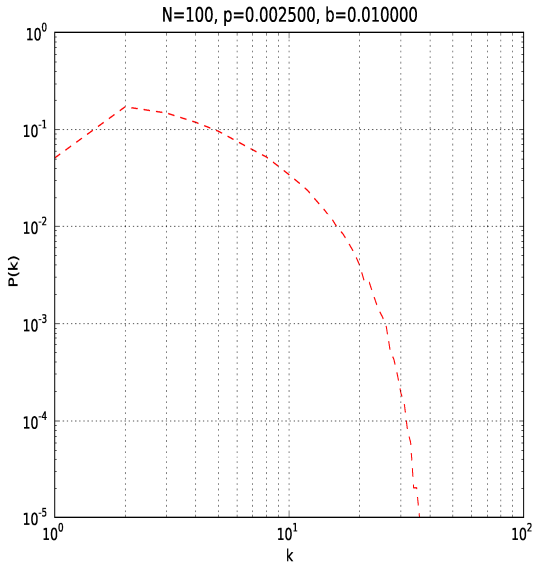
Huang, Tsai and Sun (2010) used three rules to generate friend-making networks—friend making, joining and leaving, and friendship updates—until each network reached a statistically stationary state. Taking a bottom-up, network-oriented simulation approach to modeling reflects the evolutionary mechanism of real-world social networks. They built on insights from previous studies (e.g., Davidsen, Ebel & Bornholdt, 2002) to apply local and interactive rules to acquaintance network evolution. Findings from this approach can be used to explore human activity in specific social networks—for example, rumor propagation and disease outbreaks (Figure 2.6).



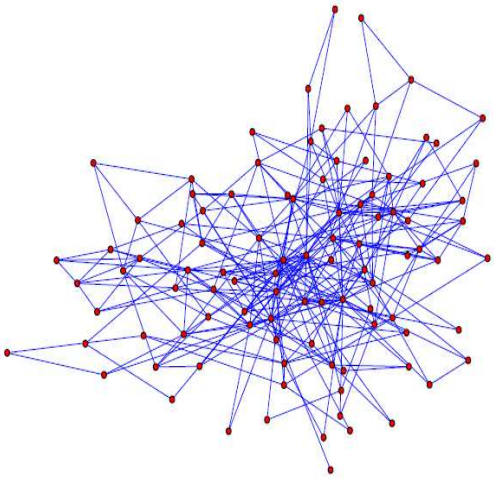
(*a*)



(*b*)



(*c*)



(*d*)

Figure 2.6. Comparison of node degree distributions and network structures between Davidsen et al.'s two-rule model (a and b) and our proposed three-rule model (c and d).

Communities, cities, and countries—even the entire planet—can be defined in terms of complex networks consisting of large-scale nodes and links. Each node

represents one individual with status-determining attributes (often referred to as node-related local information) such as epidemiological progress, contagiousness, and immunization (Huang et al., 2005; Xu et al., 2007). Connections between individuals are referred to as links, with different links representing different interpersonal relationships (Pastor-Satorras & Vespignani, 2001b). In HIV/AIDS epidemic simulations they represent sexual relationships, and in SARS epidemic simulations they represent close physical proximity (Huang et al., 2004, 2005). The states of all network nodes change simultaneously during each time step. The state of an individual node is determined by its original state, its linked neighbor's state, and a set of interaction rules.

Past epidemiological research has focused on the transmission dynamics and spreading situations of biologically contagious diseases. A growing number of research efforts are focusing on non-biological and intangible concepts such as computer viruses, cultural influences, rumors, ideas, and beliefs that exist in social networks and on the Internet. In these kinds of spreading scenarios, cultural influences move ideas and beliefs between transmitters and receivers, eventually making the majority of receivers behave in the same manner as transmitters (Huang et al., 2005; Lynch, 1996; Rogers, 2003). Researchers have recently looked at epidemic dynamics and critical thresholds in homogeneous networks (e.g., Erdös and Renyi's random and Watts and Strogatz's

small-world) and heterogeneous networks (e.g., Barabási and Albert’s scale-free).

Chapter 3. Analysis of Epidemiological Transmission in Theoretical Complex Networks

Avian influenza, a flu that originally spread only among birds but is now found among birds and humans, is a likely candidate to become an epidemic or pandemic disease. Another epidemic, the 1918 influenza outbreak in North America, is one of the most studied by epidemiologists. Nine decades later, Watts (1998) described his proposed small-world property in complex networks, which has strongly influenced research involving human networks. Later, Pastor-Satorras and Vespignani (2001a) combined epidemic dynamics and complex networks to propose an epidemic model indicating that according to a scale-free network created by Barabási and Albert (1999), an epidemic threshold tends toward 0 as long as the network is sufficiently large. Based on their model (which I will refer to as the P-V model in this dissertation), Huang and Tsai proposed a modified model containing *resource limitations* and *transmission costs* for analyzing epidemic thresholds (Huang et al., 2010; Tsai and Huang, 2010). We used computer simulations to verify the model, as well as to show its practical applications.

3.1. Motivation

Researchers who take network-oriented approaches to analyzing contagious disease diffusion processes note that the topological features of social networks exert considerable influence on transmission dynamics and spreading situations associated with epidemics (Newman, 2003; Newman & Watts, 1999). Unlike non-network approaches, they support subtle analyses of epidemic dynamics (Pastor-Satorras & Vespignani, 2001a, 2001b, 2002, 2003; Huang et al., 2004, 2005). Researchers of epidemic dynamics and critical thresholds in scale-free networks consistently conclude that regardless of transmission capability, all contagious diseases have high probabilities of stable spreading and survival in scale-free networks (Xu et al., 2007).

According to Pastor-Satorras and Vespignani (2001b), a positive critical transmission threshold does not exist for the spreading of contagious diseases in scale-free social networks. In other words, even contagious diseases with tiny transmission capabilities survive in such networks. Pastor-Satorras and Vespignani's proposed spreading dynamic is expressed as follows:

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k[1-\rho_k(t)]\theta\left[\{\rho_k(t)\}\right] \tag{3.1}$$

where $\rho_k(t)$ is the density of infected nodes with k connections, λ a constant infection rate, and $\theta\left[\{\rho_k(t)\}\right]$ the probability that any given individual will become

linked to an infected individual, with θ assumed to be a function of the partial densities of infected individuals $\{\rho_k(t)\}$. Eq. (3.1) states that during each time step, infected individuals who have *k* connections will recover, yet continue to infect other individuals according to four parameters: infection rate, connectivity, number of healthy individuals, and probability $\theta[\{\rho_k(t)\}]$. Pastor-Satorras and Vespignani defined ρ_k as the steady state of $\rho_k(t)$, and observed that ρ_k is a function of λ in a steady state, therefore θ is a function of λ, such that $\theta(\lambda) = \frac{1}{\langle k \rangle}\sum_k kP(k)\rho_k$, with $P(k)$ representing connectivity distribution. Furthermore, when considering the stationary condition $d\rho_k(t)/dt = 0$ within a scale-free network in which $P(k) = 2m^2k^{-3}$ with minimum degree m, the critical epidemic threshold λ_c has the property $\lambda_c = \langle k \rangle / \langle k^2 \rangle \to 0$ as $k \to \infty$. Accordingly, for infinite size networks, either no epidemic threshold exists, or the threshold approaches 0.

New contagious diseases are constantly emerging in different parts of the world, but very few reach epidemic proportions or even survive in social networks; the majority of diseases die almost immediately following their appearance. This observation serves as our motivation to take a more detailed look at limitations in transmission and interaction processes rather than the topological features of social networks—the focus of many epidemiological studies published in the past decade. Two important factors associated with face-to-face interactions and daily contacts have

been understudied: *resource limitations* and *transmission costs*. The term *resource* in this situation is defined as what is consumed by individuals during the spreading process of a contagious disease. There are five properties associated with resources: (a) they can be visible (e.g., seminal fluid, physical power) or invisible (e.g., time, energy); (b) individual resources are finite and can be temporarily exhausted; (c) the use of one type of resource entails the consumption of smaller quantities of other types of resources, thereby reducing the total available resource amount; (d) individual resources can recover or regenerate after a period of time; and (e) they are non-reproducible. Contagious carriers who apply resources to specific recipients cannot reuse the same resources on other recipients; conversely, recipients cannot reuse resources spent on individual carriers. We acknowledge the importance of Pastor-Satorras and Vespignani's (2001a) work on the topological power-law features of scale-free social networks, especially since their ideas have inspired numerous studies on critical thresholds and immunization strategies. However, such assumptions may be unrealistic and inaccurate when applied to biologically contagious diseases spread via face-to-face interactions and daily contacts. A closer inspection of their mathematical analyses and numerical simulations reveal what we believe to be incorrect assumptions that daily interaction processes are cost-free, and that the impacts of resource limitations and transmission costs are minimal.

3.2. A Contagious Epidemiological Model under Resource Limitations and Transmission Cost Considerations

Our mathematical model is based on the epidemic simulation model shown in Eq. (3.1) as proposed by Pastor-Satorras and Vespignani (2001b). However, this model neglects individual access to energy, time, and other finite resources. Therefore, we propose a model under different infection rate-to-link degree assumptions.

To incorporate individual access to energy, time, and other finite resources, we modified the model to consider resource limitations and transmission costs using two different approaches, as shown in Eq. (3.2).

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda S_k[1-\rho_k(t)]\theta\left[\{\rho_k(t)\}\right], \text{ where } S_k = \min(\frac{R}{c}, k). \quad (3.2)$$

According to the term S_k (with R representing average resources and c transmission costs), the spreading of each infection is proportional to the minimum value of each active node's available resources (R/c) and number of links.

Using the mean field method, we let the stationary condition $d\rho_k(t)/dt = 0$, obtaining

$$\rho_k = \frac{\lambda S_k \theta(\lambda)}{1+\lambda S_k \theta(\lambda)} \quad (3.3)$$

where ρ_k is the steady state of $\rho_k(t)$. Substituting $\theta(\lambda)$ in Eq. (3.3), we get

$$\theta = \frac{1}{\langle k \rangle} \sum_k kP(k) \frac{\lambda S_k \theta}{1 + \lambda S_k \theta} \tag{3.4}$$

Note that the right side of Eq. (3.4) is concave at about θ (i.e., the second derivative is no larger than zero), and that $\theta = 0$ is considered a trivial solution. Since it is possible for θ to have a non-singular solution, we derived the inequality

$$\frac{d}{d\theta}\left(\frac{1}{\langle k \rangle} \sum_k kP(k) \frac{\lambda S_k \theta}{1 + \lambda S_k \theta} \right)\Bigg|_{\theta=0} \geq 1. \tag{3.5}$$

Differentiating Eq. (3.5) and substituting 0 for θ we get

$$\frac{1}{\langle k \rangle} \sum_k kP(k) \lambda S_k \geq 1 \text{ or } \lambda \leq \frac{\langle k \rangle}{\sum_k kP(k) S_k}. \tag{3.6}$$

Accordingly, critical threshold λ_c is defined as the maximal λ, resulting in

$$\lambda_c = \frac{\langle k \rangle}{\sum_k kP(k) S_k} \tag{3.7}$$

Since $S_k = \min(R/c, k)$, the denominator can be divided into two parts, obtaining

$$\lambda_c = \frac{\langle k \rangle}{\sum_{k \leq \frac{R}{c}} k^2 P(k) + \sum_{k > \frac{R}{c}} \frac{R}{c} kP(k)}. \tag{3.8}$$

According to the first term in the Eq. (3.8) denominator, *k* is smaller than *R*/*c*, therefore substituting *R*/*c* for *k* makes the first term larger. Similarly, according to the second term, the summation is smaller than the entire scope of *k*, therefore substituting *k* for the entire scope also makes the second term larger. Thus,

$$\lambda_c \geq \frac{\langle k \rangle}{\sum_{k \leq \frac{R}{c}} \left(\frac{R}{c}\right)^2 P(k) + \sum_k \frac{R}{c} kP(k)} \tag{3.9}$$

Using the same method, another substitution on the left side of the Eq. (3.9) denominator results in

$$\lambda_c \geq \frac{\langle k \rangle}{\sum_k \left(\frac{R}{c}\right)^2 P(k) + \sum_k \frac{R}{c} kP(k)}. \tag{3.10}$$

Since $\sum_k P(k) = 1$, we arrive at

$$\lambda_c \geq \frac{\langle k \rangle}{\left(\frac{R}{c}\right)^2 + \frac{R}{c}\langle k \rangle} = \frac{1}{\left(\frac{\left(\frac{R}{c}\right)^2}{\langle k \rangle}\right) + \frac{R}{c}}. \tag{3.11}$$

and observe that as $\langle k \rangle \to \infty$, λ_c is at minimum equal to c/R.

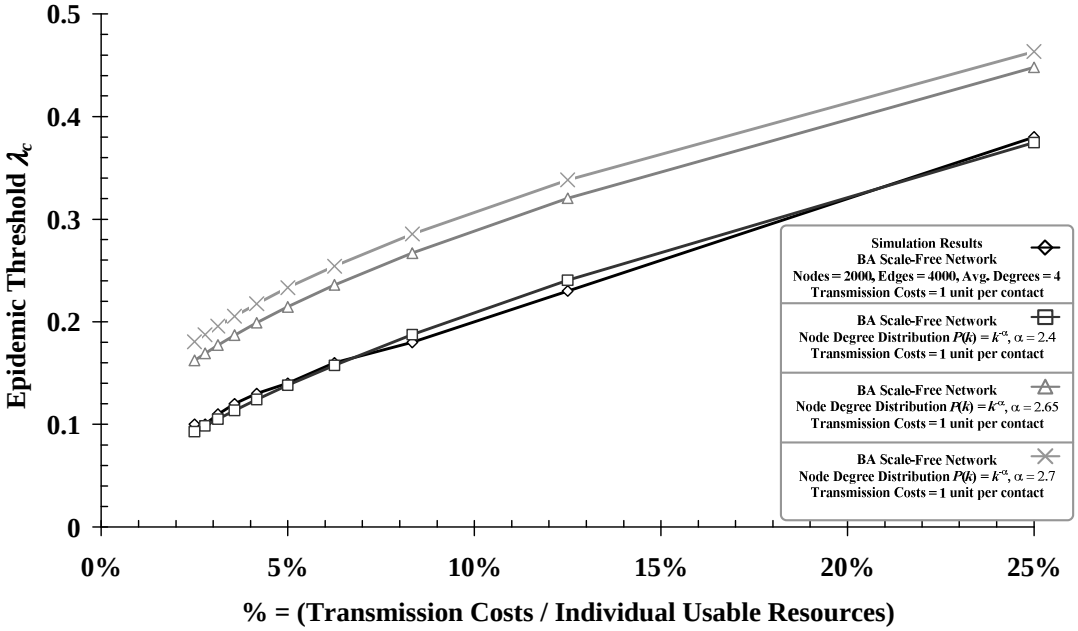


Figure 3.1. Critical Threshold λ_c is a function of the ratio of transmission costs to individual resources (c / R) in scale-free networks. We used it to analyze results from our simulation experiments and three mathematical analyses.

As shown in Figure 3.1, the mathematical results are consistent with the simulation result. The results indicate that when resources and transmission costs are taken into consideration, a significant critical threshold (above which a contagious disease exceeds control and becomes epidemic, and below which a contagious disease disappears) exists when a contagion event occurs in a scale-free network—in short, a non-zero critical threshold exists in scale-free networks. Our results also indicate that the appearance of a critical threshold is tied to a ratio of transmission costs to available resources. In summary, the lower bound of λ_c becomes larger whether transmission cost c increases or the average resource R decreases. Accordingly, an individual's available resources expand when c/R is large, thereby decreasing that individual's ability to contact almost all other personal network nodes. Since λ_c represents the threshold at which a contagious disease exceeds control and become epidemic, managing the value of λ_c is the primary concern of epidemiologists and public health officials. The result supports what we know about immunization: appropriately restricting one's resources increases the critical threshold. Neglecting one's resources makes R infinitely large, meaning we can treat those resources as inexhaustible, and

that the critical threshold λ_c will approach 0 as long as the size of the average number of links is large enough. The model thus becomes identical to Pastor-Satorras and Vespignani's model (Eq. 3.1), in which a disease has the potential to become epidemic even when the number of infected nodes is very small.

3.3. Epidemic Effect of Limited Resources/Transmission Cost Ratio

One scenario to which Eq. (3.11) can be applied is a network attack spread via the Internet—an example of a scale-free network (Barabási et al., 2002). Although spreading time is short, affected areas can be very large, with disastrous results in terms of lost data, work time, and money. One suggested strategy for controlling computer network attacks is placing restrictions on downloads from web services (e.g., a maximum of one gigabyte per day)—in other words, a time resource limitation to raise the outbreak critical threshold λ_c. Another potential strategy is charging downloading fees—that is, raising transmission costs to increase outbreak critical thresholds. The algorithm Barabási and Albert (1999) introduced to build their model (which I will refer to as the BA model in this dissertation) is based on a concept common to networks such as the Internet, the World Wide Web, and social networks—that is, for each node there is a large probability of connecting to other nodes that are already linked to still other nodes. According to the algorithm, we take m_0 disconnected nodes, steadily add new nodes, and connect the new nodes to existing m nodes at a probability of $P(k_i) = \frac{k_i}{\sum_j k_j}$, where k_i is the degree of the i-th node. The algorithm stops and a graph is created when the number of nodes reaches N. In this model it is easy to see that nodes

already having large numbers of links will gain even more—for instance, already popular web pages have high probabilities of becoming even more popular.

We applied this logic to the BA model using a connectivity distribution of $P(k) = 2m^2 k^{-3}$, and an expected number of connecting degrees of $\langle k \rangle = \int_m^\infty kP(k)dk = 2m$. Note that the probability of infected nodes can be expressed as

$$\theta = \frac{1}{2m} \int_m^\infty k 2m^2 k^{-3} \frac{\lambda S_k \theta}{1 + \lambda S_k \theta} dk$$

When k is smaller than R/c, then k can be used as a substitute for S_k ; otherwise, R/c can serve as the substitute. As a result,

$$\theta = m\lambda\theta \int_m^{\frac{R}{c}} \frac{1}{k} \frac{1}{1 + \lambda k \theta} dk + \frac{mR}{c} \frac{\lambda\theta}{1 + \lambda \frac{R}{c} \theta} \int_{\frac{R}{c}}^\infty \frac{1}{k^2} dk \tag{3.12}$$

Eliminating θ from both sides and reducing Eq. (3.12) produces

$$1 = m\lambda \ln \frac{\frac{1}{m} + \lambda\theta}{\frac{c}{R} + \lambda\theta} + \frac{m\lambda}{1 + \lambda \frac{R}{c} \theta} \tag{3.13}$$

Next, the expected density of infected nodes is given by $\rho = \sum_k P(k)\rho_k$. By substituting $P(k) = 2m^2 k^{-3}$ and ρ_k into this equation, we get

$$\rho = \int_m^\infty 2m^2 k^{-3} \frac{\lambda S_k \theta}{1 + \lambda S_k \theta} dk \tag{3.14}$$

Similarly, by dividing Eq. (3.14) and integrating we obtain

$$\rho = 2m^2 \lambda^2 \theta^2 \ln \frac{\frac{mc}{R} + \lambda m\theta}{1 + \lambda m\theta} - 2\frac{c}{R} m^2 \lambda\theta + 2m\lambda\theta + \frac{m^2 c}{R} \frac{\lambda\theta}{1 + \frac{R}{c} \lambda\theta} \tag{3.15}$$

Substituting Eq. (3.13) into Eq. (3.15),

$$\rho = 2m\lambda\theta\left\{\left[\frac{m\lambda\theta + m\frac{c}{2R}}{1+\frac{R}{c}\lambda\theta}\right] - \theta - \frac{mc}{R} + 1\right\} \tag{3.16}$$

Note that whenever $\lambda \le \frac{\theta + \frac{mc}{2R} - 1}{\frac{R}{c}\theta(1-\theta)}$, the term within the braces of Eq. (3.16) is smaller than 0, therefore so is ρ. In other words, the numerator is a criterion for a threshold to occur; if larger, threshold λ will be far from 0. We found that Eq. (3.17) is the upper bound of that epidemiological threshold:

$$\frac{R}{c} < \frac{m}{2(1-\theta)} \tag{3.17}$$

Note that when R/c is smaller than $m/[2\times(1-\theta)]$, the density of ρ infected nodes in the stationary state cannot be larger than 0. The distance of the critical threshold from 0 is decided by each node's available resources. Furthermore, when the infection probability θ or the minimum degree of each node m is fixed, the smaller R/c makes the difference between the two sides of Eq. (3.17) larger, and moves the critical threshold even further from 0. On the other hand, if R/c is fixed, the higher the infection probability θ or the lower the minimum of each degree, the higher the network's critical threshold. In other words, the critical threshold is simply determined by differences among available resources R/c, minimum degree m, and infection probability θ. The spreading of a disease in a BA network can be controlled by efficiently managing these differences.

Chapter 4. Effects of Individual Diversity on Epidemic Modeling in Realistic Social Networks

In this chapter I will describe an innovative simulation framework that combines daily commuting network data with a commonly used population-based transmission model to assess the impacts of various interventions on epidemic dynamics in Taiwan. Called the Multilayer Epidemic Dynamics Simulator (MEDSim), the proposed framework has four contact structures: within age group, between age groups, daily commute, and nationwide interaction. To test model flexibility and generalizability, outbreak locations and intervention scenarios were simulated for the 2009 swine-origin influenza A (H1N1) epidemic. Results indicate that lower transmission rates and earlier intervention activation times did not reduce total numbers of infected cases, but did delay peak times. When transmission rate was decreased by a minimum of 70%, significant epidemic peak delays were observed when interventions were activated before new case number 50; no significant effects were noted when the transmission rate was decreased by less than 30%. Observed peaks occurred more quickly when initial outbreaks took place in urban rather than rural areas. MEDSim apparently

provides insights that reflect the dynamic processes of epidemics under different intervention scenarios, thus clarifying the effects of complex contact structures on disease transmission dynamics.

4.1 Motivation

After emerging in Mexico in April of 2009, the swine-origin H1N1 influenza virus rapidly spread worldwide. In June of that year, the World Health Organization issued its highest possible pandemic alert—level 6 (Smith et al., 2009). Influenza researchers and epidemiologists have generally focused on two spreading factors: age group (determining post-infection symptoms) (Feng et al., 2005; Inaba, 2007; Langlais & Naulin, 2003; Shim et al., 2006; Supriatna et al., 2008; Wang & Zhao, 2005) and adult travel (determining routes by which viruses spread). Since individuals in the same age group tend to have similar epidemic characteristics, age group has been proposed as a distinguishing condition in terms of population compartmentalization (Feng et al., 2005; Inaba, 2007; Langlais & Naulin, 2003; Shim et al., 2006; Supriatna et al., 2008; Yang et al., 2009). Most children and adolescents have better resistance to contagious diseases than individuals age 65 and older. However, the Mexican population segment that was most affected by the H1N1 virus consisted of youth below the age of 15; of all individuals affected by the first infection wave, 61% were children and 29% adults (Fraser et al., 2009). Since novel influenza viruses are known to cause greater morbidity among children (Belshe et al., 1992), the youngest age group served as the main focus of H1N1 intervention efforts.

Many researchers have used age structure to capture heterogeneity when modeling epidemic dynamics (Inaba, 2007; Shim et al., 2006; Wang & Zhao, 2005), with some integrating compartmental models consisting of different age groups to identify the potential impacts of specific populations and temporal epidemic trends (Wang & Zhao, 2005). Childhood diseases such as rotavirus infections have been used to assess the efficacy potential of various vaccination strategies (Shim et al., 2006), and transmission threshold and stability have been the focuses of epidemic simulations involving specific age structures (Inaba, 2007).

Another important factor in modeling epidemic dynamics is population movement. Over the past three decades Taiwan has experienced a rapid increase in the number of commuters for work and other purposes, especially among young adults (Marsh, 1996)—a phenomenon perceived as supporting the spread of viruses over long distances within the country (Yang et al., 2009). Commuting is marked by strong spatial-temporal regularity: regardless of travel distance or time, most commuters follow simple and repetitive patterns (González et al., 2008). These patterns are receiving considerable attention from researchers studying the spreading dynamics of diseases and viruses (Huang et al., 2005), the clustering characteristics of epidemic diseases at the beginning of an outbreak (Riley, 2007; Tomlinson & Cockram, 2003), and the targeting of vaccinations, quarantining, and other public health policies

(Epstein, 2009; Keeling & Rohani, 2007; Liu, Takeuchi & Iwami, 2008; Pastor-Satorras & Vespignani, 2002).

A multi-layer simulation framework that combines daily commuting networks and a commonly used population-based transmission model for simulating epidemic dynamics was proposed, using the 2008-2009 seasonal influenza A and 2009 swine-origin influenza A (H1N1) outbreaks to estimate model parameters. The potential impacts of different outbreak locations and interventions on the Taiwan-wide epidemic dynamics of swine-origin influenza A were assessed, including intervention timing and different combinations of public health responses.

4.2 A Multilayer Epidemiological Model Integrating Human Commuting Networks

To analyze the spreading dynamics of epidemic diseases in detail, a top-down simulation framework was established, along with a prototype of the Multilayer Epidemic Dynamics Simulator (MEDSim), which integrates population-based and network-oriented approaches to capturing complex demographic, geographic, and biological properties, including human movement patterns and disease progression (Figure 4.1). Based on the observation that epidemic dynamics in large populations are similar to those found in deterministic systems (Keeling & Rohani, 2007), a deterministic framework was established for the MEDSim model. As shown in Figure 4.2, layer 1 individuals within the same location are organized according to age group; a population-based approach was used to model the transmission dynamics of each group. The layer 2 focus is on contact patterns and interactions between different age groups within the same location. The effects of regional interactions on human travel networks are added to layer 3 by incorporating population density and commuting volume between any two locations. In layer 4, a network-oriented approach was used to incorporate a geographic information system (GIS) for constructing human travel networks on a national scale, with nodes representing locations on commuting routes,

and links representing movement between them.

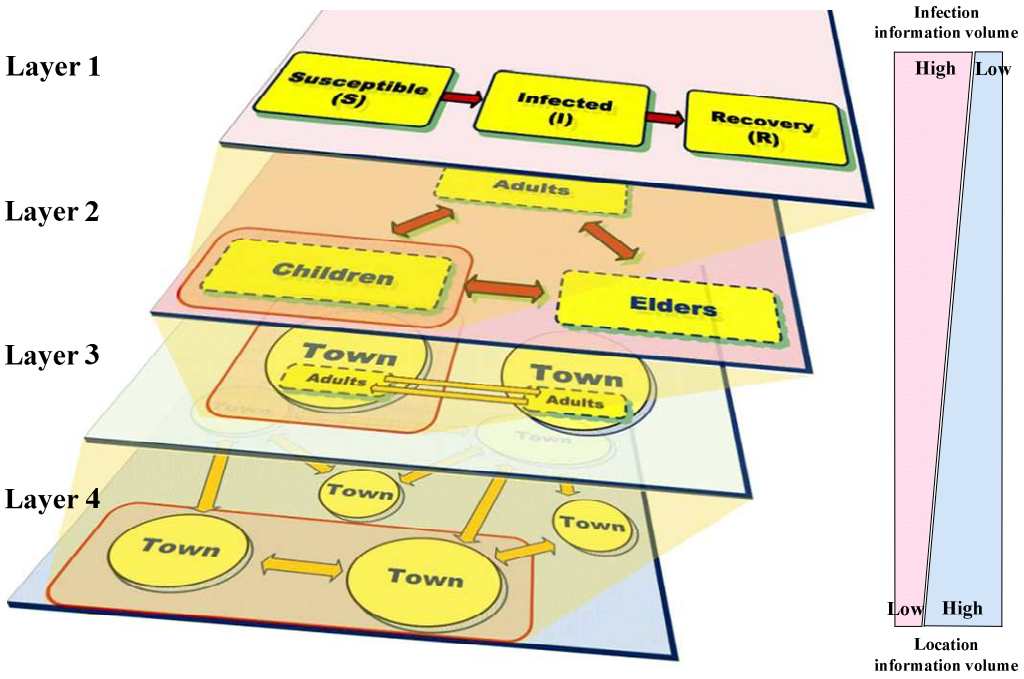
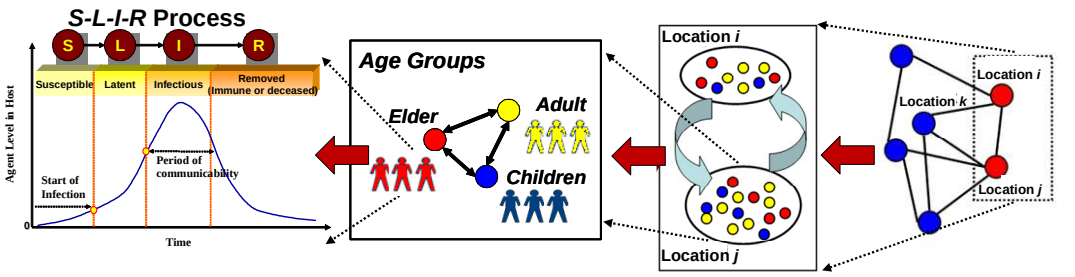


Figure 4.1. Multilayer Epidemic Dynamics Simulator (MEDSim) concept. Infection information usage is highest in Layer 1 and lowest in Layer 4, the opposite of location information.

