國立交通大學

資訊工程系

博士論文

小世界流行病學建模與公共衛生政策評估：利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

Small-World Epidemiological Modeling and Public Health Policy Assessment: Using the Social Mirror Identity Concept and Local Information for Network-based Epidemic Simulations

研究生：黃崇源

指導教授：孫春在　教授

中華民國九十四年十一月

小世界流行病學建模與公共衛生政策評估：利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

Small-World Epidemiological Modeling and Public Health Policy Assessment: Using the Social Mirror Identity Concept and Local Information for Network-based Epidemic Simulations

研究生：黃崇源　　Student：Chung-Yuan Huang

指導教授：孫春在　　Advisor：Chuen-Tsai Sun

國立交通大學
資訊工程系
博士論文

A Dissertation
Submitted to
Department of Computer Science
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in

Computer Science

November 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年十一月

國立交通大學

博碩士論文全文電子檔著作權授權書

（提供授權人裝訂於紙本論文書名頁之次頁用）

本授權書所授權之學位論文，為本人於 國立交通 大學 資訊工程 系所 組， 九十四 學年度第 一 學期取得 博 士學位之論文。

論文題目：小世界流行病學建模與公共衛生政策評估：利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

指導教授：孫春在教授

■ 同意 □不同意

本人茲將本著作，以非專屬、無償授權國立交通大學與台灣聯合大學系統圖書館：基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立交通大學及台灣聯合大學系統圖書館得不限地域、時間與次數，以紙本、光碟或數位化等各種方法收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。

論文全文上載網路公開之範圍及時間：	
本校及台灣聯合大學系統區域網路	■ 中華民國　年　月　日公開
校外網際網路	■ 中華民國　年　月　日公開

授 權 人：黃崇源　　學號：8923802

親筆簽名：

中華民國 九十四 年 十一 月 十五 日

國立交通大學

博碩士紙本論文著作權授權書

（提供授權人裝訂於全文電子檔授權書之次頁用）

本授權書所授權之學位論文，為本人於國立交通大學資訊工程系所 ______組，九十四 學年度第 一 學期取得博士學位之論文。

論文題目：小世界流行病學建模與公共衛生政策評估：利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

指導教授：孫春在教授

■ 同意

本人茲將本著作，以非專屬、無償授權國立交通大學，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立交通大學圖書館得以紙本收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行閱覽或列印。

本論文為本人向經濟部智慧局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為：____________________，請將論文延至____年____月____日再公開。

授 權 人：黃崇源

親筆簽名：____________________

中華民國 九十四 年 十一 月 十五 日

小世界流行病學建模與公共衛生政策評估：
利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

學生：黃崇源　　　　　　　　　　　　　　　　指導教授：孫春在博士

國立交通大學資訊工程學系（研究所）博士班

摘　　要

本論文的主旨在探討電腦建模與模擬在社會科學領域裡的重要性、優缺點、運用方式、各項驗證標準與評估程序。本論文擬以電腦建模與模擬為核心重點，詳細介紹二項原創性的社科模擬研究。首先，第一項研究運用電腦建模與模擬來探討流行性傳染病的傳播動態。此研究提出一套新穎的社會分身點概念來表示長距離移動、日常定點活動與多個活動點等重要的社會現象。並結合細胞自動機，建構一個能夠描述日常生活接觸與人際互動模式的小世界社會網絡模擬模型。此研究模擬2003年初在台北、新加坡與多倫多（加拿大）三地的SARS疫情，模擬結果與實際的疫情資料具有高度的一致性，充分吻合當地疫情的發展趨勢與特色。另外，此研究也證明該模擬模型非常適合探討與流行性傳染病有密切相關的公衛政策與防疫策略等。其次，本論文在第二項研究中，針對近來非常熱門的小世界社會網絡模擬模型，提出一個傳播問題模型，並設計一個敏感度分析實驗，來探討數種區域差異對於模擬歷程與結果的影響，並藉此檢視各種區域差異的影響力與敏感度。最後，本論文期望透過上述二項研究來充分掌握電腦建模與模擬的特性與潛力，釐清它的操作程序與潛力，進而探討它的意義、限制與應用範疇。

Small-World Epidemiological Modeling and Public Health Policy Assessment: Using the Social Mirror Identity Concept and Local Information for Network-based Epidemic Simulations

student：Chung-Yuan Huang Advisors：Dr. Chuen-Tsai Sun

Department（Institute）of Computer Science
National Chiao Tung University

ABSTRACT

The purpose of this dissertation is to explore the importance, advantages, applications, validation standards, and evaluation procedures of computational modeling and simulation in the social sciences. I will present two original social simulation studies in detail. In the first, computational modeling and simulation are used to investigate the transmission dynamics of epidemics and to evaluate the effectiveness of various public health policies and epidemic prevention strategies. A novel social mirror identity concept is proposed to represent social phenomena such as individual long-distance movement, daily visits to fixed locations, and multiple activity locations. The cellular automata concept is also utilized to construct a small-world social network model that represents human interactions and daily contacts. To test these concepts, I simulated the 2003 SARS outbreaks that occurred in Taipei, Singapore, and Toronto. The simulation results, which were highly consistent with actual epidemic data, corresponded with local outbreak trends and features. The simulation model was shown to be suitable for investigating public health policies and epidemic prevention strategies. In the second study I investigated the influence of local information on social simulations based on a small world model. I introduced a cellular automata-based variation with added shortcuts as a test platform for simulating the spread of an epidemic disease, then examined the influences of various local information factors on the results. It is my hope that these efforts will help future researchers determine appropriate simulation parameters, clarify operational procedures, and access meaning from simulations.

Acknowledgements

I am deeply indebted to my advisor, Professor **Chuen-Tsai Sun**, for his continuous support and intellectual suggestions. All my life, I will remember my research result from his considerate instruction. I am also very grateful to my dear friends in the Laboratory for Learning Sciences and Technology, **Ji-Lung**, who makes up a creative studying environment and really help me overcome many difficulties. Finally, I would like to thank my mother, **Xue-Yun Kuo**, and my lovely fiancée, **Chuan-Chuan Shih**. For their love and encouragement throughout the years, I can express only an inadequate acknowledgement of my appreciation. I dedicate this dissertation to them.

Contents

List of Tables

List of Figures

國家圖書館博碩士論文電子檔案上網授權書

ID:GT008923802

本授權書所授權之論文為授權人在國立交通大學資訊學院資訊工程系所＿＿＿＿＿組九十四學年度第一學期取得博士學位之論文。

論文題目：小世界流行病學建模與公共衛生政策評估：利用社會分身點概念與區域資訊建構社會網路式流行病學電腦模擬

指導教授：孫春在教授

茲同意將授權人擁有著作權之上列論文全文（含摘要），非專屬、無償授權國家圖書館，不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

※ 讀者基於非營利性質之線上檢索、閱覽、下載或列印上列論文，應依著作權法相關規定辦理。

授權人：黃崇源

親筆簽名：＿＿＿＿＿＿＿＿

民國 94 年 11 月 15 日

1. 本授權書請以黑筆撰寫，並列印二份，其中一份影印裝訂於附錄三之二(博碩士紙本論文著作權授權書)之次頁；另一份於辦理離校時繳交給系所助理，由圖書館彙總寄交國家圖書館。

Chapter 1. Introduction

Social scientists primarily use computers in two ways: as work platforms for programs that facilitate general tasks (e.g., word processing and statistical analysis) and as gateways to the “information superhighway” that has greatly increased the speed and volume of scholarly discourse. Until very recently, social scientists have not taken full advantage of computers, especially as tools for research or data analysis (Garson, 1994). On the other hand, a small number of social scientists have applied computational modeling and simulation to study social issues since the beginning of the computer revolution. Those who initially published interesting and important research results include Abelson and Carroll (1965), Coleman (1962), Forrester (1969), Guetzkow and Bowes (1957), Hermann and Hermann (1967), McPhee (1963), and Meadows et al. (1972). During the past decade, a different generation of social scientists has used computational modeling and simulation to explore social issues; its members include Epstein and Axtell (1996), Garson (1994), Gilbert and Troitzsch (1999), Hannon and Ruth (1994), Hoover and Perry (1990), Kheir (1988), Leik and Meeker (1995), and Whicker and Sigelman (1991). Three contributing reasons for this renewed interest are:

1. Modern computers possess much stronger computational capabilities and

greater reliability at a much lower price. Today's personal computers (PCs) possess many times more computational power and memory than "supercomputers" of decades past, yet are cheap enough for students and families to purchase.

2. The combination of advanced technology, operating system interfaces, software packages, and computer programming languages have made current computer systems very stable and easy to operate. Early social scientists who used computational modeling and simulation had to be well versed in assembly languages to build simulation models, and were occasionally required to manually replace PCBs or circuits. Today's computer systems and programming tools are no longer considered cryptic, allowing social scientists with rudimentary programming skills to apply specific software and development packages to establish models.

3. Computational modeling and simulation methods and technology are rapidly evolving—for example, genetic algorithms (GAs) and small-world networks (SWNs) can now be used to analyze how social culture evolves and adjusts to various factors. Social scientists are thus able to make up for the inadequacy of traditional research approaches in examining social phenomena and processes.

1.1. Models, Computational Modeling, and Simulation Models

Defining "model" is important for facilitating later discussion. In this dissertation I will view a model as a representation of a theory—that is, a scientific way to describe a real-world phenomenon. Models are integral to the development of theoretical understanding in the sense that they not only fill gaps between theory and empirical data, but also enhance interactions and influences between the two. Models allow researchers to utilize empirical data regularities to guide theory development (induction) and to apply theory to empirical analysis (deduction).

Three languages are used in the social sciences to express theories or build models: natural language, mathematics (including statistics and logic), and computational modeling and simulation (Ostrom, 1988). Turing (1950) originally claimed that all formal languages are equivalent (i.e., that natural language, classical mathematics, and computational modeling are interchangeable), but in the 1950s it was impossible to predict how computational modeling and simulation would one day be used to explore social issues. Today, certain social issues are better dealt with using mathematical tools (e.g., proving theorems), meaning that computational modeling and simulation can be used to express social science theories. Modeling and simulations are thus gaining greater acceptance.

Simulation models are models that have been transformed into computer programs (also called simulation systems). The majority of computational models and simulations entail five steps: a) developing a theory, b) transforming the theory into a simulation model, c) validating the simulation model, d) executing the model and observing its processes and outcomes, and e) regularly refining the simulation model and theory. In the system that is the focus of this dissertation, Huang, Sun, Hsieh, and Lin (2004) established the social mirror identity concept to describe long-distance movement, daily visits to fixed locations, and multiple activity locations to simulate the transmission dynamics of SARS outbreaks at different locations in early 2003. I will describe how a combination of the social mirror identity concept, cellular automata, and small world characteristics were used to construct a simulation model that fully depicts daily contact and human interactions within social networks.

1.2. Why Use Computational Models for Social Science Simulations?

Social science analytical tools include qualitative case studies, quantitative data analyses, and mathematical modeling. In all cases, research approaches must be applied with flexibility in order to accumulate scientific knowledge (King, Keohane, and Verba 1994). Until recently, it was not considered possible to use a single research approach to address all social science research problems; instead, social scientists have had to integrate multiple approaches according to the situation and the collected data. In section 1.2 I will examine whether the combination of computational modeling and simulation actually constitutes a new social science research approach (Byrne 1997; Conte & Gilbert 1995; Halfpenny 1997; Hanneman & Patrick 1997; Leik & Meeker 1995), what advantages it offers, and why it is considered “the third scientific discipline” of the social sciences (Axelrod 1997; Ilgen & Hulin 2000).

Using formal models instead of natural language models to investigate social issues has several advantages that include a) better definitional and conceptual precision, b) assumption clarity, c) ease of determining internal or logical validity, d) formal deduction power, and e) reduced ambiguity in formal communication among researchers. The major constraint of formal models is that their creators often have to simplify them in order to construct models that are sufficiently easy to control and analyze; as a result, their models can become too impractical or optimistic. In order to apply formal models to investigate certain social processes or issues, social scientists have to make trade-offs between “ease of analysis” and “accuracy of prediction.” In this dissertation I will assume that both are of equal importance.

In many ways, computational modeling and simulation allows researchers to have their cake (in terms of ease of analysis) and eat it, too (in terms of prediction accuracy). These tools allow social scientists to establish simulation models that are the equivalent of formal theoretical models. Not only are the advantages of formal models retained in computational models, they also avoid many of the constraints just mentioned—for instance, sacrificing ease of analysis for accuracy of prediction or vice versa (Johnson 1999). However, computational modeling and simulation should not be blindly applied to all social processes and issues; there are instances where formal models or other approaches are sufficient for problem representation. Note also that even though simulation models constructed via computational modeling convey more complex and complete information on social phenomena and characteristics, they are still simplifications of complex social issues.

1.3. Advantages of Social Science Computational Modeling and Simulation

As Hastie (1988) and Taber and Timpone (1994) have noted, there are many advantages to using computational modeling and simulation to study social processes. Similar to mathematical modeling and deduction, computational modeling and simulation demands accuracy and clarity. In theory, both approaches allow models to deal with inaccuracy, but mathematical models are less flexible when it comes to dealing with the inaccuracy and randomness that constantly occur in the real world. In contrast, researchers who use computational modeling and simulation approaches can add random components (i.e., hypothesize several scenarios, embed the scenarios in the simulation model, and observe their running processes or outcomes) at a low cost in terms of computing and time resources. Moreover, qualitative concepts that are difficult to express in a typical mathematical model can be demonstrated using various data structures and simulation programming techniques. Most concepts and relations that can be expressed in mathematical models are easily transferred to simulation models; however, the opposite is not true, since it requires a sharp increase in the number of required statements.

While I believe that social scientists should utilize computational modeling and simulation to study social issues, I do not believe that social scientists should only use computational modeling and simulation approaches to study real world problems (King et al. 1994). In other words, computational modeling and simulation can also be used to investigate issues that are counterintuitive but helpful for clarifying facts. In other cases they can be applied to study operational procedures and social issues in virtual reality (Epstein & Axtell, 1996)—for instance, conflicts and compromises between an agent's public and private interests. Furthermore, social scientists can use simulation models to test "what-if" questions that cannot be verified in real life.

The final advantage I will mention here is that computational modeling and simulation have strong integrative and deductive capacities (Hastie, 1988). Traditional research approaches make it difficult or impossible to a) compile empirical data with different formats into coordinated data sets or b) use diverse forms of data sets for inference and deduction. In contrast, computational modeling and simulation possess strong capacities for integrating and transforming data; combining empirical data from different sources; and performing simulation and various types of dynamic analyses in which static equilibrium, system dynamics, and emerging characteristics can be observed (Holland & Miller, 1991).

Other research approaches can deal with uncertainty or with social issues that are counterintuitive or involve virtual reality, but they often come with a requirement to loosen or simplify assumptions and outcomes; as a result, models can become enormous, complex, and unwieldy. From another perspective, computational modeling and simulation use programming skills and data structures for defining social concepts, relations, and operating procedures, thus allowing for the exploration of various hypotheses according to running processes and outcomes (Gilbert, 1999; Johnson, 1999). While traditional mathematical models can decrease deductive complexity by loosening single assumptions or initial conditions, they cannot handle complex interactions among multiple components.

In this dissertation I will discuss two reasons for applying computational modeling and simulation to social science problems. First, they combine the precision associated with quantifying analyses with the abundance associated with quantitative case analyses. Second, they are capable of representing social structures and processes without losing internal logic or accuracy. I will also argue that researchers in such fields as economics, psychology, sociology, and political science will realize more benefits from computational modeling and simulation than those working in other fields (e.g., History, Anthropology. Over time, researchers in all topic areas will eventually develop computational modeling and simulation skills that meet their specific requirements (Fiorina, 1975).

1.4. General Stages of Social Science Computational Modeling and Simulation

In an earlier section I listed the five major steps for completing a social science research project using computational modeling and simulation. Before introducing specific methods and validation procedures, I will describe the life cycle of an ideal research project using this approach. As shown in Figure 1.4.1, the five major steps can be divided into a series of sub-steps that are connected to each other so that they form a development cycle. However, actual projects are rarely as simple as the one shown in the figure; in many cases, development can appear to be a confusing mix of repetitive, overlooked, and simplified sub-steps.

Theory Development

For many social scientists, social phenomena are best viewed as events or behaviors that occur with a regularity that can be explained by theory. Existing theory cannot always explain certain social phenomena or produce sufficient interpretations. Such a situation will later turn to our intent to apply existent social scientific theories and novel hypotheses to construct more complete and satisfying explanations. Social science theories are not generated in isolated, remote states, but require the extensive understanding of previous theories or models that have been used to address the research topic in question. During the early stages of theory development, modelers often need to statistically analyze such characteristics as the randomness and regularity of empirical data and review large bodies of relevant literature.

A conclusion drawn from a social science study might entail one or a combination of the following: a) a simple theory (possibly already expressed as a mathematical model) that appears to explain the original observation; b) a reasonable but underdeveloped general theory that offers a promising explanation that seems too complex for formal analysis; c) several unconnected theoretical snippets (perhaps expressed as mathematical models), many of which find some empirical support but none of which seem capable of explaining the observation on its own; d) many separate quantitative empirical results (perhaps generated by "black box" models), none of which are capable of explaining the observation; or e) multiple qualitative studies with little attempt made at developing rigorous theory, perhaps because the underlying processes appear to be too complex for existing theory-building tools. Computational modeling and simulation is a suitable research approach in all but the first scenario.

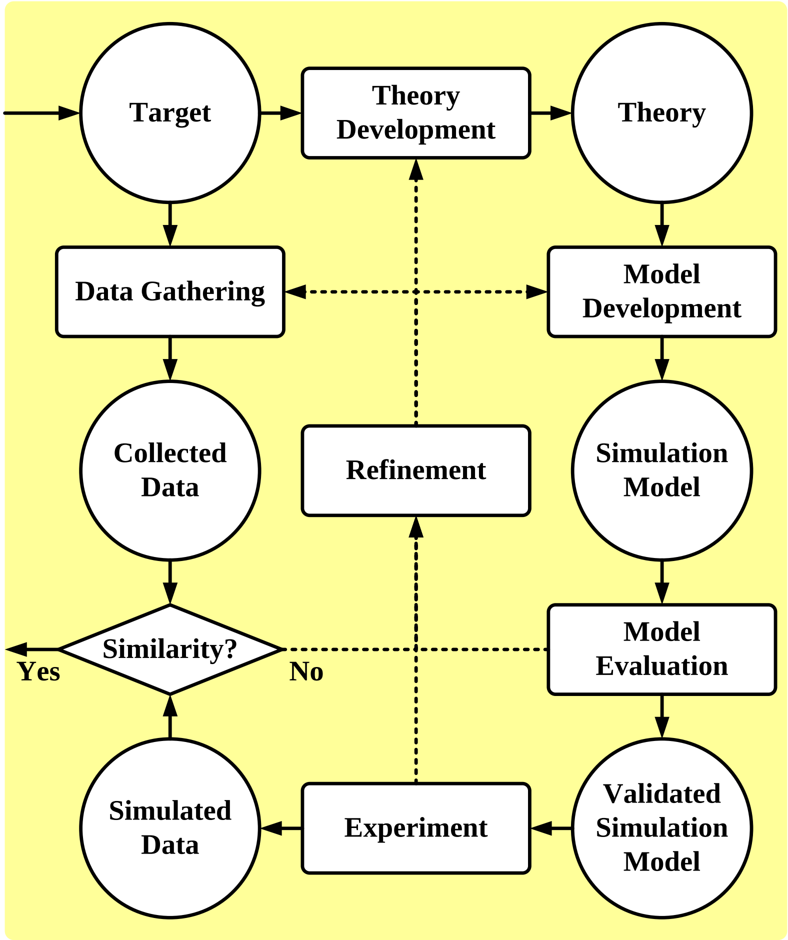


Figure 1.4.1. Computational modeling and simulation stages.

Model Development

For computational modeling and simulation studies, this stage entails two steps: establishing a transitional medium model and using the transitional model to transform theories into simulation models. Transforming theories into simulation models too soon is likely to cause mistakes in the transformation process that make it impossible to use simulation techniques to express them. The first step of the model development stage is to construct a transitional medium model based on the theory that is established in the preceding stage. A modeler needs to perform a concept inventory of major and corresponding theory variables, assumptions, initial conditions, and relations in order to list closely related concepts before constructing a complex simulation model. When performing an inventory, modelers must use qualitative descriptions for each concept (i.e., loose descriptions to indicate how a concept is operated) and introduce its quantitative characteristics in as much detail as possible (e.g., whether the concept is a continuous or discrete variable). Modelers must also establish interactive scenarios among the internal components of a simulation model based on the foundational theory. Starting from the black-box model (input and output variables only), modelers must define all factors, relations, and operating details between input and output via interactive scenarios generated from the concept inventory. When depicting operational procedures, modelers should list as many detailed processes of each component in the simulation model as possible.

Once the bridge between theory and the simulation model is built, the next step is to choose an appropriate computer programming language (e.g., high-level structured programming languages such as Java and C++ or the fourth-generation artificial intelligence languages such as Lisp and Prolog) or modeling and simulation software tool (e.g., DYNAMO, GPSS, or STELLA). Choices should be based on the needs and characteristics of the research project in question prior to transforming the theory into a simulation model. It is very important to determine the necessary methods, skills, and development tools before attempting to build a simulation model that is fully representative of a theory.

Model Evaluation

The first model evaluation step (also referred to as the verification step) consists of carefully eliminating coding errors before executing the simulation model so that no mistakes or misunderstandings occur during the transformation process. The model verification step is followed by simulation model validation; this process will be described in detail in the next chapter, but here I will focus on the major issues of using truth and beauty guidelines for evaluation purposes. Truth can be divided into internal validity, outcome validity, and (in many, but not all cases) process validity. Internal validity determines if the internal logic of a simulation model correctly expresses the theory or hypotheses upon which it based. Outcome validity measures the correspondence between outcome data and real-world empirical data. Process validity uses face validity, assumption tests, and multilevel outcome validation procedures to evaluate whether the simulation model’s operating procedures are consistent with those of the theory.

Even after a simulation model passes all of the validation guidelines in the second step, modelers need to examine the model’s running process. During a simulation, social scientists can explore in depth those problems and phenomena that counter social intuition but fulfill theoretical regulations. For example, what outcomes are produced in a simulated world where assumptions and initial conditions differ from gathered empirical data? For social science researchers, such a hypothetical analysis can increase the reliability and application of both a theory and a simulation model. The evaluation and validation procedures mentioned above should be repeated until the validity of the simulation model is ensured or until it reaches a pre-determined level of professional standards.

Refinement

The final step entails a discussion of theory and model refinement. As shown in Figure 1.4.1, theory development, transformation, simulation, understanding, and refinement occur as parts of an ongoing and ever-changing process, meaning that model refinement can take place during any sub-step of any stage. If a simulation model is found to misrepresent a theory during the third stage, modelers may need to extensively revise or even discard a simulation model, but when a validated simulation model is found to precisely represent a theory, the two are regarded as equivalent, with one's success or failure affecting the other and refinements in one implying a need to refine the other. However, if a simulation is identified as only a loose representation of its theoretical foundation, the failure of the simulation model might not indicate failure on the part of the theory. Such a situation is unwelcome; it is instead suggested that relations or similarities between a model and theory should not be too loose, since it makes it more difficult to ascertain the outcome, process, or internal validity of the simulation model.