Due to its ability to comprehensively integrate multilayer structures to generate dynamic spatial and temporal processes, Mathworks MATLAB was used to implement the MEDSim framework as a numerical computation kernel. By using Microsoft Excel to organize census and transportation data, policy makers, health professionals, and others who have less experience with specialized computer software will be able to generate simulation scenarios with minimal assistance.



Layer 1: **Contact between individuals in same age groups.**

Layer 2: **Contact between individuals in different age groups in the same location.**

Layer 3: **Contacts between locations.**

Layer 4: **Nationwide commuting network.**

Figure 4.2. MEDSim framework.

Layer 1: Within an age group

A four-state SLIR epidemiological model was used to represent different infection stages among individuals in the same age group in the same location. Individual epidemic status was initially set at Susceptible, followed by *L*atent, *I*nfectious, and *R*emoved. The numbers of pathogens that Susceptible-to-Latent hosts carry are insufficient for active transmission to other Susceptible hosts, but these numbers eventually reach levels where hosts become Infectious, begin to infect other Susceptible hosts, and eventually move toward a Removed status. The dynamics of the four epidemic states over time are expressed as Eqs. 4.1a-d, which have the following features:

1. At time t, the population of interest is divided into four compartments ($S(t)$, $L(t)$, $I(t)$ and $R(t)$) that correspond to the SLIR model's four epidemic states. Since the SLIR model is a closed system, $S(t)+L(t)+I(t)+R(t)=N$, with N a constant representing the entire population.

2. Transmission rate β is a constant representing how fast Susceptible individuals become infected and acquire a Latent status.

3. Latent rate θ is a constant used to determine transformation speed from Latent to Infected.

4. Removed rate α is a constant used to determine transformation speed from Infected to Recovered.

Ordinary differential equations can be used to express the SLIR model as

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) / N \quad (4.1a)$$

$$\frac{dL(t)}{dt} = -\theta L(t) + \beta S(t)I(t) / N \quad (4.1b)$$

$$\frac{dI(t)}{dt} = -\alpha I(t) + \theta L(t) \quad (4.1c)$$

$$\frac{dR(t)}{dt} = \alpha I(t) \quad (4.1d)$$

Figure 4.3a and b illustrate the concept and flowchart of the model's first layer, respectively. Note the modification to take self-motivated hospitalization into consideration (i.e., those individuals who visit hospitals or clinics during an influenza outbreak regardless of their infection status). Depending on diagnostic accuracy, some are confirmed as infectious and receive medical treatment in advance, thus altering transmission and removed rates for certain populations. To integrate this factor into the model, an additional three features are proposed: (a) an investigation constant s representing the percentage of a population that goes to a hospital or clinic in advance

of becoming ill; (b) a detection constant c, used to determine the percentage of a population confirmed as infectious; and (c) the time delay constant T, used to represent time between a patient with symptoms visiting a hospital or clinic and the time that his or her infection is confirmed.

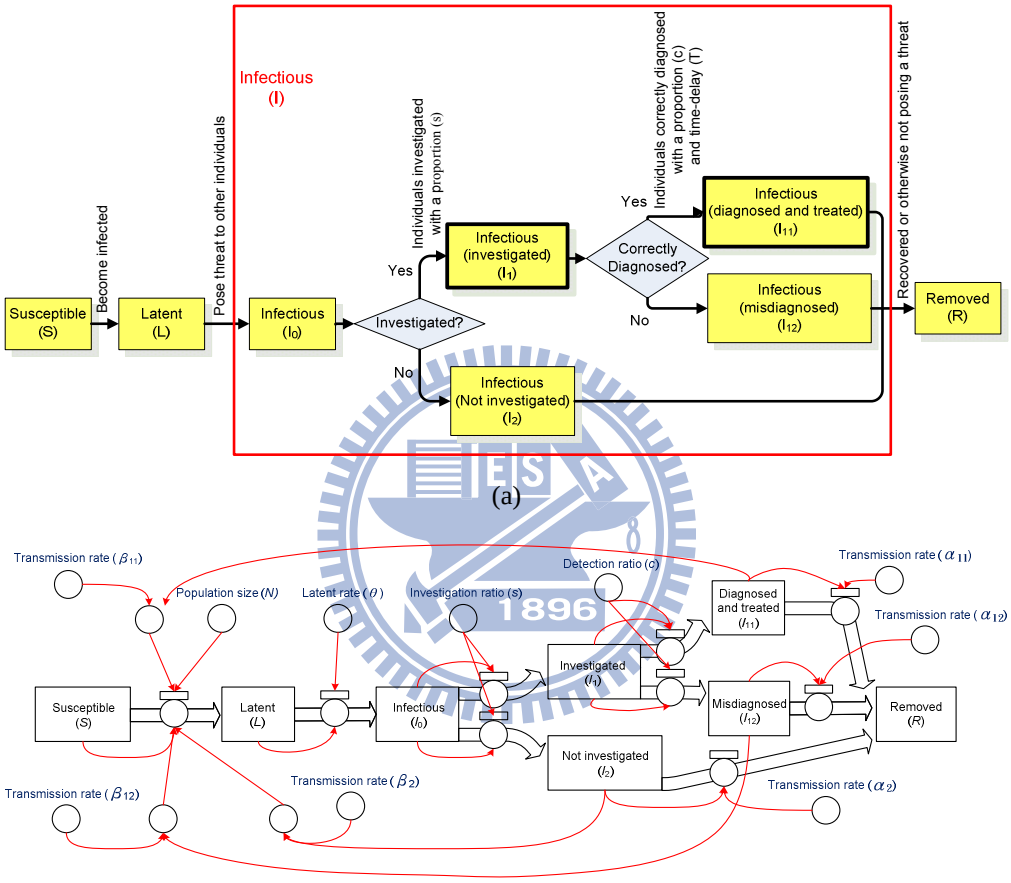


Figure 4.3. (a) Modified SLIR model layer 1 concept. (b) Modified SLIR model layer 1 flowchart.

Also in consideration of preventive health care actions among individuals, a feature was added in which individuals with an L status are moved to either an I_1 or I_2 status according to whether or not they visit a hospital or clinic (expressed as investigation proportion s). I_1 individuals are identified as I_{11} or I_{12} depending on

whether or not they are correctly diagnosed as infectious (expressed as detection proportion c). Note that regardless of positive or negative diagnosis, a period of time T must elapse prior to confirmation. The difference between state I_{11} and either I_2 or I_{12} is the transmission rate. I_{11}, I_{12} and I_2 all eventually change to state R.

This extended SLIR model is expressed as

$$\frac{dS(t)}{dt} = -S(t)(\beta_2 I_2(t) + \beta_{11} I_{11}(t) + \beta_{12} I_{12}(t)) / N \quad (4.2a)$$

$$\frac{dL(t)}{dt} = -\theta L(t) + S(t)(\beta_2 I_2(t) + \beta_{11} I_{11}(t) + \beta_{12} I_{12}(t)) / N \quad (4.2b)$$

$$\frac{dI_0(t)}{dt} = -I_0(t) + \theta L(t) \quad (4.2c)$$

$$\frac{dI_1(t)}{dt} = -I_1(t - T) + sI_0(t) \quad (4.2d)$$

$$\frac{dI_2(t)}{dt} = -\alpha_2 I_2(t) + (1 - s)I_0(t) \quad (4.2e)$$

$$\frac{dI_{11}(t)}{dt} = -\alpha_{11} I_{11}(t) + cI_1(t - T) \quad (4.2f)$$

$$\frac{dI_{12}(t)}{dt} = -\alpha_{12} I_{12}(t) + (1 - c)I_1(t - T) \quad (4.2g)$$

$$\frac{dR(t)}{dt} = \alpha_{11} I_{11}(t) + \alpha_{12} I_{12}(t) + \alpha_2 I_2(t) \quad (4.2h)$$

Layer 2: Among age groups

Individuals in different age groups have different infection properties, expressed in terms of epidemic parameters such as transmission and removed rates. Two age-related features were considered: (a) the transmission rates β_{11pq}, β_{12pq}, and β_{2pq}, which represent cross-age group infections; and (b) the relative percentage χ_p of age level, which affects the potential for cross-age infections. To distinguish among

parameters for individuals in different age groups, a subscript was added to each parameter in Eqs. 4.2a-h (with the exception of T)—for example, parameter $S(t)$ was changed to $S_p(t)$ for age level p. Three age levels were assumed when analyzing H1N1: children (birth to 14), adults (15 to 64), and seniors (65 and older). To capture the complexity of infections across age groups, transmission rates between different age levels were differentiated. Two transmission rate subscripts were added ($\mu_{p,q}$), the first for the age of an infectious individual, and the second for the age of the individual being infected (Figure 4.4). Epidemic parameters used in population-based compartmental models were also used to model infections across age groups. Three transmission rates and three removed rates were used, based on the number of individuals seeking medical attention.

To construct the layer 2 model, Eqs. 4.2a and 4.2b were respectively revised to 4.3a and 4.3b, without making any other changes to the Eq. (4.2) sub-equations.

$$\begin{aligned}\frac{dS_p}{dt} = & -S_p\chi_p\chi_p(\beta_{2pp}I_{2p} + \beta_{11pp}I_{11p} + \beta_{12pp}I_{12p}) / N_p \\ & - S_p\sum_{q\neq p}\chi_q\chi_p(\beta_{2qp}I_{2p} + \beta_{11qp}I_{11p} + \beta_{12qp}I_{12p}) / N_q\end{aligned} \tag{4.3a}$$

$$\begin{aligned}\frac{dL_p}{dt} = & -\theta_p L_p + S_p\chi_p\chi_p(\beta_{2pp}I_{2p} + \beta_{11pp}I_{11p} + \beta_{12pp}I_{12p}) / N_p \\ & + S_p\sum_{q\neq p}\chi_q\chi_p(\beta_{2qp}I_{2p} + \beta_{11qp}I_{11p} + \beta_{12qp}I_{12p}) / N_q\end{aligned} \tag{4.3b}$$

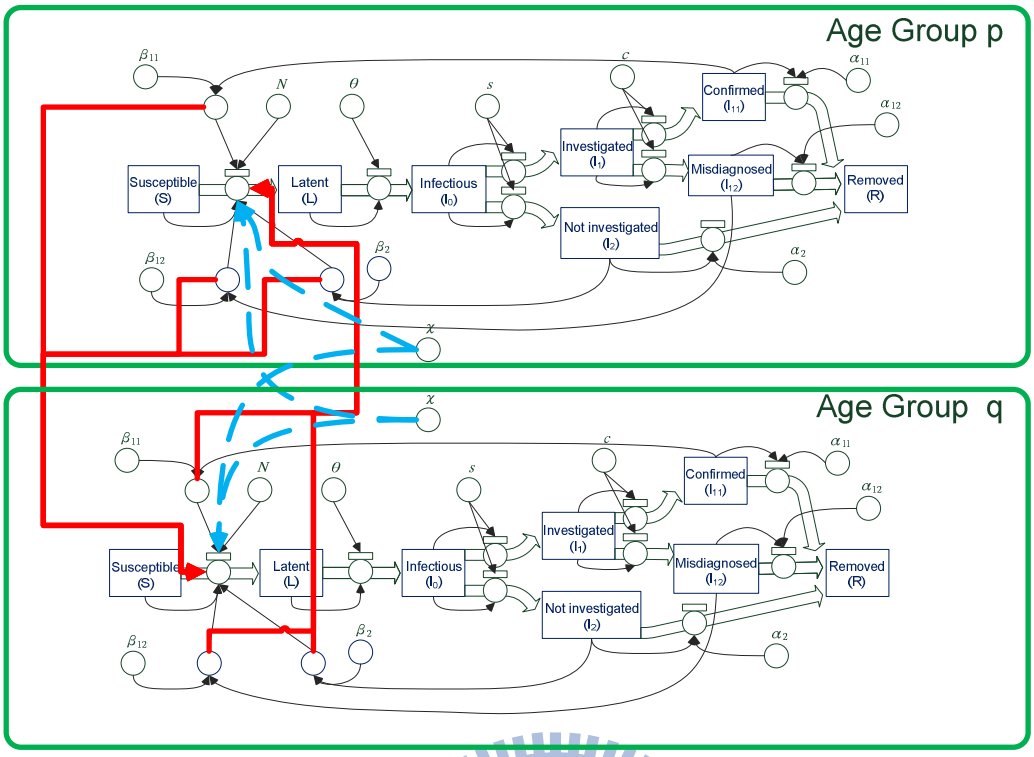


Figure 4.4. MEDSim layer 2 architecture flowchart. Thick solid lines indicate parameters for other (non-p and non-q) age groups. Thick dashed curves indicate relative percentages of each age group within the total population of each location.

Layer 3: Commuting

This study focused on the impacts of daily commuting networks on the spreading of an influenza virus. Since influenza viruses are transmitted via air-borne droplets, commuter hosts are capable of infecting other individuals along their standard routes. The layer 3 model reflects two assumptions regarding hosts with jobs: they commute over longer distances than individuals who stay at home or travel to local centers such as schools, and they tend to come into contact with individuals in the same age group along their routes and at their destinations. Higher contact frequencies among

individuals in more densely populated areas were assumed. Accordingly, the layer 3 model considers four features associated with travel between population centers (locations):

1. $\sigma(p)$, a binary value representing whether age level p is the commuter age level—that is, $\sigma(p) = \begin{cases} 1 & \text{if } p = \text{commutable age level} \\ 0 & \text{otherwise} \end{cases}$. It was assumed that children and seniors are less likely than adults to commute on a daily basis, making adults the most likely carriers of pathogens between locations.

2. $w_{j,i}$, indicating how many individuals commute from location *j* to location *i* on a daily basis.

3. η^i, a weighting factor representing average contacts among individuals in location *i* on a daily basis.

4. d^i, a normalized population density value for location *i*.

For all i and j locations in a w commuting network, the following geodemographic weight was used to measure the effects of commuting on *i* and *j* population interactions.

$$\text{Geodemographic weight}(j,i) = \sigma(p) d^i \eta^i \frac{W_{j,i}}{\sum_{k \neq j} W_{j,k}}$$

As shown in Figure 4.5, the commuting population age level in this example is 15-64 years. The $\sigma(p)$ function represents whether age group *p* is a traveling population. For all i locations in the commuting network, the term $N(i)$ represents the set of

locations connected to location i within commuting network W. The term $w_{j,i} / \sum_{k \neq j} w_{j,k}$ is the ratio of commuters between locations j and i to commuters between location j and all other locations. If location i is a large urban center, $w_{j,i} / \sum_{k \neq j} w_{j,k}$ will be large; if i is a suburb or rural location, it will be small. Public health policies involving transportation can be tested by changing contact rates among population centers in the layer 3 model.

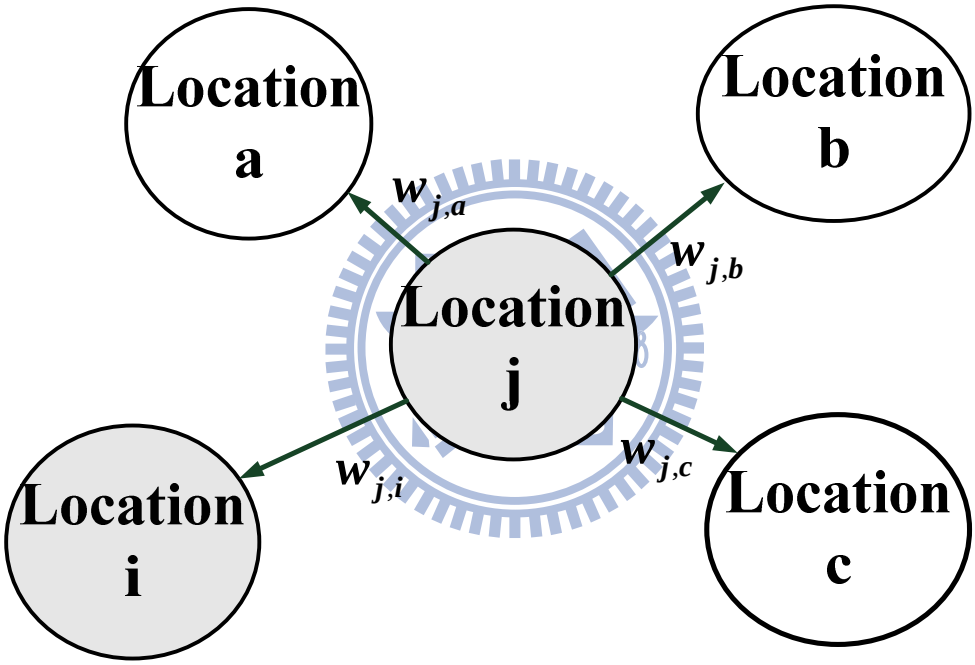


Figure 4.5. Potential movement of infectivity between locations i and j.

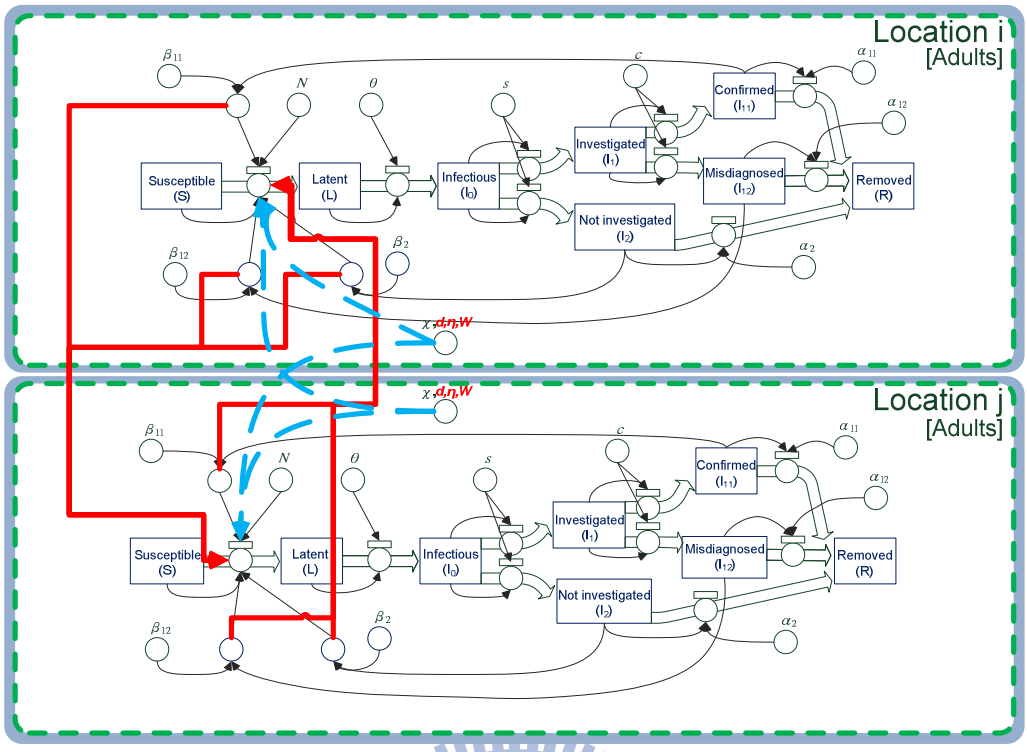


Figure 4.6. MEDSim layer 3 architecture flowchart. Properties associated with commuting between two locations are indicated by thick solid lines. Additional location properties are indicated by thick dashed lines.

The layer 3 framework is presented in Figure 4.6. To construct the layer 3 model, Eqs. 4.3a and 4.3b were respectively revised to Eqs. 4.4a and 4.4b. Note the addition of the geodemographic weight on the third line of each equation. All other Eq. 2 sub-equations are the same.

$$\begin{aligned}\frac{dS_p^i}{dt} &= -S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i) / N_p^i \\ &- S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2p}^i + \beta_{11qp}^i I_{11p}^i + \beta_{12qp}^i I_{12p}^i) / N_q^i \\ &- \sum_{j \in N(i)} \sigma(p) d^i \eta^i S_p^i \frac{W_{j,i}}{\sum_{k \neq j} W_{j,k}} (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j) / N_p^j\end{aligned} \tag{4.4a}$$

$$\frac{dL_p^i}{dt} = -\theta_p^i L_p^i + S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i) / N_p^i$$
$$+ S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i) / N_q^i \quad (4.4b)$$
$$+ \sum_{j \in N(i)} \sigma(p) d^i \eta^i S_p^i \frac{W_{j,i}}{\sum_{k \neq j} W_{j,k}} (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j) / N_p^j$$

Layer 4: Nationwide interactions

Taiwan's national travel network and the commuting weight $w_{j,i}$ were used to simulate individual movement within regions (layer 3). Nodes represent locations, and edges represent commuting weights between locations. Once transportation data are obtained, nodes can represent any scale—for instance, a building for city simulations and a town for regional or national simulations. In this research, each node represents an individual town. Layer 4 of the model consists of 409 towns and 19,014 links (Figure 4.7) representing Taiwan's national commuting network, which can be manipulated to determine the effects of various movement policies and commuting restrictions.

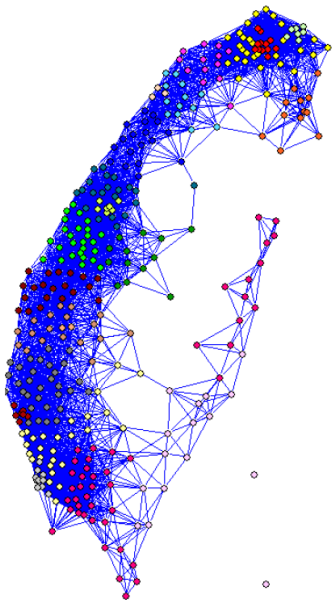


Figure 4.7. Taiwan's nationwide commuting network.

After combining the four layers, the complete MEDSim model is expressed as

$$\begin{aligned}\frac{dS_p^i}{dt} &= -S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i) / N_p^i \\ &- S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i) / N_q^i \\ &- \sigma(p) d^i \eta^i S_p^i \sum_{j \in N(i)} \frac{W_{j,i}}{\sum_{k \neq j} W_{j,k}} (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j) / N_p^j \end{aligned} \tag{4.5a}$$

$$\begin{aligned}\frac{dL_p^i}{dt} &= -\theta_p^i L_p^i + S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i) / N_p^i \\ &+ S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i) / N_q^i \\ &+ \sigma(p) d^i \eta^i S_p^i \sum_{j \in N(i)} \frac{W_{j,i}}{\sum_{k \neq j} W_{j,k}} (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j) / N_p^j \end{aligned} \tag{4.5b}$$

$$\frac{dI_{0p}^i}{dt} = -I_{0p}^i + \theta_p^i L_p^i \tag{4.5c}$$

$$\frac{dI_{1p}^i}{dt} = -I_{1p}^i (t - T) + s_p^i I_{0p}^i \tag{4.5d}$$

$$\frac{dI_{2p}^i}{dt} = -\alpha_{2p}^i I_{2p}^i + (1 - s_p^i) I_{0p}^i \tag{4.5e}$$

$$\frac{dI_{11p}^{i}}{dt} = -\alpha_{11p}^{i} I_{11p}^{i} + c_{p}^{i} I_{1p}^{i}(t-T) \quad (4.5f)$$

$$\frac{dI_{12p}^{i}}{dt} = -\alpha_{12p}^{i} I_{12p}^{i} + (1-c_{p}^{i}) I_{1p}^{i}(t-T) \quad (4.5g)$$

$$\frac{dR_{p}^{i}}{dt} = \alpha_{11p}^{i} I_{11p}^{i} + \alpha_{12p}^{i} I_{12p}^{i} + \alpha_{2p}^{i} I_{2p}^{i} \quad (4.5h)$$

The parameters used in Eqs. 4.5a-h are listed in Table 4.1.

Table 4.1. MEDSim Parameters

Category	Layer	Attribute	Symbol	Description
Epidemic	**1**	*Transmission rate*	β_{11pp}^{i}	Transmission rate from investigated/diagnosed/treated age group *p* to same age group in town *i*
			β_{12pp}^{i}	Transmission rate from investigated/misdiagnosed age group *p* to same age group in town *i*
			β_{2pp}^{i}	Transmission rate from non-investigated age group *p* to same age group in town *i*
		Latent rate	θ_{p}^{i}	Latent rate of age group *p* in town *i*
		Removed rate	α_{11p}^{i}	Removed rate of investigated/diagnosed/treated age group *p* in town *i*
			α_{12p}^{i}	Removed rate of misdiagnosed age group *p* in town *i*
			α_{2p}^{i}	Removed rate of non-investigated age group *p* in town *i*
		Investigation ratio	s_{p}^{i}	Investigated proportion of age group *p* in town *i* (Default: 0.6)
		Detection ratio	c_{p}^{i}	Correctly diagnosed proportion of age group *p* in town *i* (Default: 0.6)
		Delay time	T	Time between investigation and correct diagnosis (Default: 3)
	2	*Transmission rate*	β_{11xy}^{i}	Transmission rate from investigated/diagnosed/treated age group *p* and same age group *q* in town *i*
			β_{12xy}^{i}	Transmission rate from misdiagnosed age-group *p* and same age group *q* in town *i*
			β_{2xy}^{i}	Transmission rate from non-investigated age group *p* and same age group *q* in town *i*
Location	**2**	*Relative percentage*	χ_{p}^{i}	Age group *p* as a percentage of town *i* population. (Source: ROC Interior Ministry)
	3	*Determination*	$\sigma(p)$	Binary value for commutable age level (Default: adult)
		Relative density	d^{i}	Population of town *i* as a percentage of the largest town's population (Source: ROC Interior Ministry)
		Commuting weight	$W_{j,i}$	Number of commuters from town *j* to town *i* (Source: ROC Institute of Transportation)
	4	*Intercity rate*	η^{i}	Average number of daily contacts among individuals in location *i* (Default: 0.8)

Technological Framework

Figure 4.8 shows the MEDSim technological framework, including a simulation flowchart, census databases, and the relationship between the four MEDSim layers and the databases. The first step is to manually create an Excel data set for the user's chosen scenario—for example, determining breakout locations or public health policies. Most data sets consist of spatial locations and census information, which are used to establish geographic and demographic categories; each MEDSim parameter belongs to at least one of these. Since the layer 1 focus in this example is on disease progression on an individual level, standard expert-based parameters in compartmental models associated with epidemics were used instead of transportation or census databases (Keeling & Rohani, 2007). In layer 2, percentages of individuals in each age group were determined from census data, and the numbers of individuals in each location were gathered from transportation databases. In layer 3, transportation databases were used to gather information on the numbers of individuals traveling between towns on a daily basis. In layer 4, transportation data were used to establish the underlying national travel network.

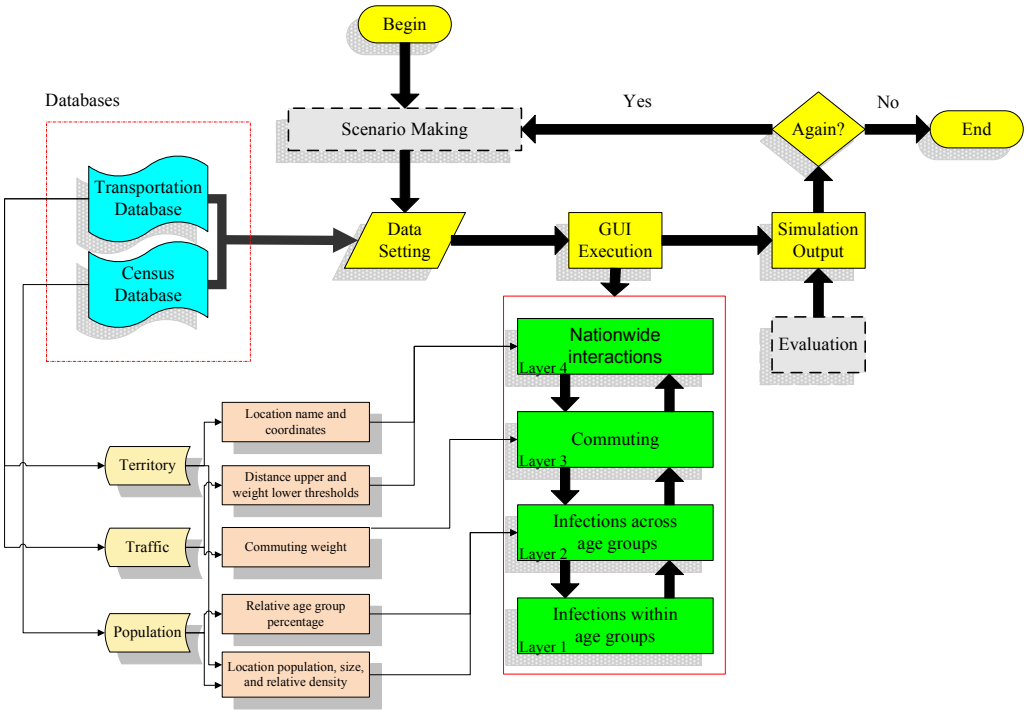


Figure 4.8. MEDSim simulation tool framework.