When simulation model outcomes are not consistent with real world data, modelers must revise their assumptions, flow simulation model, and/or theory or start all over again. Regardless of the situation, modelers can benefit from the experience and use the same data. Furthermore, a comparative analysis of simulation model outcomes and real-world data can produce ideas for developing better theories and simulation models. Ideally, each cycle will give social scientists additional understanding and control over theory and simulation model development so that the theory is continually enriched and the simulation model made more precise and efficient.

1.5. Dissertation Overview

In Chapter 2 I will introduce simulation model validation standards and evaluation procedures, then discuss various challenges to model validation. In Chapter 3 I will give a brief overview of the transition from the dynamic simulation approach used for many years in the social sciences to the computational modeling and simulation approach that makes use of small-world networks (SWNs), scale-free networks (SFNs), random networks (RNs), and other small-world models.

In Chapter 4 I will present the details of an original research project conducted by Huang et al. (2004a) to use computational modeling and simulation to examine the transmission dynamics of disease epidemics and to evaluate the effectiveness of various public health policies and epidemic prevention strategies.

In Chapter 5 I will describe an original research project involving sensitivity analysis (Huang et al., 2004c). The study was designed to investigate the influence of local information on social simulations that are performed using a small world model. A cellular automata variation with added shortcuts was used as a test platform for simulating the spread of an epidemic and for examining the influences of various factors. The results of the study are offered as a means of helping researchers determine appropriate simulation parameters. In Chapter 6 I offer a conclusion and suggestions for future research directions.

Chapter 2. Validating Social Simulation Models

Validating a simulation model is a critical step that not only improves model accuracy, but also increases outcome reliability (Lave & March, 1993). As stated in Chapter 1, since many simulation model types are capable of representing (at least to a certain degree) a particular social science theory or hypothesis, an effective validation procedure is required. In this dissertation I will introduce computational modeling and simulation methods and techniques for use in the social sciences, but first I will address an important question: How can a researcher distinguish between superior and inferior simulation models? According to Lave and March (1993), truth (whether a simulation model can precisely express its target theory or hypothesis) and beauty (a combination of an aesthetic standard and model practicality) are two useful guidelines for validating simulation models. In this chapter I will explain how these guidelines work (Fig. 2.1).

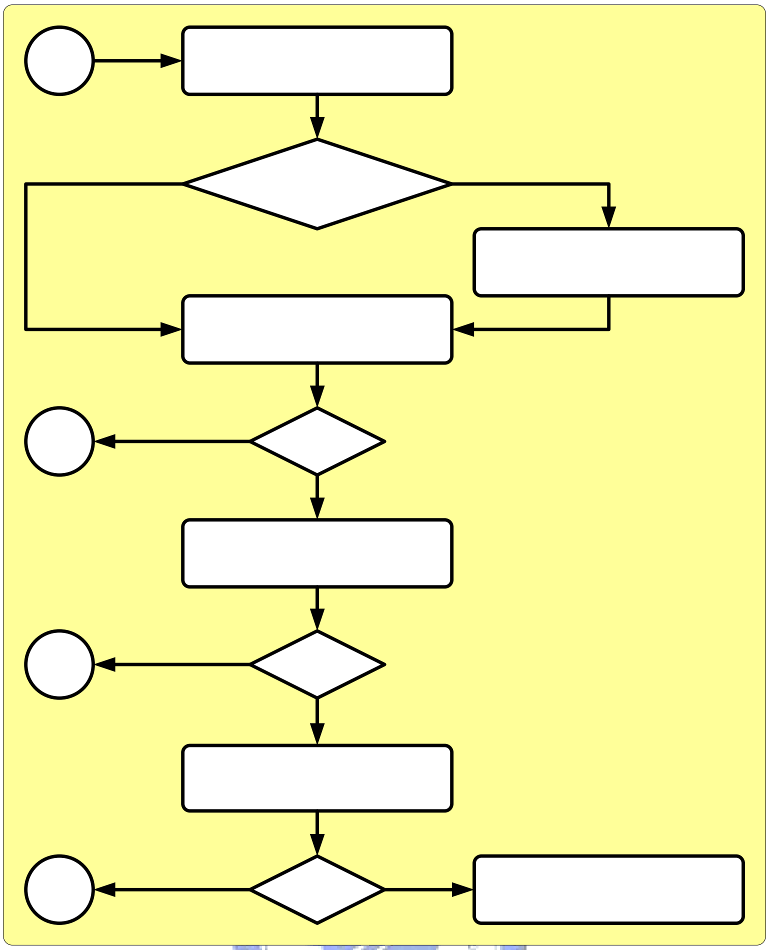


Figure 2.1. Computational modeling and simulation evaluation stages.

A

A

2.1. Truth

Validating truth is the most important yet most difficult task for any simulation model. Model evaluation requires the establishment of various baselines for comparison; examples include comparing outcomes with expert expectations or comparing multiple outcomes and their effectiveness in competitive models. In this section I will illustrate various validity dimensions—outcome, process, internal, and reliability and sensitivity analysis. Outcome validity is used to evaluate consistency between the outcomes/predictions of a simulation model and real-world data. Process validity assesses the running process of a simulation model and how it matches actual procedures. Internal validity focuses on how a simulation model's internal logic can precisely represent its underlying theory or hypothesis. Reliability and sensitivity analyses examine the robustness of a stochastic simulation model and test the degree of model sensitivity to various parameters.

Many social simulation model designers believe that process validity is vital to examining the compatibility of a model's running process and actual procedures. There are two exceptions. First, process validity is not suitable for some simulation models under certain circumstances. For instance, genetic algorithms (GAs) (Holland, 1975), which are based on the Darwinist concept of the "survival of the fittest," are often applied to scheduling, classification, or optimization problems (Goldberg, 1989) regardless of how well they fit in terms of process validity. A second example is artificial neural networks (ANNs), which are often used to represent complete organisms but not individual intelligence systems (Schrodt, 1995). The second exception is the impossibility of social science simulation models being completely accurate; it is meaningless and inappropriate to ask a simulation model to perfectly represent real-world operating procedures.

A single statistical method is never enough for social simulation researchers and modelers. The best goal is to find the most suitable statistical method in terms of research needs and simulation model characteristics (e.g., analysis of variance, analysis of linear or nonlinear regression, tests for means comparisons). As with all social scientists, social simulation researchers and modelers must be familiar with a variety of statistical methods and techniques.

Outcome Validity

Almost without exception, simulation model validation methods focus on whether predicted outcomes correspond to real-world data. While I used the term prediction above, the term that is commonly used in social simulations is postdiction, indicating that simulation outcomes generally correspond to things that have already happened in the real world rather than in the unknown future. Of course, it is insufficient to measure simulation model validity simply by comparing outcomes and real-world data, since models can be very complex and encompass many elements, parameters, variables, rules, components, and relations. A single predictive result is not enough to support a claim that some element, component, or simulation model is unreliable. In many situations, when a prediction fails or an outcome does not fulfill expectations, a modeler will look for missing procedures, improper control processes, or mistakes in parameters sets before abandoning a model. To date, there is no clear and distinctive standard for determining whether or not a prediction is acceptable. Modelers must therefore compare any simulation model with its competitors and use simple statistical tests to determine whether it does indeed produce better results.

Provided that simulation model components remain the same, a complex simulation model can be taken apart to test the outcome validity of its component parts at a lower level first, at upper levels second, and as a whole system third. In all cases, sub-components should be broken down until they cannot be divided into further sub-components. The goal of any analysis should be to investigate whether component interactions are correct and whether simulation outcomes correspond to real-world data. In contrast, interaction and integration tests proceed in a bottom-up fashion, with sub-components examined individually before being connected with other sub-components for further study. Since this process is very detailed, it is impractical for examining the largest and most complex simulation models.

When conducting a social simulation study, data should be divided into a minimum of two parts—training and test samples. Training samples help modelers establish the initial values of elements, parameters, variables, components, rules, and relations required to create a simulation model. They are also useful for model training via adjustments made according to data provided by preceding training samples. Test samples focus on examining correspondence between simulation results and test samples. At this stage it is important to prevent high levels of similarity between training and test samples, since every simulation model should be treated as a concrete way to represent the training samples. If training samples resemble test samples, it could interfere with determining simulation outcome validity—a critical step for any simulation model.

Validity can be determined using a simple statistical test to check the degree of correspondence between simulation results and test samples. In the absence of a high degree of correspondence, the simulation model needs to be revised or discarded. The next step entails a comparison of prediction performances between the simulation model in question and its competitors. In theory, simulation models that make the cut are the strongest performers in their field; if the predictive performance of one simulation model outweighs all others, there is a very good chance that the simulation model will pass most outcome validity examinations. In some cases several simulation models will be equivalent in terms of performance, yet a modeler will chose one over the others because an important component of the chosen model produces more precise predictions compared to the others. For this reason, the best test is a multi-level examination that not only investigates the overall outcomes and predictive performances of a simulation model as a whole, but also checks the outcomes and predictive performances of its sub-components.

Process Validity

Process validity measures the correspondence between a simulation model's running processes and the courses of action of the social issue being simulated. Unfortunately, there is no simple method for evaluating the process validity of a simulation model. In the absence of a direct method, researchers can use an indirect method that addresses three process validity examination requirements:

It examines face validity. In other words, it addresses the issue of whether the model is reasonable and acceptable in the minds of experts or scholars who are familiar with the issue being studied. Face validity is the most commonly applied means for examining a simulation model. A stricter standard entails an extended version of the Turing test (Turing, 1950), which is based on the claim that if individuals using independent judgment cannot tell the difference between the operating processes or outcomes of a simulation model and real systems, the simulation model can be said to exhibit process validity. Sensitivity analysis (to be discussed in a later section) can also be used to verify face validity by systematically changing the value of one important parameter and holding the others at a fixed value, then examining whether or not the simulation model behavior matches the expectations of an expert or modeler. Of course, such an examination requires that modelers collect the required data inputs, behavioral modes, and outputs for certain conditions. Although such a method requires prior knowledge, there is no need to have a comprehensive understanding of all behavioral modes and outputs under all conditions.

It directly examines simulation model hypotheses and uses those that are proven to be correct to justify the model's operating processes and degrees of acceptance. Supposing that certain hypotheses cannot be justified after evaluation and analysis (since certain simulation models are overly simplified and exaggerated representations of real world scenarios), modelers must cautiously evaluate the degree to which inaccurate hypotheses affect behavioral modes and outcomes. Decisions need to be made as to whether inaccurate hypotheses should be revised or discarded. Again, sensitivity analysis is very useful for this function. Modelers can use it to analyze which elements, variables, parameters, rules, or relations affect behavioral modes and outcomes. According to causal theory, knowing how many errors an invalid hypothesis can cause helps to determine how simulation model process validity is impaired.

It uses outcome validity (especially multi-level) tests to examine the process validity of a simulation model. Horizontal, vertical, or other classifications (e.g., object-oriented, even-driven, or multi-agents) can be used to break down the running process into several sub-processes, and each sub-process can be further divided into more detailed sub-processes. Theoretically, a simulation model can be broken down to the most basic level, with each sub-process corresponding to an independent and indivisible component. Outcome validity tests can then be used to determine whether or not sub-process outcomes are correct and fulfill expectations. Although such tests are not completely equivalent to actual examinations of simulation model running processes, they do increase operating process validity. However, tests that examine individual components are very costly in terms of computation time and resources, and require the collection of large real-world samples in order to fulfill test requirements. Furthermore, even when test sample accuracy requirements are very demanding, collected samples may contain extra information to the extent that it cannot be used to evaluate simulation model sub-processes. In spite of these problems, such examinations are considered the most effective way to evaluate the process validity of a simulation model.

Internal Validity, Reliability Analysis, and Sensitivity Analysis

As mentioned above, social simulation projects require at least five steps: a) proposing a theory or hypothesis to explain a social phenomenon, b) developing a formal model of said theory or hypothesis, c) using a computer language or modeling development tool to construct a simulation model or simulation system that corresponds to the formal model, d) verifying and validating the simulation model, and e) executing the simulation model and collecting and analyzing the generated data. Note that the simulation model/system, theory/hypothesis and formal model are considered equivalent and interchangeable—in other words, a simulation model or system should faithfully and reliably represent the theory or hypothesis and formal model. If not, the simulation model is meaningless. Experience dictates that the more complex the theory or hypothesis, the greater the likelihood of making mistakes when establishing a formal or simulation model, which leads to inconsistency among the three elements. Therefore, it is necessary to examine the internal validity of a simulation model before examining its outcome and process validity. However, this raises an important issue: how to evaluate a simulation model that can fully represent its corresponding theory/hypothesis and formal model.

The first step in evaluating internal validity is one that every modeler should use during a simulation model's development stage: determining if the basic structure of the model (i.e., elements, variables, parameters, relations, rules, and components) are equivalent to the theory, hypothesis, and formal model that it intends to represent. To give an example, cognitive psychologists have traditionally depicted human cognitive activity as an information processing system consisting of a long-term memory component and a working memory component. Simulating human cognitive activity requires the application of a data structure or another method (abstract or practical) to represent the two memory components; otherwise, one cannot claim that the system corresponds to cognitive theory. If a modeler claims that he or she has created an information process system based on cognitive theory, how should we evaluate the model to support or refute that claim?

There are two ways to solve this problem. First, a modeler can use face validity to evaluate the simulation model or models and periodically consult with experts or scholars familiar with such a theory to determine if the model actually captures the spirit of the theory. Second, a modeler can use existing cases to examine simulation models and to determine whether or not (under specific conditions) the simulation model running process and outcomes fulfill the expectations of an expert or scholar. Cases used to examine models must be valid, but they do not necessarily have to be accurate or complete. Furthermore, they may represent a theory or hypothesis in a very simplified or specialized format.

Applying existing cases to examine simulation models is similar to another method of examining internal validity: reliability analysis, which focuses on whether or not outcomes produced by repetitive simulation executions are consistent according to simple statistical tests. In other words, for simulation models that involve stochastic uncertainty, model designers must run simulation models many times (depending on the statistical method employed) to examine outcome consistency. From the perspective of reliability, the value of a simulation model with high stochastic uncertainty should be measured according to whether or not there is a sufficient number of executions to determine outcome consistency. If a modeler wants to examine a simulation model that can produce output with probabilistic outcomes, it is a much easier task: all the modeler needs to do is systematically compare the outcomes of each execution until certain robust estimate values converge. In order to prove that their model is stable and reliable, Jones, Radcliff, Taber, and Timpone (1995) applied every possible set of initial parameters to their simulation model and performed one million executions per parameter set. This example also serves as an example of a clear examination method that is worth noting for future reference: no matter how many stochastic components a simulation model owns, the model can be executed repeatedly on a computer until the modelers make an evaluation. Modelers need to take advantage of their computing tools to perform many executions, revising or controlling certain components or procedures while holding the others at fixed values in order to examine how revisions affect a model. In short, reliability analysis gives social simulation modelers and researchers the ability and tools to understand whether a vital stochastic component in a simulation model is interactive or additive.

A third means of examining the internal validity of a simulation model is sensitivity analysis. Converting a theory, hypothesis, or formal model into a simulation model entails deciding the domain of each parameter, the scope of initial conditions, the sampling technique for stochastic elements, and probability distributions. During sensitivity analysis, modelers systematically change parameter values or other component settings (e.g., probability distribution) and examine how simulation model performance or outcomes are affected. This type of analysis allows modelers to distinguish between two kinds of parameters: when model performance or outcomes are influenced by slight changes in a parameter value, this is referred to as a sensitive parameter. Other simulation models are not influenced, regardless of how much a parameter value is changed (within a reasonable range).

As long as the internal validity of a simulation model is assured, modelers can use sensitivity analysis to examine a simulation model hypothesis. Simulation models are usually viewed as simplifications of social issues for exploration purposes, in which hypotheses are proposed and tested. Under such circumstances, sensitivity analysis can be used to examine how different degrees of simplification affect the model. If a simulation model's constraints are loosened but its qualitative analysis conclusion remains the same or stays consistent, then the hypothesis is plausible; if not, the hypothesis requires special attention. During the last stage of simulation model development, sensitivity analysis can serve as a warning mechanism or guidance procedure. As mentioned earlier in this chapter, sensitivity analysis is also suitable for examining outcome and process validity.

2.2. Beauty

Although some argue that the standards of a "beautiful" simulation model vary from person to person, we should not ignore aesthetic criteria when evaluating a simulation model. According to Lave and March (1993), three aesthetic characteristics need to be acknowledged: simplicity, fertility, and surprise. These characteristics are not only a matter of personal taste, but also ones that any researcher who uses computational modeling and simulation to explore social issues should consider seriously.

Simplicity

Computational modeling generally focuses on simplified understanding through exploring systems designed to explain the real world. However, simulation models try to filter the real world and address a pivotal theory by eliminating unnecessary details according to the Occam's razor recommendation that "Things should not be multiplied without good reason" (Starfield, Smith, and Starfield 1990). Applied to computational modeling and simulations, concise and succinct simulation models should be accurate and valid.

In physics and other physical sciences, succinctness is regarded as a useful exploration principle, but some biologists (including Crick, 1988) and social scientists view an emphasis on succinctness as misleading. This leads to the question, "Is it reasonable to apply Occam's razor to social simulation models?" If a simulation model is simplified too much, its process or outcomes may become so confusing that they cannot be properly examined, thus making the associated theory or system inaccurate or unhelpful in understanding real world phenomena. On the other hand, some simulation models are too complex to be useful in understanding a system or theory. The following principle can be inferred from this situation: if a researcher cannot clearly track the operation or outcomes of a simulation model, the model will create more problems than it solves. According to Occam's razor, evaluating the conciseness of a simulation model requires a determination of whether or not it precisely expresses all of a theory's important procedures, followed by a determination of whether any procedures can be removed without affecting proper model operation. In other words, simplicity has value in terms of both aesthetics and practicality.

Fertility

Fertility refers to the implications that a simulation model conveys, how many theories it covers, and how broadly it applies. When two models are used to test the same theory, the model that generates a greater number of predictions is more highly valued than the one that generates fewer. The simulation model that generates more predictions is said to have a richer framework for theoretical deductions, thus making it easier to investigate its predictions. Fertility is also related to simplicity in terms of strictly controlling the number of assumptions during evaluation—that is, a complex simulation model with more assumptions is only considered fertile if it generates more predictions or implications.

Furthermore, a complex simulation model is said to generate more in the same manner that merchandise quality increases with price. A very complex simulation model that generates only a few predictions is generally viewed with suspicion; on the other hand, after a simulation model is simplified it should be examined in terms of whether or not it maintains a high level of prediction quality. A complex model that produces more detailed predictions is preferable to a simple model that generates predictions that are not as well defined. It is also important to consider how broadly a model's predictions or implications can be applied; a model that accounts for a larger number of scenarios is more valuable than one that explains only a few. Again, aesthetic standards and practicality merge within the fertility criterion.

Surprise

Another important aspect of fertility is surprise. An effective simulation model often produces unexpected but conceptually applicable and easily examined predictions. A prediction with strong implications may be surprising in situations where a researcher does not expect an outcome or outcomes to be generated from a simulation model, yet it produces data that fits well with facts or other evidence. In other situations, an outcome may contradict a researcher's intuition or appear to be estranged from facts or other evidence, but the outcome turns out to be correct based on a logical deduction or analysis. Precision and surprise often coincides in social simulations—that is, when a theory can be correctly expressed in a simple simulation model, precise and surprising predictive results are sometimes generated. Compared to complex models, simple simulation models are less likely to produce surprising predictions because their outcomes cannot be directly applied without further transformation and explanation. Theories that require computational modeling and simulation are usually more complex, and therefore have greater potential to produce surprising conclusions that extend our knowledge.

2.3. Summary

A perfect simulation model should be valid and beautiful, but very few achieve these ideals. Researchers should instead look for a balance between perfection and real-life obstacles. Although many model designers are familiar with the examination procedures and standards discussed in this chapter, they still face the challenge of properly applying them to test different simulation models with applications, purposes, and methods. In this chapter I have proposed several principles in terms of timing and examination methods that other researchers may find useful when applying computational modeling and simulation techniques to social science issues.

Truth can be divided into internal, outcome, and process validity. When examining internal validity, sensitivity analysis can be used to examine simulation models, especially more complex models. Outcome validity relies on traditional quantitative concepts to test simulation models, but the question of how to quantify a simulation model remains, as well as the question of judging results after quantification. Based on the existing literature, I have described several examining tools to help modelers test their simulation models, and suggested other tools (e.g., face validity, directly testing hypothesis validity, sensitivity analysis, and multi-level tests borrowed from outcome validity) for inspecting the process validity of a simulation model.

Simplicity has traditionally been the most important criterion for simulation model beauty. Conciseness is still a desirable goal, but when the theories to be tested by a simulation model are very complex, modelers need to achieve truth before pursuing conciseness and guarantee that model operating procedures and outcomes are both correct and precise. Fertility and surprise are two reasons why social scientists adopt computational modeling and simulation, since they are more likely than other formal models to produce surprising implications.

Chapter 3. Related Computational Models and Concepts

3.1. Dynamic Simulation

Dynamic simulation is one of the earliest computational modeling and simulation methods in the social sciences (Huckfeldt et al. 1982). Famous research projects that applied this method to explore social phenomena include urban systems of Forrester (1969), the global population of Meadows et al. (1972), and electoral systems of McPhee (1963). Dynamic simulation is the process of constructing a mathematical model of some real-world system and analyzing its behaviors and results through computer-based experiments. In part because of the availability of special simulation software (e.g., STELLA or GPSS), dynamic simulation remains one of the most popular and productive computational modeling and simulation methods.

As note, dynamic simulation refers to the construction of and experimentation with a computational model of a dynamic system. For example, to save time and money, a network engineer might propose a network model with a novel connection topology, using computers to construct a virtual network environment model under certain conditions and simulating their research results. Likewise, an epidemiologist may examine the transmission dynamics and growth tendency of an infected population through a simulation model. As a result, the network engineer saves time and money; while the epidemiologist gains experimental control and the ability to manipulate the theoretical world.

For more detailed introductions to dynamic simulation, we recommend Forrester (1980). This book includes many modeling techniques and research examples, as well as introducing DYNAMO and its applications. In addition, please refer to the researches of Kheir (1988), Hoover and Perry (1990), Whicker and Sigelman (1991), and Hannon and Ruth (1994). In Garson (1994), there are many introductions and evaluations regarding the application of dynamic simulation.

An Example of Dynamic Simulation – Compartmental SIR Models

In epidemiology, the compartmental model (Kermack and McKendrick 1927; Edeletein-Keshet 1988) is a type of dynamic simulation system applied by mathematical epidemiologists to estimate the overall trend of epidemic outbreaks. Until recently it was still a prototype of main epidemiological models. One of the most famous and typical models is the SIR model proposed by Kermack and McKendrick (1927) (Fig. 3.1 and equation 3.1). Using the SARS outbreak last year as an example, many researchers applied the compartmental model to roughly estimate the transmission dynamics as well as future development tendencies of SARS, and approximately explained its spread phenomenon, for example, the super-spreader events (SSEs) during the outbreaks (Chowell et al. 2003; Donnelly et al. 2003; Lipsitch et al. 2003; Ng et al. 2003; Riley et al. 2003). However, such a simulation model only roughly calculates the total amount and the increasing or decreasing change of infected population at every discrete time step, and applies differential equations to generate pivotal parameters during the simulation, such as the basic case reproduction number R_0 (Anderson and May 1982), that are supposed to be essential for public health specialists and epidemiologists. To produce more accurate simulation results, recently researchers have further divided each population into different sub-groups according to age, residency, infectious rate, and various characteristics that interest epidemiologists (Lipsitch et al. 2003; Ng et al. 2003; Riley et al. 2003). However, the entire simulation model becomes complicated and enormous as a result, and can only be applied under specific circumstances. No matter which ways of categorization and what characteristics there are, however, such simulation models ignore the fact that social phenomena emerge only when many people contact, interact with, and change each other. In other words, the typical compartmental model focuses merely on the characteristics of epidemics themselves (transmission rate, patients' mortality rate, and recovery rate) without taking into account important social characteristics such as population structure, space, heterogeneity, localization, and interaction. Furthermore, due to the absence of these social characteristics, such simulation models cannot be applied to explore important issues in public health policies and disease prevention strategies.

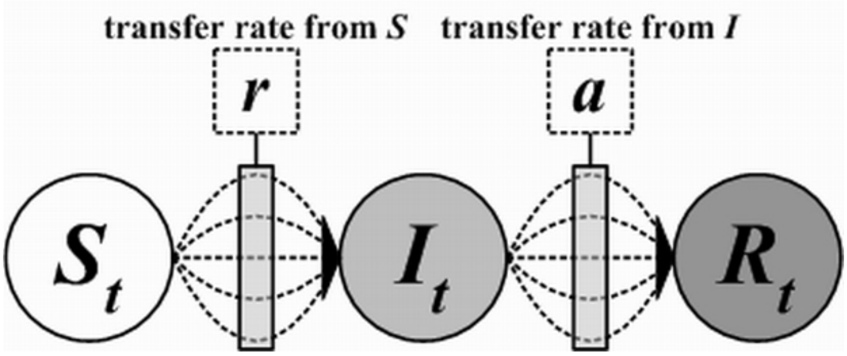


Figure 3.1. General SIR model transfer diagram. *S*, susceptible class; *I*, infective class; *R*, recovered class

A typical SIR model divides a social population into three mixed groups: a ***S***usceptible group prone to infection but not yet infected, an ***I***nfectious group of individuals who have the disease and are capable of infecting susceptible individuals, and a ***R***emoved group of fully recovered or deceased individuals who cannot infect others. A set of differential equations such as the one shown in Equation 3.1 is used to trace the dynamic process of individual movement within groups and to calculate increases, decreases, and total numbers of individuals for all groups in discrete time steps.

$$\begin{cases} \dfrac{dS}{dt} = -rS(t)I(t) \\ \dfrac{dI}{dt} = rS(t)I(t) - \alpha I(t) \\ \dfrac{dR}{dt} = \alpha I(t) \\ N = S(t) + I(t) + R(t) \end{cases}$$

$r = \gamma \cdot \beta \cdot \frac{I(t)}{N}$ is the infection rate, α is the removal rate, and N is population size

(3.1)

For epidemiologists, the basic case reproduction number R_0 is an index parameter with an important reference value, which is the number of people infected by a patient before she recovers or dies. When R_0 is greater than 1, not only does the number of infected patients keep increasing, but the transmission rate soars because of the increase of the newly infected. As a result, an inevitable outbreak is doomed to happen. When R_0 is equal to 1, the spread of disease is claimed to be stable, i.e. each patient transmits her virus to one healthy person on average. The value 1 is thus called the threshold value for plague. To prevent an epidemic from becoming a plague, R_0 has to be controlled such that it is lower than the threshold value. When R_0 is smaller than 1, a patient does not necessarily have a chance to transmit her viruses to another, and since the recovery rate is higher than the increase of patients, the number of infected patients will decrease eventually, and ultimately disappear.

3.2. Simple Social Network Models

To fully utilize computational modeling and simulation to explore interesting social processes and issues, social scientists need a simple social network that properly represents interpersonal relationships to be the fundamental structure of a simulation model. There are two ways to construct a simple social network model: the first, presented in Figure 3.2.1(a), uses low-dimensional lattices to represent a social network (Ahmed and Elgazzar 2001; Boccara et al. 1994; Koopman 2004), for instance, one-dimensional ring-shaped periodic lattices or two-dimensional toric periodic lattices. In such a network model, since every node is connected to its closest nodes while the number and the nodes connected remain the same, the model is also called a regular network model. The second, as shown in Figure 3.2.1 (b), uses a random network (Erdös 1959) to represent a social network. This type of network model and the compartmental model mentioned earlier are largely equivalent; that is, many overall characteristics of a network can be represented by statistics. Moreover, the random network model is a beautiful and effective network that is used to act as a representative of the complicated, chaotic and unpredictable society.

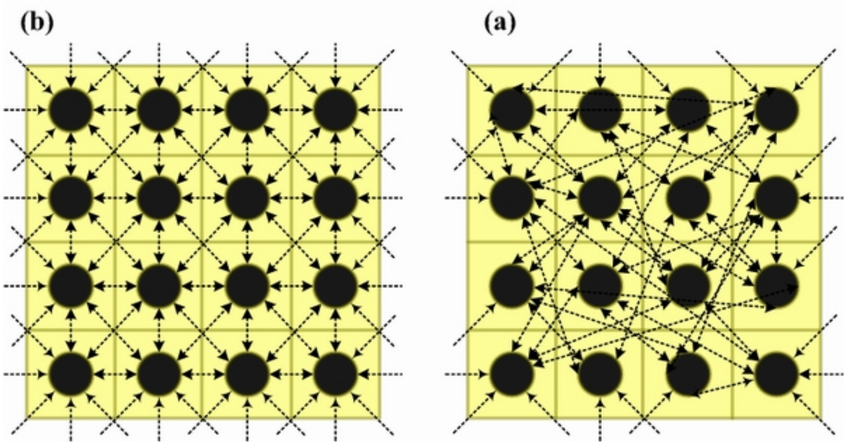


Figure 3.2.1. Simple social network models. (a) Regular network model. (b) Random network model

For network models, each community, city, country, or the entire world is a social network, while each node represents an individual with various attributes that accounts for her state, such as epidemiological progress, gender, age and immunization. The edge between two nodes represents connections between two individuals. Different epidemics can be presented by different edges to demonstrate different interpersonal relationships. For example, when examining AIDS, lines represent sexual relationships; when examining SARS, they represent individuals' close contact. In each discrete time step, the state of every node in the network changes simultaneously. The set of all of the node states represent the overall transmission of an epidemic.

An Example of Simple Social Network Model – Cellular Automata

Cellular automata (sometimes called lattice models) is not only a simple social network model but also considered to be a subset of dynamic simulations that is specifically used to model discrete dynamic systems composed of many interacting units. In cellular automata, every unit (also called lattices or cells, equivalent to nodes in a network) contains several discrete attributes and a set of very simple rules that define the change of attributes and the interaction between this unit and its neighbor units. Modelers usually apply cellular automata to construct the non-linear system of a simulation model, where the lower-level components require intensive interaction. One of the examples includes applying cellular automata to simulate the formation, spread and change of public opinions. Although the dynamic of cellular automata is complex, unpredictable and chaotic, each unit contributing to such phenomenon is simple and can be precisely described (Cowen and Miller 1990; Haken 1983). As shown in Figure 3.2.2, cellular automata represent a social network model. A limited number of discrete attributes is employed to represent its statues. The simplest discrete attribute contains only binary values, such as “on/off,” “black/white,” or “infectious/healthy.” At each discrete time step, the new state of each unit is determined by its state transition function.

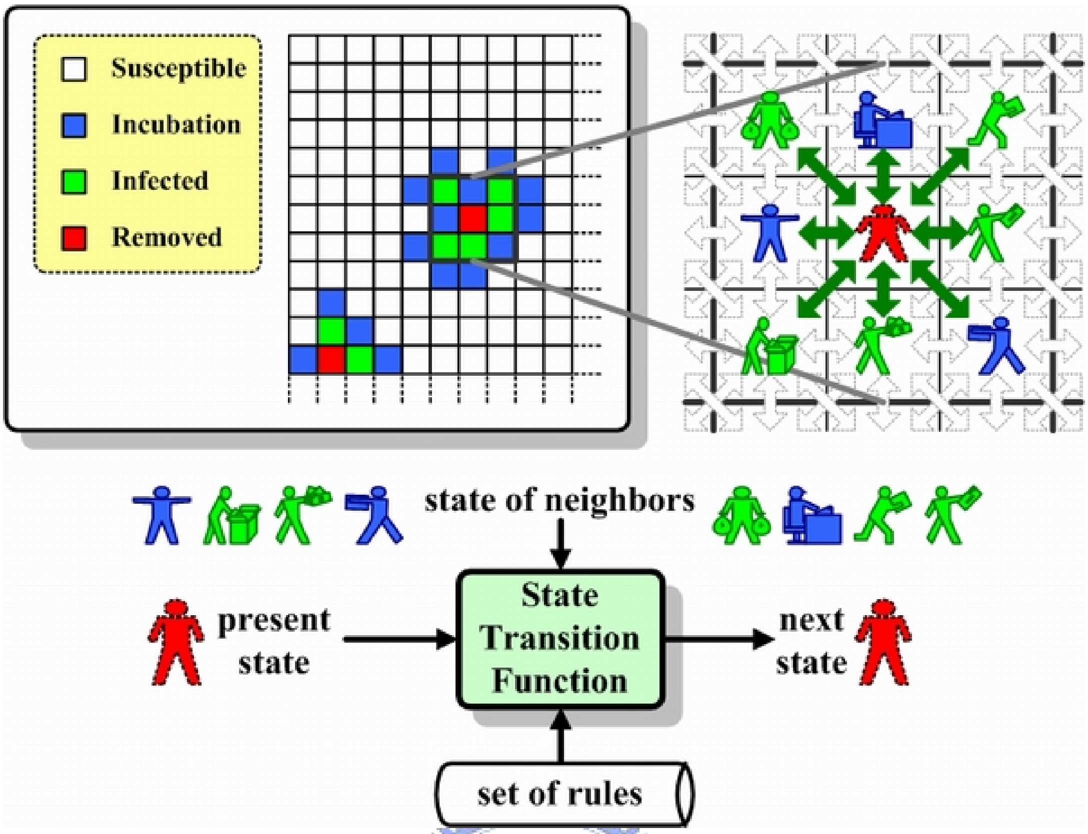


Figure 3.2.2. Cellular Automata, state transition function, and Moore neighborhood concept

When using cellular automata, modelers have to first define (1) the dimension of cellular automata—one- or two-dimensional; (2) the length of every dimension, i.e., how many units there are in each dimension. Suppose every dimension of two-dimensional cellular automata contains 100 units, 10,000 units are included in the cellular automata; (3) the number of attributes of every unit, and the discrete attribute values included (such as two state values—life and death); (4) a function defining which units are neighbors; (5) a state transition function for calculating new states based on the states of neighboring units (as shown in Formulas 3.2.1 and 3.2.2). Because the conditions above are completely defined by modelers, cellular automata are considered to be a particularly flexible computational modeling and simulation.

$$\pi_{i,j}(t)=\begin{cases}S, & \text{if } \tau_{i,j}(t)=0,\\ E, & \text{if } \tau_{i,j}(t)\in(1,\tau_E),\\ I, & \text{if } \tau_{i,j}(t)\in(\tau_E+1,\tau_I),\\ R, & \text{if } \tau_{i,j}(t)\in(\tau_I+1,\tau_R).\end{cases}$$

(3.2.1)

$$\tau_{i,j}(t+1)=\begin{cases}0 & \text{if } \tau_{i,j}(t)=0\wedge \text{not infection}(c_{i,j}),\\ 1 & \text{if } \tau_{i,j}(t)=0\wedge \text{infection}(c_{i,j}),\\ \tau_{i,j}(t)+1 & \text{if } 1\le\tau_{i,j}(t)<\tau_R,\\ 0 & \text{if } \tau_{i,j}(t)=\tau_R.\end{cases}$$

(3.2.2)

Social scientists Thomas Cusack and Richard Stoll (1990) have proposed a fascinating application of the cellular automata approach. They successfully applied cellular automata to model the realist theory of international politics. The entire international system (cellular automata) consists of 98 countries (units or lattices, i.e., territories). Every country has three characteristics: territory (that is, besides the unit belonging to the country, how many other units it also possesses), power and basic preferences. Each iteration of the simulation model includes five phases of behavior: (a) empire (comprised of many countries) faces civil war such that it might collapse and become a group of smaller countries with smaller territories; (b) faces conflicts with neighboring countries and occupying the neighboring countries; (c) the conflicts aggravate, or suddenly disappear because alliances are formed; (d) one or more wars erupt because of the conflicts; (e) to respond to the results of wars, the power of participating countries are adjusted accordingly. From the simulation processes and results, Cusack and Stoll find that, at a certain level, many unbelievable conditions have to exist at the same time to ensure the stability of an international system. In other words, unless there are strict initial conditions and hypotheses, wars and the expansion of an empire will appear repeatedly. For readers interested in cellular automata, please refer to the researches of Farmer et al. (1984), Rietman (1989) and Gutowitz (1991).

Recently, researchers have applied two-dimensional cellular automata to investigate the local transmission mechanism and phenomena of epidemics (Ahmed and Agiza 1998; Ahmed and Elgazzar 2001; Benyoussef et al. 2003; Yacoubi and Jai 2002; Fuentes and Kuperman 1999; Martins et al. 2001; Boccara et al. 1993). For epidemiologists and social scientists, cellular automata is a concrete regular network model with social features such as population structure, local aggregation, social space, heterogeneity, and interaction, which are essential to various epidemiological and communication issues. Therefore, the transmission of epidemics can be observed easily in such a simulation model. However, such a simple social network model lacks the important "small-world" social feature that is essential to the study of epidemics, so that it fails to represent the low degrees of separation among individuals in a real society (Wang and Chen 2003). Moreover, cellular automata cannot be applied to effectively investigate public health policies and epidemic prevention strategies that are closely related to epidemic issues.

3.3. Small-World Social Network Models

Creating the small-world method and conducting letter delivery experiment, Milgram (1967) proposed “six degrees of separation” to explain a tricky fact of the world: in reality, human beings interact frequently, forming one group after another, but their degrees of separation are surprisingly low—everybody in the world is separated by only six other people. Not until Watts and Strogatz (1998) proposed the first small-world network model (Fig. 3.3.1), verifying that it has characteristics of high clustering and yet low degrees of separation, did many researchers realize that topological networks and structures are ubiquitous in the real world. In particular, the small-world social network deeply influences the development of many social issues and their consequences (Moore and Newman 2000; Comellas and Sampels 2002; Newman 2000; Wang and Chen 2003; Watts 1999). From then on, whether a social simulation model can portray characteristics of high clustering and low degrees of separation has become an important index when examining social network models. Since social individuals are characterized by long-distance movements, daily visits to fixed locations, multiple activity locations and local clustering, the average distance between any two individuals is shortened. Consequently, the actual geographical locations and distance become only secondary factors in causing epidemic outbreaks.

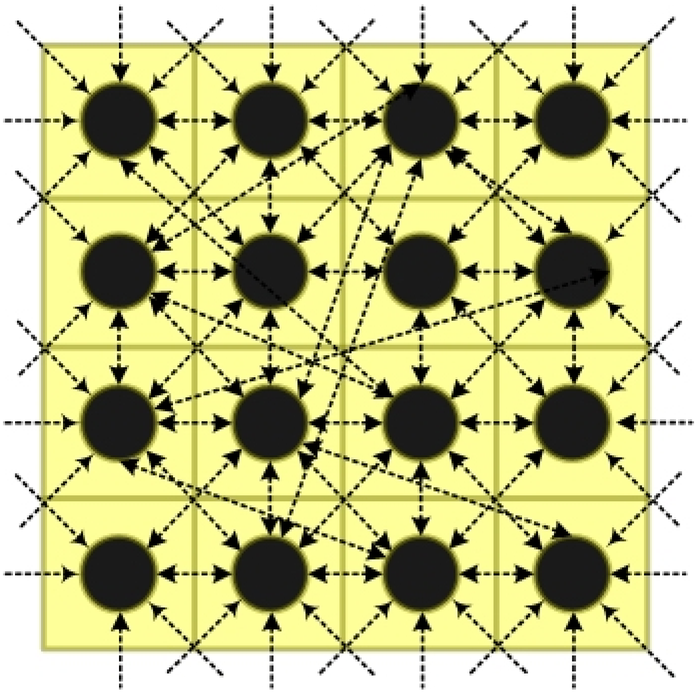


Figure 3.3.1. Small-world social network models

Watt and Strogatz's model triggered a large number of proposals for other models (Albert and Barabási 2002; Barabási and Albert 1999; Barabási et al. 1999; Erdös and Renyi 1959; Newman 2000; Newman and Watts 1999) that also exhibit small world phenomena. As shown in Figure 3.3.2, small world models can be categorized as small-world networks (SWNs) (Watts and Strogatz 1998; Newman and Watts 1999), scale-free networks (SFNs) (Albert and Barabási 2002; Barabási et al. 1999; Barabási and Albert 1999), and random networks (RNs) (Erdös and Renyi 1959). To generate a SWN (Fig. 3.3.2(a)), start with an *n*-dimension regular graph in which each node is connected to a *z* quantity of neighbors (usually, $z \geq 2n$) (Wang and Chen 2003; Watts and Strogatz 1998). Each edge of the graphic is then randomly rewired (i.e., one end of a connection is shifted to a new node chosen at random) with probability *p*. In a variation of the original model proposed by Newman and Watts (Newman and Watts 1999), long-range links (referred to as “shortcuts” in this dissertation) are inserted between pairs of randomly chosen nodes. This variation exhibits such small world phenomena as clustering, and is therefore considered very similar to human social networks.

To generate a SFN (Fig. 3.3.2(b)), start with a small number m_0 of nodes (Albert and Barabási 2002; Barabási et al. 1999; Barabási and Albert 1999). At each iteration, a new node is introduced and connected to $m \leq m_0$ preexisting nodes with a probability that depends on the vertex degree of each node. New nodes are preferentially attached to existing nodes that have large numbers of connections. This type of model exhibits small world phenomena and clustering among small numbers of nodes with very large vertex degrees. These are similar to the hyperlinks used in the World Wide Web (Albert et al. 1999; Faloutsos et al. 1999; Medina and Matta 2000).

Finally, RNs (Fig. 3.3.2(c)) can be generated by adding a number of links betweens pairs of randomly chosen nodes (Erdös and Renyi 1959). RNs are capable of exhibiting small world phenomena if enough links are added, but without (or with very little) clustering—an unusual situation in the real world.

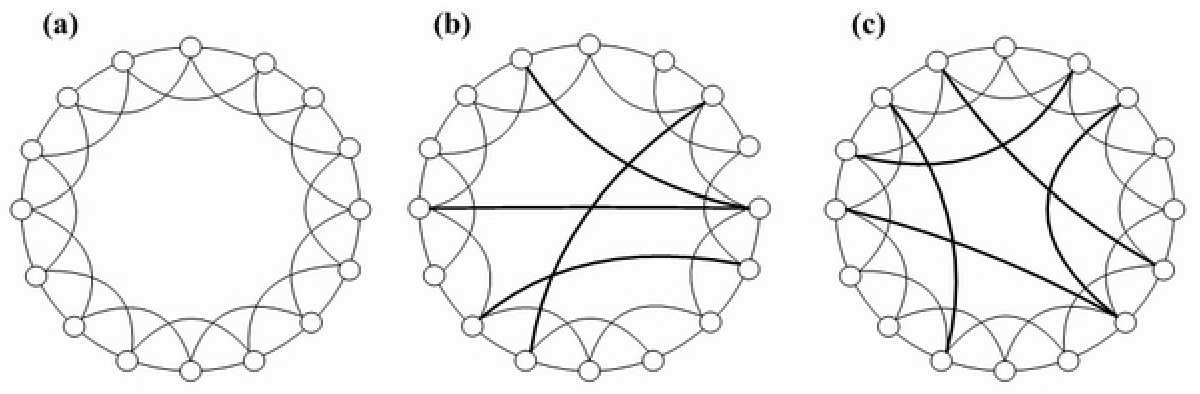


Figure 3.3.2. Three small-world models; a, small-world network (SWN); b, scale-free network (SFN); c, random network (RN)

The Small-World Phenomenon

To examine whether a social network model is in effect a small-world network model, two validation indices—the clustering coefficient and the separation coefficient—are required. The clustering coefficient evaluates the degree of connection between two neighboring nodes. As shown in equation 3.3, graph G represents a social network, v_i is a node in graph G, the vertex degree of node v_i is k_i, the edges that actually exist among these k_i nodes is E_i ($E_i \le k_i \times (k_i - 1) / 2$), and clustering coefficient is C_i. Thus, the clustering coefficient $C(G)$ of the whole social network is the average of the clustering coefficients at every node.

$$C_i = \frac{2 \times E_i}{k_i \times (k_i - 1)} \qquad (3.3)$$

The separation coefficient evaluates the shortest distance between two random nodes. The separation coefficient of the whole social network, therefore, is the average of the shortest distances between any two nodes. Another important characteristic of the small world is that, when the number of individuals in a society increases, the average separation coefficient between any two individuals does not increase proportionally but instead logarithmically (Newman 2000).

The Triadic Closure Concept

The triadic closure concept was first proposed by the mathematician Rapoport (1957). In reality, human beings are “birds of a feather,” which brings about human interactions in a social network that far beyond the illustration of the random network model mentioned earlier. Employees in the same company, classmates in the same school, and regulars at a Starbucks, for example, have a much bigger chance to know each other than two random strangers. These people are acquainted with each other not because of random probabilities but because of what they have in common. Later, Rapoport proposed a “triadic closure concept” (Fig. 3.3.3) that is even more fundamental than the concept of “birds of a feather:” two strangers with a common friend might know each other after a certain period of time and might even become friends. Suppose that Dick and the owner of the grocery store next door, Frank, are buddies, and Dick’s wife Ella is pals with Grace, who owns a fruit shop across the back yard. In this case, Dick is very likely to be introduced to Grace, while Ella is likely to be friendly with Frank. This example shows that triadic relations are the fundamental unit in a group structure, indicating that the progress of a social network is not a random network without social rules tied up to the connection among individuals, but instead a triadic closure relation. If the relation keeps going, e.g., Frank knows Grace through Ella, and then other longer closure relations might follow and ultimately become a tightly connected group. Because there exist tight connections among human beings, whenever an epidemic outbreaks in a certain area, the healthy but susceptible locals are most likely to be infected or badly ill, for they have formed triadic or polygonal closure relations with many infectious patients.

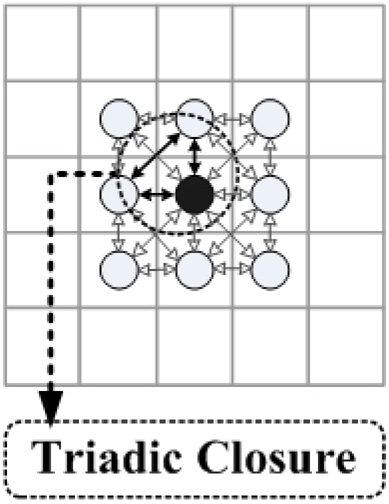


Figure 3.3.3. The triadic closure relationship in social network model

Chapter 4. A Novel Small-World Model with Social Mirror Identity Concept for Epidemic Simulations

The author proposes a novel small-world model that makes use of cellular automata with the mirror identities of daily-contact social networks to simulate epidemiological scenarios. We established the mirror identity concept (a miniature representation of frequently visited places) to acknowledge human long-distance movement and geographic mobility. Specifically, the model was used to a) simulate the dynamics of SARS transmission in Singapore, Taipei, and Toronto and b) discuss the effectiveness of the respective public health policies of those cities. We believe the model can be applied to influenza, enteroviruses, AIDS, and other contagious diseases according to the various needs of health authorities.