Figure 4.9 presents a screenshot of a MEDSim GUI. Multilayer epidemic model parameters are initialized at the beginning of each simulation. Model parameters requiring setup are (a) initial outbreak conditions, including the name of the town and number of infected persons in an age group identified by the surveillance system; (b) disease transmission parameters at different layers, including transmission, latent, and removed rates according to the SLIR process for each age group, contact rates between age groups, and regional contact probabilities between towns; and (c) output maps and charts of towns of interest and severity indicators to be monitored (e.g., daily infected cases, daily new cases, and epidemic velocity and acceleration). Daily epidemic progress can be monitored in terms of sizes and locations of red circles on maps (infected individuals), epidemic curves on time charts, and output panels (numbers of

infected individuals at different times in different locations). Regarding kernel execution, MEDSim models can be used for computing epidemic dynamics. Simulation results can be shown as graphical curves, or expressed and recorded as numerical files. Last, simulation results are evaluated by users, who can repeat steps as required.

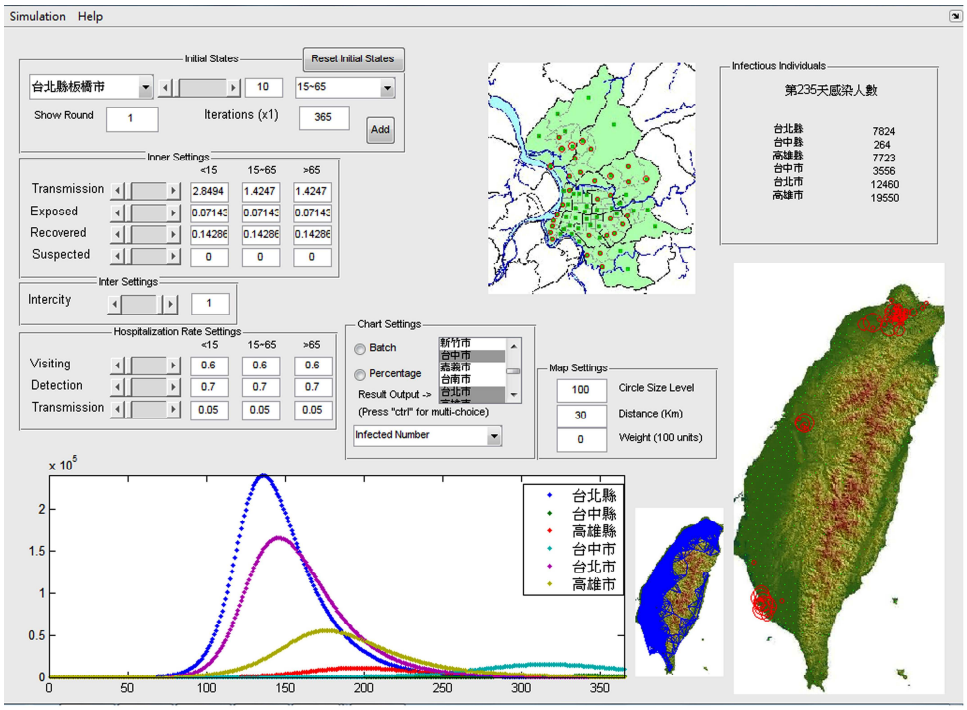


Figure 4.9. MEDSim implementation GUI.

Statistical Analysis for Model Validation

Two indices for comparing simulated and actual numbers of infected individuals were used to test the reliability and validity of time-series MEDSim data: correlation coefficient (CC) and coefficient of efficiency (CE), respectively expressed as Eqs. (4.6)

and (4.7) (Huang et al., 2004). $\{X_t \mid t = 1, 2, ..., n\}$ represents the number of actual infected individuals, and $\{Y_t \mid t = 1, 2, ..., n\}$ the number calculated by MEDSim. In both sets, t denotes time step in 1-week units; a total of n weeks is represented by each set. $\overline{X}$ and $\overline{Y}$ denote the means of X_t and Y_t, respectively. The CC test measures data distance: higher positive values indicate positive correlations, and lower negative values indicate negative correlations. The CE test is used to measure the level of accuracy between two data sets; higher values indicate greater accuracy.

$$CC = \frac{\sum_{t=1}^{n}(X_t - \overline{X})(Y_t - \overline{Y})}{\sqrt{\sum_{t=1}^{n}(X_t - \overline{X})^2}\sqrt{\sum_{t=1}^{n}(Y_t - \overline{Y})^2}} \in [-1,1] \tag{4.6}$$

$$CE = 1 - \left[\frac{\sqrt{\sum_{t=1}^{n}(X_t - Y_t)^2}}{\sqrt{\sum_{t=1}^{n}(X_t - \overline{X})^2}}\right] \in [0,1] \tag{4.7}$$

4.3 Simulating the 2009 Novel H1N1 Influenza

MEDSim reliability (in terms of parameter calibration and model fit) was tested using actual epidemic curves. Public health policies were tested and compared based on the above parameters, and MEDSim was used to simulate the influenza A (H1N1) virus and to determine the effects of various policies. Population data from the Republic of China (ROC) Ministry of the Interior and transportation data from the ROC Transportation Institute (2001) were used to establish simulation parameter settings.

Parameterization

The seasonal influenza A and swine-origin influenza A (H1N1) viruses were used to perform parameterization. Default parameter values are shown in Table 4.1. Parameters for both viruses were systematically calibrated to create a small range, based on parameters normally used with standard SLIR settings (Keeling & Rohani, 2007). Summaries of MEDSim attribute settings and values are given in Table 4.2 and Table 4.3. The transmission rates β^{i}_{11pp}, β^{i}_{12pp} and β^{i}_{2pp} were directional between age groups. Individual age group targets are presented in the form of sub-columns. Experimental results from applying MEDSim using the Table 4.2 and Table 4.3 parameter values for the two influenza viruses are shown in Figures 10a and b, respectively. Actual and simulated case data for both influenzas are shown in weekly

units.

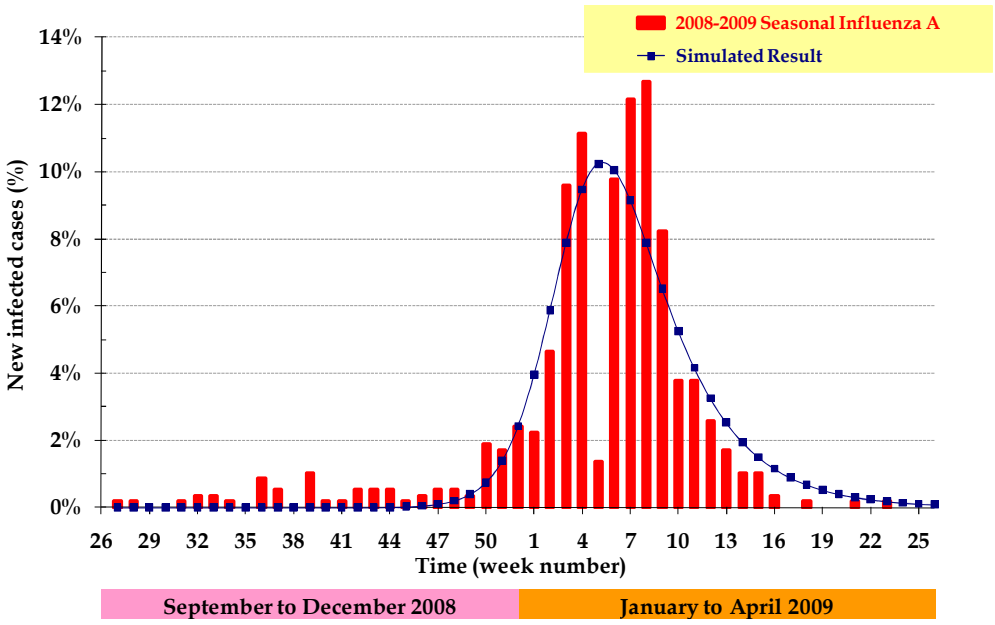
The CC and CE results for the two influenza epidemics are 0.86 and 0.74 for seasonal, and 0.77 and 0.36 for swine-origin H1N1, respectively. Figure 4.10a shows the plotting of fractions of new infected cases of seasonal influenza A in Taiwan between September 2008 and April 2009, normalized to total cases. Higher CC and CE values for seasonal influenza explain the similarities between the two curves. Figure 4.10b shows the plotting of fractions of new infected cases for the swine-origin influenza A virus in Taiwan from week 25 to week 52 in 2009, also normalized to total cases. As shown, the number of actual cases decreased between weeks 37 and 48, followed by an increasing trend, resulting in a lower CE value. This two-wave pattern is very similar to global diffusion patterns associated with international travel. Since international travel was not incorporated at this stage, the swine-origin H1N1 model failed to capture the second wave; however, it did capture the peak time for the first (primary) wave (Figure 4.10b).

Table 4.2. MEDSim parameters used for fitting simulation curves with actual seasonal influenza A curves in Taiwan between September 2008 and April 2009

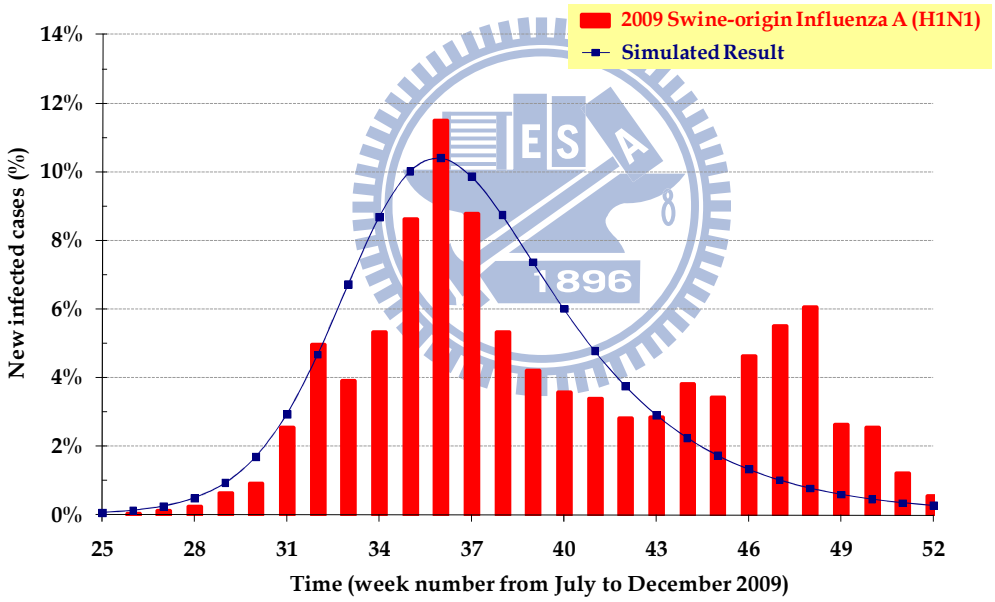
Layer	Attribute	Value					
		Children		Adults		Seniors	
1	β^i_{11pp}	1.3333		0.6667		0.6667	
	β^i_{12pp}	3.3333		1.6667		1.6667	
	β^i_{2pp}	3.3333		1.6667		1.6667	
	θ^i_p	0.0714					
	α^i_{11p}	0.1429					
	α^i_{12p}	0.2500					
	α^i_{2p}	0.1429					
2	Target	Adults	Seniors	Children	Seniors	Children	Adults
	β^i_{11xy}	0.6667	0.6667	1.3333	0.6667	1.3333	0.6667
	β^i_{12xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667
	β^i_{2xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667

Table 4.3. MEDSim parameters used for fitting simulation curves to actual swine-origin influenza A (H1N1) curves in Taiwan from week 25 to week 52

Layer	Attribute	Value					
		Children		Adults		Seniors	
1	β^i_{11pp}	2.6667		1.3333		1.3333	
	β^i_{12pp}	3.3333		1.6667		1.6667	
	β^i_{2pp}	3.3333		1.6667		1.6667	
	θ^i_p	0.0714					
	α^i_{11p}	0.3333					
	α^i_{12p}	0.1429					
	α^i_{2p}	0.1667					
2	Target	Adults	Seniors	Children	Seniors	Children	Adults
	β^i_{11xy}	1.3333	1.3333	2.6667	1.3333	2.6667	1.3333
	β^i_{12xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667
	β^i_{2xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667



(a)



(b)

Figure 4.10. Comparison of weekly new infected cases between actual and simulated results normalized for (a) seasonal influenza A and (b) swine-origin H1N1 influenza A.

Intervention Policy Evaluation

Different public health policies were tested and compared using the above-described parameters. In addition, the effects of medical advice quality and

number of commuters were simulated, and original epidemic curves were compared with those following the implementation of the public health policies. Special emphases were placed on peak numbers of infected cases and peak infection days, since the goals of public health officials are to reduce the peak number (since it has a direct effect on social costs such as drugs and hospital beds), and to delay the peak day.

Figure 4.11 has two parts, one addressing the impacts of transmission rate reduction, and one on the effects of various intervention policies. Results from simulated observation indices for different transmission rates are shown in Table 4.4 . According to the Figure 4.11 data for weekly fractions of new infected cases, both curve peak and height were negatively affected by decreased transmission rates. According to the Figure 4.12 data on the cumulative number of new infections at different transmission rates, that number decreased as transmission rate decreased. In Figure 4.13, two observation indices were used to distinguish between the epidemic curve produced by the highest transmission rate, and the curves shown in Figure 4.11. According to the first observation index (fraction of new infected cases at epidemic curve peak), the largest transmission intensity of an epidemic disease infects a population over a period of one week, which affects the quantity of available public health resources; a decreasing value reduces the burden on those resources. The second index (epidemic curve peak week number) indicates the severity and urgency of an

epidemic, which affects deadlines for initiating public health policies; a higher value indicates more time for making policy decisions.

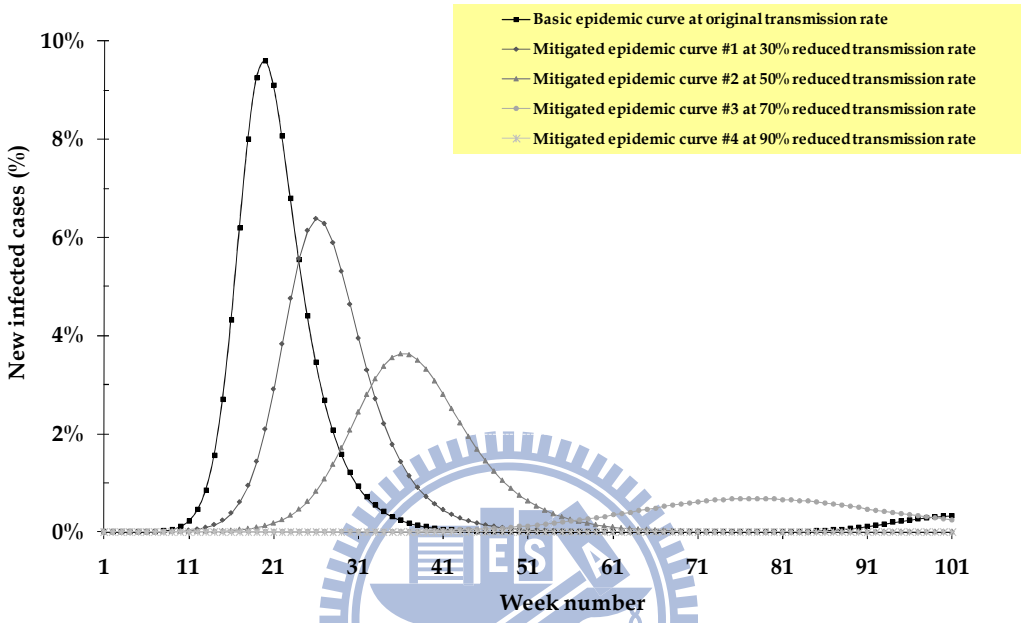


Figure 4.11. New infected cases per week at different transmission rates.

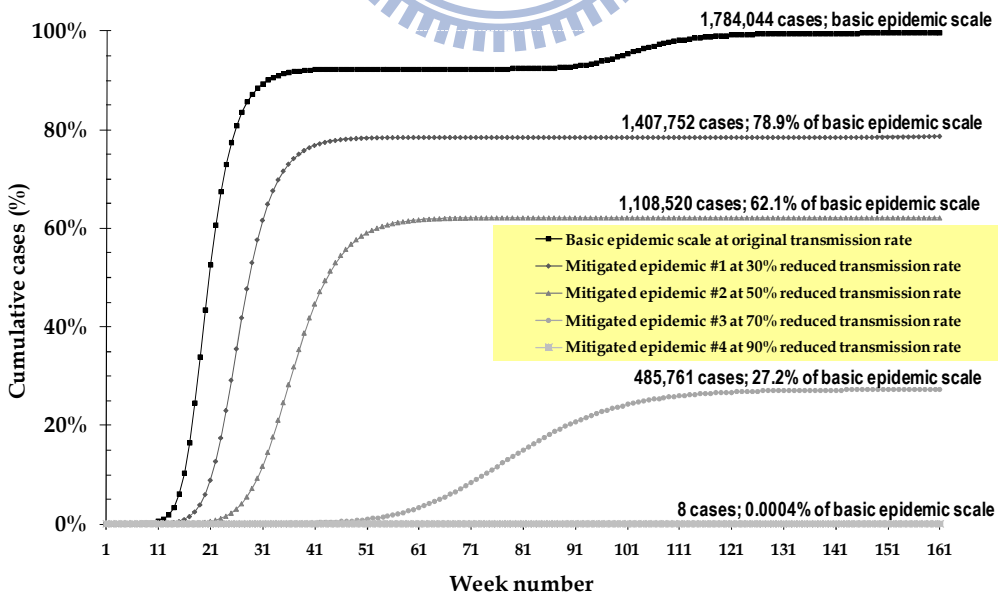


Figure 4.12. Cumulative new infected cases at different transmission rates.

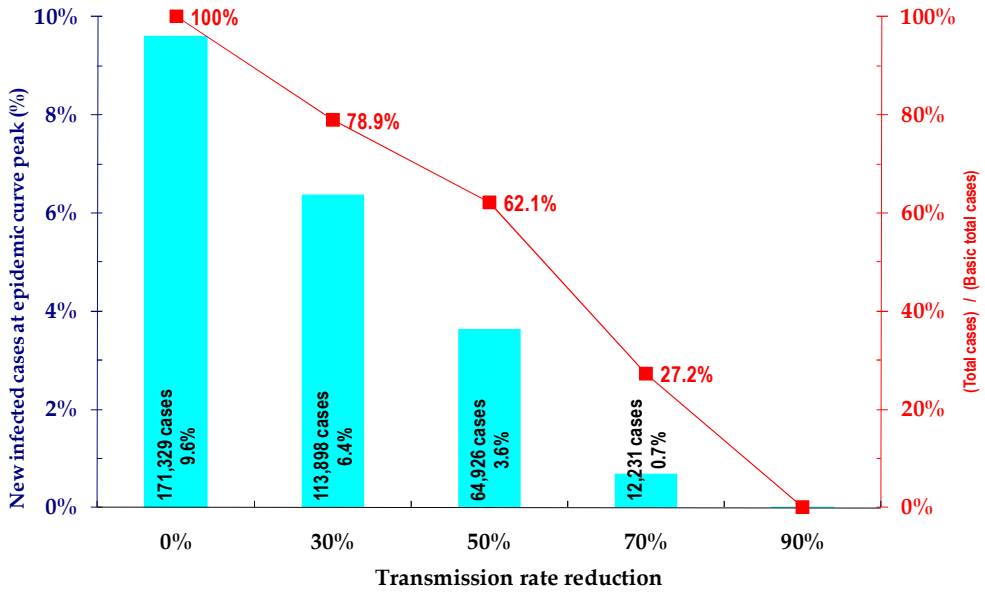


Figure 4.13. Basic epidemic curve at a 0% reduced transmission rate expressed according to two observation indexes.

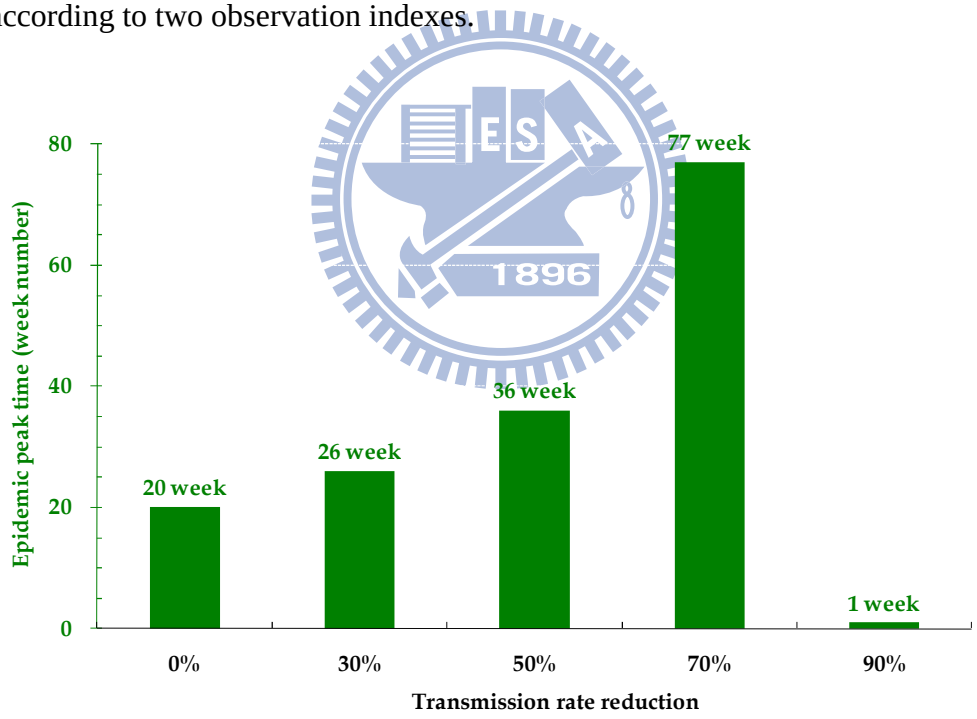


Figure 4.14. Comparison of new infected cases at epidemic curve peak at different transmission rates.

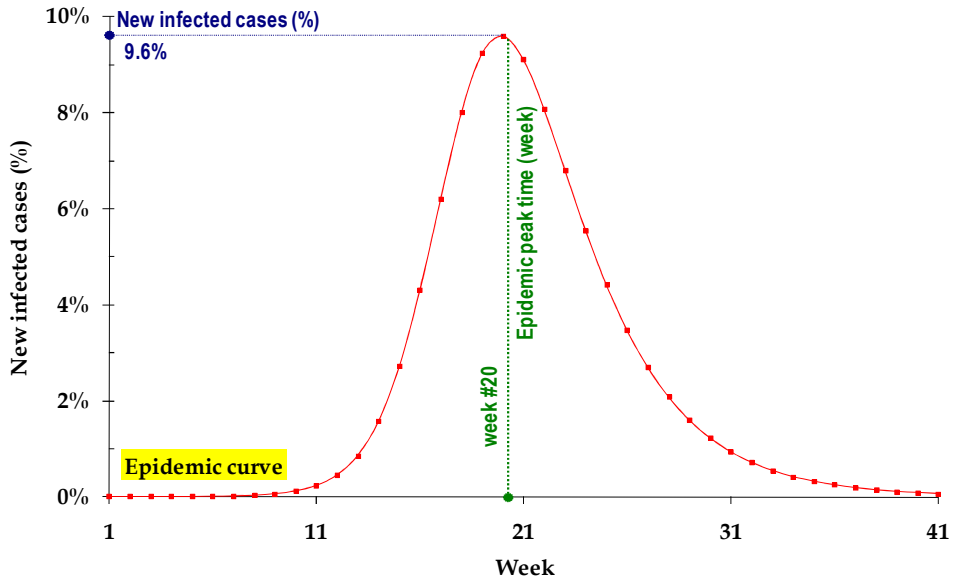


Figure 4.15. Weekly new cases at curve peak at different transmission rates.

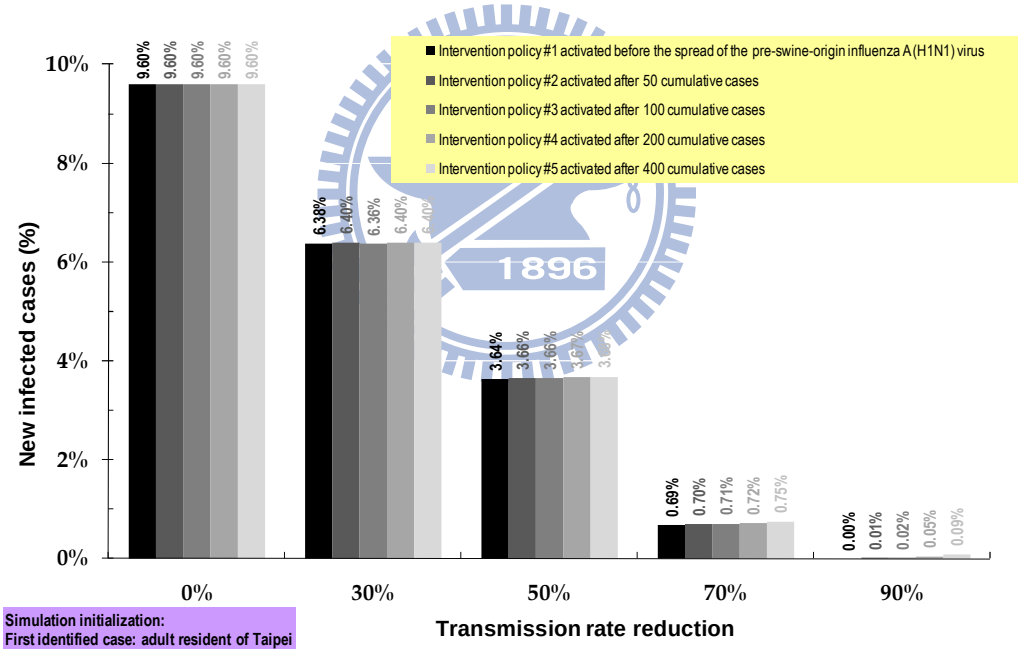


Figure 4.16. New infected cases at epidemic curve peak according to various intervention policy scenarios.

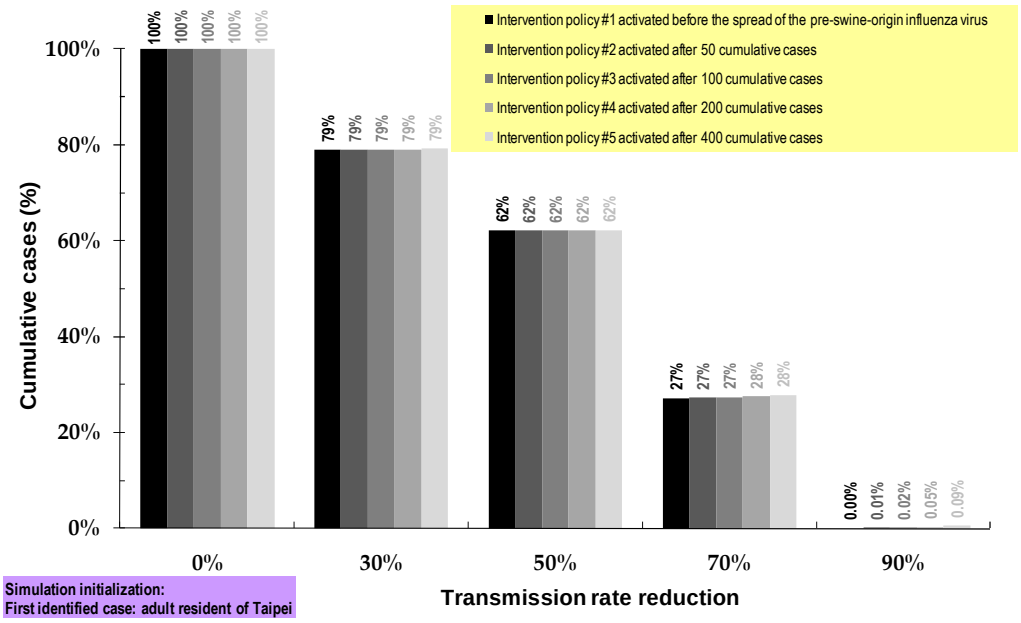


Figure 4.17. Numbers of infected cases according to various intervention policy scenarios.