4.1. Motivation

In anticipation of the next outbreak of Severe Acute Respiratory Syndrome (SARS) (Peiris et al. 2003), molecular biologists, epidemiologists, sociologists, private laboratories, and public health agencies are committing considerable amounts of time and resources to confirming viral structure, developing vaccines and antidotes, establishing faster inspection methods, and revising public health policies (Anand et al. 2003; Chowell et al. 2003; Donnelly et al. 2003; Guan et al. 2003; Lipsitch et al. 2003; Marra et al. 2003; Ng et al. 2003; Nishiura et al. 2003; Riley et al. 2003; Rota et al. 2003). The last topic on this lists—specifically, the efficacy of various public health policies—is the focus of the present chapter.

Identifying the best possible suite of public health policies requires detailed knowledge of SARS transmission dynamics based on the limited amount of data collected during the 2002-2003 SARS outbreak (Sebastian and Hoffmann 2003; World Health Organization [WHO] 2003). This information can be used to establish a SARS transmission model (Dye and Gay 2003) for balancing the social costs and resource expenditures required for controlling future outbreaks (WHO 2003). Policies that were implemented in 2002-2003 included the wearing of masks (by the general public or by health care/hospital workers), hand washing, quarantining, restrictions on hospital visitations, and wide-scale efforts to take the body temperatures of individual citizens. Unfortunately, improper implementation and inappropriate timing occasionally produced such secondary impacts as disease concealment, social discrimination against SARS patients and health care workers, and the panic buying of masks.

Computational modeling and simulation is increasingly being used to match public health policies with the characteristics of local populations. In addition to information on disease transmission, suitable SARS simulation models require accurate data on how social networks operate in modern societies (Dye and Gay 2003)—for instance, human clustering behavior, the potential for multiple contacts, and long-distance movement. The model that we will describe in this chapter uses a combination of cellular automata (for the direct simulation of individual interactions) (Boccara et al. 1994) and a concept that we have developed and named *mirror identities*, which allows the model to consider low degrees of separation, long-distance movement, and daily visits to fixed locations. Combined, these factors assist in the creation of a realistic SARS simulation platform with small-world characteristics; we believe the model also has potential utility for simulating other infectious diseases (e.g., influenza, enteroviruses, and HIV/AIDS) as well as social issues (e.g., communication problems).

4.2. The Proposed Model

Our proposed model consists of two layers (Fig. 4.2.1). The upper layer is a multi-agent system used to simulate real-world heterogeneous cohorts. The lower layer consists of two-dimensional cellular automata (i.e., two-dimensional toric periodic lattices) used to demonstrate real-world activity spaces. The mirror identity concept connects the two layers, resulting in a small-world network model for analyzing the transmission dynamics of epidemic diseases and social issues.

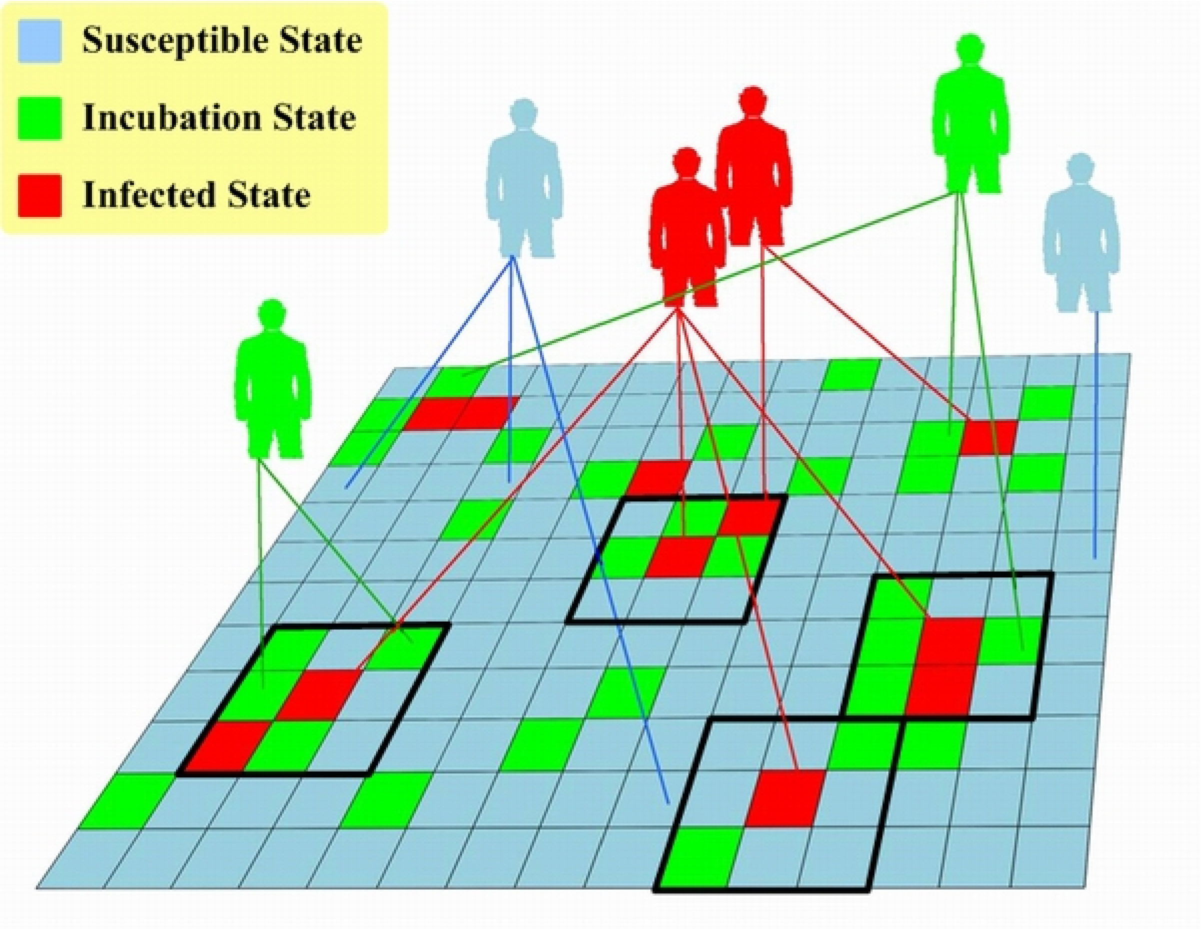


Figure 4.2.1: Cellular automata with mirror identity model (CAMIM).

Cellular Automata with Mirror Identities Model (CAMIM)

The abstract mirror identity concept is based on human interactions and daily routines within the confines of a modern society. It specifies an individual's social attributes—for instance, long-distance movement, daily visits to fixed locations, and multiple-activity locations. In our proposed model, individuals are viewed as single agent entities of an upper-layer multi-agent system; the places that an individual visits on a regular basis are defined as mirror identities.

We will use Andy, a retired senior citizen who lives alone, as an example. Every morning at 8:00 a.m. he rides his motor scooter to a suburban nursing home, where he serves as a volunteer. He helps a nurse named Cindy to provide care for three residents named Bob, Dick, and Eric. Every evening at 6:00 p.m. he eats at an inexpensive Japanese restaurant, where he usually chats with the owner, the chef, and several other regular customers. After dinner, he goes home, changes clothes, then goes to a neighborhood tavern to spend some time with his friends Frank and Gerry. Andy rarely deviates from this routine. According to our proposed model, Andy, Bob, Cindy, Dick and Eric are upper-layer agents, and Andy's home, the nursing home, the Japanese restaurant and the tavern are lower-layer mirror identities. Note that his motor scooter is considered an extension of his home instead of an activity node, since he rarely rides with others.

Each agent in the upper layer has a set of attributes describing its epidemiological progress and social mobility states (Table 4.2.1 and Fig. 4.2.2). Each mirror identity, which can freely access agent attributes, has a group of private attributes that represent its current status and local data (Table 4.2.2). Agents can freely access the attributes of any mirror identity they are connected to. Furthermore, agents can use their mirror identities to form clusters with other agents. For example, Andy belongs to three groups—one each at the nursing home, the tavern, and the Japanese restaurant. In formal terms, all of an agent's mirror identities are connected through that agent; they form a star-shaped topology with the agent at the center and the mirror identities at the vertices.

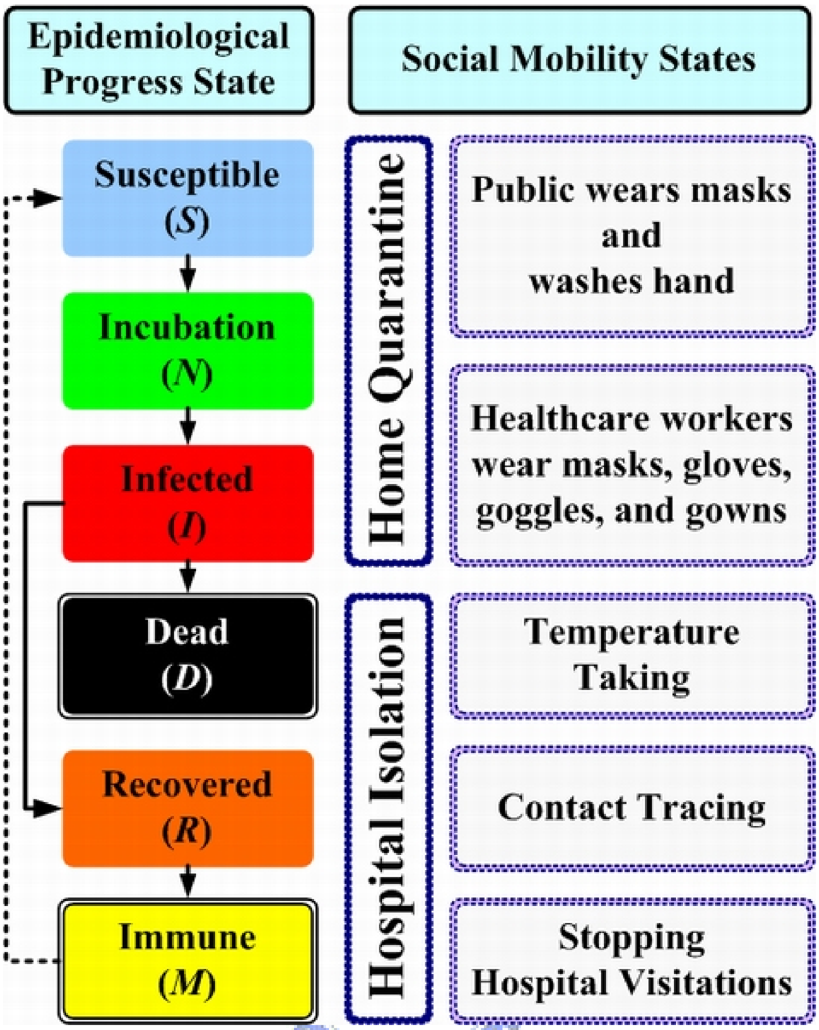


Figure 4.2.2: Epidemiological states and two social mobility states.

Table 4.2.1: Agent Attributes.

Attribute	**Type**	**Description**	**Value**
ID	Integer	Unique serial number that identifies virtual society agent; sequence not considered critical.	1~*P*
E	Symbol	When the $Population_{Agent}$ parameter is set, the configured $Rate_{ForeverImmune}$ determines the rate of agents classified as *M* (Immune) in the epidemiological progress state *E*—that is, the population of permanently immune agents. All remaining agents are classified as *S* (Susceptible), meaning “not yet infected but prone to infection.”	**Susceptible**, Incubation, Infected, Recovered, **Immune**, Dead
Mobility	Symbol	When the $Population_{Agent}$ parameter is set, the *Mobility* status of every agent is preset as “free”—that is, these agents have no restrictions in terms of interacting with the mirror identities of neighboring agents. When an agent is placed under home quarantine or hospital isolation, its *Mobility* status is respectively changed to Quarantined or Isolated. This means the agent is restricted to its rooted mirror identity (e.g., home, hospital, or dormitory), and the activities of all mirror identities are temporarily suspended.	**Free**, Quarantined, Isolated
Count	Integer	Records the number of an agent’s mirror identities; every agent has a minimum of one and a maximum of *M*. These numbers are normally distributed.	1~*M*
MirrorIdentity	Set	Data structure for containing mirror identities; each contains at least one.	
Age	Symbol	In the proposed model, agents are divided according to three age levels: young (1 to 20), prime (21 to 60); and old (61 and higher). When a simulation system is initiated, agent ages are randomly set based on the $Rate_{Young}$, $Rate_{Prime}$, and $Rate_{Old}$ parameters.	**Young**, **Prime**, **Old**
Super	Boolean	Denotes whether an agent is a super-spreader. If yes, set *Super* to “true”; if no, to “false”. When simulation system is initiated, the $Rate_{Super}$ parameter is used to determine which agents are super-spreaders.	**true**, **false**
$Immunity_{Permanent}$	Boolean	Denotes whether an agent is permanently immune. If yes, set $Immunity_{Permanenty}$ to “true”; if no, to “false”. When simulation system is initiated, the $Rate_{PermanentImmunity}$ parameter is used to determine which agents are permanently immune.	**true**, **false**
Day	Integer	Number of days for the three epidemiological progress states. If an infected agent has not yet recovered, *Day* is used to indicate the number of infected days; for recovered agents, *Day* is used to indicate the number of days since full recovery. If a recovered agent has temporary antibodies, *Day* is used to indicate the number of immune days.	
$Rate_{Contact}$	Real	Rate of contact with other agents. For all agents, $Rate_{Contact}$ is normally distributed.	0~1
WearingMask	Boolean	Denotes whether agent wears a mask. If yes, set *WearingMask* to “true”; if no, to “false”. When simulation system is initiated, the *WearingMask* attribute for all agents is preset to “false”. When a mask wearing policy is enacted (for the general public or healthcare workers), the $Rate_{Participation}$ parameter is used to determine which agents wear masks.	true, **false**
MaskType	Real	Represents average prevention grade of agent masks. The higher the number (close to 1), the greater the efficacy.	0~1
$Quarantined_{Day}$	Integer	Number of home quarantine days, with a range of 0 to $Policy.Parameter.Day_{Quarantined}$.	

Table 4.2.2: Mirror Identity Attributes.

Attribute	**Type**	**Description**	**Value**
Root	Boolean	Each agent has one mirror identity whose *Root* = true; for all other mirror identities, *Root* = false. The root mirror identity is used to mimic special activity locations—e.g., homes, hospitals, and dormitories.	**true**, **false**
Suspend	Boolean	When the simulation system is initiated, *Suspend* = false for all agent mirror identities, denoting that all mirror identities are free to move about without any restrictions. Except for its rooted mirror identity, *Suspend* = true for all of the mirror identities of an agent in home quarantine or hospital isolation. This represents the idea that the agent cannot move about until the end of the home quarantine or recovery period. If the agent dies, *Suspend* = true for all mirror identities (including the rooted mirror identity), representing the idea that the agent can no longer visit any other locations.	true, **false**
Location	(Integer, Integer)	The first number represents the x-axis coordinate and the second number the y-axis coordinate for the location of a mirror identity in the two-dimensional lattice. Each mirror identity is mapped to a single coordinate location; in other words, each coordinate location contains a single mirror identity of only one agent.	
Neighbor	Set	Represents the coordinate locations of an agent's eight mirror identities. Moore or von Neumann neighborhood relationships are used in most simulation systems. Under the Moore system, each mirror identity is defined as having eight neighbor agents; under the von Neumann system the number is four. We adopted the Moore neighborhood definition for our SARS simulation experiments.	

The majority of agents have between 2 and 5 mirror identities, with the number of mirror identities connected to an agent representing a normal distribution. In our proposed model, the more mirror identities an agent has, the larger the number of activity nodes and the greater the agent's influence. Using an epidemic disease as an example, the greater the number of lattices connected to an agent, the greater the chances that the agent will become infected and/or transmit the disease to other agents. Lattices that surround each other in cellular automata represent neighbors—for example, the mirror identities of Andy, Cindy, Bob, Dick, and Eric are adjacent to each other, and Andy's tavern mirror identity is adjacent to Frank and Gerry's.

In our model, one discrete time step is the equivalent of one day in the real world. The states of agents and their mirror identities change simultaneously during each discrete time step, and each agent’s mirror identity comes into contact with its surrounding mirror identities. The attributes of the agent, its mirror identity, and surrounding mirror identities vary according to the interaction rules described in sections 4.3 and 4.4, simulation and epidemic parameters (Table 4.2.3), public health policy parameters (Table 4.2.4), input data tables (Table 4.2.5), and various random values. Accordingly, our combination of cellular automata and mirror identities is capable of displaying multiple social network characteristics: fixed locations visited daily, long-distance movement, local clustering, high degrees of clustering, and low degrees of separation.

Table 4.2.3: Simulation System and Epidemic Infection Parameters.

Attribute	**Type**	**Description**
$Population_{Agent}$	Set	Stores total agent population in the simulation system; maximum capacity is P agents.
P	Integer	Total number of agents.
M	Integer	Upper limit of an agent's mirror identities.
H	Integer	Height of the two-dimensional lattice used in the cellular automata.
W	Integer	Width of the two-dimensional lattice used in the cellular automata.
N	Integer	Total number of usable lattice H × W in the cellular automata.
$Day_{Incubation}$	Integer	Average number of incubation days.
$Day_{Infectious}$	Integer	Average number of infectious days.
$Day_{Recovered}$	Integer	Average number of recovered days.
Day_{Immune}	Integer	Temporarily immune to the disease; average number of incubation days.
$Rate_{Super}$	Real	Percentage of super-spreaders among total population.
$Rate_{Young}$	Real	Percentage of young (0 to 20 years) agents in total population.
$Rate_{Prime}$	Real	Percentage of prime (21 to 60 years) agents in total population.
$Rate_{Old}$	Real	Percentage of old (60 years and above) agents in total population.
$Rate_{ForeverImmunity}$	Real	Percentage of permanently immune agents in total population.
$Rate_{Infection}$	Real	Average infection rate.
$Rate_{Death}$	Real	Average death rate.

Table 4.2.4: Public Health Policy Parameters.

Policy	**Attribute**	**Type**	**Description**	**Value**
WearingMaskInGP	$Rate_{Participation}$	Real	Policy participation rate.	0~1
	$Rate_{Prevention}$	Real	Infectious disease prevention rate.	0~1
WearningMaskInHW	$Rate_{Participation}$	Real	Policy participation rate.	0~1
	$Rate_{Prevention}$	Real	Infectious disease prevention rate.	0~1
TemperatureMeasuring	$Rate_{Detection}$	Real	Fever detection success rate.	0~1
	$Rate_{Participation}$	Real	Measurement participation rate.	0~1
HomeQuarantine	*Class*	Symbol	A- and B-class quarantines.	A, B
	$Day_{Quarantined}$	Integer	Number of home quarantine days.	0~1
	$Rate_{Participation}$	Real	Policy participation rate	0~1
RestrictingAccessToHospitals	$Rate_{Participation}$	Real	Policy participation rate.	0~1
ReducingPublicContact	$Rate_{Participation}$	Real	Policy participation rate.	0~1

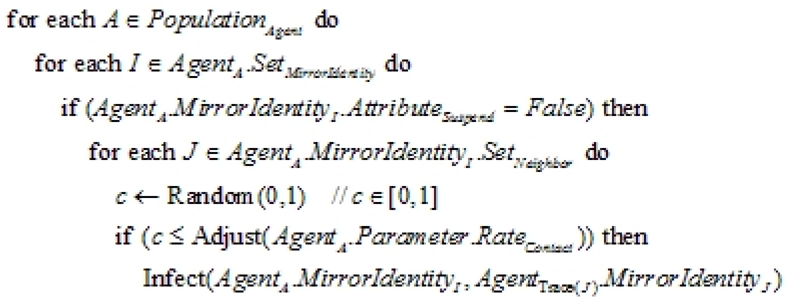
Table 4.2.5: Input Data for Simulating SARS Epidemic Curves in Taiwan, Singapore, and Toronto.