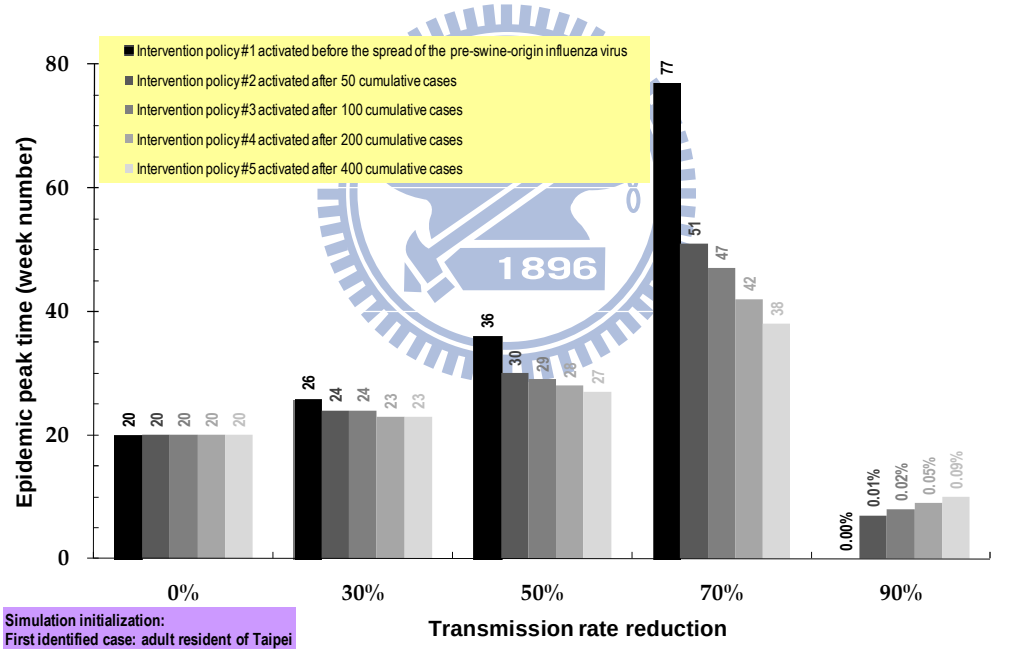


Figure 4.18. Week numbers of epidemic curve peaks according to various intervention policy scenarios.

Table 4.4 Observation index values according to different transmission rates.

Observation Index	Transmission Rate Reduction (%)				
	0%	30%	50%	70%	90%
Total cases.	1,784,044	1,407,752	1,108,520	485,761	8
New infected cases at epidemic curve peak.	171,329	113,898	64,926	12,231	8
Week number of epidemic curve peak.	20	26	36	77	∞
Fraction of new infected cases at epidemic curve peak.	9.6%	6.4%	3.6%	0.7%	0%
$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	78.9%	62.1%	27.2%	0.%

Results from comparisons of epidemic curve peaks at different reduced transmission rates are shown in Figure 4.13. The basic fraction of new infected cases at curve peak (noted as 100%, with a transmission rate of 1.0) is shown in the leftmost part of the graph. The relative total number of cases (red line) consists of two line segments, one from 1.0 to 0.5, and the other from 0.5 to 0.1. According to this result, transmission rate should be reduced by at least 50% to obtain better peak number suppression. An obvious decrease in peak number occurs when the transmission rate is reduced to 70%.

Curve peak week numbers at different transmission rates are shown in Figure 4.14. Note that week number increased as transmission rate decreased—a positive result for public health policy makers. Results from simulations of various long- and short-term intervention policy activation scenarios are shown in Figure 4.16,Figure 4.17 and Figure 4.18. No differences in numbers of infected cases were observed for different intervention policy activation times (Figure 4.16 and Figure 4.17). However, epidemic peak was delayed from week 55 to week 71 when intervention policy activation time

was set at 50 with a 70% reduction in transmission rate (Figure 4.18). Activation time exerted a much weaker effect on peak timing at a 30% transmission rate reduction. According to these results, while time of intervention policy activation did not significantly reduce the number of infected cases, it did exert an obvious effect in terms of delaying peak time—a positive result in terms of public health policy determination and preparation.

Next, differences in swine-origin H1N1 influenza A starting locations in Taiwan were compared in terms of their effects on the subsequent spreading of the disease (Figure 4.19, Table 4.5 and Table 4.6). Taipei was labeled a high-density area and Taichung a low-density area. In the first (pre-swine origin virus) scenario, case number peaked much earlier in Taipei (20) than in Taichung (61). When the transmission rate was reduced to 30%, the Taichung peak was significantly delayed. When comparing numbers of infected cases at the curve peak, both locations had approximately the same number of new cases, but Taipei had a much larger number of total cases. After reducing the transmission rate from 50% to 30%, Taichung had a much later peak week compared to Taipei, with no effect of intervention policy activation time on total number of cases or newly infected cases in either location. The results suggest that less densely populated starting locations are more sensitive to intervention policy activation time—that is, the combination of early activation time and low transmission rate

significantly delays epidemic curve peaks in less densely populated locations.

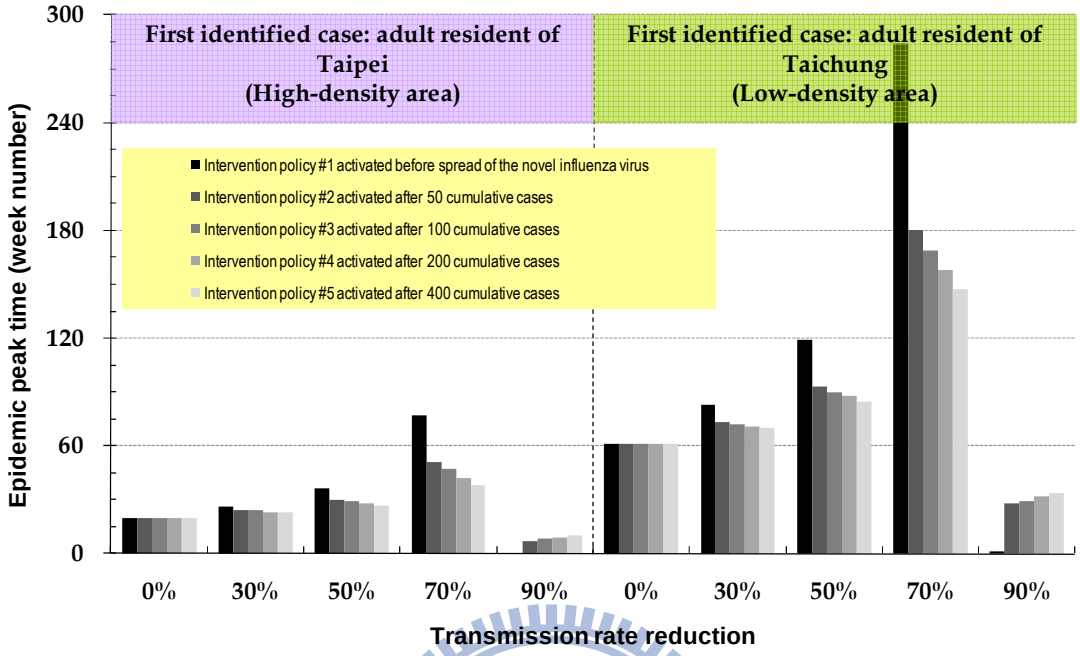


Figure 4.19. Epidemic peak week numbers for urban and rural areas.

Table 4.5. Observation index values according to different policy activation scenarios during swine-origin influenza A (H1N1) outbreak in Taipei.

Policy Activation Time	Observation Index	Transmission Rate Reduction				
		0%	30%	50%	70%	90%
Scenario #1 **Before the swine-origin influenza A (H1N1) virus emerges.**	Total cases.	1,784,044	1,407,752	1,108,520	485,761	8
	New infected cases at epidemic curve peak.	171,329	113,898	64,926	12,231	8
	Week number of epidemic curve peak.	20	26	36	77	0
	Fraction of new infected cases at epidemic curve peak.	9.60%	6.38%	3.64%	0.69%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	78.90%	62.14%	27.23%	0.00%
Scenario #2 **After 50 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	1,784,044	1,409,827	1,108,794	487,425	855
	New infected cases at epidemic curve peak.	171,329	114,120	65,235	12,468	155
	Week number of epidemic curve peak.	20	24	30	51	7
	Fraction of new infected cases at epidemic curve peak.	9.60%	6.40%	3.66%	0.70%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	79.02%	62.15%	27.32%	0.05%
Scenario #3 **After 100 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	1,784,044	1,410,263	1,108,993	488,900	1991
	New infected cases at epidemic curve peak.	171,329	113,532	65,314	12,604	349
	Week number of epidemic curve peak.	20	24	29	47	8
	Fraction of new infected cases at epidemic curve peak.	9.60%	6.36%	3.66%	0.71%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	79.05%	62.16%	27.40%	0.11%
Scenario #4 **After 200 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	1,784,044	1,410,782	1,109,355	491,563	4599
	New infected cases at epidemic curve peak.	171,329	114,191	65,442	12,883	818
	Week number of epidemic curve peak.	20	23	28	42	9
	Fraction of new infected cases at epidemic curve peak.	9.60%	6.40%	3.67%	0.72%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	79.08%	62.18%	27.55%	0.26%
Scenario #5 **After 400 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	1,784,044	1,411,273	1,109,893	496,246	10000
	New infected cases at epidemic curve peak.	171,329	114,185	65,669	13,408	1680
	Week number of epidemic curve peak.	20	23	27	38	10
	Fraction of new infected cases at epidemic curve peak.	9.60%	6.40%	3.68%	0.75%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	79.11%	62.21%	27.82%	0.56%

Table 4.6. Observation index values according to different policy activation scenarios during swine-origin influenza A (H1N1) outbreak in Taichung.

Policy Activation Time	Observation Index	Transmission Rate Reduction				
		0%	30%	50%	70%	90%
Scenario #1 **Before the swine-origin influenza A (H1N1) virus emerges.**	Total cases.	2,190,247	1,672,733	1,112,428	485,801	8
	New infected cases at epidemic curve peak.	172,083	114,556	64,551	12,186	8
	Week number of epidemic curve peak.	61	83	119	284	1
	Fraction of new infected cases at epidemic curve peak.	7.86%	5.23%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.37%	50.79%	22.18%	0.00%
Scenario #2 **After 50 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	2,190,247	1,672,266	1,117,265	487,030	767
	New infected cases at epidemic curve peak.	172,083	113,760	64,598	12,200	120
	Week number of epidemic curve peak.	61	73	93	180	28
	Fraction of new infected cases at epidemic curve peak.	7.86%	5.19%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.35%	51.01%	22.24%	0.04%
Scenario #3 **After 100 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	2,190,247	1,671,019	1,120,702	488,492	1723
	New infected cases at epidemic curve peak.	172,083	113,672	64,430	12,194	273
	Week number of epidemic curve peak.	61	72	90	169	29
	Fraction of new infected cases at epidemic curve peak.	7.86%	5.19%	2.94%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.29%	51.17%	22.30%	0.08%
Scenario #4 **After 200 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	2,190,247	1,674,627	1,125,289	491,418	3668
	New infected cases at epidemic curve peak.	172,083	113,592	64,556	12,198	520
	Week number of epidemic curve peak.	61	71	88	158	32
	Fraction of new infected cases at epidemic curve peak.	7.86%	5.19%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.46%	51.38%	22.44%	0.17%
Scenario #5 **After 400 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed.**	Total cases.	2,190,247	1,677,338	1,132,127	49,486	7424
	New infected cases at epidemic curve peak.	172,083	112,155	64,605	12,188	1057
	Week number of epidemic curve peak.	61	70	85	147	34
	Fraction of new infected cases at epidemic curve peak.	7.86%	5.12%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.58%	51.69%	22.67%	0.34%

The proposed model is capable of providing insights that reflect the dynamic processes of epidemics according to various intervention scenarios involving outbreak location, intervention timing, and different policy suites. I view this multilayer approach as both convenient and effective for public health practitioners and administrators responsible for initiating early responses to potential pandemics, and for assessing intervention strategies in outbreak locations.

This particular part of this dissertation has several limitations, such as the lack of confirmed numbers of H1N1 influenza A cases in Taiwan—at this time it is not a notifiable disease in this country. The data used for model validation reflect severe and hospitalized cases, which were assumed as having the same proportions as non-severe cases per time unit. Differences between actual and simulated cases can be significantly reduced when using appropriate parameter values in terms of investigation and detection proportions. Second, since the SLIR model is imprecise in terms of its *Removed* designation, it was not possible to address the number of H1N1-related deaths in any discussion of peak time delay. In real-world scenarios involving pandemic diseases with high death rates, peak time delays are very important for disease prevention policy decisions. Third, due to the limited scope of this study, it was not possible to gather and organize the exceptionally large amounts of available data on all areas represented by network nodes (e.g., workplaces, houses and schools). Instead,

location and age were used for purposes of population grouping, due to their similarities in responses to epidemic-related factors. Furthermore, other attributes such as income level and number of social groups per individual were not addressed when determining transmission rate, removed rate, or other parameters.

Chapter 5. Simulation Architecture for Studying Network-based Computational Epidemiology Issues and for Public Health Education Purposes

This chapter looks at issues tied to predicting epidemic outbreaks on a national scale, and testing the efficacies of different combinations of epidemic intervention policies. Network-based simulations have been proven as useful approaches for epidemiologists to address epidemic dynamics. In addition, investigations of complex public health issues are easy to simulate, therefore universities and research institutes are now using network-based simulations as teaching tools for epidemiology and public health students. However, instructors have reported that the process of constructing appropriate network-based epidemic models and running simulations is difficult, especially when the modeling of individual movement and contact patterns is involved.

I worked with three other authors to establish a four-category framework based

on geodemographic properties for use by undergraduate students and novice researchers wanting to construct network-based simulation models. The framework was evaluated with two infectious disease scenarios in Taiwan—HIV and influenza. Results indicate that the framework significantly improved student efforts to learn epidemic transmission principles, and to analyze the efficacies of various public health policies. To construct a multi-scale contact network, the proposed framework can be used to build geodemographic commuting or travel networks.

5.1 Motivation

A wide variety of network-based simulations are currently being used to model epidemic dynamics and to evaluate combinations of epidemic intervention policies. Due to their capability of modeling the movement and contact behavior of individuals, network-based simulations are being used by a growing number of researchers to explore epidemic dynamics (see, for example, Alfonseca, Martinez-Bravo & Torrea, 2000; Axelrod, 1997; Barrett et al., 2005; Boccara & Cheong, 1993; Ferguson et al., 2005; Gilbert & Troitzsch, 2005; Hsieh, Huang, Sun & Chen, 2005; Hsieh, Sun, Kao & Huang, 2006; Huang et al., 2005; Schneeberger et al., 2004; Stroud, Del Valle, Sydoriak, Riese & Mniszewski, 2007; Sumodhee et al., 2005).

To implement network-based simulations in the modeling of epidemic dynamics, we constructed social network simulations for modeling the transmission dynamics of HIV, SARS, and influenza in Taiwan (Hsieh et al., 2005; Hsieh et al., 2006; Huang et al., 2004, 2005; Sumodhee et al., 2005). For purposes of training students and novice epidemiologists, instructors from many disciplines are collaborating with simulation researchers to recreate the transmission dynamics of infectious diseases, and to improve general understanding of public health policy efficacies (Hsieh et al., 2006; Huang et al., 2005). However, computational epidemiology researchers and instructors

are still addressing individual problems involving movement and contact patterns among millions of people of different ages and with different professions, educational levels, marital/partner statuses, and levels of epidemiological resistance (Barrett et al., 2005; Boccara & Cheong, 1993; Huang et al., 2004, 2005). In addition, emerging and re-emerging infectious disease outbreaks can develop randomly and unexpectedly, depending on the breadth of early stage outbreaks, numbers of randomly imported cases, the responses of infected individuals, and contacts with other susceptible individuals (Barrett et al., 2005; Huang et al., 2004, 2005). Public health policies executed by health authorities also directly and indirectly affect epidemic dynamics and spreading situations (Hsieh et al., 2005). Furthermore, improper implementation and the inappropriate timing of public health policy activation occasionally produce such secondary impacts as disease concealment and social discrimination against infected patients and the health care employees who provide their care (Huang et al., 2004). In spite of these factors, most students and novice researchers in public health and related disciplines still use questionnaires or field investigation techniques when studying epidemic outbreaks—a process that prevents many from gaining a macro view of epidemic dynamics, or from assessing the potential efficacies of public health policies for prevention and control.

In this chapter I will discuss ways of applying network-based simulation

approaches to undergraduate and novice researcher education, and describe simulations of the transmission dynamics of two infectious disease scenarios in Taiwan—HIV and influenza. The goals are to clearly illustrate existing challenges to building network-based epidemic simulations, and to assist epidemiology students and novice researchers in their efforts to predict the transmission dynamics of emerging and re-emerging infectious diseases.

5.2 Potential Benefits in Learning Through Epidemic Simulations

Hands-on learning through the use of epidemic simulations has at least three potential benefits:

1. Operational. Epidemiology problems often require examinations of the influences of various public health policies in specific environments. Using the SARS outbreak of 2003 as an example, epidemiologists would be interested in measuring the potential impacts of public health policies, but running real-world experiments would be impossible in such a context. With simulation tools, epidemiology instructors and students can examine the influences of different public health policies in different regions, and execute "what-if" experiments to study the emerging behaviors of infections when irrelevant health policies are temporarily removed. In short, simulations can be optimized for learning (Bertsche, Crawford & Macadam, 1996).

2. Observational. Users can take epidemic simulation processes and adjust their scales for observation purposes, slow them down, or speed them up (Sumodhee et al., 2005). Epidemic simulations not only allow novice researchers to practice professional skills without having to invest large amounts of resources, but are also recognized as an efficient approach to reviewing or proving epidemiological concepts.

This protects learners from having to jump into high-risk situations. In classrooms, post-simulation reports allow teachers to determine which concepts their students have mastered (Hargrave & Kenton, 1999; Klein, Berlin, Kostolansky & Del Palacio, 2004; Levy, Levy & Solomon, 1995).

3. Construction. Epidemic simulations can be used to create or explore environments. Using public health policy assessments as an example, learners can practice predicting developments that might result from different combinations of public health policies. In classrooms, epidemiology instructors can exert relatively precise control over knowledge construction and accumulation (Hargrave & Kenton, 1999).

In response to a wide variety of geographic and demographic restrictions, we divided all network-based epidemic simulations into four categories. The first category reflects the use of real contact tracing for constructing small-scale individual-to-individual contact networks. Using the 2003 SARS outbreak as an example, health authorities in Taiwan and Singapore attempted to construct contact histories for all infected individuals in order to quarantine anyone who had come into contact with a carrier. The second category consists of individuals and locations. For example, saunas and bars frequented by homosexuals can be viewed as activity locations bridging susceptible individuals with HIV carriers; for illegal drug users,

infection locations include syringes and chemicals used for drug dilution. To construct social networks for illegal drug users, epidemiologists must determine how many times a user shares a syringe with other users during one week or month, or how many users share the same diluting agent in a single session.

The third category reflects individual neighborhood concepts using statistical geographic properties. In the absence of real contact data, epidemiologists may need to build a specific and customized social network using well-constructed and appropriate interaction and contact assumptions. An example is social mirror identities that connect two layers in the Cellular Automata with Social Mirror Identity Model (CASMIM), a small-world social network that preserves the properties of individuals who interact with their neighbors within two-dimensional geographic spaces (Huang et al., 2004, 2005). The properties reflect such activities as long-distance movement and daily visits to fixed locations. The fourth category frequently requires significant support in the form of demographic or geographic data. For example, Ferguson et al.'s (2005) Southeast Asian influenza simulation used statistical data for group density, household size, age distribution, school and workplace size, and individual travel information. The spread of HIV among homosexuals serves as a negative example—that is, movement, location, and means of sexual contact are less obvious, making it more difficult to build a network-based

HIV epidemic simulation (Sumodhee et al., 2005).

To build a network-based epidemic model, four properties are considered exceptionally important:

1. Time scale. In the case of HIV diffusion via heterosexual contact, frequency distributions of sexual behaviors over one month or one year show power-law distribution features (Schneeberger et al., 2004), but the same is not true when the time scale is reduced to one day or one week. It is also important to remember that different diseases have different incubation periods (e.g., 5 days for SARS versus 6 months to 20 years for HIV) and immunization time frames.

2. Geographic scale. Care must be taken when selecting this scale. For example, a new form of influenza tends to be expressed as a large-scale epidemic, therefore models for countries that have multiple regions require the consideration of cross-border transportation networks. Building a social network for any modern city with an established mass transportation system must assume a strong and varied mix of human movement, which can affect considerations of inter-regional transportation.

3. Data dependency. Data granularity determines the best method for building a network model. Using homosexual HIV diffusion as an example, any situation in which data are limited to frequency distributions of sexual contact restricts modelers to using abstract von Neumann and Moore neighborhood concepts (Huang et al.,

2005). However, if movement within a high-risk contact group can be traced, modelers can create simulations capable of predicting further development.

4. Extendability. Due to the diversity of data collected for epidemiology issues, simulations of specific infectious diseases often require modifications to existing network models. For example, the homogeneous mixing hypothesis used in random networks assumes that all members of a group are well-mixed (i.e., equal probabilities exist for contact between any two members), but data on sexual contact or needle sharing do not support this hypothesis. Therefore, extendibility is a major concern when applying an existing network model to new epidemic simulations.

Using computer simulations as a pedagogical tool is now common in many technology training programs, as well as in the teaching of science concepts (Colpitts, 2002; Hsieh et al., 2006; Liao & Sun, 2001). Computer simulations are also being used in epidemiology disciplines to support educational and training efforts based on constructivist learning principles. In addition to mitigating learner obsession with the minutiae of complex procedures described in epidemiology textbooks (Wenglinsky, 1998), simulations provide multiple opportunities for “learning by doing” (Oehme & Seitzer, 2000). Constructivists believe that learners draw upon prior knowledge when forming new schema via discovery learning (Bruner, 1977). When confronted with a new stimulus, learners apply their own knowledge bases to accommodate new

information and to alter their existing schema (Piaget, 1977). When constructive learning processes are embedded in epidemic simulations, students can learn by doing, have more and better opportunities for discovering interesting primary and secondary epidemic issues, and gain hands-on experience for dealing with real-world public health issues.

Originally developed for medical education in the early 1970s, *problem-based learning* is now considered a core teaching model in over 60 medical schools (Savery & Duffy, 1996). The use of simulations for learning and teaching has two characteristics that make it compatible with the theoretical foundations of problem-based learning:

1. Engagement. Students often request epidemic simulations to assist with learning and to gain a sense of engagement with real-world epidemiology problems. This allows for the introduction of related concepts to the learning process. There is no “perfect” simulation, but simulations can still support meaningful learning experiences as long as scenario limitations are taken into account (Aldrich, 2003).

2. Interaction flexibility. Epidemic simulations can be used with interaction and feedback methods to illustrate how infectious diseases are spread under different conditions and circumstances (Aldrich, 2003). Epidemiology problems are usually complex, and rarely have single “correct” answers, which encourages learners to

repeatedly manipulate parameters. With sufficient practice, learners or novice researchers can learn how to transfer their new knowledge to real-world infectious diseases.

5.3 Teaching Computational Modeling and Simulation

Processes and goals associated with learning via epidemic simulations differ from those associated with traditional classroom and textbook-centered learning. Epidemic simulation scenarios are often open-ended and poorly defined (Hsieh et al., 2005), and problems frequently arise after simulations are started. We therefore suggest that novices be required to use instruction-based manuals to run epidemic simulations and to create professional quality reports or presentations of their learning results. Teacher preparation time will vary depending on the required epidemiology background, scenario construction requirements, and necessary instruction to help learners formulate problem statements, collect data, run simulations, and create reports. Evaluative techniques for learning results also differ from those used in traditional classroom settings, and require some training on the part of instructors. In light of the amount of required background knowledge (Hargrave & Kenton, 1999), we suggest using pre-instructional time to teach public health policy assessment and epidemic outbreak prediction skills, and post-instructional time to teach skills in epidemic simulation construction and analysis. Both are appropriate for learning-by-doing experiences.

We designed a five-step epidemiology teaching process: (a) introducing epidemiology knowledge and background scenarios; (b) preparing a pre-test for guiding students to key properties of an epidemiological issue; (c) creating step-by-step instruction-based epidemic simulations with appropriate sample data, user manuals for operating epidemic simulations, and experiment design examples; (d) unrestrained operating time, which allows students to construct and develop their own experiments; and (e) post-tests or final presentations to evaluate student understanding. Since disease scenarios often have no single or absolute approach, it is difficult to evaluate how well novice learners understand the operational aspects of simulations. One potential solution is to design constructive pre-tests and post-tests. Using epidemic simulations associated with public health policies as an example, novice learners may be asked to compare the efficacies of different combinations of public health policies before and after an epidemic simulation is run. In addition, we have observed that novice learners exhibit wide differences in terms of controlling simulation parameters (Hsieh et al., 2006), and therefore suggest that parameters be used as an evaluation criterion.

Developing appropriate prevention and control measures entails making and monitoring the results of multiple rules and decisions made at different points during an epidemic. Conflicts among decision criteria for different strategies are inevitable,

and our proposed multi-scale simulation framework can help decision makers test and refine different strategies at different layers. For example, Layer 1 can be used to simulate and evaluate a vaccination policy by changing transmission rates among groups at greater risk of infection (e.g., children or seniors). This would allow for the testing of social distance measures such as school closures. Layer 2 can be used to evaluate quarantine strategies by changing contact rates among different age groups, layer 3 can be used to evaluate travel restrictions by changing regional contact rates among cities, and layer 4 can be used for the same purpose by changing the structure of the commuting network. By analyzing multi-scale interactions, decision makers can prepare themselves for making rapid proactive intervention decisions in response to clearly identified outbreak transmission pathways. Furthermore, our simulation framework can provide additional geospatial insight into epidemiological processes underlying control measures. Spatial orientation and visualization are necessary when monitoring disease progression and generating potential control strategies. We incorporated a geographic information system (GIS) into our multi-scale simulation framework in order to capture spatial variation in disease transmission throughout Taiwan. GIS supports a visual analysis of the spatial impacts of individual control measures. Combining multi-scale simulations, spatial visualization, and geographic information can clarify spatial and temporal characteristics in support of potential

pandemic preparation and control measures.

Chapter 6. Conclusions

The goal of this dissertation was to integrate realistic human social networks into a standard epidemiological disease transmission model. Toward that goal, I presented the potential benefits of network-based computational epidemiological simulations. Starting from theoretical complex network topology, I gave a possible explanation for why infectious diseases are extinguished at small transmission rates, even in scale-free networks. The study results suggest the possibility of controlling the spread of epidemics in scale-free networks by manipulating resources and costs associated with an infection event. I then proposed a multilayer network-based computational epidemiological framework called MEDSim, whose development I assisted with, to integrate realistic social networks into traditional epidemiological models. To demonstrate and test model flexibility and generalizability, the 2009 A/H1N1 influenza epidemic was used to compute outbreak locations and to simulate intervention scenarios. Results indicate that the proposed MEDSim framework can help public health organizations decide when to implement intervention strategies by simultaneously analyzing multilayer interactions. For novices studying computational epidemiology and public health principles, I worked with three other authors to describe an instruction program for building network-based epidemic models. The

goal is to help individuals with less advanced computing skills to build epidemiological models, determine appropriate simulation parameters, and construct operational procedures.

To build on this positive beginning, in the future I will work with the researchers cited in this dissertation to expand the multilayer framework in order to make it suitable for other acute diseases, and to make it responsive to complex human contact structures. I have five goals:

1. To model future disease spreading activity, I will work on modifying MEDSim parameters to include dynamic variables that change over time.
2. To account for vaccinations—specifically among school-age children, but also among other age groups—I will add one more state to the first MEDSim layer.
3. I will add transportation routes (e.g., highways, railways, air routes) to the fourth MEDSim layer.
4. I will work on adding an "international layer" to MEDSim in order to model cross-border epidemic dynamics.
5. I will work on extending MEDSim for use as a general purpose disease modeling framework—for example, modifying contact structures such as human-mosquito contact in order to model vector-borne diseases such as

dengue fever and malaria, and human-animal contact to model zoonotic diseases such as rabies and Japanese encephalitis.