Category	**Attribute**	**Type**	**Description**	**Value**
Imported Cases	*Time Point*	Date	Date when the imported case occurred.	
	Amount	Integer	Number of patients.	0~999
	Phase	Symbol	Imported during incubation or illness period	Incubation, **Infected**
	Super-spreader	Boolean	Determine whether the imported patient is a super-spreader.	true, **false**
Public Health Policy	*Related Attributes*	See Table 4		
Run	*Day*	Integer	Number of execution days.	0~99

4.3. Modeling Epidemiological Features

From Contact to Infection to Symptom

Based on an adjusted contact rate (*Agent.Parameter.Rate*$_{Contact}$) and a random number *c*, the mirror identities of each agent determines whether or not it will interact individually with the mirror identities of eight adjacent neighbors. If the *c* is lower than the contact rate, the mirror identity of agent A comes into contact with the mirror identity of neighbor agent B. The contact rate *Agent.Parameter.Rate*$_{Contact}$ depends on whether a “reducing public contact” policy or other parameter settings have been enacted. Throughout this section, we will express these concepts in pseudo-code; here the pseudo-code is



```
for each $A \in Population_{Agent}$ do
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    if $(Agent_A.MirrorIdentity_I.Attribute_{Suspend} = False)$ then
      for each $J \in Agent_A.MirrorIdentity_I.Set_{Neighbor}$ do
        $c \leftarrow \mathrm{Random}(0,1)$   // $c \in [0,1]$
        if $(c \leq \mathrm{Adjust}(Agent_A.Parameter.Rate_{Contact}))$ then
          $\mathrm{Infect}(Agent_A.MirrorIdentity_I, Agent_{\mathrm{Trace}(J)}.MirrorIdentity_J)$
```

Assume that agent A has a mirror identity that is adjacent to a mirror identity of agent B, that agent A has been infected and is contagious, and that agent B is both susceptible and prone to infection. When the two agents come into contact, a combination of infection rate (*System.Parameter.Rate*$_{Infection}$) and a random number *n* determines whether or not agent B is infected by agent A. If the *n* is lower than the infection rate, agent B's epidemiological state becomes *N* (incubation) and the period attribute (*Agent.Attribute*$_{Day}$) becomes 1 (denoting that symptoms have not appeared and that agent B cannot transmit the disease). The infection rate *System.Parameter.Rate*$_{Infection}$ is determined by such factors as immunity rate—that is, whether agent A is a super-spreader (CDC 2003b; Sebastian and Hoffmann 2003), in home quarantine, in hospital isolation, etc.

```
if (Contact($Agent_A.MirrorIdentity_I, Agent_B.MirrorIdentity_I$)) then
  if ($Agent_A.Attribute_E = I \wedge Agent_B.Attribute_E = S$) then
    $n \leftarrow$ Random(0,1)   // $n \in [0,1]$
    if ($n \leq$ Adjust($System.Parameter.Rate_{Infection}$)) then
      $Agent_B.Attribute_E \leftarrow N$   // $N$ means incubation
      $Agent_B.Attribute_{Day} \leftarrow 1$
```

Agent A's epidemiological state will automatically change from *N* to *I* (Infected) once it has exceeded the incubation period *System.Parameter.Day*$_{Incubation}$.

```
if ($Agent_A.Attribute_E = N$) then
  if ($Agent_A.Attribute_{Day} > System.Parameter.Day_{Incubation}$) then
    $Agent_A.Attribute_E \leftarrow I$   // $I$ means Infectious
```

When agent A's epidemiological state is *I* and it has exceeded the infectious period *System.Parameter.Day*$_{Infectious}$, a combination of the adjusted death rate (*System.Parameter.Rate*$_{Death}$) and a random number *d* determines whether the agent enters the *D* (Death) or *R* (Recovered) state. Death rates are determined by such factors as age, whether the agent was placed under home quarantine throughout its incubation and infective periods, whether it received treatment in hospital isolation, and its public activities (if any) during the period of illness.

```
if ($Agent_A.Attribute_E = I$) then
    if ($Agent_A.Attribute_{Day} > System.Parameter.Day_{(Incubation+Infectious)}$) then
        $d \leftarrow$ Random(0,1)    // $d$ means death rate
        if ($d \leq$ Adjust($System.Parameter.Rate_{Death}$)) then
            $Agent_A.Attribute_E \leftarrow D$    // $D$ means Died
            $Agent_A.Attribute_{Day} \leftarrow 0$
        else
            $Agent_A.Attribute_E \leftarrow R$    // $R$ means Recovered
            $Agent_A.Attribute_{Day} \leftarrow 1$
```

When agent A’s epidemiological state is *R* and it has exceeded recovery period *System.Parameter.Day*$_{Recovered}$, it automatically enters an *M* (Immune) state.

```
if ($Agent_A.Attribute_E = R$) then
    if ($Agent_A.Attribute_{Day} > System.Parameter.Day_{Recovered}$) then
        $Agent_A.Attribute_E \leftarrow M$    // $M$ means Immune
```

In the *M* state, the *Agent.AttributeForever*$_{Immune}$ parameter is used to determine whether agent A’s immunity is permanent or temporary—that is, whether complete recovery or renewed susceptibility occurs following *System.Parameter.Day*$_{Immune}$.

```
if ($Agent_A.Attribute_E = M \wedge$ not $Agent_A.Attribute_{ForeverImmune}$) then
    if ($Agent_A.Attribute_{Day} > System.Parameter.Day_{Immune}$) then
        $Agent_A.Attribute_E \leftarrow S$    // $S$ means Susceptible
        $Agent_A.Attribute_{Day} \leftarrow 0$
```

Families and Hospitals

Our proposed model can also be used to represent such concepts as homes, dormitories, and hospitals. As shown in Table 2, all mirror identities have two private attributes: *root* and *suspend*. For most agents, one mirror identity’s root attribute is designated as true but the root attributes of its other mirror identities are designated as false. In contrast, the suspend attributes of all mirror identities of an agent are designated as false. To facilitate later discussion, we will assume the presence of a rooted mirror identity–that is, a mirror identity whose root attribute is always designated as true. The rooted mirror identity can be used to represent such unique (e.g., one-of-a-kind) units as homes, dormitories, and hospitals.

```
for each $A \in Population_{Agent}$ do
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow False$
    $Agent_A.MirrorIdentity_I.Attribute_{Root} \leftarrow False$
  $n \leftarrow$ Random(1, Count($Agent_A.Set_{MirrorIdentity}$))
  $Agent_A.MirrorIdentity_{Index(n)}.Attribute_{Root} \leftarrow True$
```

If a health authority enforces a home quarantine of agent A, then the suspend attributes of all its mirror identities (workplace, school, bus stations, and so on) are marked as true; the one exception is agent A's rooted mirror identity—that is, its home. The lattice points surrounding agent A's rooted mirror identity represent the mirror identities of the agent's family members or cohabitants. Once the home quarantine is lifted, the suspend attributes of these mirror identities (except for that of the rooted mirror identity) return to false, indicating a resumption of normal agent A activities.

```
if (IsQuarantine($Agent_A$)) then
  $Agent_A.Attribute_{Mobility} \leftarrow Quarantined$
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    if ($Agent_A.MirrorIdentity_I.Attribute_{Root} = False$) then
      $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow True$
if (not IsQuarantine($Agent_A$)) then
  $Agent_A.Attribute_{Mobility} \leftarrow Free$
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow False$
```

We believe another advantage of the model is that it does not require fixed areas of lattice points representing hospitals. Assume that agent B, with a confirmed epidemiological state of *I*, enters isolation voluntarily. Similar to the preceding example, the suspend attributes of all agent B mirror identities are changed to true, with the exception of its rooted mirror identity. This represents a scenario where agent B is receiving treatment in hospital isolation, and where the entire agent's outside activities cease. The lattice points surrounding agent B's rooted mirror identity represent medical staff, nurses, healthcare workers, and perhaps family members. If agent B recovers, the suspend attributes of the affected mirror identities return to false, indicating a resumption of agent B's normal activities. If the agent dies, the suspend attributes of all agent B mirror identities (including its root mirror identity) are permanently changed to false, indicating the permanent cessation of all of the agent's activities.

```
if (Isolated($Agent_A$)) then
  $Agent_A.Attribute_{Mobility} \leftarrow Agent_A.Attribute_{Mobility} + Isolated$
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    if ($Agent_A.MirrorIdentity_I.Attribute_{Root} = False$) then
      $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow True$
if (not Isolated($Agent_A$) $\wedge$ $Agent_A.Attribute_E = R$) then
  $Agent_A.Attribute_{Mobility} \leftarrow Agent_A.Attribute_{Mobility} - Isolated$
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow False$
if (not Isolated($Agent_A$) $\wedge$ $Agent_A.Attribute_E = D$) then
  $Agent_A.Attribute_{Mobility} \leftarrow Agent_A.Attribute_{Mobility} - Isolated$
  for each $I \in Agent_A.Set_{MirrorIdentity}$ do
    $Agent_A.MirrorIdentity_I.Attribute_{Suspend} \leftarrow True$
```

4.4. Modeling Public Health Policies

Mask Policy—General Public vs. Healthcare Workers

A mask-wearing policy for the general public has two parameters: participation rate and prevention efficiency. Participation rate refers to the percentage of individuals in the total population who actually wear masks, and prevention efficiency represents the protection grade of the masks being used. Both parameters are adjustable. When this policy is enacted, the simulation system uses the participation rate to randomly assign a number of individuals who abide by wearing masks. If agent A in a simulation system has an *S* status but wears a mask, its infection probability decreases in accordance with the prevention efficiency parameter. The chances of an *I*-status agent A infecting others decreases if the simulated agent wears a mask before and after the outbreak of symptoms; this potential is also affected by the prevention efficiency parameter.

```
if ($\mathrm{On}(Policy_{WearingMaskInGP}) \vee \mathrm{Change}(Policy_{WearingMaskInGP})$) then
  if ($Policy_{WearingMaskInGP}.Parameter.Rate_{Participation} > 0$) then
    for each $A \in Population_{Agent}$ do
      $n \leftarrow \mathrm{Random}(0,1)$   // $n \in [0,1]$
      if ($n \leq Policy_{WearingMaskInGP}.Parameter.Rate_{Participation}$) then
        $Agent_A.Attribute_{WearingMask} \leftarrow True$
        $Agent_A.Attribute_{MaskType} \leftarrow Policy_{WearingMaskInGP}.Parameter.Rate_{Prevention}$
      else
        $Agent_A.Attribute_{WearingMask} \leftarrow False$
```

The same process used to represent hospitals can also be used to simulate a mask-wearing policy. Once the policy is enacted, agents surrounding the rooted mirror identity of agents in hospital isolation either wear or don’t wear masks based on the participation rate parameter; the prevention efficiency parameter also determines whether or not the infection probability of neighboring agents is reduced.

```
if (On($Policy_{WearingMaskInHW}$) $\vee$ Change($Policy_{WearingMaskInHW}$)) then
  when event (Isolated($Agent_A$)) do
    for each $N \in Agent_A.MirrorIdentity_{Root}.Set_{Neighbor}$ do
      $n \leftarrow$ Random(0,1)   // $n \in [0,1]$
      if ($n \leq Policy_{WearingMaskInHW}.Parameter.Rate_{Participation}$) then
        $Agent_{Trace(N)}.Attribute_{WearingMask} \leftarrow True$
        $Agent_{Trace(N)}.Attribute_{MaskType} \leftarrow Policy_{WearingMaskInHW}.Parameter.Rate_{Prevention}$
      else
        $Agent_{Trace(N)}.Attribute_{WearingMask} \leftarrow False$
```

Taking Body Temperature

If a temperature measurement policy is enforced, the mirror identities of each agent will be claimed by its surrounding agents collectively whether it should taken body temperature before it comes into contact with them. This decision is made based on a combination of a participation rate parameter and a random number *n*. An *n* that is lower than the participation rate means that neighboring agents are abiding by the policy of measuring the temperatures of agents that want to come into contact with them. Results depend on the detection rate parameter—in other words, the higher the detection rate and the more accurate the thermometers being used, the lower the rate of spreading the disease.

Reducing Public Contact

At the end of the 2002-2003 SARS epidemic, there were many reports (e.g., Sebastian and Hoffmann 2003; WHO 2003) describing the reduction of public contact as an effective means of controlling the spread of the disease. After this policy was enacted in our simulation, the combination of the participation rate parameter and a random number *n* determined whether or not the mirror identities of two agents interacted before coming into physical contact. An *n* higher than the participation rate indicated that either an agent had decided against coming into contact with agents surrounding a particular mirror identity, or simply had no reason for mirror identity interactions.

A/B Class Home Quarantines

According to an A-class home quarantine policy, if agent C is identified as being ill after such a policy is enacted, all agents surrounding agent C's mirror identities must decide whether they should go into home quarantine based on the participation rate parameter. As in the hospital isolation example described above, all mirror identities of neighboring agents that decide to enter home quarantine immediately stop all activities until the separation period is completed, as determined by a public health policy parameter. This requirement does not apply to rooted mirror identities, which are still allowed to come into contact with other agents.

A B-class home quarantine policy is similar to an A-class policy, but it affects a slightly larger number of agents than the A-class policy. If one mirror identity of agent C is adjacent to a particular mirror identity of agent D (e.g., agents C and D are a cohabiting couple), this represents one degree of separation; if one mirror identity of agent D is adjacent to a particular mirror identity of agent E (e.g., coworkers in the same office), this represents two degrees of separation between agents E and C. Accordingly, when agent C is diagnosed with the disease, both D and E face the risk of infection, meaning that both D and E must enter home quarantine.

Controlling Hospital Access

During the actual SARS epidemic, Singaporean and Taiwanese health authorities imposed strict rules concerning hospital visitations (Sebastian and Hoffmann 2003); we simulated this "controlling hospital access" policy using our proposed model. We assumed that agent A showed symptoms of the disease and was admitted to a hospital for treatment in isolation. If agent B's rooted mirror identity is adjacent to agent A's rooted mirror identity, it indicates that agent B may be a member the hospital staff, a nurse, a healthcare worker, or a very close relative; if agent C's non-rooted mirror identities are adjacent to agent A's rooted mirror identity, it indicates that agent C is a distant relative, friend, classmate, or coworker. If a strict visitation policy is enacted, agent B is allowed to visit agent A, but agent C is not.

4.5. Simulating SARS with CAMIM

After initializing the model and establishing parameters according to SARS disease information disseminated by the Centers for Disease Control (CDC)(2003a, 2003c, 2003f, 2003h) and World Health Organization (WHO)(2003) (Table 4.2.5), we ran simulations of SARS transmission dynamics in various geographic areas and compared the effectiveness of various public health policies and disease prevention strategies (Figs. 4.5.1 and 4.5.2). SARS originated in Guangdong, in southern China, therefore in all other countries it is considered an imported virus. We therefore used imported cases announced by health authorities as our model's simulation trigger (Appendix A). For each simulation we included the number of infectious people who entered a country, the discrete time step during which they entered, and whether or not they were exposed or infected as they entered a country. We triggered various public health policies according to the actual announcements of local health authorities, and adjusted our simulation environment, epidemic, and public health policy parameters according to actual disease information presented by the CDC (2003b, 2003d, 2003e, 2003g) and Sebastian and Hoffmann (2003). In other words, our model makes use of actual epidemic parameter values from the CDC, WHO, and the health authorities of affected countries, thus avoiding the use of derived or estimated data.

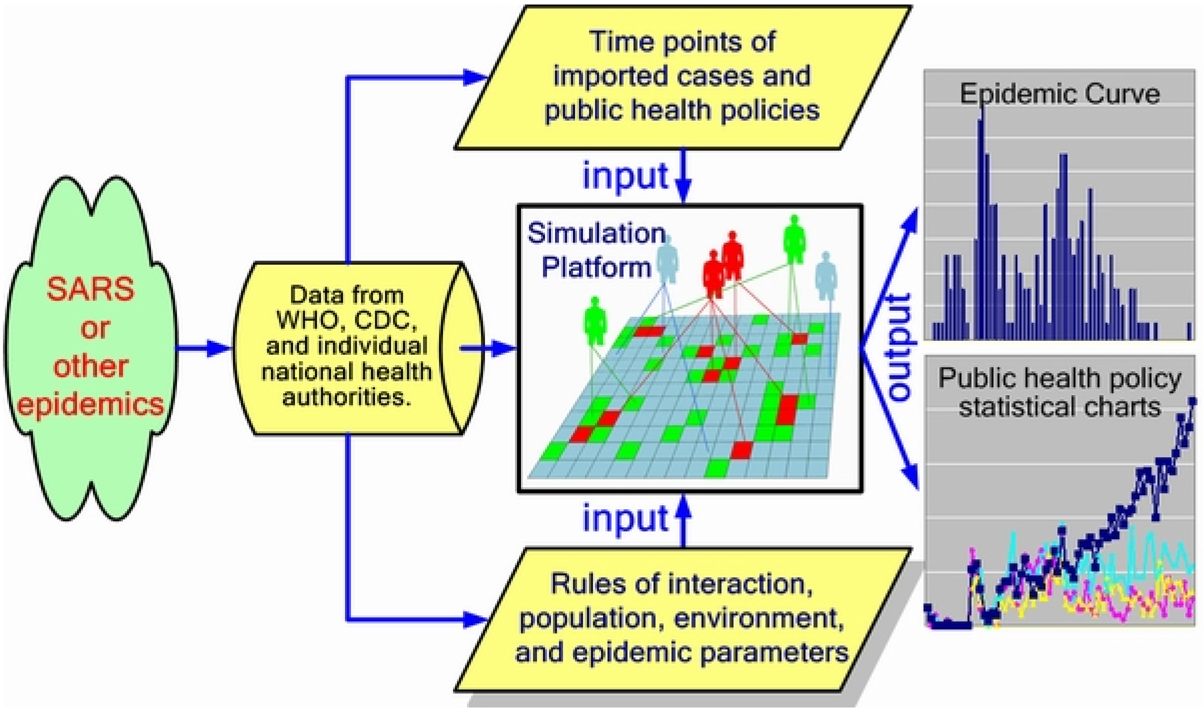
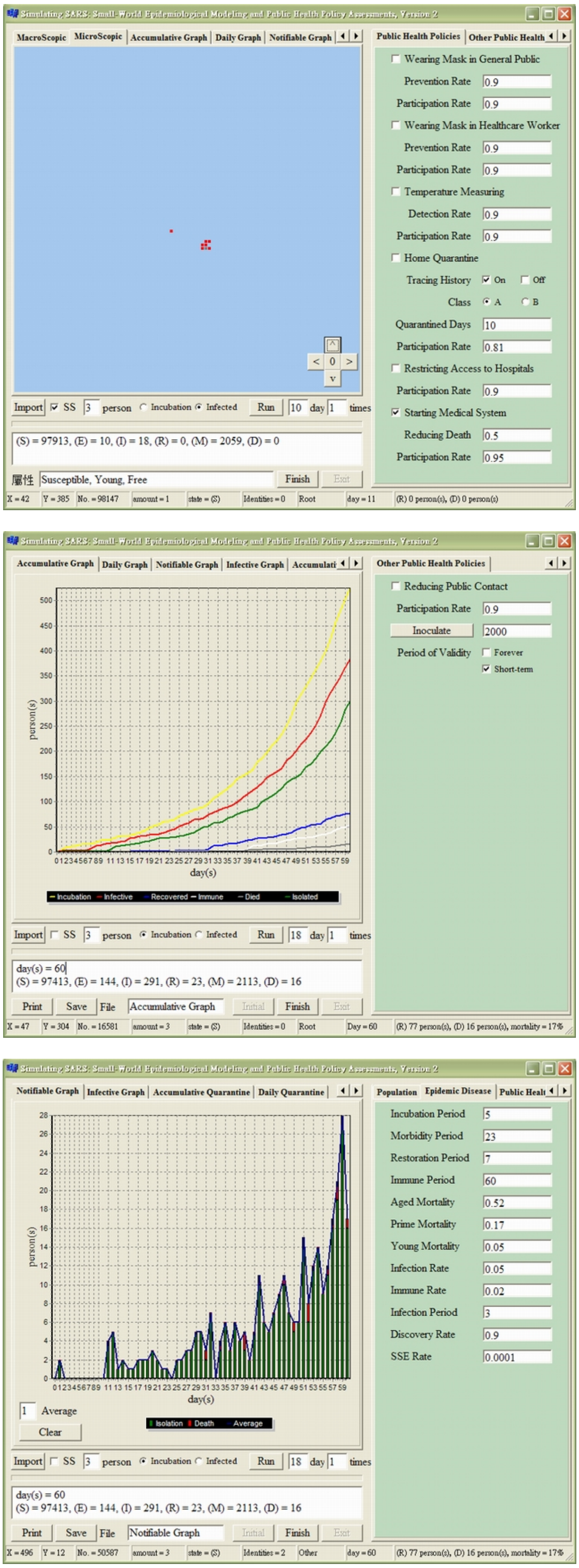


Figure 4.5.1: Simulation framework. Data on reported cases came from the World Health Organization (WHO) and health authorities in Singapore, Taiwan, and Toronto. Input data was distributed into three categories: epidemic parameters (e.g., average incubation period, infection rate, distribution among age groups, mortality); imported cases (e.g., time point, amount, imported during incubation or illness period); and public health policies, activated according to data from individual nations (e.g., number of quarantine days, efforts to take body temperatures, restricting access to hospitals). Simulation output includes cellular automata states and various statistical charts.



☐ Wearing Mask in General Public
Prevention Rate 0.9
Participation Rate 0.9
☐ Wearing Mask in Healthcare Worker
Prevention Rate 0.9
Participation Rate 0.9
☐ Temperature Measuring
Detection Rate 0.9
Participation Rate 0.9
☐ Home Quarantine
Tracing History ☑ On ☐ Off
Class ● A ○ B
Quarantined Days 10
Participation Rate 0.81
☐ Restricting Access to Hospitals
Participation Rate 0.9
☑ Starting Medical System
Reducing Death 0.5
Participation Rate 0.95

☑ SS 3 person ○ Incubation ● Infected 10 day 1 times

(S) = 97913, (E) = 10, (I) = 18, (R) = 0, (M) = 2059, (D) = 0

屬性 Susceptible, Young, Free

X = 42 Y = 385 No. = 98147 amount = 1 state = (S) Identities = 0 Root day = 11 (R) 0 person(s), (D) 0 person(s)

☐ Reducing Public Contact
Participation Rate 0.9
Inoculate 2000
Period of Validity ☐ Forever ☑ Short-term

☐ SS 3 person ● Incubation ○ Infected 18 day 1 times

day(s) = 60
(S) = 97413, (E) = 144, (I) = 291, (R) = 23, (M) = 2113, (D) = 16

File Accumulative Graph

X = 47 Y = 304 No. = 16581 amount = 3 state = (S) Identities = 0 Root Day = 60 (R) 77 person(s), (D) 16 person(s), mortality = 17%

Incubation Period 5
Morbidity Period 23
Restoration Period 7
Immune Period 60
Aged Mortality 0.52
Prime Mortality 0.17
Young Mortality 0.05
Infection Rate 0.05
Immune Rate 0.02
Infection Period 3
Discovery Rate 0.9
SSE Rate 0.0001

1 Average

☐ SS 3 person ● Incubation ○ Infected 18 day 1 times

day(s) = 60
(S) = 97413, (E) = 144, (I) = 291, (R) = 23, (M) = 2113, (D) = 16

File Notifiable Graph

X = 496 Y = 12 No. = 50587 amount = 3 state = (S) Identities = 2 Other day = 60 (R) 77 person(s), (D) 16 person(s), mortality = 17%

Figure 4.5.2: Simulation platform for contagious infection.

Singapore SARS Outbreak

According to the comparison of actual and simulated SARS cases in Singapore shown in Figure 4.5.3, the simulated curve has a very close fit with data published by the city-state’s health authority for the two outbreaks that occurred between February 25 and May 5 of 2003 (CDC 2003b; Sebastian and Hoffmann 2003; WHO 2003). The first outbreak was attributed to imported cases, and emergency public health policies were not activated. The second was attributed to the compound effects of secondary infections, and several emergency policies were put into effect on March 24 (e.g., a ban on visits to patients in hospitals or under home quarantine). The number of new cases dropped dramatically at the beginning of June, and soon afterwards the World Health Organization (WHO) announced that the disease was under control.

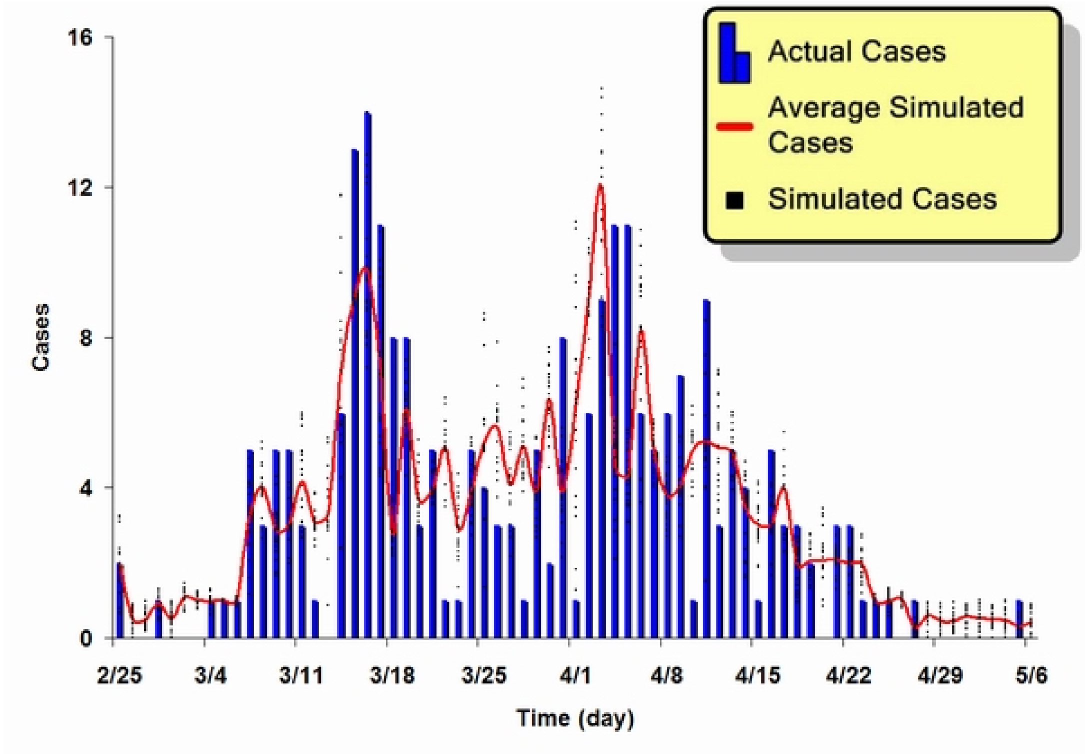


Figure 4.5.3: A Comparison of actual and simulated epidemic results for the SARS outbreak in Singapore. Blue bars represent actual reported cases, red line represents an average of 20 simulation results, and black dots represent 20 simulation results.