Bibliography

Ahmed, E., & Agiza, H. N. (1998). On modeling epidemics Including latency, incubation and variable susceptibility. *Physica A, 253*, 347-352. doi:10.1016/S0378-4371(97)00665-1

Aldrich, C. (2003). *Simulations and the future of learning: An innovative (and perhaps revolutionary) approach to e-learning*. San Diego: Pfeiffer & Co.

Alfonseca, M., Martinez-Bravo, M. T., & Torrea, J. L. (2000). Mathematical Models for the Analysis of Hepatitis B and AIDS Epidemics. *Simulation*, *74*(4), 219-226. doi:10.1177/003754970007400403

Anderson, R. M., & May, R. M. (1991). *Infectious diseases of humans: dynamics and control*. New York: Oxford: University Press Oxford.

Axelrod, R. (1997). Advancing the art of simulation in the social sciences. *Complexity*, *3*(2), 16-22.

Bailey, N. T. J. (1950). A simple stochastic epidemic. *Biometrika*, *37*(3-4), 193-202. doi:10.1093/biomet/37.3-4.193

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509-512.

Barabási, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002).

Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, *311*(3-4), 590-614.

Barrett, C. L., Eubank, S. G., & Smith, J. P. (2005). If Smallpox Strikes Portland... *Scientific American*, *292*(3), 54–61.

Barrett, C., Eubank, S., & Marathe, M. (2006). Modeling and Simulation of Large Biological, Information and Socio-Technical Systems: An Interaction Based Approach. In *Interactive Computing: A new Paradigm* (Ed. D. Goldin, S. Smolka and P. Wegner., pp. 353-394). Berlin: Springer Verlag.

Barthelemy, M., Barrat, A., Pastor-Satorras, R., & Vespignani, A. (2004). Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Physical review letters*, *92*(17), 178701.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., & Vespignani, A. (2005). Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, *235*(2), 275-288. doi:10.1016/j.jtbi.2005.01.011

Bartlett, M. S. (1956). Deterministic and stochastic models for recurrent epidemics. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability* (Vol. 4, pp. 81-109).

Belshe, R. B., Swierkosz, E. M., Anderson, E. L., Newman, F. K., Nugent, S. L., & Maassab, H. F. (1992). Immunization of infants and young children with live

attenuated trivalent cold-recombinant influenza A H1N1, H3N2, and B vaccine. *The Journal of infectious diseases*, *165*(4), 727-732.

Bertsche, D., Crawford, C., & Macadam, S. E. (1996). Is simulation better than experience. *The McKinsey Quarterly*, (1), 50-58.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: structure and dynamics. *Physics Reports*, *424*(4-5), 175-308.

Boccara, N., & Cheong, K. (1993). Critical behaviour of a probabilistic automata network SIS model for the spread of an infectious disease in a population of moving individuals. *Journal of Physics A: Mathematical and General*, *26*(15), 3707-3717.

Boguñá, M., & Pastor-Satorras, R. (2002). Epidemic spreading in correlated complex networks. *Physical Review E*, *66*(4), 47104.

Bruner, J. S. (1977). *The process of education*. Cambridge: Harvard University Press.

Colizza, V., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences*, *103*(7), 2015-2020.

Colpitts, B. (2002). Teaching transmission lines: a project of measurement and simulation. *Education, IEEE Transactions on*, *45*(3), 245-252.

doi:10.1109/TE.2002.1024617

Davidsen, J., Ebel, H., & Bornholdt, S. (2002). Emergence of a small world from local interactions: Modeling acquaintance networks. *Physical Review Letters*, *88*(12), 128701.

Davis, G. F., Yoo, M., & Baker, W. E. (2003). The small world of the American corporate elite, 1982-2001. *Strategic organization*, *1*(3), 301-326.

Dezső, Z., & Barabási, A. L. (2002). Halting viruses in scale-free networks. *Physical Review E*, *65*(5), 55103.

Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, *28*(4), 365-382.

Diosan, L., & Dumitrescu, D. (2007). Evolutionary coalition formation in complex network. *Studia Universitatis Babes-Bolyai*, Informatica, *LII*(2), 115-129.

Directorate-General of Budget, Accounting, and Statistics. (2001). A preliminary report on year 2000 population and household census in Taiwan. Taipei, Taiwan.

Draief, M. (2006). Epidemic processes on complex networks. *Physica A: Statistical Mechanics and its Applications*, *363*(1), 120-131.

Draief, M., Ganesh, A., & Massoulié, L. (2008). Thresholds for virus spread on networks. *Annals of applied probability*, *18*(2), 359-378.

Edmunds, W. J., O'Callaghan, C. J., & Nokes, D. J. (1997). Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proceedings of the Royal Society B: Biological Sciences*, *264*(1384), 949-957.

Edwards, K. L., & Clarke, G. P. (2009). The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. *Social Science & Medicine*, *69*(7), 1127-1134. doi:10.1016/j.socscimed.2009.07.037

Epstein, J. M. (2009). Modelling to contain pandemics. *Nature*, *460*(7256), 687. doi:10.1038/460687a

Erdos, P., & Renyi, A. (1960). On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Science*, *5*, 17-61.

Estrada, E., & Hatano, N. (2008). Communicability in complex networks. *Physical Review E*, *77*(3), 036111. doi:10.1103/PhysRevE.77.036111

Feng, Z., Huang, W., & Castillo-Chavez, C. (2005). Global behavior of a multi-group SIS epidemic model with age structure. *Journal of Differential Equations*, *218*(2), 292-324.

Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., et al. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, *437*(7056), 209-214.

Flache, A., & Hegselmann, R. (2001). Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics. *Journal of Artificial Societies and Social Simulation*, *4*(4).

Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., Griffin, J., et al. (2009). Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science*, *324*(5934), 1557-1561. doi:10.1126/science.1176062

Fuentes, M. A., & Kuperman, M. N. (1999). Cellular automata and epidemiological models with spatial dependence. *Physica-Section A*, *267*(3), 471-486.

Gilbert, G. N., & Troitzsch, K. G. (2005). *Simulation for the social scientist*. United Kingdom: Open University Press.

González, M. C., Hidalgo, C. A., & Barabási, A. L. (2008). Understanding individual human mobility patterns. *Nature*, *453*(7196), 779-782.

Grais, R. F., Hugh Ellis, J., & Glass, G. E. (2003). Assessing the impact of airline travel on the geographic spread of pandemic influenza. *European journal of epidemiology*, *18*(11), 1065-1072.

Handcock, M. S., & Jones, J. H. (2006). Interval estimates for epidemic thresholds in two-sex network models. *Theoretical Population Biology*, *70*(2), 125-134.

Hargrave, C. P., & Kenton, J. M. (1999). Preinstructional simulations: Implications for science classroom teaching. *Journal of Computers in Mathematics and Science Teaching*, *19*(1), 47-58.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM review*, *42*(4), 599-653.

Hsieh, J. L., Huang, C. Y., Sun, C. T., & Chen, Y. M. A. (2005). Using the CAMIM small-world epidemic model to analyze public health policies. In *Proceedings of Western Simulation Multiconference on Health Sciences Simulation* (pp. 63-69). Presented at the Western Simulation MultiConference 2004 (WMC 2004), New Orleans.

Hsieh, J. L., Huang, C. Y., Sun, C. T., & Tsai, Y. S. (2009). Learning to build network-oriented epidemic simulation models in epidemiology education. *International Journal of Simulation and Process Modelling*, *5*(1), 31-41.

Hsieh, J. L., Sun, C. T., Kao, G. Y. M., & Huang, C. Y. (2006). Teaching through simulation: epidemic dynamics and public health policies. *Simulation*, *82*(11), 731-759.

Huang, C. Y., Sun, C. T., Hsieh, J. L., Chen, Y. M. A., & Lin, H. (2005). A novel

small-world model: Using social mirror identities for epidemic simulations. *Simulation*, *81*(10), 671-699.

Huang, C. Y., Sun, C. T., Hsieh, J. L., & Lin, H. L. (2004). Simulating SARS: Small-world epidemiological modeling and public health policy assessments. *Journal of Artificial Societies and Social Simulation*, *7*(4).

Huang, C. Y., Sun, C. T., & Lin, H. C. (2005). Influence of local information on social simulations in small-world network models. *Journal of Artificial Societies and Social Simulation*, *8*(4).

Huang, C. Y., Tsai, Y. S., & Sun, C. T. (2009). Effects of resource and remembering on social networks (pp. 837-842). Budapest, Hungary: International Foundation for Autonomous Agents and Multiagent Systems.

Huang, C. Y., Tsai, Y. S., & Sun, C. T. (2010). Effects of friend-making resources costs and remembering on acquaintance networks. *Physica A: Statistical Mechanics and Its Applications*, *389*(3), 604-622.

Huang, C. Y., Tsai, Y. S., Sun, C. T., Hsieh, J. L., & Cheng, C. Y. (2010). Influences of Resource Limitations and Transmission Costs on Epidemic Simulations and Critical Thresholds in Scale-free Networks. *Simulation*, *85*(3), 205-219.

Huang, C. Y., Tsai, Y. S., & Wen, T. H. (2010a). Simulations for Epidemiology and Public Health Education. *Journal of Simulation*, *4*(1), 1-13.

Huang, C. Y., Tsai, Y. S., & Wen, T. H. (2010b). A Network-based Simulation Architecture for Studying Epidemic Dynamics. *Simulation*. doi:10.1177/0037549709340733

Hwang, W., Kim, T., Ramanathan, M., & Zhang, A. (2008). Bridging centrality: Graph mining from element level to group level. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 336-344). Presented at the Knowledge Discovery and Data Mining Conference (KDD '08), Las Vegas, USA.

Hyman, J. M., & Stanley, E. (1988). Using mathematical models to understand the AIDS epidemic. *Mathematical Biosciences, 90*(1-2), 415-473.

Inaba, H. (2007). Age-structured homogeneous epidemic systems with application to the MSEIR epidemic model. *Journal of Mathematical Biology*, *54*(1), 101-146.

Jeger, M. J., Pautasso, M., Holdenrieder, O., & Shaw, M. W. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, *174*(2), 279-297.

Kan, C. C., Lee, P. F., Wen, T. H., Chao, D. Y., Wu, M. H., Lin, N. H., Huang, S. Y. J., et al. (2008). Two Clustering Diffusion Patterns Identified from the 2001-2003 Dengue Epidemic, Kaohsiung, Taiwan. *American Journal of Tropical Medicine and Hygiene*, *79*(3), 344-352.

Kaza, S., Xu, J., Marshall, B., & Hsinchun Chen. (2009). Topological Analysis of Criminal Activity Networks: Enhancing Transportation Security. *Intelligent Transportation Systems, IEEE Transactions on, 10*(1), 83-91. doi:10.1109/TITS.2008.2011695

Keeling, M. J., & Grenfell, B. T. (2000). Individual-based perspectives on R0. *Journal of Theoretical Biology, 203*(1), 51-61.

Keeling, M. J., & Rohani, P. (2007). *Modeling infectious diseases in humans and animals* (First edition.). Princeton: Princeton University Press.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A, 115*, 700-721.

Klein, C. A., Berlin, L. S., Kostolansky, T. J., & Del Palacio, J. R. (2004). *Stock simulation engine for an options trading game*. Google Patents.

Kuiken, T., Rimmelzwaan, G. F., Van Amerongen, G., & Osterhaus, A. (2003). Pathology of human influenza A (H5N1) virus infection in cynomolgus macaques (Macaca fascicularis). *Veterinary Pathology Online*, *40*(3), 304-310.

Langlais, M., & Naulin, J. M. (2003). An age structured SI epidemic problem in a heterogeneous environment. *Evolution Equations: Applications to Physics, Industry, Life Sciences and Economics*, *55*, 307-321.

Larsen, J., Axhausen, K. W., & Urry, J. (2006). Geographies of social networks: meetings, travel and communications. *Mobilities*, *1*(2), 261-283.

Levin, S. A., & Durrett, R. (1996). From individuals to epidemics. *Philosophical Transactions: Biological Sciences, 351*(1347), 1615-1621.

Levy, M., Levy, H., & Solomon, S. (1995). Microscopic Simulation of the Stock Market: the Effect of Microscopic Diversity. *Journal de Physique I*, *5*(8), 1087-1107. doi:10.1051/jp1:1995183

Li, G., & Jin, Z. (2005). Global stability of a SEIR epidemic model with infectious force in latent, infected and immune period. *Chaos, Solitons and Fractals*, *25*(5), 1177-1184.

Liao, Y. H., & Sun, C. T. (2001). An educational genetic algorithms learning tool. *IEEE Transactions on Education*, *44*(2), 210-219. doi:10.1109/13.925863

Liu, Q. X., & Jin, Z. (2005). Cellular automata modelling of SEIRS. *Chinese Physics*, *14*, 1370-1377.

Liu, X., Takeuchi, Y., & Iwami, S. (2008). SVIR epidemic models with vaccination strategies. *Journal of Theoretical biology*, *253*(1), 1-11.

Liu, X., Xia, L. I., YEH, A. G., JinQiang, H. E., & Jia, T. A. O. (2007). Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Sci China Ser D-Earth Sci*, *50*(10), 1578-1588.

Liu, Z., Lai, Y. C., & Ye, N. (2003). Propagation and immunization of infection on general networks with both homogeneous and heterogeneous components. *Physical Review E*, *67*(3), 31911.

Lloyd, A. L., & May, R. M. (2001). EPIDEMIOLOGY: How viruses spread among computers and people. *Science*, *292*(5520), 1316-1317. doi:10.1126/science.1061076

Lynch, A. (1996). *Thought contagion: How belief spreads through society* (First edition.). New York: Basic Books.

Mao, L., & Bian, L. (2010). Spatial-temporal transmission of influenza and its health risks in an urbanized area. *Computers, Environment and Urban Systems*, *34*(3), 204-215.

Marsh, R. M. (1996). *The great transformation: social change in Taipei, Taiwan since the 1960s*. New York: M.E. Sharpe.

May, R. M., & Lloyd, A. L. (2001). Infection dynamics on scale-free networks. *Physical Review E*, *64*(6), 66112.

Menard, A. (2008). VecGCA: a vector-based geographic cellular automata model allowing geometric transformations of objects. *Environment and Planning B: Planning and Design*, *35*, 647-665.

Meyers, L. A., Newman, M. E. J., Martin, M., & Schrag, S. (2003). Applying network

theory to epidemics: control measures for mycoplasma pneumoniae outbreaks. *Emerging Infectious Diseases*, *9*(2), 204-210.

Mikler, A. R., Venkatachalam, S., & Abbas, K. (2005). Modeling infectious diseases using global stochastic cellular automata. *Journal of Biological Systems*, *13*(4), 421-440.

Molinari, N. M., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E., & Bridges, C. B. (2007). The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine*, *25*(27), 5086-5096. doi:10.1016/j.vaccine.2007.03.046

Montoya, J. M., & Solé, R. V. (2002). Small world patterns in food webs. *Journal of theoretical biology*, *214*(3), 405-412.

Moore, C., & Newman, M. E. J. (2000). Epidemics and percolation in small-world networks. *Physical Review E*, *61*(5), 5678-5682. doi:10.1103/PhysRevE.61.5678

Morris, M. (1997). Sexual networks and HIV. *Aids - London - Current Science Then Rapid Science Publishers Then Lippincott Raven*, *11*, 209-216.

Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Physical Review E*, *66*(1), 16128.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM*

Review, *45*, 167-256.

Newman, M. E. J., & Watts, D. J. (1999). Scaling and percolation in the small-world network model. *Physical Review E*, *60*(6), 7332-7342. doi:10.1103/PhysRevE.60.7332

Oehme F., & Seitzer D. (2000). Learn by doing: how to include new requirements of research in engineering education. *European Journal of Engineering Education*, *25*, 131-137.

Ortiz-Pelaez, A., Pfeiffer, D. U., Soares-Magalhães, R. J., & Guitian, F. J. (2006). Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Preventive veterinary medicine*, *76*(1-2), 40-55.

Parham, P. E., & Ferguson, N. M. (2006). Space and contact networks: capturing the locality of disease transmission. *Journal of the Royal Society Interface*, *3*(9), 483-493.

Pastor-Satorras, R., & Vespignani, A. (2002). Immunization of complex networks. *Physical Review E*, *65*(3), 36104.

Pastor-Satorras, R., & Vespignani, A. (2001a). Epidemic dynamics and endemic states in complex networks. *Physical Review E*, *63*(6), 066117. doi:10.1103/PhysRevE.63.066117

Pastor-Satorras, R., & Vespignani, A. (2001b). Epidemic spreading in scale-free networks. *Physical Review Letters*, *86*(14), 3200-3203. doi:10.1103/PhysRevLett.86.3200

Pastor-Satorras, R., & Vespignani, A. (2002). Epidemic dynamics in finite size scale-free networks. *Physical Review E*, *65*(3), 035108. doi:10.1103/PhysRevE.65.035108

Pastor-Satorras, R., & Vespignani, A. (2003). Epidemics and immunization in scale-free networks. In *Handbook of Graphs and Networks: From the Genome to the Internet* (p. 417). German: Wiley-VCH.

Piaget, J. (1977). The development of thought: Equilibration of cognitive structures (trans. A. Rosin). *New York*: *Viking*.

Pourbohloul, B., Meyers, L. A., Skowronski, D. M., Krajden, M., Patrick, D. M., & Brunham, R. C. (2005). Modeling control strategies of respiratory pathogens. *Emerging Infectious Diseases*, *11*(8), 1249-1256.

Rae, A. (2009). From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census. *Computers, Environment and Urban Systems*, *33*(3), 161-178.

Riley, S. (2007). Large-scale spatial-transmission models of infectious disease. *Science*, *316*(5829), 1298-1301.

Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., Hedley, A. J., Leung, G. M., et al. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science, 300*(5627), 1961-1966.

Rogers, E. M. (2003). *Diffusion of innovations* (Origin edition.). Washinton: Free press.

Rollett, A. P. (1945). Mathematical models and constructions. *The Mathematical Gazette*, *29*(287), 181-192.

Savery, J. R., & Duffy, T. M. (1996). Problem based learning: An instructional model and its constructivist framework. *Constructivist learning environments: Case studies in instructional design*, 135-148.

Sawyer, R. K. (2003). Artificial societies: Multiagent systems and the micro-macro link in sociological theory. *Sociological Methods & Research*, *31*(3), 325-363.

Schneeberger, A., NAT, D. R., MERCER, C. H., GREGSON, S. A., FERGUSON, N. M., NYAMUKAPA, C. A., ANDERSON, R. O., et al. (2004). Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sexually transmitted diseases, 31*(6), 380-387.

Shim, E., Feng, Z., Martcheva, M., & Castillo-Chavez, C. (2006). An age-structured

epidemic model of rotavirus with vaccination. *Journal of Mathematical Biology*, *53*(4), 719-746.

Shirley, M. D., & Rushton, S. P. (2005). The impacts of network topology on disease spread. *Ecological Complexity*, *2*(3), 287-299.

Silva, S. L., Ferreira, J. A., & Martins, M. L. (2007). Epidemic spreading in a scale-free network of regular lattices. *Physica A: Statistical Mechanics and its Applications*, *377*(2), 689-697.

Sirakoulis, G. C., Karafyllidis, I., & Thanailakis, A. (2000). A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecological Modelling*, *133*(3), 209-223.

Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., Ma, S. K., et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*, *459*(7250), 1122-1125.

Stroud, P., Del Valle, S., Sydoriak, S., Riese, J., & Mniszewski, S. (2007). Spatial Dynamics of Pandemic Influenza in a Massive Artificial Society. *Journal of Artificial Societies and Social Simulation*, *10*(4).

Sumodhee, C., Hsieh, J. L., Sun, C. T., Huang, C. Y., & Chen, A. Y. M. (2005). Impact of Social Behaviors on HIV Epidemic: A Computer Simulation View. In *Computational Intelligence for Modelling, Control and Automation, 2005 and*

International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on (Vol. 2, pp. 550-556). Presented at the Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on. doi:10.1109/CIMCA.2005.1631526

Supriatna, A. K., Soewono, E., & Van Gils, S. A. (2008). A two-age-classes dengue transmission model. *Mathematical Biosciences*, *216*(1), 114-121.

Sypsa, V., Pavlopoulou, I., & Hatzakis, A. (2009). Use of an inactivated vaccine in mitigating pandemic influenza A(H1N1) spread: a modelling study to assess the impact of vaccination timing and prioritisation strategies. *Eurosurveillance*, *14*(41), 19356.

Thakur, R., & Sharma, M. (2009). GIS and Challenges to Planning and Development Applications in Peripheral Regions. In *Planning and Socioeconomic Applications* (Vol. 1, pp. 125-138). Berlin: Springer-Verlag.

Tomlinson, B., & Cockram, C. (2003). SARS: experience at Prince of Wales Hospital, Hong Kong. *The Lancet*, *361*(9368), 1486-1487.

Tsai, Y. S., Huang, C. Y., Wen, T. H., Sun, C. T., & Yen, M. Y. (2010). Integrating epidemic dynamics with daily commuting networks: building a multilayer framework to assess influenza A (H1N1) intervention policies. *Simulation*.

doi:10.1177/0037549710379481

Tsai, Y. S., & Huang, C. Y. (2010). Effects of Resources and Costs on Diffusion Dynamics. *International Journal of Intelligent Information and Database Systems*, *4*(1), 19-42.

Tsai, Y. S., Sun, C. T., & Huang, C. Y. (2008). Epidemic Dynamics and Thresholds in Agent-Based Simulations under Realistic Resources and Cost Conditions (Vol. 2, pp. 65-70). Presented at the Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on. doi:10.1109/WIIAT.2008.11

Tsai, Y. S., Sun, C. T., & Huang, C. Y. (2010). Response to Wilson's Note on 'Influences of Resource Limitations and Transmission Costs on Epidemic Simulations and Critical Thresholds in Scale-Free Networks'. *Simulation*. doi:10.1177/0037549709351543

Vázquez, A., Flammini, A., Maritan, A., & Vespignani, A. (2003). Modeling of protein interaction networks. *Complexus*, *1*, 38-44.

Von Neumann, J., & Burks, A. W. (1966). *Theory of self-reproducing automata*. Chicaco: University of Illinois.

Wang, J. (2006). *Spatial analysis*. Beijing, China: Science Press.

Wang, W., & Ruan, S. (2004). Simulating the SARS outbreak in Beijing with limited

data. *Journal of theoretical biology*, *227*(3), 369-379.

Wang, W., & Zhao, X. Q. (2005). An age-structured epidemic model in a patchy environment. *SIAM Journal on Applied Mathematics*, *65*, 1597-1614.

Wang, X. F. (2002). Complex networks: topology, dynamics and synchronization. *International journal of bifurcation and chaos in applied sciences and engineering*, *12*(5), 885-916.

Watts, D. J. (2003). *Six degrees: The science of a connected age*. New York: WW Norton & Company.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440-442. doi:10.1038/30918

Wenglinsky, H. (1998). *Does It Compute? The Relationship between Educational Technology and Student Achievement in Mathematics.* Princeton: Educational Testing Service, Princeton, NJ. Policy Information Center.

White, S. H., del Rey, A. M., & Sánchez, G. R. (2007). Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, *186*(1), 193-202.

Wylie, J. L., Shah, L., & Jolly, A. (2007). Incorporating geographic settings into a social network analysis of injection drug use and bloodborne pathogen prevalence. *Health & place*, *13*(3), 617-628.

Xu, X., Zhang, X., & Mendes, J. F. F. (2007). Impacts of preference and geography on

epidemic spreading. *Physical Review E*, *76*(5), 056109. doi:10.1103/PhysRevE.76.056109

Yang, R., Wang, B. H., Ren, J., Bai, W. J., Shi, Z. W., Wang, W. X., & Zhou, T. (2007). Epidemic spreading on heterogeneous networks with identical infectivity. *Physics Letters A*, *364*(3-4), 189-193.

Yang, Y., Atkinson, P., & Ettema, D. (2008). Individual space-time activity-based modelling of infectious disease transmission within a city. *Journal of The Royal Society Interface*, *5*(24), 759-772.

Yang, Y., Sugimoto, J. D., Halloran, M. E., Basta, N. E., Chao, D. L., Matrajt, L., Potter, G., et al. (2009). The Transmissibility and Control of Pandemic Influenza A (H1N1) Virus. *Science*, *326*(5953), 729-733. doi:10.1126/science.1177373

Zhong, S. B., Huang, Q. Y., & Song, D. J. (2009). Simulation of the spread of infectious diseases in a geographical environment. *Science in China Series D: Earth Sciences*, *52*(4), 550-561.

Zhou, C., Sun, Z., & Xie, Y. (1999). *Geographical Cellular Automata*. Beijing, China: Science Press.

Zhou, T. (2008). Mixing navigation on networks. *Physica A: Statistical Mechanics and Its Applications*, *387*(12), 3025-3032.

Zhou, T., Yan, G., & Wang, B. (2005). Maximal planar networks with large clustering

coefficient and power-law degree distribution. *Physical Review E*, *71*(4), 046141. doi:10.1103/PhysRevE.71.046141

Bio

Yu-Shiuan Tsai received his B.S. degree in Mathematics (2002) and M.S. in Mathematics (2005) from National Taiwan University. He recently was awarded a Ph.D. in Computer Science by National Chiao Tung University. His current research interests include computational epidemiology, complex networks and systems, social science research, and mathematic analysis.

Integrating epidemic dynamics with daily commuting networks: building a multilayer framework to assess influenza A (H1N1) intervention policies

Yu-Shiuan Tsai, Chung-Yuan Huang, Tzai-Hung Wen, Chuen-Tsai Sun and Muh-Yong Yen

SIMULATION published online 7 September 2010
DOI: 10.1177/0037549710379481

Simulation

Simulation: Transactions of the Society of Modeling and Simulation International
0(00) 1–21

DOI: 10.1177/0037549710379481
sim.sagepub.com

Integrating epidemic dynamics with daily commuting networks: building a multilayer framework to assess influenza A (H1N1) intervention policies

Yu-Shiuan Tsai[1], Chung-Yuan Huang[2], Tzai-Hung Wen[3], Chuen-Tsai Sun[1] and Muh-Yong Yen[4]

Abstract

We describe an innovative simulation framework that combines daily commuting network data with a commonly used population-based transmission model to assess the impacts of various interventions on epidemic dynamics in Taiwan. Called the *Multilayer Epidemic Dynamics Simulator* (MEDSim), our proposed framework has four contact structures: within age group, between age groups, daily commute, and nationwide interaction. To test model flexibility and generalizability, we simulated outbreak locations and intervention scenarios for the 2009 swine-origin influenza A (H1N1) epidemic. Our results indicate that lower transmission rates and earlier intervention activation times did not reduce total numbers of infected cases, but did delay peak times. When the transmission rate was decreased by a minimum of 70%, significant epidemic peak delays were observed when interventions were activated before new case number 50; no significant effects were noted when the transmission rate was decreased by less than 30%. Observed peaks occurred more quickly when initial outbreaks took place in urban rather than rural areas. According to our results, the MEDSim provides insights that reflect the dynamic processes of epidemics under different intervention scenarios, thus clarifying the effects of complex contact structures on disease transmission dynamics.

Keywords

computer simulation, epidemic dynamics, geographic information system, multilayer model, travel network

1. Introduction

After emerging in Mexico in April of 2009, the swine-origin H1N1 influenza virus rapidly spread worldwide. In June of that year, the World Health Organization issued its highest possible pandemic alert: level 6.[1] Influenza researchers and epidemiologists have focused on two spreading factors: age group (determining post-infection symptoms)[2–7] and adult travel (determining routes by which viruses spread). Since individuals in the same age group tend to have similar epidemic characteristics, age group has been proposed as a distinguishing condition in terms of population compartmentalization.[2–6,8] Children and adolescents generally have better resistance to contagious diseases than individuals age 65 and older. However, the Mexican population segment that was most affected by the H1N1 virus consisted of youths below the age of 15; of all individuals affected by the first infection wave, 61% were children and 29% adults.[9] Since novel influenza viruses are known to cause greater morbidity among children,[10] the youngest age group served as the main focus of H1N1 intervention efforts.

[1]Department of Computer Science, National Chiao Tung University, Taiwan.
[2]Department of Computer Science and Information Engineering and Research Center for Emerging Viral Infections, Chang Gung University, Taiwan.
[3]Department of Geography and Infectious Disease Research and Education Center, National Taiwan University, Taiwan.
[4]Infectious Disease Section, Taipei City Hospital, Taiwan.

Corresponding author:
Tzai-Hung Wen, Department of Geography and Infectious Disease Research and Education Center, National Taiwan University, 1 Sec. 4, Roosevelt Road, Taipei 10617, Taiwan.
Email: wenthung@ntu.edu.tw

Many researchers have used age structure to capture heterogeneity when modeling epidemic dynamics,[3,5,7] with some integrating compartmental models consisting of different age groups to identify potential impacts of specific populations and temporal epidemic trends.[7] Childhood diseases, such as rotavirus infections, have been used to assess the efficacy potential of various vaccination strategies,[5] and transmission threshold and stability have been the focuses of epidemic simulations involving specific age structures.[3]

Another important factor in modeling epidemic dynamics is population movement. Over the past three decades Taiwan has experienced a rapid increase in the number of commuters for work and other purposes, particularly among young adults[11] – a phenomenon perceived as supporting the spread of viruses over long distances within the country.[8] Commuting is marked by strong spatial-temporal regularity: regardless of travel distance or time, most commuters follow simple and repetitive patterns.[12] These patterns are receiving considerable attention from researchers studying the spreading dynamics of diseases and viruses,[13] the clustering characteristics of epidemic diseases at the beginning of a breakout,[14,15] and the targeting of vaccinations, quarantining, and other public health policies.[16–19]

The two most commonly used approaches to modeling epidemic spreading dynamics are population based and network oriented. In population-based approaches, hosts sharing the same symptoms are modeled or grouped in terms of limited numbers of classes (also known as compartments) that researchers analyze and compare.[2–7] Combinations of classes are used to model and analyze population dynamics. For example, the *Susceptible*, *Latent*, *Infectious*, or *Recovered* (SLIR)[20] model gives individuals one of four infection statuses and differential equations are used to study system dynamics in terms of transitions between epidemiological phases. Depending on whether removed individuals can become susceptible a second time, diseases can be modeled as SLIR or SLIRS cycles.

Network-oriented approaches emphasize individual heterogeneity, interactions among individuals, and network structure.[21,22,40] Individuals in a network are represented as nodes, and interactions between them as links. Network nodes can be used to represent the characteristics of individuals, locations, neighborhoods, or cities, and models can incorporate the temporal dynamics of these features. Time frames for links between two nodes can be preferentially defined[23] – an approach commonly used to represent group structures for individuals exhibiting interaction or relationship patterns.[24–27] Network-oriented approaches are suitable for capturing complex contact patterns among individuals, exploring epidemic dynamics, and assessing the efficacies of public health policies.[19,22,28,29] Lattice networks have been used to determine distance relationships between individuals. In contrast, random networks support features associated with casual contacts among mobile individuals and the low degree of separation commonly observed in social networks.[30–32] Some researchers incorporate more realistic underlying networks (e.g. daily contact networks) when modeling interaction behaviors.[30–32] These approaches are viewed as reliable for investigating epidemics, with the transmission dynamics of specific network models being manipulated to investigate the spread of emerging infectious diseases.[33,39] The topological features of social networks have recently been found to exert considerable influence on the transmission dynamics and critical thresholds of infectious diseases, thus supporting the subtle analyses that network-oriented models are incapable of.[4,13,28,34]

Population-based and network-oriented approaches respectively emphasize large-scale population-level and individual-level perspectives. Each has its own limitations. Population-based approaches are suitable for discussing dynamic variation across individuals in the same compartment, but they are weak in terms of modeling individual heterogeneity and addressing human travel networks.[22,30] Since individuals are modeled as groups, any two group members are assumed to have a direct connection, which is not true in the real world. Furthermore, movement and activity are location dependent; therefore, phenomena cannot be simulated by a population-based approach that assumes a homogeneous population distribution. In contrast, network-oriented approaches may be appropriate for introducing individual heterogeneity, but they are computation intensive and time consuming when simulating the behaviors of individuals with multiple attributes in large-scale social environments.[17,30] Many efforts have been made to match individual and population behaviors with heterogeneity and computation requirements when studying epidemic dynamics.[35–37]

Here we will propose a multilayer simulation framework that combines daily commuting networks and a commonly used population-based transmission model for simulating epidemic dynamics. We used the 2008–2009 seasonal influenza A and 2009 swine-origin influenza A (H1N1) outbreaks to estimate model parameters. We then assessed the potential impacts of different outbreak locations and interventions on the Taiwan-wide epidemic dynamics of swine-origin influenza A, including intervention timing and different combinations of public health responses.

2. Multilayer epidemic dynamic simulation

To analyze the spreading dynamics of epidemic diseases in detail, we established a top-down simulation

framework and implemented a prototype of our Multilayer Epidemic Dynamics Simulator (MEDSim). The MEDSim integrates population-based and network-oriented approaches to capturing complex demographic, geographic, and biological properties, including human movement patterns and disease progression (Figure 1). Based on the observation that epidemic dynamics in large populations are similar to those found in deterministic systems,[16] we established a deterministic framework for our MEDSim model. As shown in Figure 2, layer 1 individuals within the same location are organized according to age group;

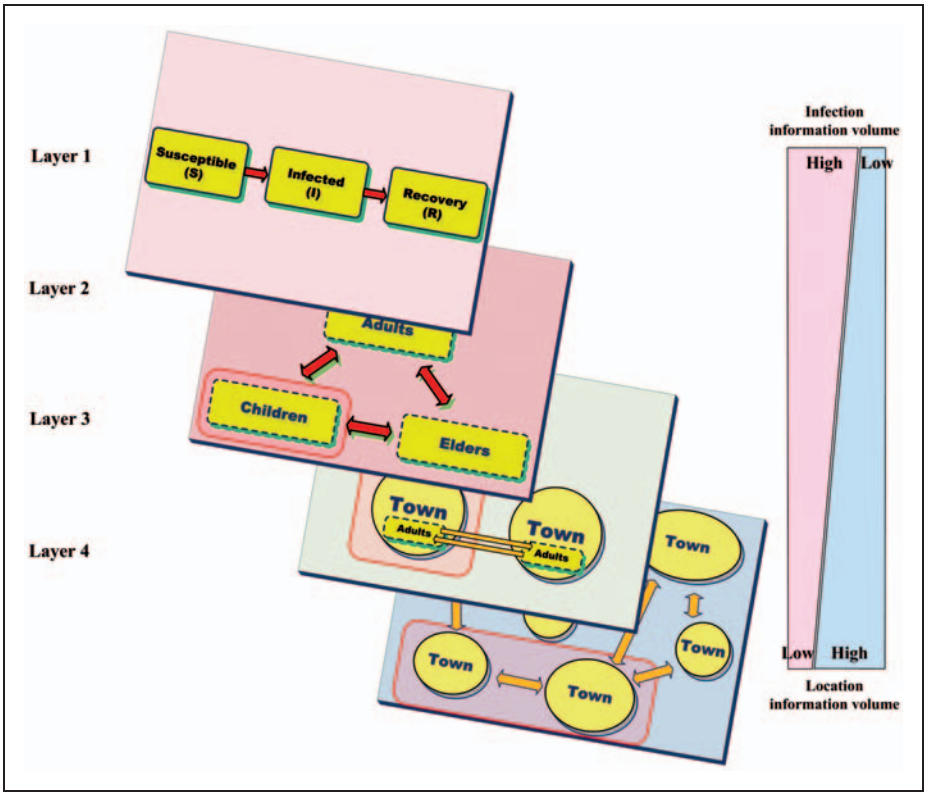


Figure 1. The MEDSim concept. Infection information usage is highest in layer 1 and lowest in layer 4, the opposite of location information.

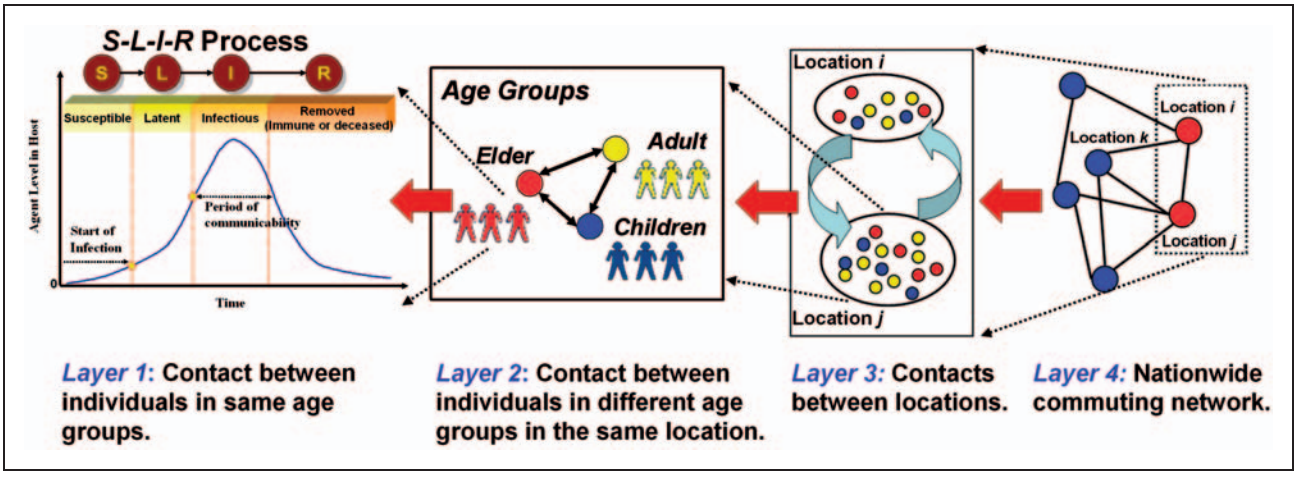


Figure 2. MEDSim framework.

a population-based approach is used to model the transmission dynamics of each group. The layer 2 focus is on contact patterns and interactions between different age groups within the same location. The effects of regional interactions on human travel networks are added to layer 3 by incorporating population density and commuting volume between any two locations. In layer 4, a network-oriented approach is used to incorporate a geographic information system (GIS) for constructing human travel networks on a national scale, with nodes representing locations on commuting routes and links representing movement between them.

Due to its ability to comprehensively integrate multilayer structures to generate dynamic spatial and temporal processes, we used Mathworks MATLAB to implement our MEDSim framework as a numerical computation kernel. By using Microsoft Excel to organize census and transportation data, policy makers, health professionals, and others who have less experience with specialized computer software will be able to generate simulation scenarios with minimal assistance.

2.1. Layer 1: Within age groups

We used the four-state SLIR epidemiological model to represent different infection stages among individuals in the same age group in the same location. Individual epidemic status is initially set at *S*usceptible (vulnerable to infection but not yet infected), followed by *L*atent (infected but unable to infect others), *I*nfectious (capable of infecting other individuals), and *R*emoved (i.e. recovered, deceased, or otherwise not posing any further threat). The numbers of pathogens that Susceptible-to-Latent hosts carry are insufficient for active transmission to other Susceptible hosts, but these numbers eventually reach levels where hosts become Infectious, begin to infect other Susceptible hosts, and eventually move toward a Removed status. The dynamics of the four epidemic states over time are expressed as Equations (1a)–(1d), which have the following features.

1. At time t, the population of interest is divided into four compartments ($S(t)$, $L(t)$, $I(t)$, and $R(t)$) corresponding to the SLIR model's four epidemic states. Since the SLIR model is a closed system, $S(t) + L(t) + I(t) + R(t) = N$, with N a constant representing the entire population.
2. Transmission rate β is a constant representing how fast Susceptible individuals become Infected and acquire a Latent status.
3. Latent rate θ is a constant used to determine transformation speed from Latent to Infected.
4. Removed rate α is a constant used to determine transformation speed from Infected to Recovered.

Ordinary differential equations can be used to express the SLIR model as follows:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t)/N \tag{1a}$$

$$\frac{dL(t)}{dt} = -\theta L(t) + \beta S(t)I(t)/N \tag{1b}$$

$$\frac{dI(t)}{dt} = -\alpha I(t) + \theta L(t) \tag{1c}$$

$$\frac{dR(t)}{dt} = \alpha I(t) \tag{1d}$$

Figures 3(a) and (b) respectively present the concept and a flowchart of our model's first layer. Note our modification in the interest of taking into consideration self-motivated hospitalization (i.e. those individuals who visit hospitals or clinics during an influenza outbreak regardless of their infection status). Depending on diagnostic accuracy, some are confirmed as infectious and receive medical treatment in advance, thus altering transmission and removed rates for certain populations. To integrate this factor into the model, we propose adding three features: (a) an investigation constant s representing the percentage of a population that goes to a hospital or clinic in advance of becoming ill; (b) a detection constant c, used to determine the percentage of a population confirmed as infectious; and (c) a time delay constant T, indicating the amount of time between a patient with symptoms visiting a hospital or clinic and the time his or her infection is confirmed. The default values of parameters s and c are both 0.6 (Table 1), meaning that 60% of the infected population is prone to visiting hospitals and/or clinics for medical advice, and 60% of those visitors are correctly diagnosed as carrying the pathogen. The default value of parameter T is 3 (Table 1), meaning that it takes three days to confirm that a hospital or clinic patient with symptoms is infected with the pathogen. In simulations, correctly diagnosed patients are equivalent to confirmed cases in real-world influenza surveillance systems.

In consideration of preventive health care, we added a feature in which individuals with an L status are moved to either an I_1 (infected and prone to visiting hospitals and/or clinics for medical advice) or I_2 (infected but not prone to visiting hospitals and/or clinics) status, based on whether or not they actually visit a hospital or clinic; this feature is expressed as investigation proportion s. I_1 individuals are identified as either I_{11} (correctly diagnosed as carrying the pathogen) or I_{12} (incorrectly diagnosed as carrying the pathogen – in other words, false negatives); this is expressed as detection proportion c. Note that regardless of

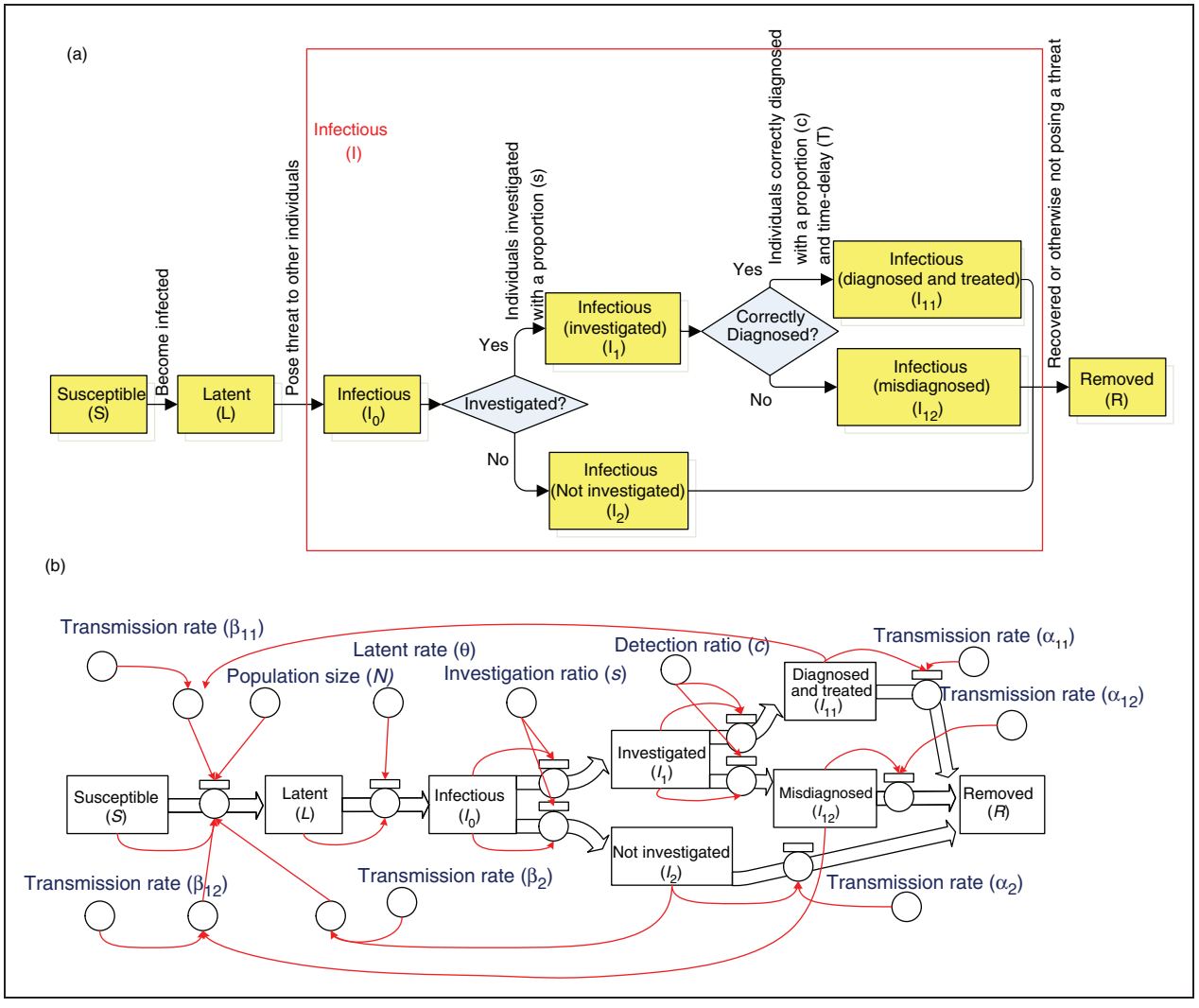


Figure 3. (a) Modified SLIR model layer 1 concept. (b) Modified SLIR model layer 1 flowchart.

positive or negative diagnoses, a T period of time must elapse prior to confirmation. The difference between state I_{11} and either I_2 or I_{12} is the transmission rate. I_{11}, I_{12}, and I_2 all eventually change to state R.

The extended SLIR model can be expressed as

$$\frac{dS(t)}{dt} = -S(t)(\beta_2 I_2(t) + \beta_{11} I_{11}(t) + \beta_{12} I_{12}(t))/N \quad (2a)$$

$$\frac{dL(t)}{dt} = -\theta L(t) + S(t)(\beta_2 I_2(t) + \beta_{11} I_{11}(t) + \beta_{12} I_{12}(t))/N \quad (2b)$$

$$\frac{dI_0(t)}{dt} = -I_0(t) + \theta L(t) \quad (2c)$$

$$\frac{dI_1(t)}{dt} = -I_1(t-T) + sI_0(t) \quad (2d)$$

$$\frac{dI_2(t)}{dt} = -\alpha_2 I_2(t) + (1-s)I_0(t) \quad (2e)$$

$$\frac{dI_{11}(t)}{dt} = -\alpha_{11} I_{11}(t) + cI_1(t-T) \quad (2f)$$

$$\frac{dI_{12}(t)}{dt} = -\alpha_{12} I_{12}(t) + (1-c)I_1(t-T) \quad (2g)$$

$$\frac{dR(t)}{dt} = \alpha_{11} I_{11}(t) + \alpha_{12} I_{12}(t) + \alpha_2 I_2(t) \quad (2h)$$

2.2. Layer 2: Among age groups

Depending on age range, individual infection properties differ in terms of epidemic parameters such as transmission and removed rates. We considered two age-related

Table 1. MEDSim parameters

Category	Layer	Attribute	Symbol	Description
Epidemic	1	*Transmission rate*	β^i_{11pp}	Transmission rate from investigated/diagnosed/treated age group p to same age group in town i
			β^i_{12pp}	Transmission rate from investigated/misdiagnosed age group p to same age group in town i
			β^i_{2pp}	Transmission rate from non-investigated age group p to same age group in town i
		Latent rate	θ^i_p	Latent rate of age group p in town i
		Removed rate	α^i_{11p}	Removed rate of investigated/diagnosed/treated age group p in town i
			α^i_{12p}	Removed rate of misdiagnosed age group p in town i
			α^i_{2p}	Removed rate of non-investigated age group p in town i
		Investigation ratio	s^i_p	Investigated proportion of age group p in town i (Default: 0.6)
		Detection ratio	c^i_p	Correctly diagnosed proportion of age group p in town i (Default: 0.6)
		Delay time	T	Time between investigation and correct diagnosis (Default: 3)
	2	*Transmission rate*	β^i_{11xy}	Transmission rate from investigated/diagnosed/treated age group p and same age group q in town i
			β^i_{12xy}	Transmission rate from misdiagnosed age-group p and same age group q in town i
			β^i_{2xy}	Transmission rate from the non-investigated age group p and same age group q in town i
Location	2	*Relative percentage*	χ^i_p	Age group p as a percentage of town i population. (*Source*: ROC Interior Ministry)
	3	*Determination*	$\sigma(p)$	Binary value for commuter age level (Default: adult)
		Relative density	d^i	Population of town i as a percentage of the largest town's population (*Source*: ROC Interior Ministry)
		Commuting weight	$w_{j,i}$	Number of commuters from town i to town j (*Source*: ROC Institute of Transportation)
	4	*Intercity rate*	η^i	Average number of daily contacts among individuals in location i (Default: 0.8)

features: the transmission rates β_{11pq}, β_{12pq}, and β_{2pq}, which represent cross-age group infections, and the relative percentage χ_p of age level, which affects the potential for cross-age infections. To distinguish among parameters for individuals in different age groups, we also added a subscript to each Equation (2) parameter (with the exception of T) – for example, we changed parameter $S(t)$ to $S_p(t)$ for age level p. We assumed three age levels when analyzing H1N1: children (from birth to 14), adults (15–64), and seniors (65 and older). Transmission rates between age levels were differentiated to capture the complexity of infections across age groups. We added two subscripts to transmission rate μ to create $\mu_{p,q}$: p for the age of an infectious individual, and q for the age of the individual being infected (Figure 4). Epidemic parameters used in population-based compartmental models were also used to model infections across age groups. We used three transmission rates and three removed rates, based on the number of individuals seeking medical attention.

To construct the layer 2 model, we revised Equations (2a) and (2b) to (3a) and (3b), respectively, without making any other changes to the Equation (2) subequations, as follows:

$$\frac{dS_p}{dt} = S_p\chi_p\chi_p(\beta_{2pp}I_{2p} + \beta_{11pp}I_{11p} + \beta_{12pp}I_{12p})/N_p - S_p\sum_{q\neq p}\chi_q\chi_p(\beta_{2qp}I_{2q} + \beta_{11qp}I_{11q} + \beta_{12qp}I_{12q})/N_p \tag{3a}$$

$$\frac{dL_p}{dt} = -\theta_p L_p + S_p\chi_p\chi_p(\beta_{2pp}I_{2p} + \beta_{11pp}I_{11p} + \beta_{12pp}I_{12p})/N_p + S_p\sum_{q\neq p}\chi_q\chi_p(\beta_{2qp}I_{2q} + \beta_{11qp}I_{11q} + \beta_{12qp}I_{12q})/N_p \tag{3b}$$

2.3. Layer 3: Commuting

For the present research we focused on the impacts of daily commuting networks on the spreading of an influenza virus. Since influenza viruses are transmitted via airborne droplets, commuter hosts are capable of

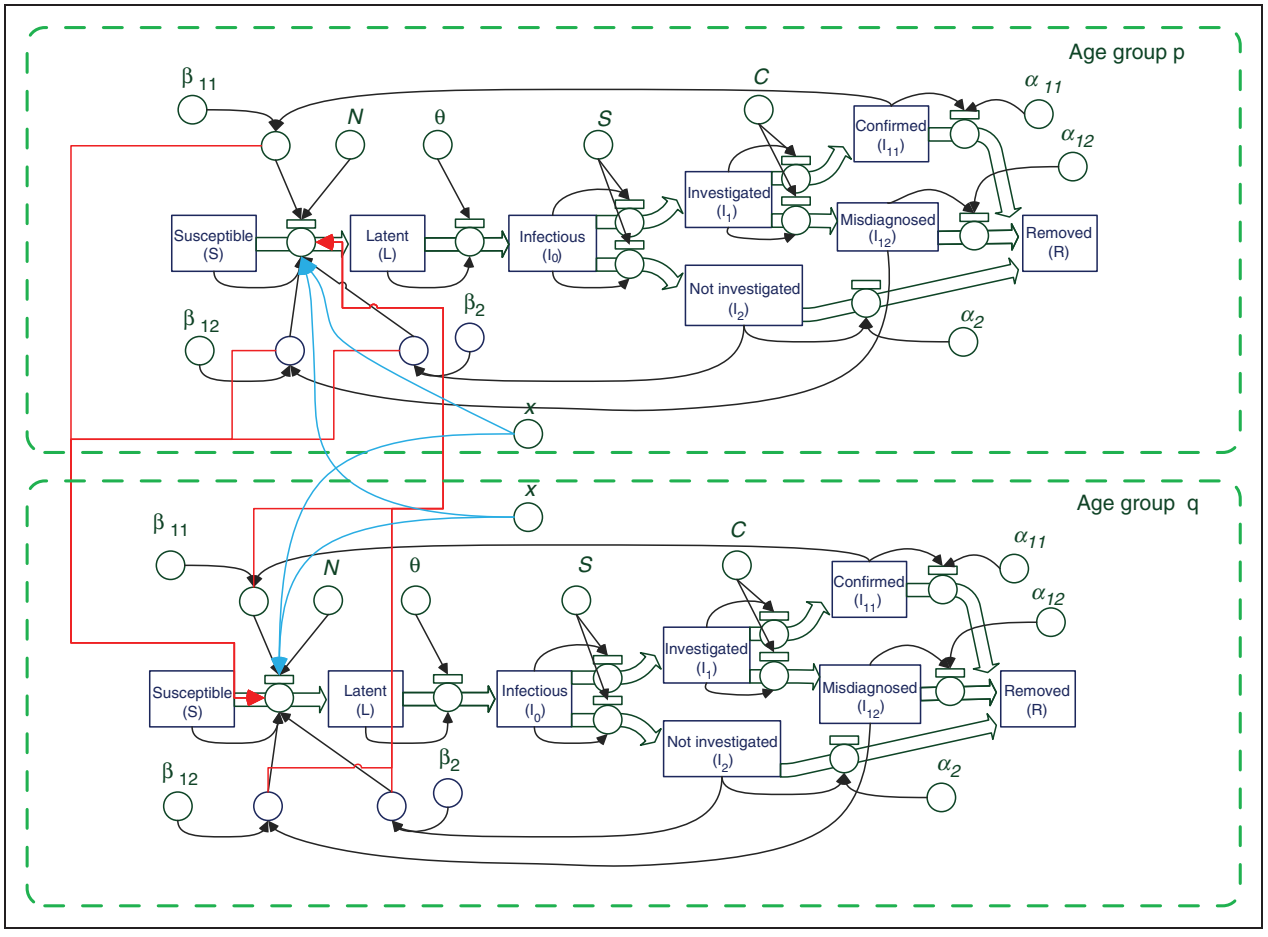


Figure 4. MEDSim layer 2 architecture flowchart. Red lines indicate parameters for other (non-p and non-q) age groups. Blue curves indicate relative percentages of each age group within the total population of each location. (Color online only).

infecting other individuals along their standard routes. The layer 3 model reflects two assumptions regarding hosts with jobs: they commute over longer distances than individuals who stay at home or travel to local centers such as schools, and they tend to come into contact with individuals in the same age group along their routes and at their destinations. We also assumed higher contact frequencies among individuals in more densely populated areas. Accordingly, the layer 3 model considers four features associated with travel between population centers (locations).

$\sigma(p)$, a binary value representing whether age level p is the commuter age level – that is, $\sigma(p) = \begin{cases} 1 & \text{if } p = \text{commutable age level} \\ 0 & \text{otherwise} \end{cases}$. We assumed that children and seniors are less likely than adults to commute on a daily basis, making adults the most likely carriers of pathogens between locations.

$w_{j,i}$, indicating how many individuals commute from location j to location i on a daily basis.

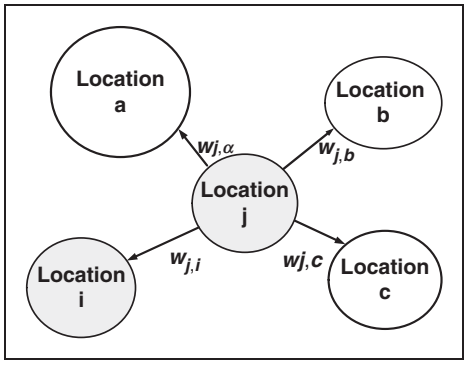


Figure 5. Potential movement of infectivity between locations *i* and *j*.

η^i, a weighting factor representing the average number of contacts among individuals in location i on a daily basis.

d^i, a normalized population density value for location i.

For all i and j locations in a w commuting network, we used the geodemographic weight shown as

Equation (4) to measure the effects of commuting on i and j population interactions:

$$\text{Geodemographic weiht}\ (j,i) = \sigma(p) d^i \eta^i \frac{w_{j,i}}{\sum_{k \neq j} w_{j,k}} \quad (4)$$

As shown in Figure 5, the commuting population age level in this example is adult (15–64). The $\sigma(p)$ function represents whether age group p is a traveling population. For all i locations in the commuting network, the term $N(i)$ represents the set of locations connected to location i within commuting network w. The term $w_{j,i}/\sum_{k \neq j} w_{j,k}$ is the ratio of commuters between locations j and i to commuters between j and all other locations. If location i is a large urban center, $w_{j,i}/\sum_{k \neq j} w_{j,k}$ will be large; if i is a suburb or rural location, it will be small. Public health policies involving transportation can be tested by changing contact rates among population centers in the layer 3 model.