Taipei SARS Outbreak

Our simulation of the Taipei situation included several public health policies enforced by that city's government, including several grades of home quarantine and a mask-wearing requirement for all bus and train travelers (CDC 2003d; CDC 2003g; Sebastian and Hoffmann 2003; WHO 2003). As shown in Figure 4.5.4, the simulated results have a close fit with the epidemic curve of probable cases published by the Taiwanese health authority on September 28, 2003—that is, a major spike followed by several smaller outbreaks. We believe the heavier concentration in the Taipei curve (compared to Singapore's) is due to several different factors, including late case discoveries, delays in seeking treatment, illness cover-ups, public interactions, and the large number of cases imported by travelers returning from Hong Kong. In Singapore, all imported cases were reported prior to the first outbreak, and the second wave resulted from compound infections. In Taiwan, the reported s-curve is more representative of a typical infection pattern.

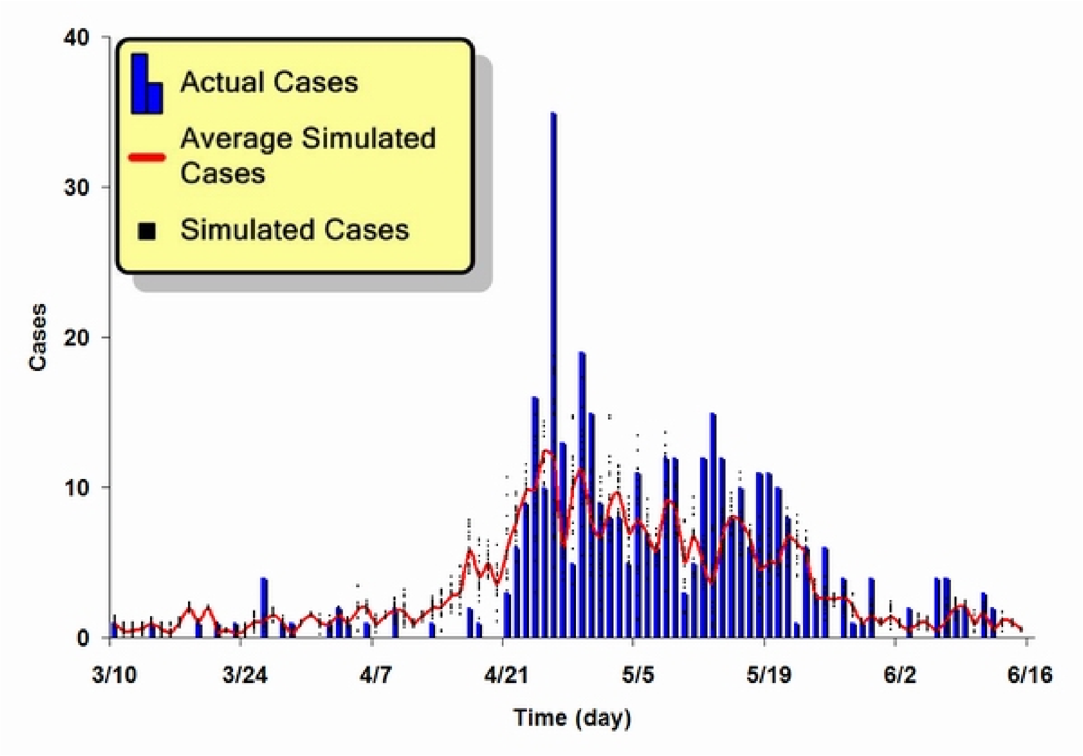


Figure 4.5.4: A comparison of actual and simulated epidemic results for the SARS outbreak in Taipei.

Toronto SARS Outbreak

In Toronto, the SARS scenario consisted of two major waves with almost no new cases in between (Fig. 4.5.5) (CDC 2003e; Sebastian and Hoffmann 2003; WHO 2003). After a re-examination of the data in August of 2003, the Canadian authorities acknowledged several additional cases during the lull period. According to our simulation, the second wave would not have been as severe if strong public health policies had been enforced for a longer period following the first wave. In our simulation, epidemic control measures—especially restricted hospital access and reduced public contact with infected persons—were relaxed after the first wave subsided. This resulted in a second spike occurring within a few days of the actual spike that was reported by Toronto health authorities. Our results matched Kamps-Hoffmann's (Sebastian and Hoffmann 2003) conclusion that the Toronto government lifted its control measures too quickly. Because of increased contact between patients and visitors and relaxed rules on the wearing of masks or respirators by health care workers, Toronto suffered a second nosocomial transmission period.

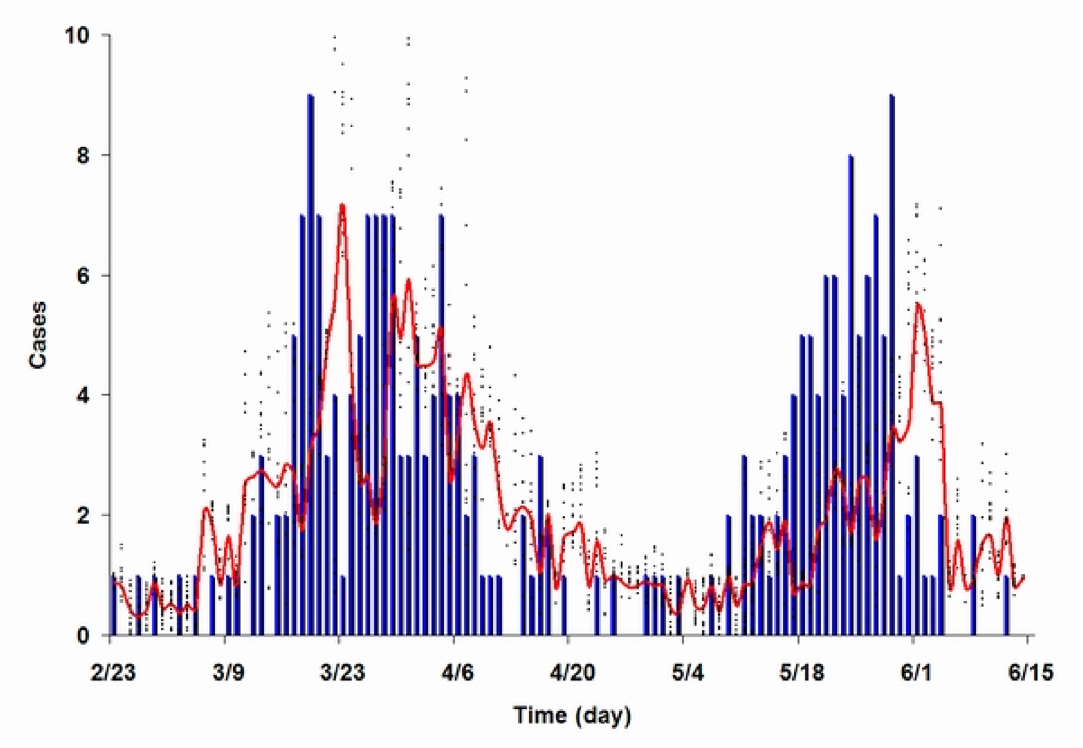


Figure 4.5.5: A comparison of actual and simulated epidemic results for the SARS outbreak in Toronto. We assumed that the second outbreak occurred because preventive policies were relaxed too soon following the first outbreak.

From the combined results of these simulations, we suggest that our proposed model is a useful tool for purposes of cross-checking hypothesized findings and for gaining insight into how infectious disease epidemics develop.

Home Quarantines

In addition to the above simulations, we tested our model using the home quarantine policy. After releasing details of the global SARS outbreak on March 12, 2003, WHO officials suggested that home quarantine periods should be at least twice as long as the then-average 4-6 day incubation period in order to suppress the spread of the disease (CDC 2003a; CDC 2003g; WHO 2003). Consequently, the governments of Singapore, Taiwan, and Canada established and enforced 10-day quarantine policies during the epidemic, and for a short period the Taiwanese government enforced a 14-day policy. According to our simulation results, a minimum 10-day quarantine period was required for suppressing the number of new cases—the same time period recommended by WHO (Fig. 4.5.6). We observed that the SARS epidemic curve slowed down considerably and that the disease became endemic when the quarantine period was a minimum 10 days, otherwise it was impossible to control the disease.

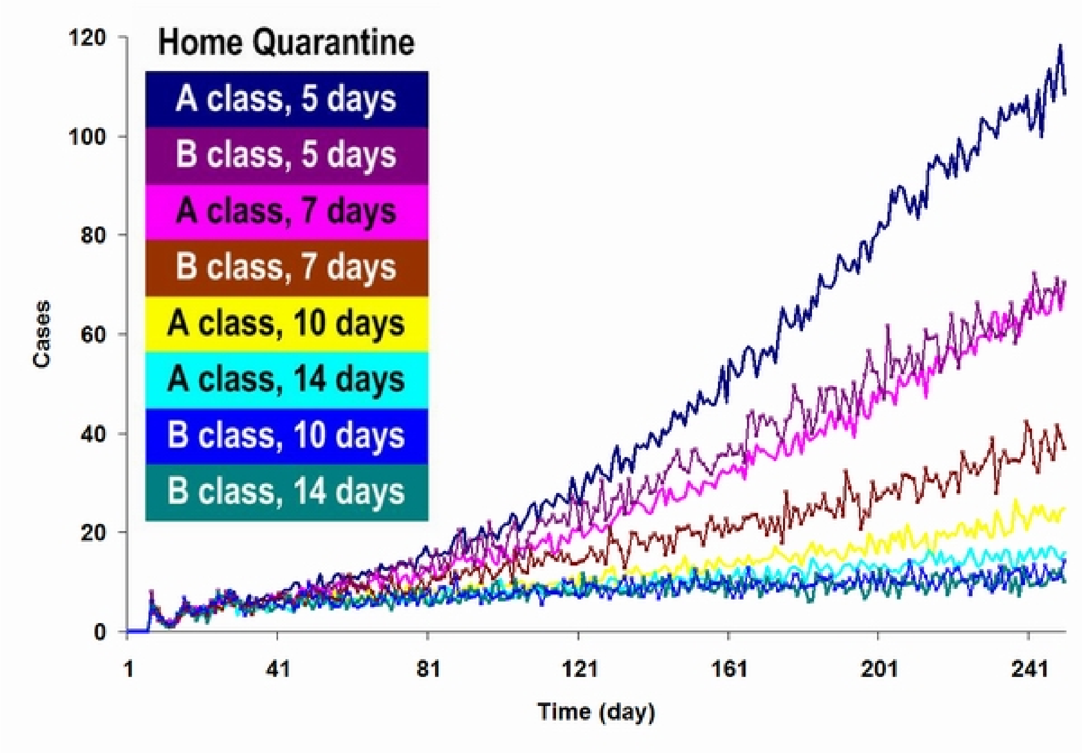


Figure 4.5.6: Results from a simulation based on various home quarantine policies. The time period for the simulation was 250 days, with a default incubation period of 5 days. The results indicate that different home quarantine restriction levels exerted different impacts on the SARS epidemic, and that a home quarantine policy by itself was insufficient for suppressing the epidemic.

4.6. Analyzing Public Health Policies

Taking Body Temperature

The Singaporean and Taiwanese governments both implemented temperature measurement policies during the epidemic, going so far as to launch national campaigns that included installing temperature-monitoring equipment and setting up manual temperature measurement stations at various government buildings, clinics, and public transportation stations (Sebastian and Hoffmann 2003; WHO 2003). According to our simulation results, when such policies are both comprehensive and compulsory, they reduce the number of feverish individuals entering public places. However, they are difficult to execute; implementation methods tend to vary, oversights are common, and an unknown number of individuals manage to evade having their temperatures taken.

The results from our simulation suggest that a participation rate of between 80 and 90 percent is required for this public health policy to have a positive effect in controlling a SARS epidemic (Fig. 4.6.1). At a rate of 65 percent or lower, the policy has little effect. In addition, the policy incurs significant social costs—providing inexpensive thermometers, setting up stations for their distribution, setting up temperature screening stations, and arranging for manual temperature measurements at various government buildings, medical clinics, and public transportation stations.

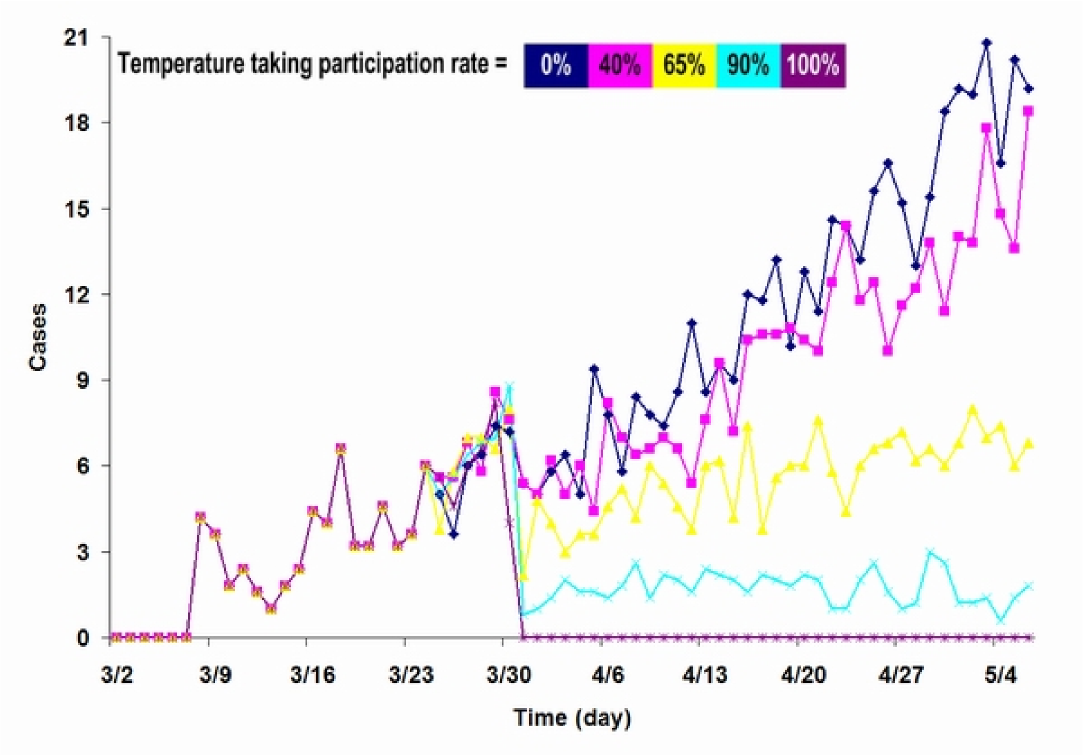


Figure 4.6.1: Results from a simulation focused on temperature measuring policy at different participation levels. We used the 8 imported case reported in Singapore to trigger the simulation. In each 66-day simulation run, the policy was activated on day 24; the goal was to compare impacts at different participation rates.

Wearing Masks with Different Protection Levels—General Public vs. Healthcare Workers

The governments of Taiwan and Hong Kong made great efforts to promote general mask-wearing policies, which led to hoarding and panic buying (Sebastian and Hoffmann 2003; WHO 2003). Masks are categorized according to grade—ordinary, surgical, N95 respirator masks, etc. In Taiwan, a serious shortage of professional masks for medical staff occurred following a mad rush by the general population to purchase masks regardless of grade; this triggered a debate on the necessity of wearing N95 respirator masks outside of hospitals and clinics.

According to the results presented in Figure 4.6.2, ordinary and surgical masks can assist in controlling an epidemic outbreak as long as wearing them becomes a strong habit for the desired time period. At a prevention efficiency of 65 percent or more (that is, the mask covers the mouth and nose), epidemics can be controlled but not eliminated. When wearing ordinary masks, medical staff members have higher infection rates (Figs. 4.6.2 and 4.6.3); these personnel clearly benefit from wearing N95 and other high-resistance masks in hospitals and other medical centers. From our simulation, we suggest that the general public does not require high-resistance masks, and that higher grade masks should be reserved for use by medical staff and healthcare workers.

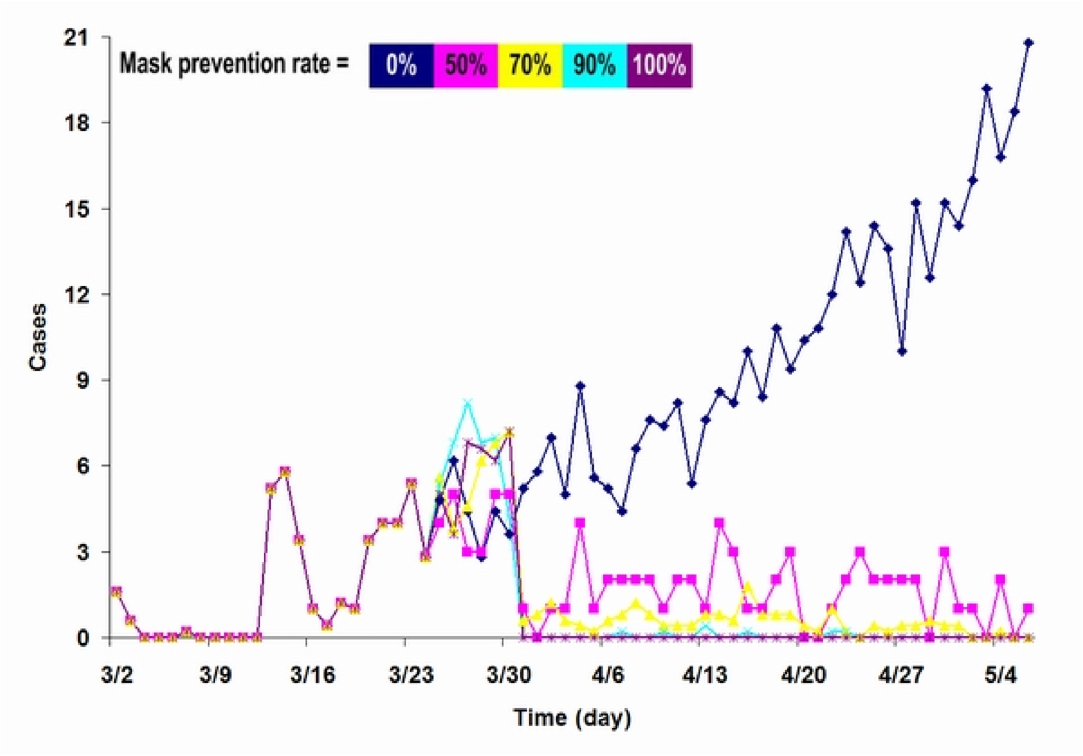


Figure 4.6.2: Results from a simulation focused on the impact of mask-wearing by the general public, comparing different mask protection levels.

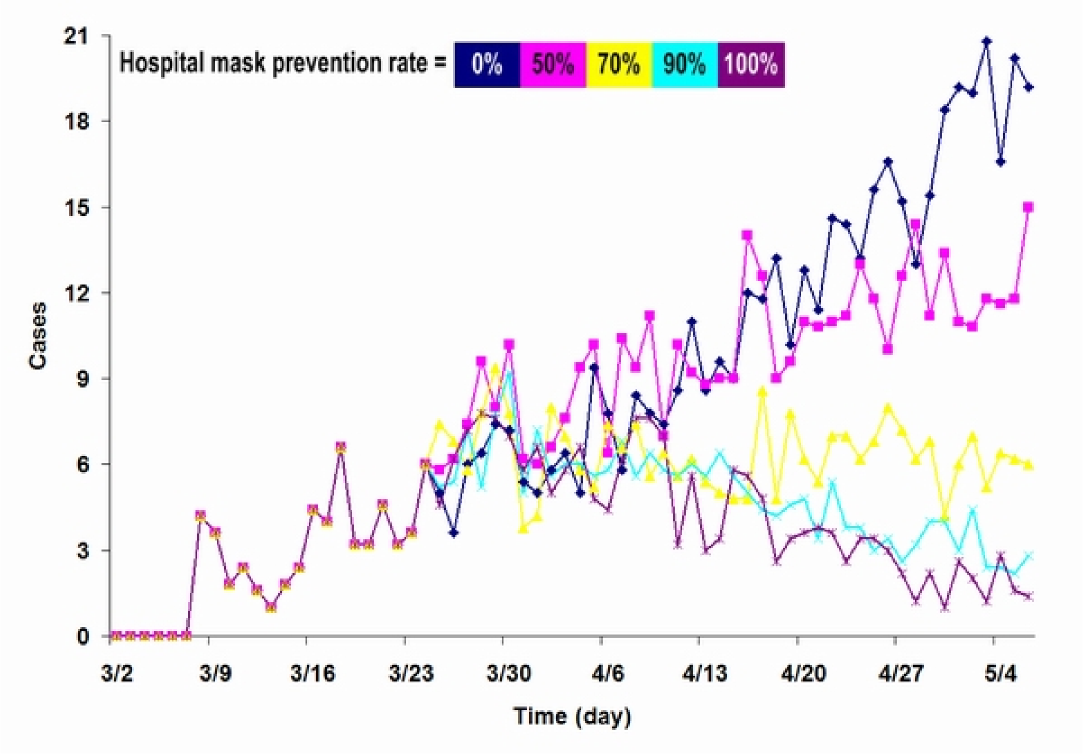


Figure 4.6.3: Results from a simulation focused on the impact of mask-wearing by healthcare workers in healthcare facilities, comparing different mask protection levels.

4.7. Assessing Public Health Suites

Different public health policies have different social costs. Home quarantining is highly effective, but it requires considerable amounts of labor and material resources compared to temperature measurement and mask-wearing policies. We ran simulations of various prevention strategies in an attempt to find an optimal combination of public health policies in terms of efficacy and cost, and found that a combination of mask-wearing by the general public and reducing contact in public places was the best combination for suppressing the spread of SARS (Fig. 4.7). Some costs are involved in mask purchases, but few costs are associated with limited public contact. In addition, mask wearing addresses an epidemic at its source—disease transmission.

The combined strategies of temperature measurements, restricted hospital visitations, and mask-wearing by healthcare workers should be considered a remedial reaction to a SARS outbreak. This strategy suite is ineffective in stopping patients in the incubation stage or patients suffering from minor symptoms from spreading the disease to others. In addition to its numerous loopholes, this suite also requires substantial amounts of labor and material resources. The combination of home quarantines and reducing contact in public places also has high social costs, yet the disease can still be transmitted if strict isolation is not observed for the time periods discussed in an earlier section. Numerous instances of intra-family infections were reported during the 2002-2003 SARS outbreaks—evidence that the combination of these prevention strategies is ineffective in controlling this kind of epidemic.