The layer 3 framework is presented in Figure 6. To construct the layer 3 model, we revised Equations (3a) and (3b) to Equations (4a) and (4b), respectively. Note the addition of a geodemographic weight on the third line of each equation. All other Equation (2) subequations are the same.

$$\begin{aligned}\frac{dS_p^i}{dt} = & -S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i)/N_p^i \\ & - S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i)/N_q^i \\ & - \sum_{j \in N(i)} \sigma(p) d^i \eta^i S_p^i \frac{w_{j,i}}{\sum_{k \neq j} w_{j,k}} \\ & \times (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j)/N_p^j \end{aligned} \quad (4a)$$

$$\begin{aligned}\frac{dL_p^i}{dt} = & -\theta_p^i L_p^i + S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i)/N_p^i \\ & + S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i)/N_q^i \\ & + \sum_{j \in N(i)} \sigma(p) d^i \eta^i S_p^i \frac{w_{j,i}}{\sum_{k \neq j} w_{j,k}} \\ & \times (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j)/N_p^j \end{aligned} \quad (4b)$$

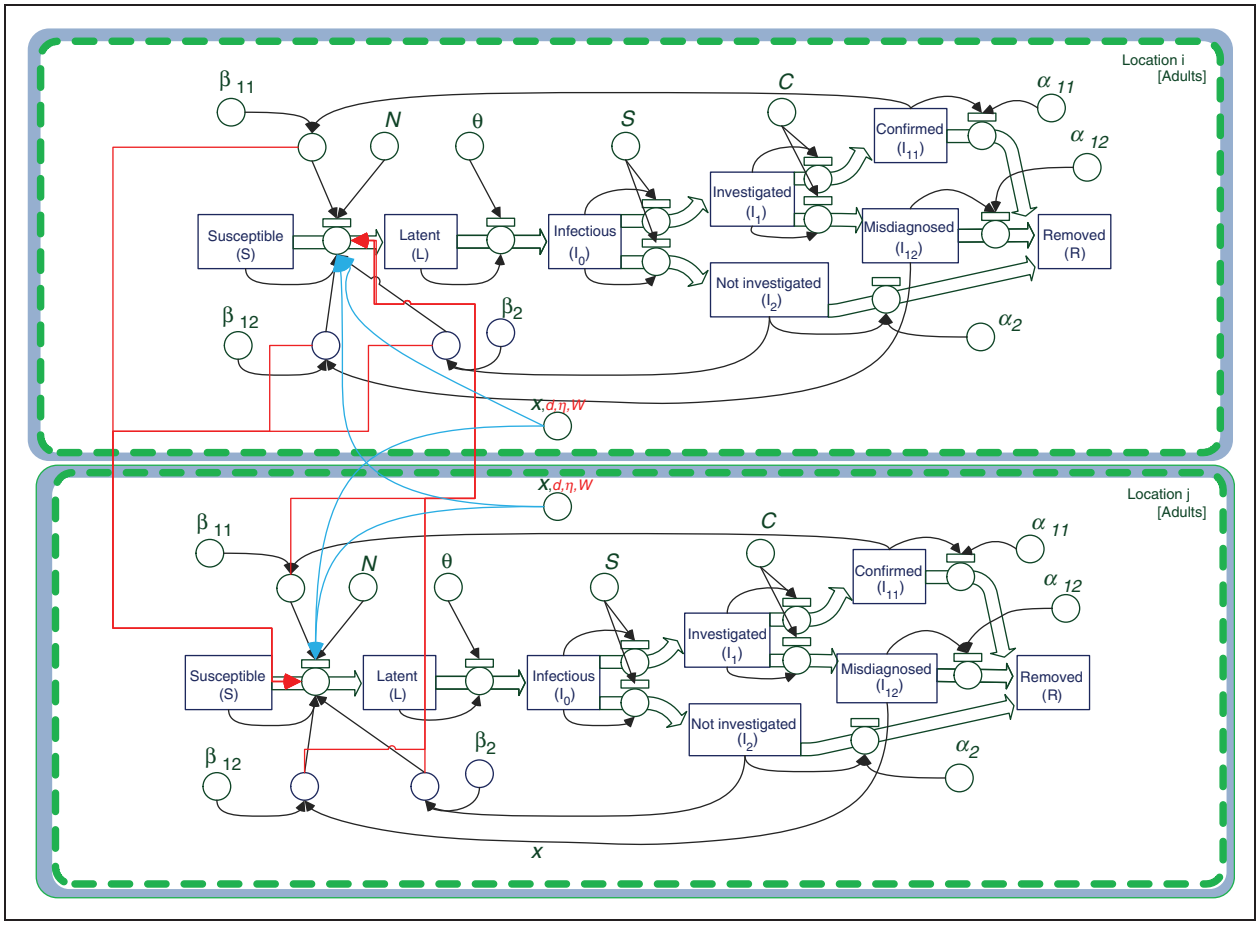


Figure 6. MEDSim layer 3 architecture flowchart. Properties associated with commuting between two locations are indicated by red lines. Additional location properties are indicated by blue lines. (Color online only).

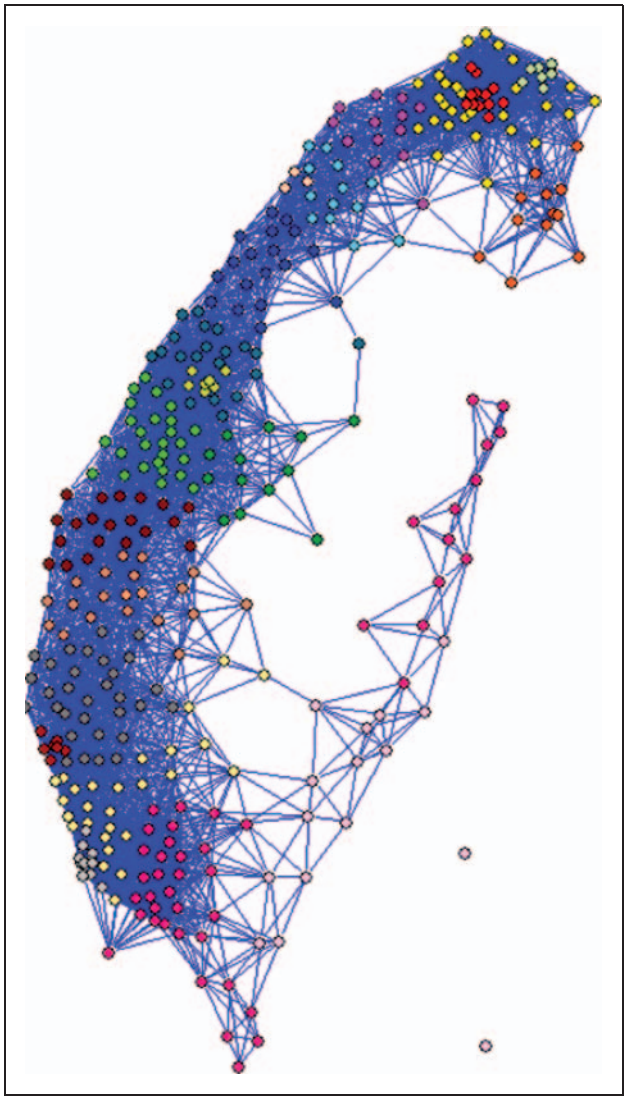


Figure 7. Taiwan's nationwide commuting network.

2.4. Layer 4: Nationwide interactions

We used Taiwan's national travel network and commuting weight $w_{j,i}$ to simulate individual movement within regions (layer 3). Nodes represent locations, and edges represent commuting weights between locations. Once transportation data are obtained, nodes can represent any scale – for instance, a building for city simulations and a town for regional or national simulations. In the present study, each node represents an individual town. Layer 4 of our model consists of 409 towns and 19,014 links (Figure 7) representing Taiwan's national commuting network, which can be manipulated to determine the effects of various movement policies and commuting restrictions.

After combining the four layers, the complete MEDSim model can be expressed as

$$\begin{aligned}\frac{dS_p^i}{dt} = & -S_p^i d^i \chi_p^i \chi_p^i (\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i)/N_p^i \\ & - S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i (\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i)/N_q^i \\ & - \sigma(p) d^i \eta^i S_p^i \sum_{j \in N(i)} \frac{w_{j,i}}{\sum_{k \neq j} w_{j,k}} \\ & \times (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j)/N_p^j \end{aligned} \tag{5a}$$

$$\begin{aligned}\frac{dL_p^i}{dt} = & -\theta_p^i L_p^i + S_p^i d^i \chi_p^i \chi_p^i \left(\beta_{2pp}^i I_{2p}^i + \beta_{11pp}^i I_{11p}^i + \beta_{12pp}^i I_{12p}^i\right)/N_p^i \\ & + S_p^i \sum_{q \neq p} d^i \chi_q^i \chi_p^i \left(\beta_{2qp}^i I_{2q}^i + \beta_{11qp}^i I_{11q}^i + \beta_{12qp}^i I_{12q}^i\right)/N_q^i \\ & + \sigma(p) d^i \eta^i S_p^i \sum_{j \in N(i)} \frac{w_{j,i}}{\sum_{k \neq j} w_{j,k}} \\ & \times (\beta_{2pp}^j I_{2p}^j + \beta_{11pp}^j I_{11p}^j + \beta_{12pp}^j I_{12p}^j)/N_p^j \end{aligned} \tag{5b}$$

$$\frac{dI_{0p}^i}{dt} = -I_{0p}^i + \theta_p^i L_p^i \tag{5c}$$

$$\frac{dI_{1p}^i}{dt} = -I_{1p}^i (t - T) + s_p^i I_{0p}^i \tag{5d}$$

$$\frac{dI_{2p}^i}{dt} = -\alpha_{2p}^i I_{2p}^i + (1 - s_p^i) I_{0p}^i \tag{5e}$$

$$\frac{dI_{11p}^i}{dt} = -\alpha_{11p}^i I_{11p}^i + c_p^i I_{1p}^i (t - T) \tag{5f}$$

$$\frac{dI_{12p}^i}{dt} = -\alpha_{12p}^i I_{12p}^i + (1 - c_p^i) I_{1p}^i (t - T) \tag{5g}$$

$$\frac{dR_p^i}{dt} = \alpha_{11p}^i I_{11p}^i + \alpha_{12p}^i I_{12p}^i + \alpha_{2p}^i I_{2p}^i \tag{5h}$$

The parameters used in Equation (5) are listed in Table 1.

2.5. Technological framework

Figure 8 shows the MEDSim technological framework, including a simulation flowchart, census databases, and relationships between the four MEDSim layers and the databases. The first step is to manually create an Excel data set for the scenario in question – for example, determining breakout locations or public health policies. Most data sets consist of spatial locations and census information, which are used to establish geographic and demographic categories; each MEDSim parameter belongs to at least one of the two.

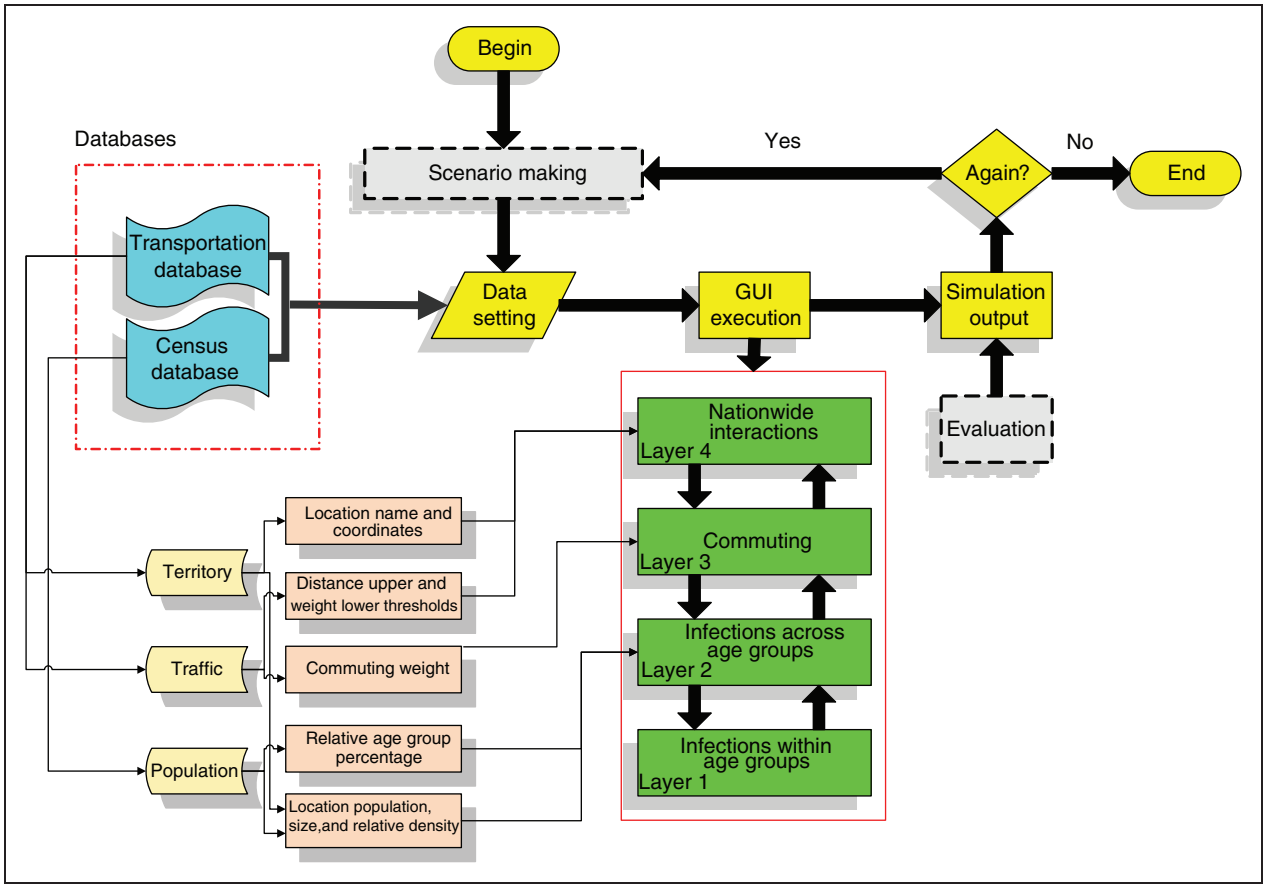


Figure 8. MEDSim simulation tool framework.

Since our layer 1 focus in this example is on disease progression at an individual level, standard expert-based parameters in compartmental models associated with epidemics were used instead of transportation or census databases.[16] In layer 2, percentages of individuals in each age group were determined from census data, and the numbers of individuals in each location were gathered from transportation databases. In layer 3, transportation databases were used to gather information on the numbers of individuals traveling between towns on a daily basis. In layer 4, transportation data were used to establish the underlying national travel network.

Figure 9 presents a screenshot of a MEDSim graphical user interface (GUI). Multilayer epidemic model parameters are initialized at the beginning of each simulation. Model parameters requiring setup are: (a) initial outbreak conditions, including the name of the town and number of infected persons in an age group identified by the surveillance system; (b) disease transmission parameters at different layers, including transmission, latent, and removed rates according to the SLIR process for each age group, contact rates between age groups, and regional contact probabilities between towns; and (c) output maps and charts for the towns of interest and severity indicators to be monitored (e.g. daily infected cases, daily new cases, and epidemic velocity and acceleration). Daily epidemic progress can be monitored in terms of sizes and locations of red dots on maps (infected individuals), epidemic curves on time charts, and output panels (numbers of infected individuals at different times in different locations). Regarding kernel execution, MEDSim models can be used for computing epidemic dynamics. Simulation results can be shown as graphical curves or expressed and recorded as numerical files. Lastly, simulation results are evaluated by users, who can repeat steps as required.

2.6. Statistical analysis for parameterization

To test the reliability and validity of time-series MEDSim data, we used two indices to compare simulated and actual numbers of infected individuals: correlation coefficient (CC) and coefficient of efficiency (CE), respectively expressed as Equations

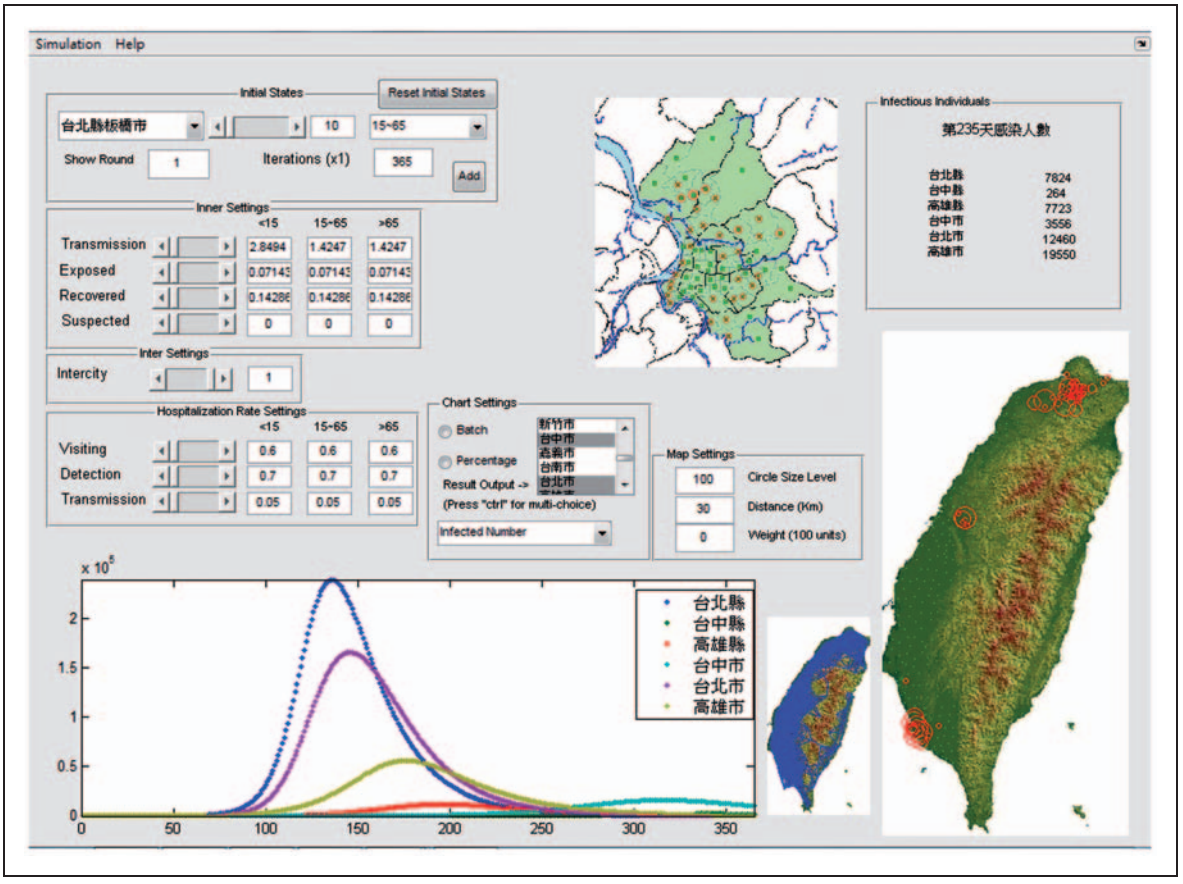


Figure 9. MEDSim implementation GUI.

(6) and (7).[22] Here $\{X_t|t = 1, 2, \ldots, n\}$ represents the number of actual infected individuals and $\{Y_t|t = 1, 2, \ldots, n\}$ the number calculated by the MEDSim. In both sets, t denotes the time step (in 1 week units); a total of n weeks is represented by each set. We use $\overline{X}$ and $\overline{Y}$ to denote the means of X_t and Y_t, respectively. The CC test measures data distance: higher positive values indicate positive correlations and lower negative values indicate negative correlations. The CE test is used to measure the level of accuracy between two data sets; higher values indicate greater accuracy:

$$CC = \frac{\sum_{t=1}^{n}(X_t - \overline{X})(Y_t - \overline{Y})}{\sqrt{\sum_{t=1}^{n}(X_t - \overline{X})^2}\sqrt{\sum_{t=1}^{n}(Y_t - \overline{Y})^2}} \in [-1, 1] \quad (6)$$

$$CE = 1 - \left[\frac{\sqrt{\sum_{t=1}^{n}(X_t - Y_t)^2}}{\sqrt{\sum_{t=1}^{n}(X_t - \overline{X})^2}}\right] \in [0, 1] \quad (7)$$

3. Results and discussion

We tested MEDSim reliability (in terms of parameter calibration and model fit) with actual epidemic curves, tested and compared public health policies based on the above parameters, and used MEDSim to simulate the influenza A (H1N1) virus and to determine the effects of the chosen policies. To establish simulation parameter settings, we used population data from the Republic of China (ROC) Ministry of the Interior and transportation data from the ROC Transportation Institute.[38]

3.1. Parameterization

We used the seasonal influenza A and swine-origin influenza A (H1N1) viruses to perform parameterization. The default parameter values are shown in Table 1. We systematically calibrated parameters for both viruses to create a small range, based on parameters normally used with standard SLIR settings.[16] Summaries of MEDSim attribute settings and values are given in Tables 2 and 3. The transmission rates β^i_{11pp}, β^i_{12pp}, and β^i_{2pp} were directional between age groups. Individual age group targets are presented in

Table 2. MEDSim parameters used for fitting simulation curves with actual seasonal influenza A curves in Taiwan between September 2008 and April 2009

		Value					
Layer	Attribute	Children		Adults		Seniors	
1	β^i_{11pp}	1.3333		0.6667		0.6667	
	β^i_{12pp}	3.3333		1.6667		1.6667	
	β^i_{2pp}	3.3333		1.6667		1.6667	
	θ^i_p			0.0714			
	α^i_{11p}			0.1429			
	α^i_{12p}			0.2500			
	α^i_{2p}			0.1429			
	Target	Adults	Seniors	Children	Seniors	Children	Adults
2	β^i_{11xy}	0.6667	0.6667	1.3333	0.6667	1.3333	0.6667
	β^i_{12xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667
	β^i_{2xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667

Table 3. MEDSim parameters used for fitting simulation curves to actual swine-origin influenza A (H1N1) curves in Taiwan from week 25 to week 52

		Value					
Layer	Attribute	Children		Adults		Seniors	
1	β^i_{11pp}	2.6667		1.3333		1.3333	
	β^i_{12pp}	3.3333		1.6667		1.6667	
	β^i_{2pp}	3.3333		1.6667		1.6667	
	θ^i_p			0.0714			
	α^i_{11p}			0.3333			
	α^i_{12p}			0.1429			
	α^i_{2p}			0.1667			
	Target	Adults	Seniors	Children	Seniors	Children	Adults
2	β^i_{11xy}	1.3333	1.3333	2.6667	1.3333	2.6667	1.3333
	β^i_{12xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667
	β^i_{2xy}	1.6667	1.6667	3.3333	1.6667	3.3333	1.6667

the form of sub-columns. Experimental results from applying the MEDSim using the Table 2 and 3 parameter values for the two influenza viruses are shown in Figures 10(a) and (b), respectively. Actual and simulated case data for both influenzas are shown in weekly units.

Our CC and CE results for the two influenza epidemics are 0.86 and 0.74 for seasonal and 0.77 and 0.36 for swine-origin H1N1. In Figure 10(a) we plotted the fractions of new infected cases of seasonal influenza A in Taiwan between September 2008 and April 2009, normalized to total cases. Higher CC and CE values for seasonal influenza explain the similarities between the two curves. In Figure 10(b) we plotted fractions of new infected cases for the swine-origin influenza A virus in Taiwan from week 25 to week 52 in 2009, also normalized to total cases. As shown, the number of actual cases decreased between weeks 37 and 48, followed by an increasing trend, resulting in a lower CE value. This two-wave pattern is very similar to global diffusion patterns associated with international travel. Because we did not incorporate international travel at this stage, our swine-origin H1N1 model failed to capture the second wave; however, it did capture the peak time for the first (primary) wave (Figure 10(b)).

3.2. Intervention policy evaluation

We tested and compared different public health policies using the above-described parameters, simulated the

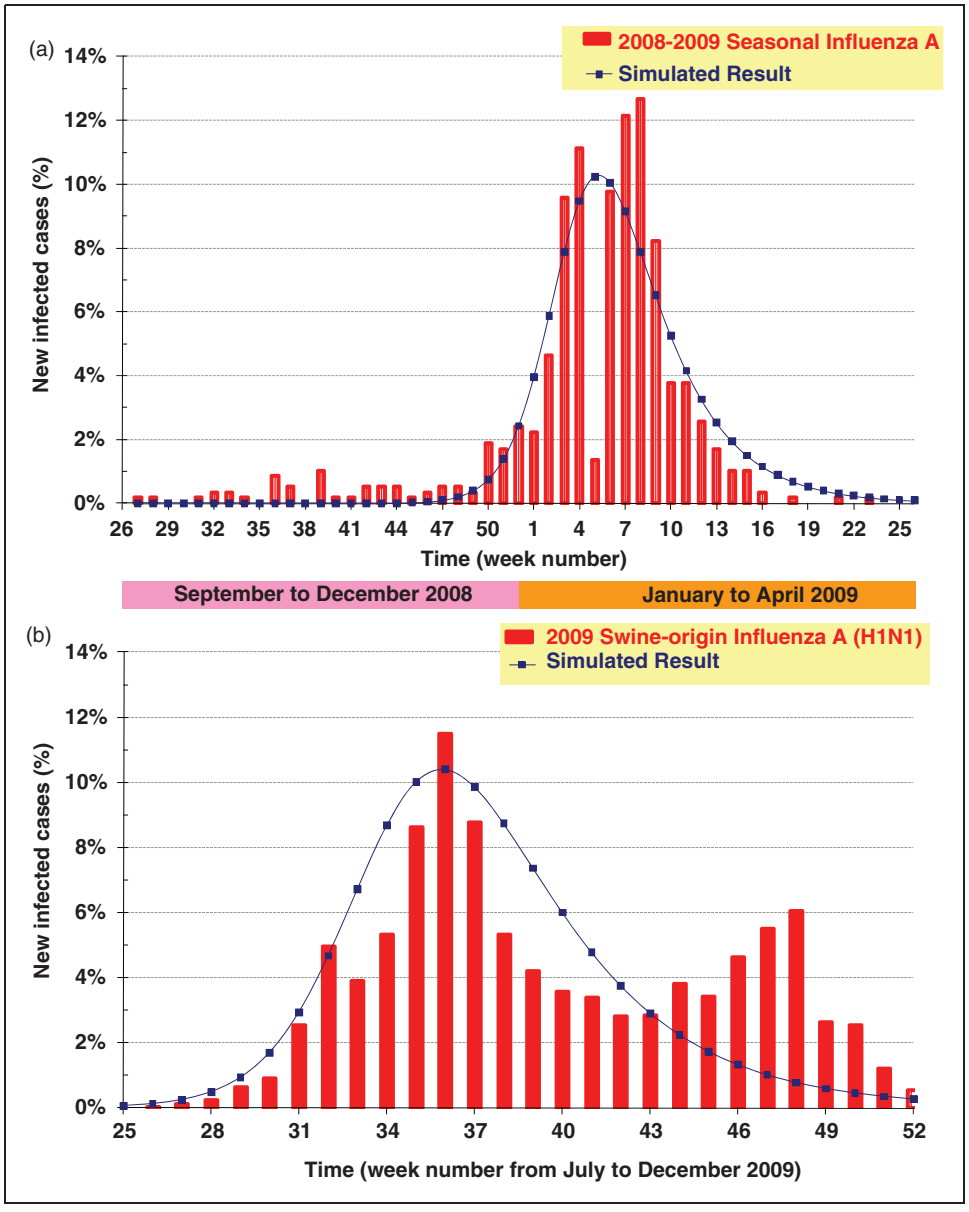


Figure 10. Comparison of weekly new infected cases between actual and simulated results normalized for (a) seasonal influenza A and (b) swine-origin H1N1 influenza A.

Table 4. Observation index values according to different transmission rates

Observation index	Transmission rate reduction (%)				
	0%	30%	50%	70%	90%
Total cases.	1,784,044	1,407,752	1,108,520	485,761	8
New infected cases at epidemic curve peak	171,329	113,898	64,926	12,231	8
Week number of epidemic curve peak	20	26	36	77	∞
Percentage of new infected cases at epidemic curve peak	9.6%	6.4%	3.6%	0.7%	0%
$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	78.9%	62.1%	27.2%	0.%

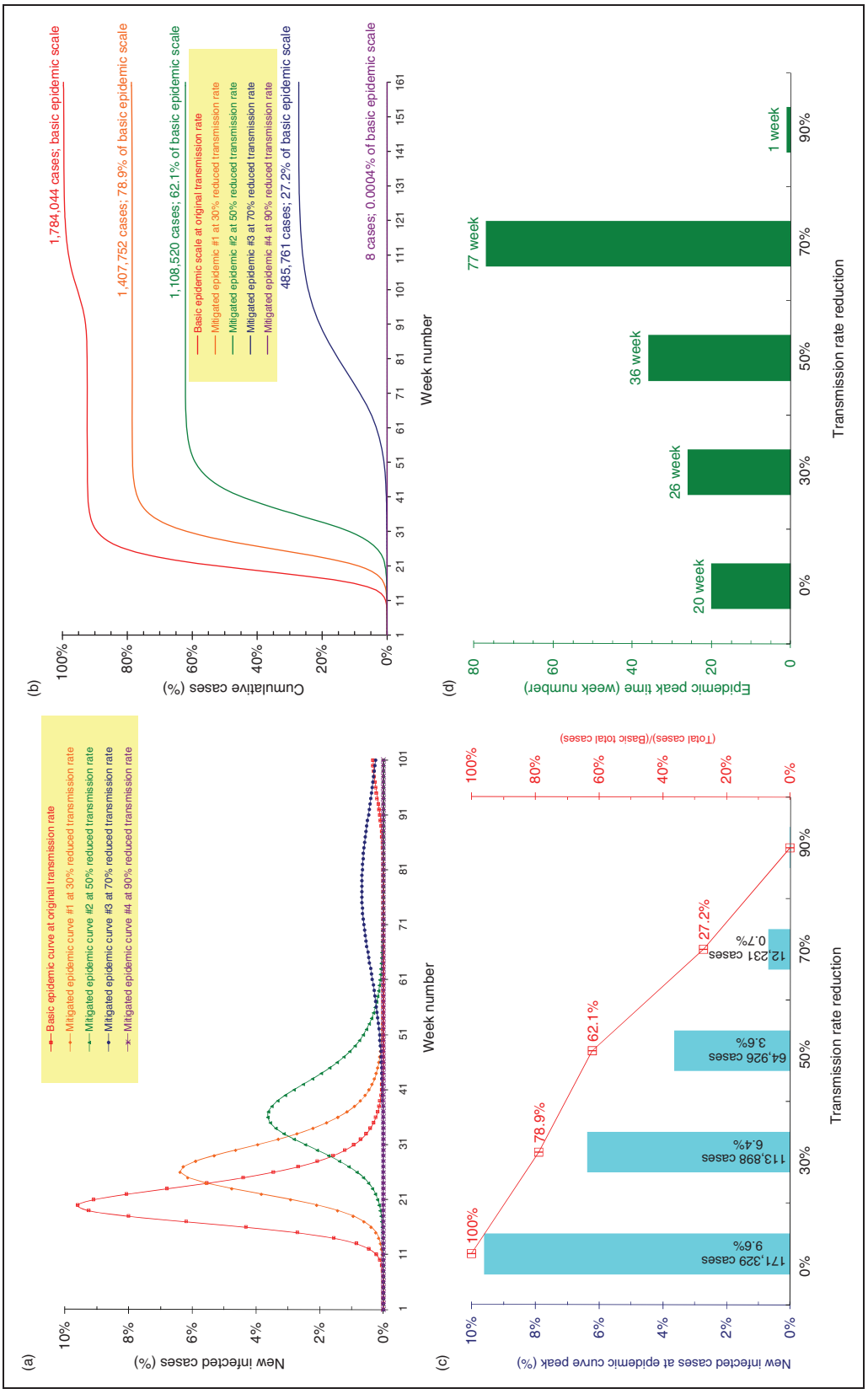


Figure 11. (a) New infected cases per week at different transmission rates. (b) Cumulative new infected cases at different transmission rates. (c) Basic epidemic curve at a 0% reduced transmission rate expressed according to two observation indexes. (d) Comparison of new infected cases at epidemic curve peak at different transmission rates. (e) Weekly new cases at curve peak at different transmission rates. (f) New infected cases at epidemic curve peak according to various intervention policy scenarios. (g) Numbers of infected cases according to various intervention policy scenarios. (h) Week numbers of epidemic curve peaks according to various intervention policy scenarios. (Color online only).

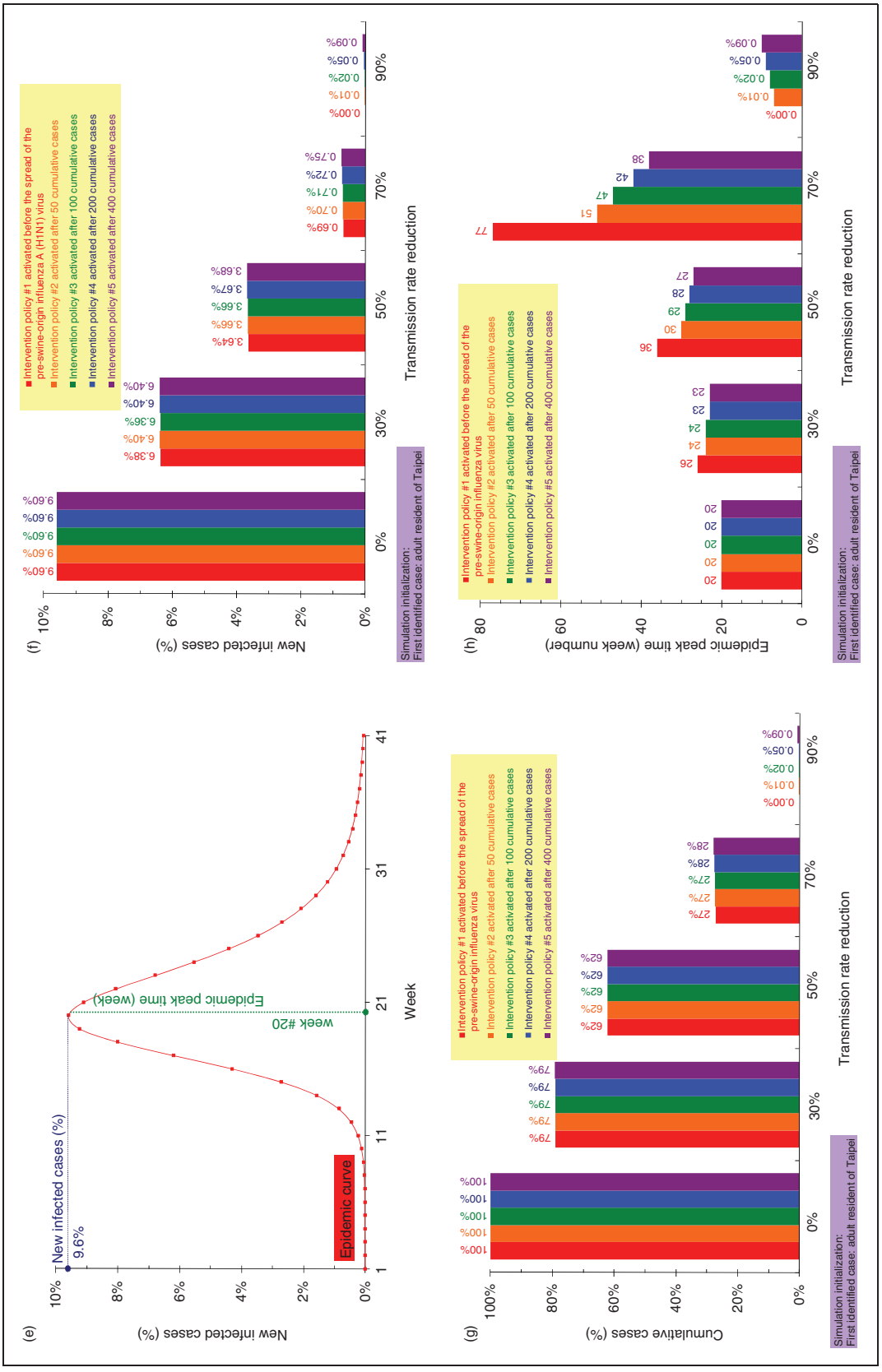


Figure 11. Continued.

effects of medical advice quality and number of commuters, and compared original epidemic curves with those following public health policy implementation. Special emphases were placed on peak numbers of infected cases and peak infection days. The goals of public health officials include reducing the peak number (since it has a direct effect on social costs, such as drugs and hospital beds) and delaying peak day.

Figure 11 has two parts, one addressing the impacts of transmission rate reduction and one the effects of various intervention policies. The results from simulated observation indices for different transmission rates are shown in Table 4. According to the Figure 11(a) data for weekly fractions of new infected cases, both curve peak and height were negatively affected by decreased transmission rate. According to the Figure 11(b) data on the cumulative number of new infections at different transmission rates, that number decreased as transmission rate decreased. In Figure 11(c) we used two observation indices to distinguish between the epidemic curve produced by the highest transmission rate and the curves shown in Figure 11(a). According to the first observation index (fraction of new infected cases at epidemic curve peak), the strongest epidemic disease transmission intensity affects a population and negatively impacts public health resources over a period of one week. The second index (epidemic curve peak week number) indicates the severity and urgency of an epidemic, thus impacting deadlines for initiating public health policies; higher values indicate more time for making policy decisions.

The results from our comparisons of epidemic curve peaks at different reduced transmission rates are shown in Figure 11(c). The basic fraction of new infected cases at curve peak (noted as 100%, with a transmission rate of 1.0) is shown in the leftmost part of the graph. The relative total number of cases (red line) consists of two line segments, one from 1.0 to 0.5 and the other from 0.5 to 0.1. According to this result, transmission rate should be reduced by at least 50% to obtain better peak number suppression. An obvious decrease in peak number occurs when the transmission rate reduction is 70%.

Curve peak week numbers at different transmission rates are shown in Figure 11(d). Note that week number increased as transmission rate decreased – a positive result for public health policy makers. The results from simulations of various long- and short-term intervention policy activation scenarios are shown in Figures 11(f)–(h). No differences in numbers of infected cases were observed for different intervention policy activation times (Figures 11(f) and (g)). However, epidemic peak was delayed from weeks 55 to 71 when intervention policy activation time was set at 50 with a 70% reduction in transmission rate (Figure 11(h)). Activation time exerted a much weaker effect on peak timing at a 30% reduction in transmission rate. According to these results, while time of intervention policy activation did not significantly reduce the number of infected cases, it did exert an obvious effect in terms of delaying peak time – a positive result for public health policy determination and preparation.

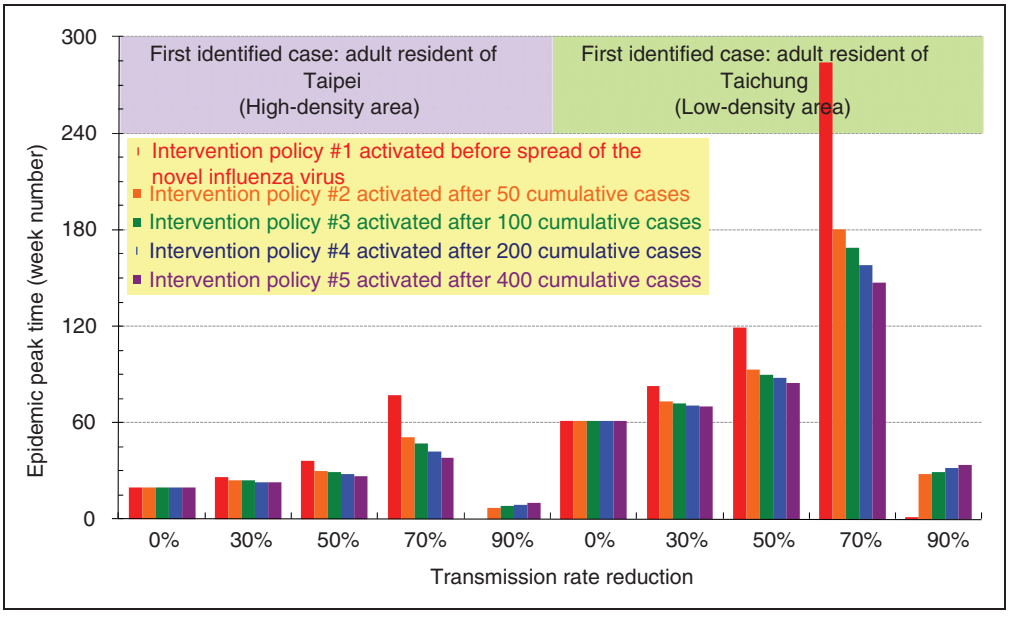


Figure 12. Epidemic peak week numbers for urban and rural areas.

Table 5. Observation index values according to different policy activation scenarios during swine-origin H1N1 influenza A outbreak in Taipei

		Transmission rate reduction				
Policy activation time	Observation index	0%	30%	50%	70%	90%
Scenario #1 Pre-virus appearance	Total cases	1,784,044	1,407,752	1,108,520	485,761	8
	New infected cases at epidemic curve peak	171,329	113,898	64,926	12,231	8
	Week number of epidemic curve peak	20	26	36	77	0
	Percentage of new infected cases at epidemic curve peak	9.60%	6.38%	3.64%	0.69%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	78.90%	62.14%	27.23%	0.00%
Scenario #2 After 50 cases are diagnosed	Total cases	Same as above	1,409,827	1,108,794	487,425	855
	New infected cases at epidemic curve peak		114,120	65,235	12,468	155
	Week number of epidemic curve peak		24	30	51	7
	Percentage of new infected cases at epidemic curve peak		6.40%	3.66%	0.70%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		79.02%	62.15%	27.32%	0.05%
Scenario #3 After 100 cases are diagnosed	Total cases	Same as above	1,410,263	1,108,993	488,900	1,991
	New infected cases at epidemic curve peak		113,532	65,314	12,604	349
	Week number of epidemic curve peak		24	29	47	8
	Percentage of new infected cases at epidemic curve peak		6.36%	3.66%	0.71%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		79.05%	62.16%	27.40%	0.11%
Scenario #4 After 200 cases are diagnosed	Total cases	Same as above	1,410,782	1,109,355	491,563	4,599
	New infected cases at epidemic curve peak		114,191	65,442	12,883	818
	Week number of epidemic curve peak		23	28	42	9
	Percentage of new infected cases at epidemic curve peak		6.40%	3.67%	0.72%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		79.08%	62.18%	27.55%	0.26%
Scenario #5 After 400 cases are diagnosed	Total cases	Same as above	1,411,273	1,109,893	496,246	10,000
	New infected cases at epidemic curve peak		114,185	65,669	13,408	1,680
	Week number of epidemic curve peak		23	27	38	10
	Percentage of new infected cases at epidemic curve peak		6.40%	3.68%	0.75%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		79.11%	62.21%	27.82%	0.56%

Next, we compared differences in swine-origin H1N1 influenza A starting locations in Taiwan and their effects on the subsequent spreading of the disease (Figure 12, Tables 5 and 6). Taipei was labeled a high-density area and Taichung a low-density area. In the first (pre-swine-origin virus) scenario, case numbers peaked much earlier in Taipei (20) than in Taichung (61). When the transmission rate was reduced to 30%, the Taichung peak was significantly delayed. When comparing numbers of infected cases at the curve peak, both locations had approximately the same number of new cases, but Taipei had a much larger number of total cases. After reducing the transmission rate from 50% to 30%, Taichung had a much later peak week compared to Taipei, with no effect of intervention policy activation time on the total number of cases or newly infected cases in either location. The results suggest that less densely populated starting locations are more sensitive to intervention policy activation time – that is, the combination of early activation time and low transmission rate significantly delays epidemic curve peaks in less densely populated locations.

Table 6. Observation index values according to different policy activation scenarios during swine-origin H1N1 influenza A outbreak in Taichung

		Transmission rate reduction				
Policy activation time	Observation index	0%	30%	50%	70%	90%
Scenario #1 Before the swine-origin influenza A (H1N1) virus emerges	Total cases	2,190,247	1,672,733	1,112,428	485,801	8
	New infected cases at epidemic curve peak	172,083	114,556	64,551	12,186	8
	Week number of epidemic curve peak	61	83	119	284	1
	Percentage of new infected cases at epidemic curve peak	7.86%	5.23%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$	100%	76.37%	50.79%	22.18%	0.00%
Scenario #2 After 50 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed	Total cases	(see above)	1,672,266	1,117,265	487,030	767
	New infected cases at epidemic curve peak		113,760	64,598	12,200	120
	Week number of epidemic curve peak		73	93	180	28
	Percentage of new infected cases at epidemic curve peak		5.19%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		76.35%	51.01%	22.24%	0.04%
Scenario #3 After 100 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed	Total cases	(see above)	1,671,019	1,120,702	488,492	1,723
	New infected cases at epidemic curve peak		113,672	64,430	12,194	273
	Week number of epidemic curve peak		72	90	169	29
	Percentage of new infected cases at epidemic curve peak		5.19%	2.94%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		76.29%	51.17%	22.30%	0.08%
Scenario #4 After 200 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed	Total cases	(see above)	1,674,627	1,125,289	491,418	3,668
	New infected cases at epidemic curve peak		113,592	64,556	12,198	520
	Week number of epidemic curve peak		71	88	158	32
	Percentage of new infected cases at epidemic curve peak		5.19%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		76.46%	51.38%	22.44%	0.17%
Scenario #5 After 400 cumulative swine-origin influenza A (H1N1) infected cases are diagnosed	Total cases	(see above)	1,677,338	1,132,127	49,486	7,424
	New infected cases at epidemic curve peak		112,155	64,605	12,188	1057
	Week number of epidemic curve peak		70	85	147	34
	Percentage of new infected cases at epidemic curve peak		5.12%	2.95%	0.56%	0%
	$\left(\frac{\text{Total cases of epidemic curve}}{\text{Total cases of basic epidemic curve}}\right)$		76.58%	51.69%	22.67%	0.34%

4. Conclusion

Our goal in this paper was to integrate complex human travel networks into a standard SLIR disease transmission model to create a four-layer simulation prototype named the MEDSim. The framework is offered to researchers interested in determining the contributions of complex human contact structures to the transmission dynamics of influenza viruses. Our proposed model is capable of providing insights that reflect the dynamic processes of epidemics according to various intervention scenarios involving outbreak location, intervention timing, and different policy suites. We view this multilayer approach as both convenient and effective for public health practitioners and administrators responsible for initiating early responses to potential pandemics, and for assessing intervention strategies in outbreak locations.

This study has several limitations, such as the lack of confirmed numbers of H1N1 influenza A cases in

Taiwan (at this time it is not a notifiable disease in this country). The data used for parameterization reflect severe and hospitalized cases, which we assume as having the same proportions as non-severe cases per time unit. Differences between actual and simulated cases can be significantly reduced when using appropriate parameter values in terms of investigation and detection proportions. Secondly, since the SLIR model is imprecise in terms of its *Removed* designation, we could not address the number of H1N1-related deaths in our discussion of peak time delay. In real-world scenarios involving pandemic diseases with high death rates, peak time delays are very important for disease prevention policy decisions. Thirdly, due to the limited scope of this study, we did not gather and organize the exceptionally large amounts of available data for all areas represented by network nodes (e.g. workplaces, houses, and schools) or network data for long-distance transportation (e.g. highway, railway, or airline). Instead, we used location and age for population grouping, based on their similarities in responses to epidemic-related factors. Furthermore, we did not address other individual attributes, such as income level or number of social groups per individual, when determining transmission rate, removed rate, or other parameters.

We believe our proposed MEDSim framework can help public health organizations decide when to implement intervention strategies by simultaneously analyzing multilayer interactions. To build on this positive beginning, we plan to expand the multilayer framework to make it suitable for other acute diseases, as well as to make it responsive to complex human contact structures. Although our focus in this pilot study was on a novel influenza epidemic in Taiwan, the general multilayer framework concepts can be transferred to other sites. The SLIR model in layer 1 can be considered a general model for all droplet-transmitted respiratory infections, and the age group and commuting interactions in layers 2 and 3 can be disassembled to meet the requirements of risk factors for other infectious diseases. Furthermore, the network topology in layer 4 can be modified to meet the needs of different scales of link-node structures as noted in an earlier section. However, when transferring the proposed multilayer framework to other sites, data on the link-node network structures and transmission parameters for the diseases being studied must be collected, organized, and verified. One of our goals is to establish a portable framework for this procedure. Our plans also include extending the MEDSim for use as a general-purpose disease modeling framework. For example, we will work on adding long-distance transportation networks to our model to determine the impacts of shutting down railway or airline systems, and on modifying different contact structures (e.g. mosquito–human) to model vector-borne diseases, such as dengue fever and malaria, as well as human–animal contact diseases, such as rabies and Japanese encephalitis.

Funding

This research was supported by a grant from the ROC National Science Council (NSC 98-2410-H-002-168-MY2 and NSC 98-2314-B-182-043). The authors are also grateful for financial support provided by the Infectious Disease Research and Education Center at National Taiwan University and Department of Health, Executive Yuan, ROC (Taiwan). The funders had no role in study design, data collection and analysis, or preparation of the manuscript.

Conflict of interest statement

The Authors declare no conflict of interest related to this work.

References

1. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 2009; 459: 1122–1125.
2. Feng Z, Huang W and Castillo-Chavez C. Global behavior of a multi-group SIS epidemic model with age structure. *J Differ Equat* 2005; 218: 292–324.
3. Inaba H. Age-structured homogeneous epidemic systems with application to the MSEIR epidemic model. *J Math Biol* 2007; 54: 101–146.
4. Langlais M and Naulin JM. An age structured SI epidemic problem in a heterogeneous environment. *Evol Equat Appl Phys Ind Life Sci Econ* 2003; 55: 307–321.
5. Shim E, Feng Z, Martcheva M and Castillo-Chavez C. An age-structured epidemic model of rotavirus with vaccination. *J Math Biol* 2006; 53: 719–746.
6. Supriatna AK, Soewono E and Van Gils SA. A two-age-classes dengue transmission model. *Math Biosci* 2008; 216: 114–121.
7. Wang W and Zhao XQ. An age-structured epidemic model in a patchy environment. *SIAM J Appl Math* 2005; 65: 1597–1614.
8. Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, Matrajt L, et al. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 2009; 326: 729–733. doi:10.1126/science.1177373.
9. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science* 2009; 324: 1557–1561.
10. Belshe RB, Swierkosz EM, Anderson EL, Newman FK, Nugent SL and Maassab HF. Immunization of infants and young children with live attenuated trivalent cold-recombinant influenza A H1N1, H3N2, and B vaccine. *J Infect Dis* 1992; 165: 727–732.
11. Marsh RM. *The great transformation: social change in Taipei, Taiwan since the 1960s*. New York: M.E. Sharpe, 1996.

12. González MC, Hidalgo CA and Barabási AL. Understanding individual human mobility patterns. *Nature* 2008; 453: 779.
13. Huang CY, Sun CT, Hsieh JL, Chen YMA and Lin H. A novel small-world model: Using social mirror identities for epidemic simulations. *Simulation* 2005; 81: 671–699.
14. Tomlinson B and Cockram C. SARS: experience at Prince of Wales Hospital, Hong Kong. *The Lancet* 2003; 361: 1486–1487.
15. Riley S. Large-scale spatial-transmission models of infectious disease. *Science* 2007; 316: 1298–1301.
16. Keeling MJ and Rohani P. *Modeling infectious diseases in humans and animals*, 1st edn. New Jersey: Princeton University Press, 2007.
17. Epstein JM. Modeling to contain pandemics. *Nature* 2009; 460: 687, doi:10.1038/460687a.
18. Liu X, Takeuchi Y and Iwami S. SVIR epidemic models with vaccination strategies. *J Theor Biol* 2008; 253(1): 1–11.
19. Pastor-Satorras R and Vespignani A. Immunization of complex networks. *Phys Rev E* 2002; 65: 036104.
20. Li G and Jin Z. Global stability of a SEIR epidemic model with infectious force in latent, infected and immune period. *Chaos Solitons Fractals* 2005; 25: 1177–1184.
21. Boguna M and Pastor-Satorras R. Epidemic spreading in correlated complex networks. *Phys Rev E* 2002; 66: 47104.
22. Huang CY, Sun CT, Hsieh JL and Lin H. Simulating SARS: small-world epidemiological modeling and public health policy assessments. *J Artif Soc Soc Simulat* 2004; 7: http://jasss.soc.surrey.ac.uk/7/4/2.html.
23. Ortiz-Pelaez A, Pfeiffer DU, Soares-Magalhães RJ and Guitian FJ. Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Prev Vet Med* 2006; 76: 40–55.
24. Newman MEJ. The structure and function of complex networks. *SIAM Rev* 2003; 45: 167–256, http://arxiv.org/abs/condmat/0303516.
25. Barabási AL and Albert R. Emergence of scaling in random networks. *Science* 1999; 286: 509.
26. Watts DJ and Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998; 393: 440–442. doi:10.1038/30918.
27. Erdos P and Renyi A. On the evolution of random graphs. *Publ Math Inst Hungarian Acad Sci* 1960; 5: 17–61.
28. Pastor-Satorras R and Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett* 2001; 86: 3200, doi:10.1103/PhysRevLett.86.3200.
29. Huang CY, Sun CT and Lin HC. Influence of local information on social simulations in small-world network models. *J Artif Soc Soc Simulat* 2005; 8: 8, http://jasss.soc.surrey.ac.uk/8/4/8.html.
30. Barrett CL, Eubank SG and Smith JP. If smallpox strikes Portland. *Sci Am* 2005; 292: 54–61.
31. Barrett CL, Eubank S, Kumar VSA and Marathe MV. Understanding large-scale social and infrastructure networks: a simulation-based approach. *SIAM News* 2004; 37(4): 1–5.
32. Barrett C., Eubank S., and Marathe M. Modeling and Simulation of Large Biological, Information and Socio-Technical Systems: An Interaction Based Approach. *Interactive Computation: The New Paradigm*, 1st edn. 353–392. Berlin: Springer-Verlag, 2006.
33. May RM and Lloyd AL. Infection dynamics on scale-free networks. *Phys Rev E* 2001; 64: 66112.
34. Draief M, Ganesh A and Massoulié L. Thresholds for virus spread on networks. *Ann Appl Probab* 2008; 18: 359–378.
35. Levin SA and Durrett R. From individuals to epidemics. *Phil Trans Biol Sci* 1996; 351: 1615–1621.
36. Davis GF, Yoo M and Baker WE. The small world of the American corporate elite, 1982–2001. *Strat Organ* 2003; 1: 301–326.
37. Sawyer RK. Artificial societies: multiagent systems and the micro-macro link in sociological theory. *Soc Meth Res* 2003; 31: 325–363.
38. Directorate-General of Budget, Accounting, and Statistics. *A preliminary report on year 2000 population and household census in Taiwan*. Taipei, Taiwan 2001.
39. Liu Z, Lai YC and Ye N. Propagation and immunization of infection on general networks with both homogeneous and heterogeneous components. *Phys Rev E* 2003; 67: 31911.
40. Huang CY, Tsai YS, Wen TH. A Network-Based Simulation Architecture for Studying Epidemic Dynamics. *Simulation* 2010; 86: 351–368.

Yu-Shiuan Tsai received his BS and MS in Mathematics from the National Taiwan University, Taiwan, in 2002 and 2005, respectively. He is currently a PhD candidate in the Department of Computer Science, National Chiao Tung University. His current research interests include complex networks and systems, social simulations, and computational epidemiology.

Chung-Yuan Huang received his MS in Computer Information and Science (2000) and his PhD in Computer Science (2005), both from the National Chiao Tung University, Taiwan. He is currently an Associate Professor in the Department of Computer Science and Information Engineering and a member of the Research Center for Emerging Viral Infections at Chang Gung University, Taiwan. His research interests include complex adaptive networks and systems, agent-based modeling and simulation for social science research, and computational epidemiology.

Tzai-Hung Wen received his PhD in Engineering (2006) from the National Taiwan University and is currently an Assistant Professor in the Department of Geography, College of Science, National Taiwan University, Taiwan. He is also a joint faculty member of the Infectious Diseases Research and Education Center at National Taiwan University and Department of Health, Executive Yuan, ROC (Taiwan). His research interests

cover applications of GISs and spatiotemporal modeling in infectious disease epidemiology and evaluation of control strategies for epidemics.

Chuen-Tsai Sun is currently a joint Professor in the Department of Computer Science and Graduate Institute of Education, National Chiao Tung University, Taiwan.

Muh-Yong Yen accomplished his training in infectious diseases at the General Veterans Hospital in 1987. He received further education at Columbia University, USA, in molecular biology, and at the National Sun Yat-Sen University, where he obtained an EMBA. From 2006 to 2009, he was the elected president of the Infection Control Society of Taiwan. He is now Deputy Superintendent of Taipei City Hospital, and Director of the Division for Disease Control and Prevention, Taipei City Government. His field of interests include emerging infectious diseases, infection control, and crisis management.