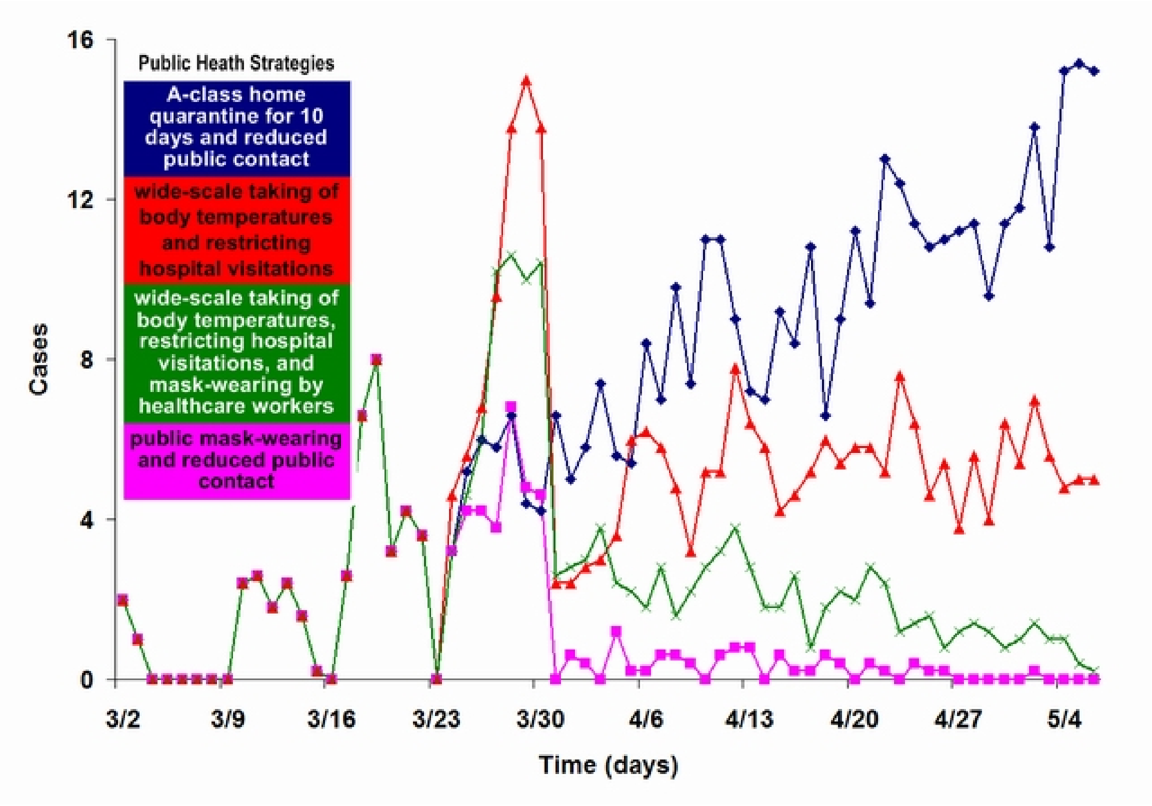


Figure 4.7: A Comparison of various public health policy suites. We used the 8 imported cases reported in Singapore to trigger the simulation. Policy suites went into effect on day 24 of the 66-day simulation. Suite 1 (cyan): A-class home quarantine for 10 days and reduced public contact; suite 2 (red): wide-scale taking of body temperatures and restricting hospital visitations; suite 3 (green): wide-scale taking of body temperatures, restricting hospital visitations, and mask-wearing by healthcare workers; suite 4 (pink): public mask-wearing and reduced public contact.

Chapter 5. Influence of Local Information on Social Simulations in Small-World Network Models

As part of Watts and Strogatz's "small world model" of disordered networks, local information mechanisms such as landscape properties are used to approximate real-world conditions in social simulations. The author investigated the influence of local information on social simulations based on the small world model, using a cellular automata variation with added shortcuts as a test platform for simulating the spread of an epidemic disease or cultural values/ideas. Our results will help researchers determine appropriate parameters for future simulations.

5.1. Motivation

“Small world” models are commonly used to study the structures of social networks—groups of individuals who exhibit interaction or relationship patterns (Albert and Barabási 2002; Milgram 1967; Barabási et al. 1999; Barabási and Albert 1999; Erdös and Renyi 1959; Watts and Strogatz 1998). Newman (2000) notes that small world models have special topological properties found in real-world human societies, including strong local clusters and small average distances between node pairs. They are therefore popular among researchers who construct social simulations of virtual societies, communication problems (especially epidemics), and the spread of cultural beliefs and influences—all of which are affected by transmission routes (Comellas et al. 2000; Keeling 1999; Moore and Newman 2000; Newman 2000; Newman 2002; Tsimring and Huerta 2001; Watts 1999; Zanette 2003; Zekri and Clerc 2001).

Factors that affect communication problem simulation results include network structure, divergence between individuals, and information-transmitting medium (Comellas 2000; Watts 1999). The degree of mitigation is tied to the type and amount of local information found on nodes and edges. Communication network structure (vertex degree information) is determined by the number of individual friendships.

Divergence (attribute information) is expressed as individual resistance to certain diseases or cultural influences. The medium (weight information) expresses transmission effectiveness. By treating these factors as local information, mechanisms can be designed for choosing the most appropriate information for social simulations.

Here we will focus on the influences of node-related local information, vertex degree, and attributes on simulating communication problems using a small world model. Building on previous research efforts involving epidemics that emphasize the influence of social network structure and divergence between individuals, we will analyze the sensitivity of those factors as well as vertex degree and attribute information. Our primary goal is to determine which type of local information exerts the greatest influence, thus requiring greater care when establishing parameters.

5.2. Adjustable Small-World Network Model

In the present study we will use Newman and Watts' 1999 SWN model to construct what we believe is a more effective method for generating virtual social networks. Our proposed method uses vertex degree information for all nodes to control network connections in the form of shortcuts, in such a manner that produces virtual social networks that meet specific simulation or research requirements. We believe our model is more practical that others because it allows for adjustments in network structure to reflect more active or conservative populations in certain regions. In other words, it produces virtual social networks that are varied in terms of interaction targets.

The SWN model on which we built our modification (Newman and Watts 1999) is itself a variation of Watts and Strogatz's (1998) original SWN model. In the original, a SWN starts with an *n*-dimensional regular graph in which each node is connected to a *z* number of its nearest nodes (Fig. 5.2.1(a)). There exists a probability *p* that each edge in the graph will be rewired, with one edge end being randomly selected and randomly reconnected to a new node (Fig. 5.2.1(b)).

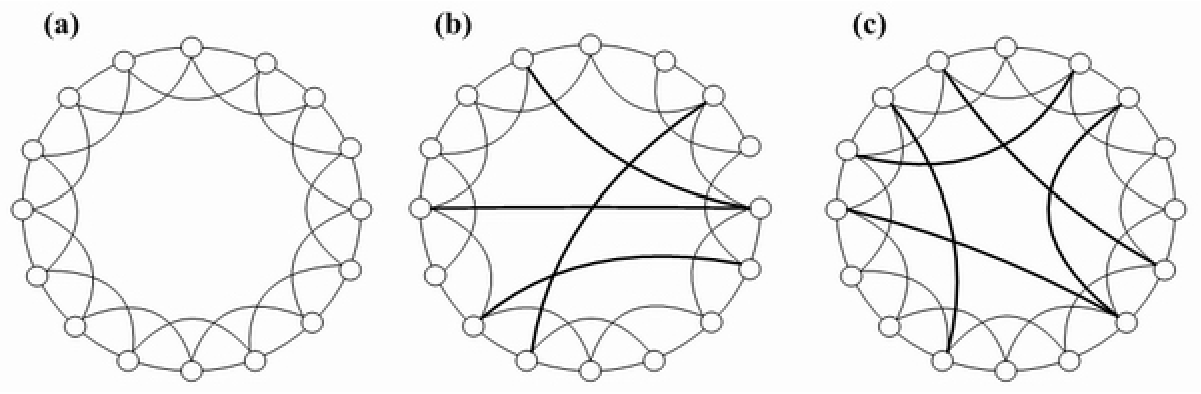


Figure 5.2.1: (a) One-dimensional regular graph with each node connected to its four adjacent nodes. (b) Watts and Strogatz's SWN model, with four rewired edges. (c) Newman and Watts' improved SWN model with five added shortcuts.

As shown in Figure 5.2.2, under certain adverse circumstances Watts and Strogatz's construction method can cause breaks in a graph (Newman 2000; Newman and Watts 1999; Wang and Chen 2003). Newman and Watts introduced a construction method that adds shortcuts instead of rewiring edges (Fig. 5.2.1(c)). According to the new method, two previously unconnected nodes are randomly selected and linked using a newly added edge, with users determining the number of new edges to be added. Newman and Watts' SWN model therefore avoids the problem of graph breakage while preserving the positive characteristic of connecting each node in the *n*-dimensional regular graph with *2n* neighboring nodes (Wang and Chen 2003). However, since there are equal probabilities of each node being chosen while shortcuts are being added, the vertex degree of each node will resemble a normal distribution, therefore failing to meet the needs of users wanting to construct SWN models that utilize network structures with different distribution methods.

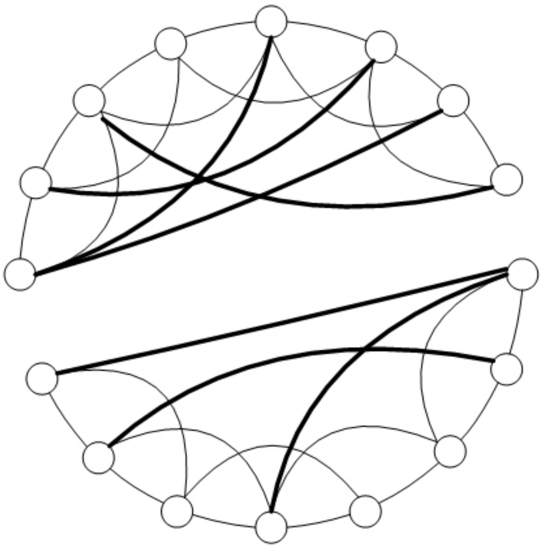


Figure 5.2.2: An example of a broken graph in Watts and Strogatz's SWN model.

In response to the restrictions of Newman and Watts' SWN model, we propose using a different construction method to change the selection process for each shortcut. As shown in Figure 5.2.3, a user must establish the weighted vertex degree $d(v_i)$ for each node v_i in the n-dimensional regular graph before constructing a SWN model. This $d(v_i)$ value must be a real number greater than 0. Our calculations for the probability $p(v_i)$ of node v_i being selected as one of the shortcut ends when a new shortcut is being added are shown in Equation 5.2.

$$p(v_i) = \frac{d(v_i)}{\sum_{v_j \in V(G)} d(v_j)} \tag{5.2}$$

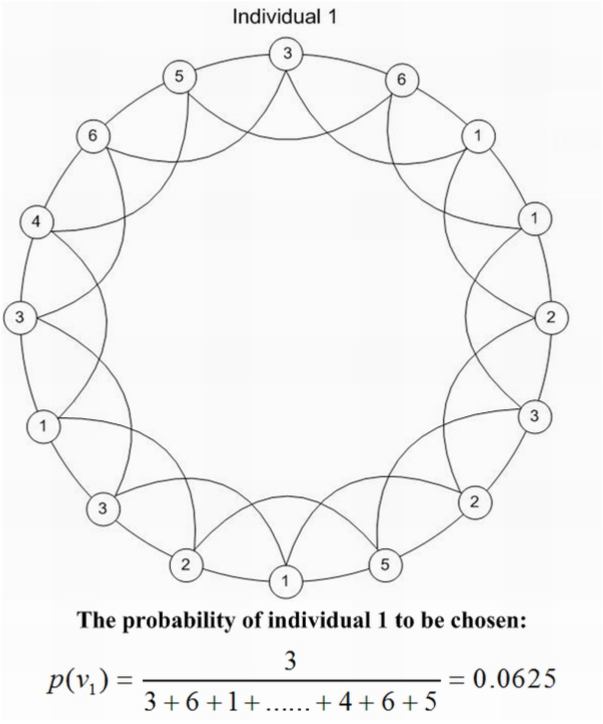


Figure 5.2.3: An example of calculating the probability of a node being selected.

According to our construction method, the higher a node's weighted vertex degree, the higher the probability it will be selected and vice versa. When all nodes have the same weighted vertex degree, the method is identical to the SWN model construction method originally proposed by Newman and Watts. The construction algorithm for our SWN model consists of

```
Step 1: for all Individuals $v_i$ in Population do
            Connect $v_i$ to $z$ nearest neighbor
            Assign vertex degree ratio information $d(v_i)$ to $v_i$

        next
Step 2: for loop 1 to shortcut number do
            label Generate shortcut:
            Individual $v_a$ ← Choose Individual by Probability $p(v_a)$
            Individual $v_b$ ← Choose Individual by Probability $p(v_b)$
            if (isLinked($v_a$, $v_b$)) then
                goto label Generate shortcut
            end if

        next
```

5.3. Communication Problem Modeling

The most basic and common communication problem model consists of a transmitter, receiver, and communication channel (Comellas 2000; Watts 1999) (Fig. 5.3.1). Using a disease epidemic as an example, diseased patients are the transmitters; newly infected individuals are the receivers; and skin contact, insects, air and water, etc. are the communication channels. A common example used by sociologists involves filmmaking, with directors as transmitters, actors and actresses the communication channels, and audiences the receivers. Directors disseminate their cultural beliefs and concepts to the public via repeated screenings.

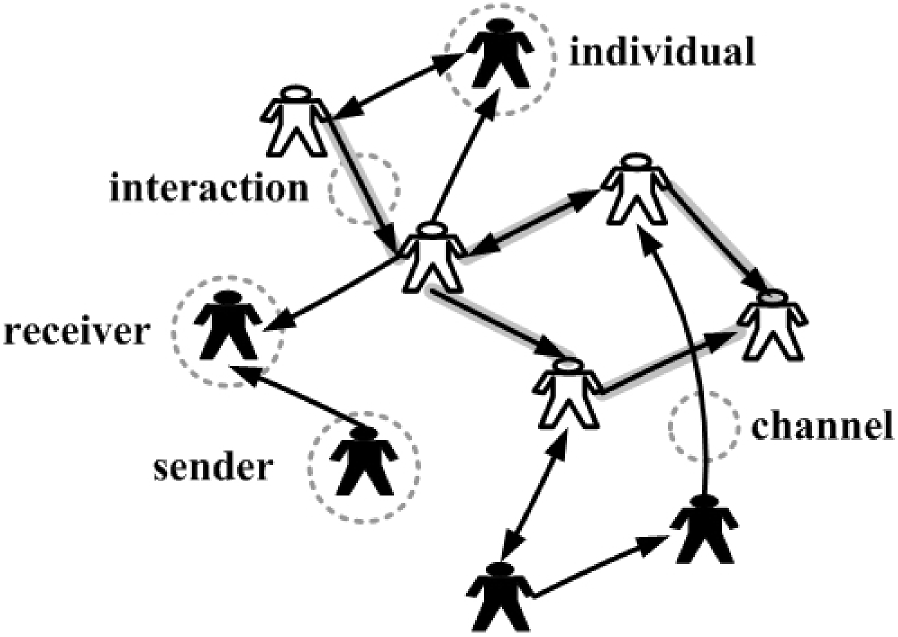


Figure 5.3.1: Communication problem model.

When designing our proposed model, we took into consideration the effects of such communication issues as network structure, differences among individuals, and media, and purposefully created a system in which individuals express a range of behavior patterns in response to various communication issues. For instance, we limited the distribution of rumors spread by word-of-mouth to small numbers of friends and neighbors within a specific time frame, but acknowledged the ability for a rumor to be spread to a wide number of recipients in a short period of time via email, television, radio, etc.

We used a simple SWN model as our basic social network framework because of its ability to reflect real-world interpersonal relationships. We also applied the state transfer concept of SIR models to simulate behavioral and transformative results from interactions among individuals—that is, when a susceptible individual interacts with an infectious individual, a certain probability exists that the status of the former will change from *S* to *I*—a probability that we refer to as $Rate_{Infect}$. Eventually the probability exists of all infectious individuals contracting the disease and being isolated from all other individuals in a society—that is, their status changes to *R* at a probability rate we refer to as $Rate_{Remove}$. In a typical SIR model, individuals in the *R* state are considered either dead or recovered; recovered individuals produce antibodies that prevent them from passing on the disease to *S* individuals.

SIR models take on new definitions in light of epidemic cultures, rumors disseminated by word-of-mouth, and other communication issues. As shown in Figure

5.3.2, the susceptible state represents an openness and willingness to accept new concepts; the infectious state represents the acceptance of a specific concept and a willingness to pass it on to other individuals; and the removed state represents a loss of interest in the originally accepted concept, meaning that R individuals will not be affected by the spreading behavior of other individuals nor actively spread the disease/concept in question to other individuals. However, SIR models entail a high possibility that an R individual will transform into an S individual. Using fashion as an example, someone with no particular liking for hip-hop pants will start wearing them due to peer pressure, then encourage others to wear them. After a certain period of time, the individual loses interest, stops wearing hip-hop pants, and doesn't notice when others keep wearing them. As more time passes, the individual has neither a strong like nor dislike of the fashion. We labeled the probability of change from R to S as a reset rate $Rate_{Reset}$.

Furthermore, we address the fact that SIR models are generally incapable of considering influences resulting from “differences between individuals.” In our proposed model, whenever susceptible individuals interact with infectious individuals, the infection rate is multiplied by attribute information (i.e., the individual’s resistance) and weight information for the communication media to determine the probability of a change in status from S to I.

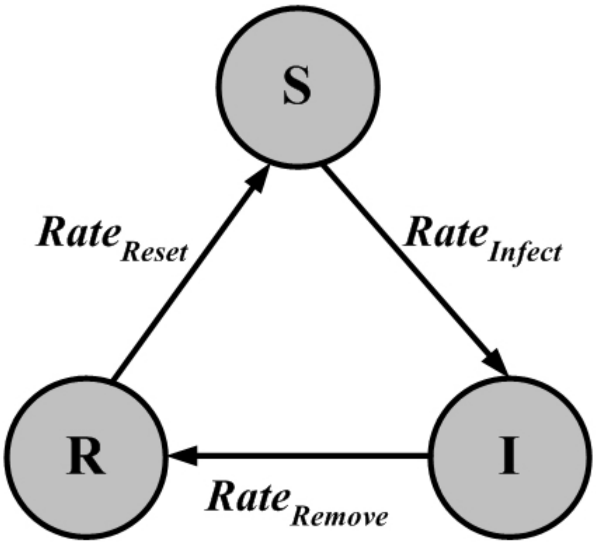


Figure 5.3.2: SIR state transfer diagram showing a repetitive cycle.

A computational simulation flowchart for our proposed model is shown in Figure 5.3.3. First, a SWN model is built using the construction method described in section 3 prior to setting the relevant parameters and attributes of the individuals involved in a problem. During simulation, SWN individuals/nodes take turns interacting with neighbors for specified time intervals. The number of interactions between any node and its neighbors can be fixed or variable, with the number of interactions randomly determined. Individual interactions do not result in immediate influences; simultaneous state changes only occur when all individuals in a SWN complete their interactions. Accordingly, interaction sequences will not influence interaction processes or results.

The following pseudo-codes were used in our communication problem simulation model.

```
for loop 1 to Time Step Limit do
    for all Individual $I_i$ in Population do
        for loop 1 to Interaction Limit do
            Individual $I_{target}$ ← Choose a Neighbor of $I_i$ by Random selection
            $SIR(I_i, I_{target})$

        next
    next
    for all Individual $I_i$ in Population do
        if $I_i.NowState = I$ then
            if random value $r < Rate_{Remove}$ then
                $I_i.NextState \leftarrow R$
            end if
        end if
        if $I_i.NowState = R$ then
            if random value $r < Rate_{Reset}$ then
                $I_i.NextState \leftarrow S$
            end if
        end if
        $I_i.NowState \leftarrow I_i.NextState$

    next
next

procedure SIR (Individual $I_a$, Individual $I_b$) is
    if $I_a.NowState = I$ and $I_b.NowState = S$ then
        if random value $r < Rate_{Infect} \times I_b.Resist$ then
            $I_b.NextState \leftarrow I$
        end if
    end if
    if $I_a.NowState = S$ and $I_b.NowState = I$ then
        if random value $r < Rate_{Infect} \times I_a.Resist$ then
            $I_a.NextState \leftarrow I$
        end if
    end if

return
```

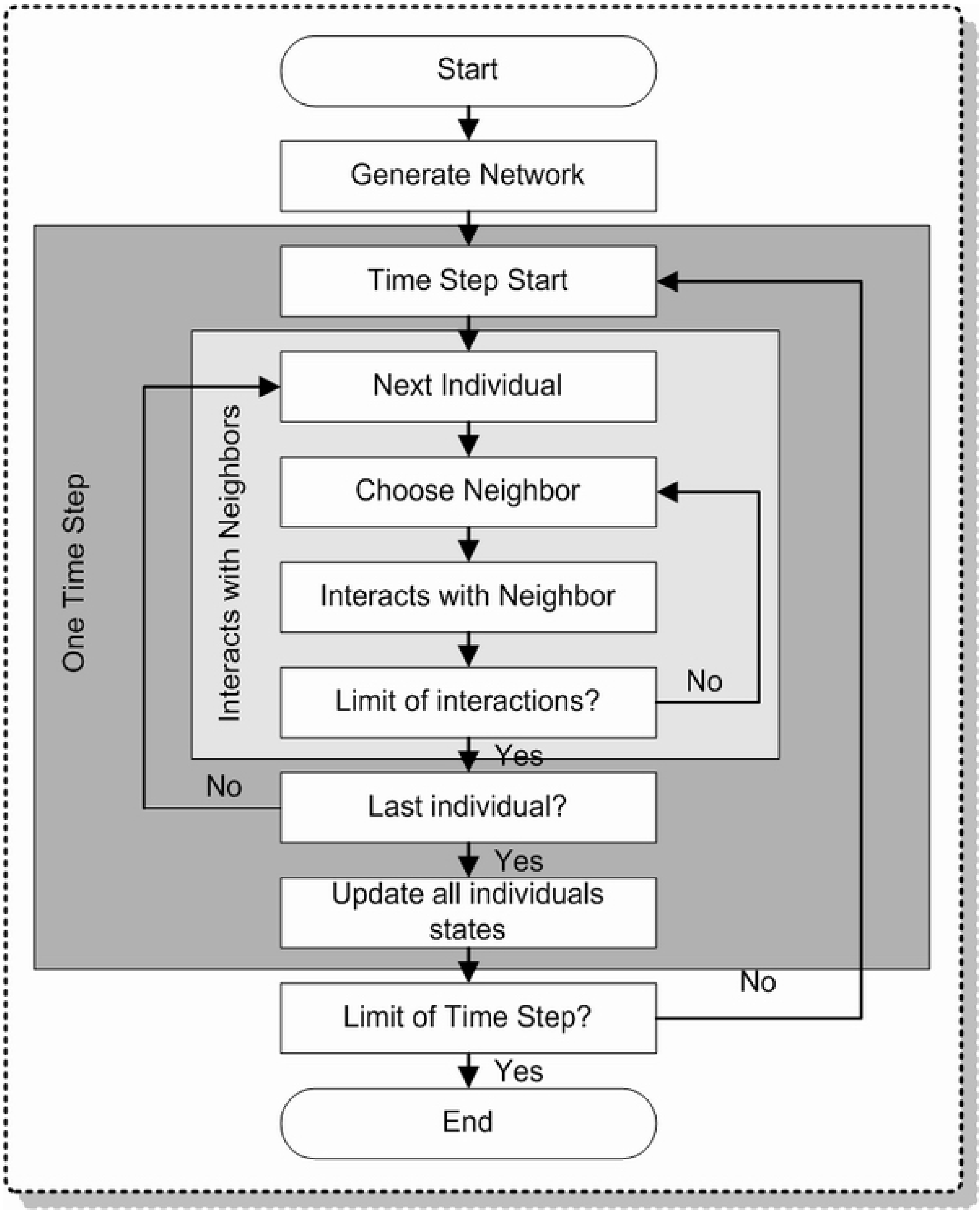


Figure 5.3.3: Simulation flowchart of the communication problem model.

5.4. Local Information Mechanisms

“Local information” refers to information that distinguishes certain individuals or channels from others. For example, everyone has two parents, therefore this number is considered global information. But there is considerable variation in numbers of brothers and sisters, making them a type of local information. Two primary local information categories are node-related (vertex degree and attribute) and edge-related (direction and weight). Vertex degree information represents the tendency of an individual to make friends; extroverts have higher vertex degrees than introverts. Attribute information, which describes individual resistance to disease, is used to express divergence. Direction refers to channel direction, categorized as either uni-directional (e.g., television, radio) or bi-directional (e.g., e-mail, telephones). Weight information represents channel effect—for instance, face-to-face communication is viewed as having greater weight than e-mail exchanges.

The most commonly used mechanism for setting local information entails random numbers that are either normally or uniformly distributed. A second mechanism entails organizing local information according to a pre-designed pattern—for example, putting all heterogeneous individuals in one location versus distributing them throughout an environment. These mechanisms are applied in a manner that allows a virtual society to approximate the real world. Our goal was to observe and identify the effects of applying local information mechanisms to node-related information.

5.5. Experiments of Sensitivity Analysis

We used SWN models (cellular automata with shortcuts) as our simulation platform for our experiments. As shown in Figure 5.5.1(a), all individuals in a typical cellular automaton have the same fixed number of neighbors, and therefore the cellular automaton is viewed as a two-dimensional toric periodic regular graph. As shown in Figure 5.5.1(b), if we use the method described in section 3 to select two nodes on a two-dimensional toric periodic regular graph and add one shortcut, the resulting automaton will meet SWN model requirements by having two small-world characteristics: a high degree of clustering and low degree of separation.

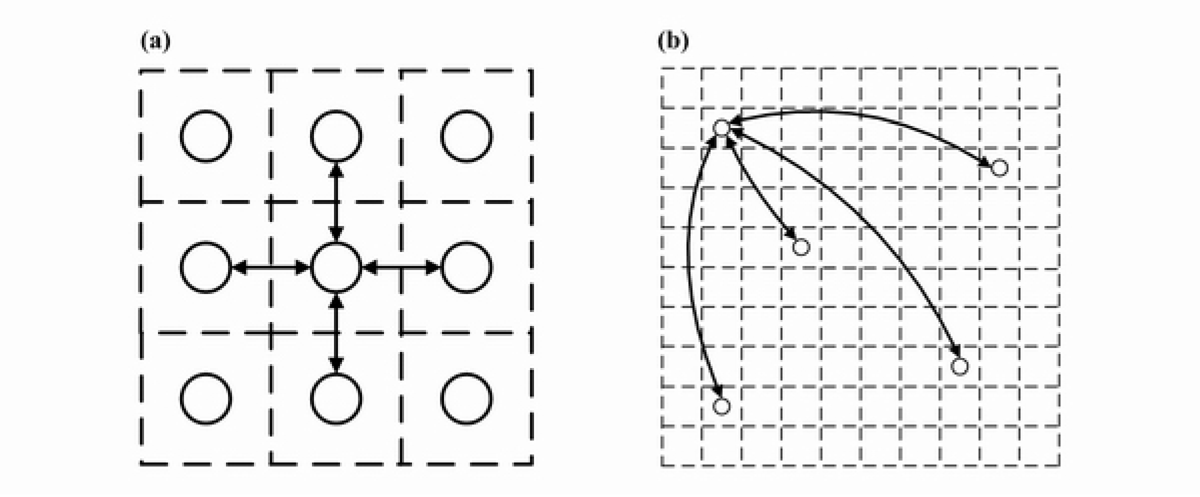


Figure 5.5.1: (a) Each individual in cellular automata has four neighbors. (b) Each individual can have 0 to n connecting shortcuts to other randomly selected individuals.

For our experiments, we used a 100 × 100 two-dimensional cellular automaton containing many shortcuts, resulting in a virtual social network consisting of 10,000 individuals (Fig. 5.5.2). This cellular automaton is capable of using a von Neumann or

Moore neighborhood; since von Neumann has been the neighborhood of choice for most researchers working with SWN models, we used it in our experiments. This means that all individuals are connected to and interact with only four surrounding neighbors plus long-distance friends connected by shortcuts.

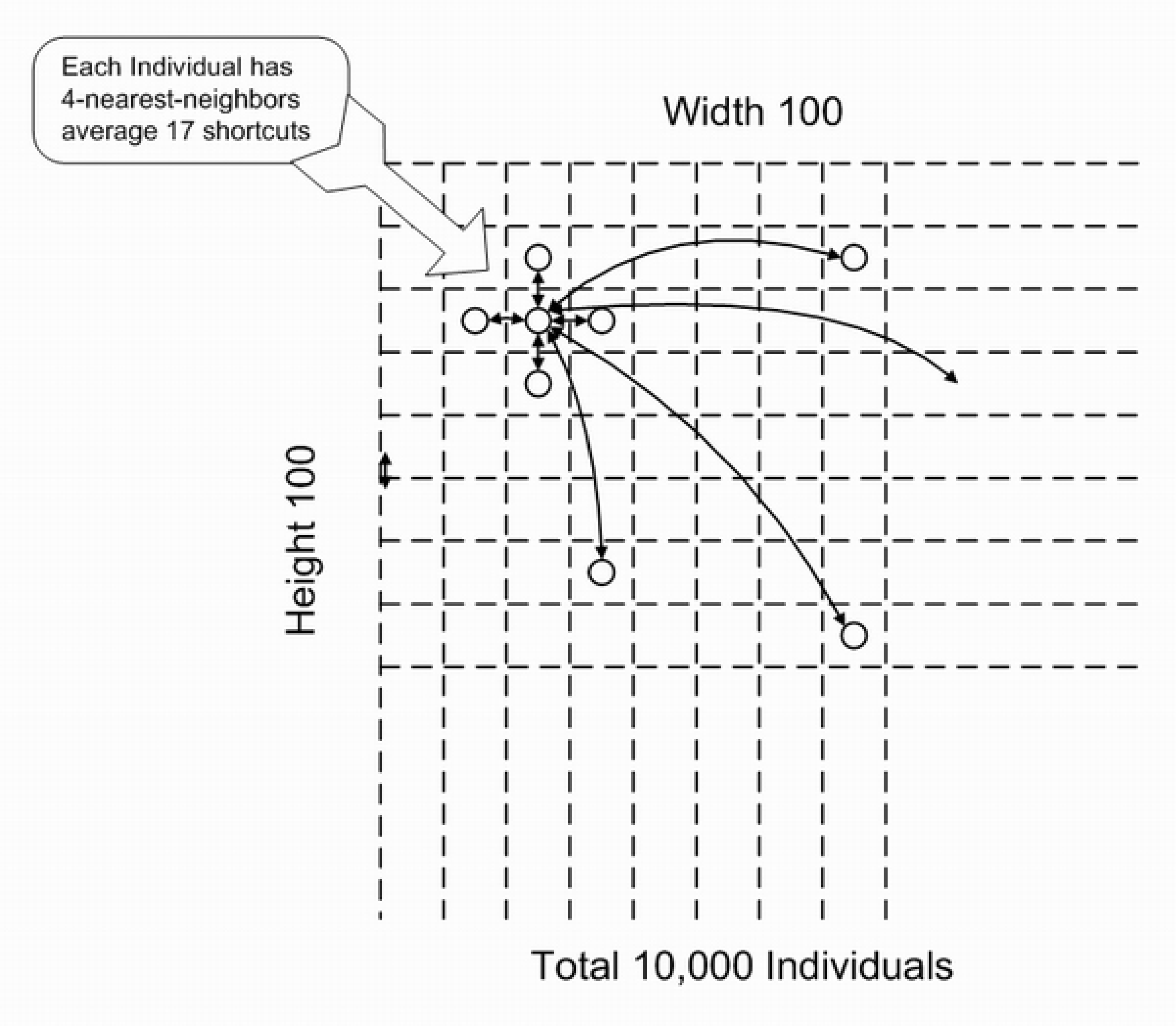


Figure 5.5.2: Test platform schematic diagram.

We were required to determine the appropriate number of shortcuts to add to our model. To maintain similarity to the real world, we used two arguments: a) a degree of separation of approximately 6 for a world population of six billion, and b) logarithmic growth between the number of SWN nodes and the average degree of separation; for a virtual social network of 10,000 individuals the average degree of separation is approximately 2.45. On average, one individual in a network of that size has 17 shortcut connections to other individuals, plus four connections with adjacent neighbors—21 in all.

Experiment 1: Vertex Degrees

Our first experiment was aimed at identifying the influences of network structure on a simulation. Differences in social networks depend on the social atmospheres in which they are formed. For instance, individuals in open societies will likely have more friends than individuals in more conservative traditional societies. Another important factor is the mix of extroverted, introverted, and in-between individuals. Sociologists are particularly interested in studying the effects of these and other factors.

We set the weighted vertex degree of each individual to meet the experimental requirement of developing different social network structures, and conducted our experiments with one of three social network distribution assumptions common to simulations. First, every individual's weighted vertex degree is identical, reflecting an average number of friends. Second, there is a uniform distribution of individual weighted vertex degrees, reflecting a society of 1/3 extroverts, 1/3 introverts, and 1/3 neither. Third, there is a normal distribution of the individual weighted vertex degrees, reflecting a population in which the majority of individuals are neither extroverted not introverted, and where extreme extroverts and introverts represent small minorities.

For the first experiment, we set the weighted vertex degree for all individuals in the first assumption group at 1. For the second group, we used a randomly selected integer between 3 and 6 from the random sequence of a uniform distribution to be used as an individual's weighted vertex degree. Each value had a 25% probability of being selected, with extroverts having twice the opportunities of introverts for making new friends. For the third group, an integer between 1 and 9 was randomly selected from a normal distribution (M = 5, SD = 1) and used as an individual's weighted vertex degree. According to the principle of normal distribution, for most individuals the weighted vertex degree was 4, 5 or 6, with very few weighted vertex degrees of 1 or 9.

Experimental results are shown in Figures 5.5.3 and 5.5.5; their respective cumulative sums based on time sequences are shown in Figures 5.5.4 and 5.5.6. With the exception of the maximum point for the first peak in the third epidemic curve in Figures 5.5.3 and 5.5.5, the rise and fall time points are very consistent. The cumulative sum curves for the three social network structures almost overlap in

Figures 5.5.4 and 5.5.6, reflecting the consistency of movement and fluctuation in the three epidemic curves shown in Figures 5.5.3 and 5.5.5.

In brief, the seemingly random dynamic process did not affect the overall development trend. We therefore conclude that adding the same number of shortcuts to the three social networks used in this experiment did not exert any influence on the movement and fluctuation of the entire epidemic curve, despite differences in the regional social network structure and a slight change in the total population at the peak of the epidemic breakout. Accordingly, it appears to be unnecessary to exert too much effort setting and adjusting fine network structures when putting such cases through additional computational simulations. Instead, one can focus on finding appropriate global information—for instance, the number of adjacent individuals to establish connections with or the number of shortcuts to add.

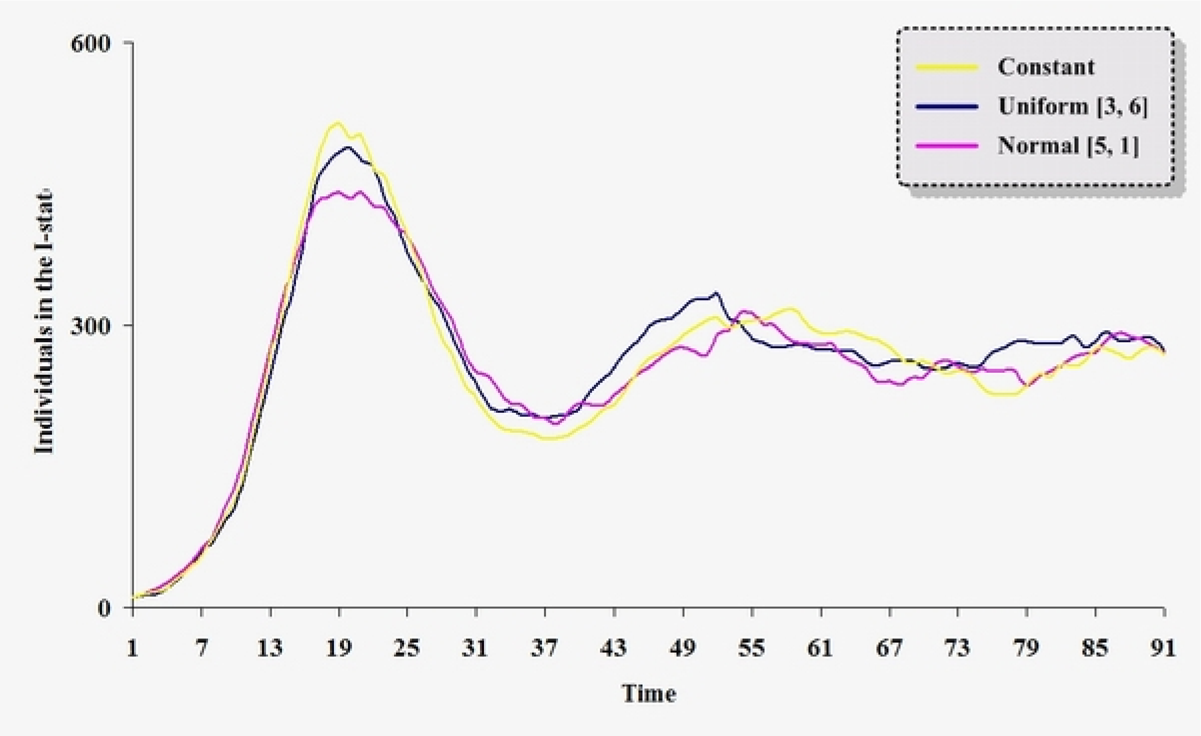


Figure 5.5.3: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

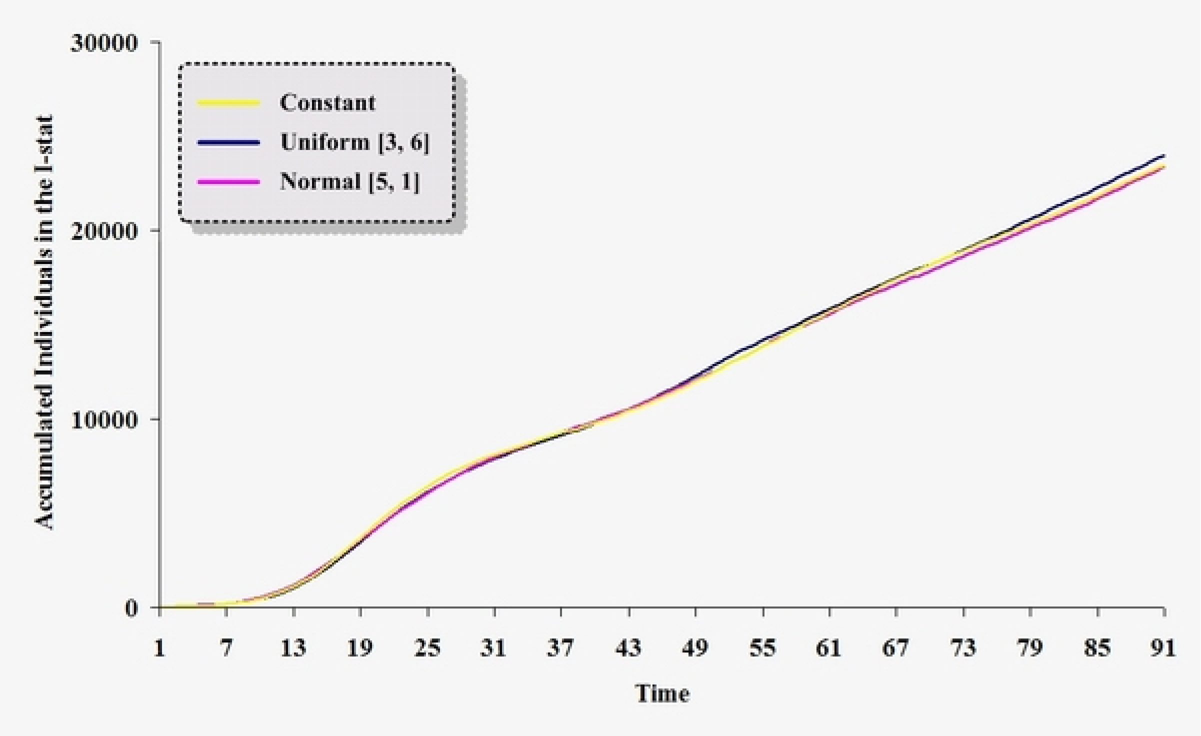


Figure 5.5.4: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

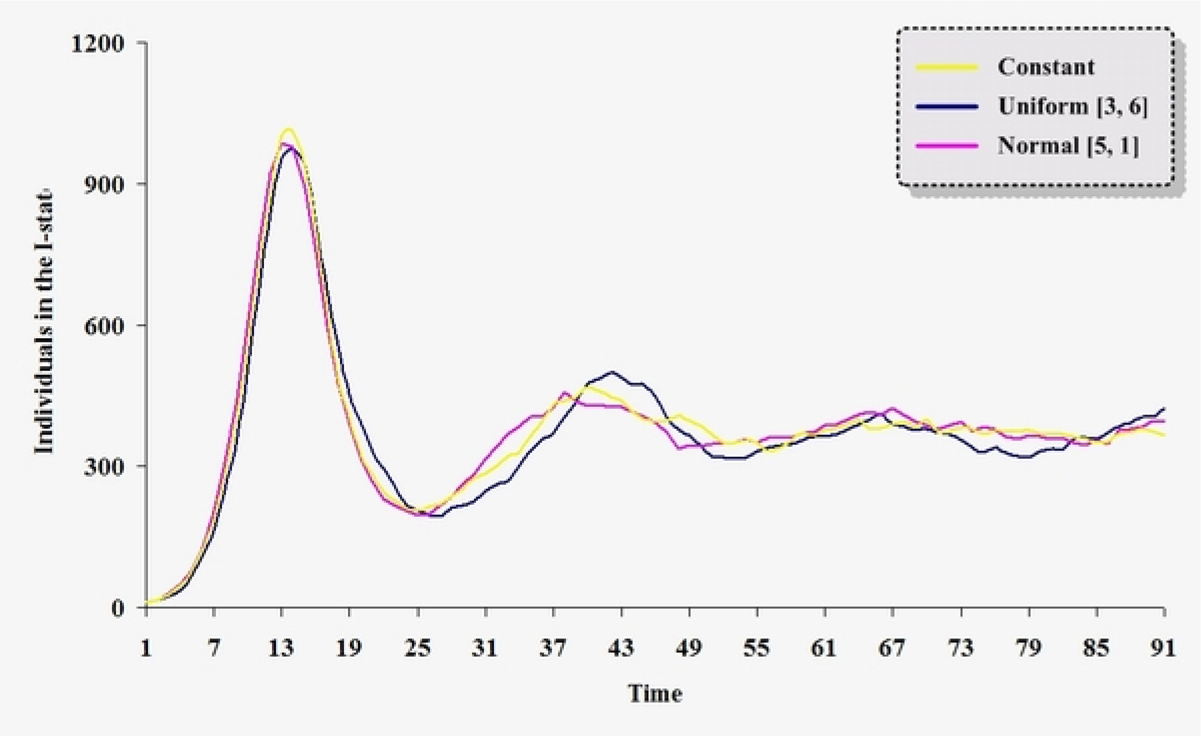


Figure 5.5.5: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

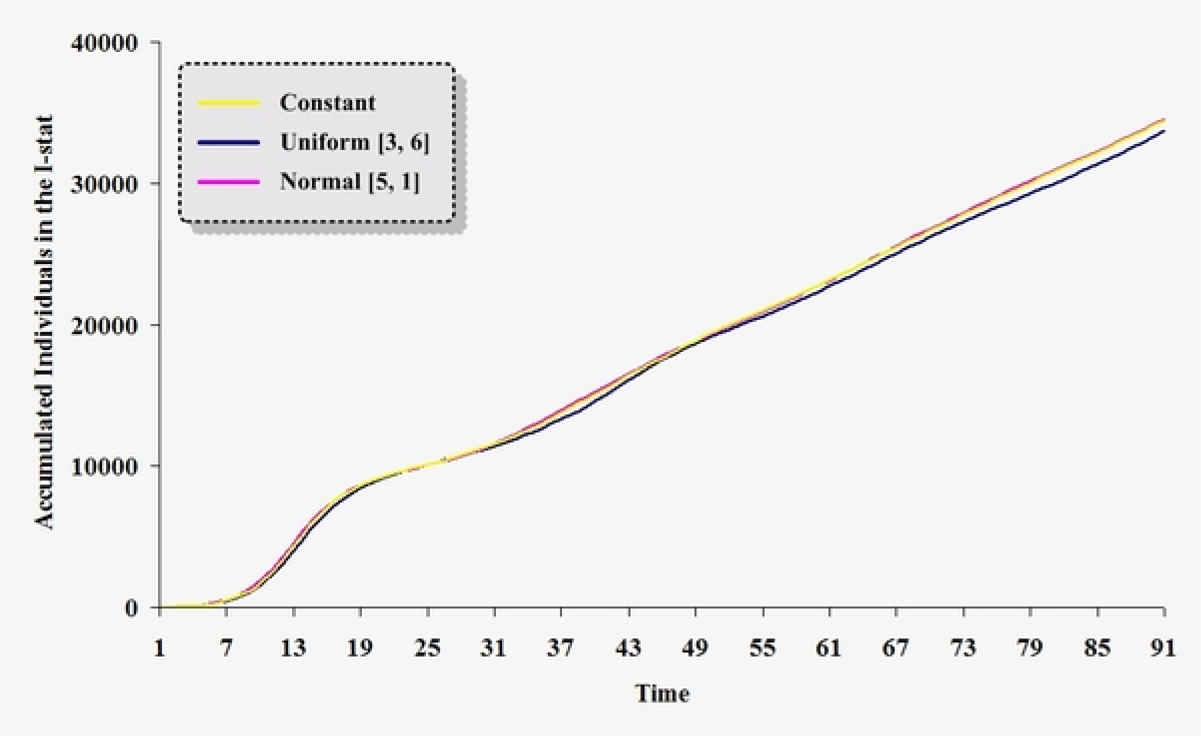


Figure 5.5.6: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

Experiment 2: Heterogeneous Proportion

The second experiment looked at the influence of heterogeneous individuals, which constitute a certain percentage of the total population. Results were expected to differ, since individuals have different levels of resistance to epidemic diseases (e.g., influenza) or cultural influences. Since the average degree of separation between any two individuals is very low in SWN models, the question of whether a large-scale epidemic will occur due to a heterogeneous minority of individuals who are especially prone to the disease was a specific focus in this experiment.

When establishing parameters, we set the resistance attribute information to represent individual differences—for instance, the resistance attribute values for heterogeneous individuals were one-half those of ordinary individuals, meaning that heterogeneous individuals had double the probability of becoming infected. Furthermore, heterogeneous and ordinary individuals were randomly distributed throughout our virtual environment. We ran six simulations with different percentages of heterogeneous individuals: 0, 1, 5, 10, 30 and 50 percent of the entire population. When the percentage of heterogeneous individuals exceeded 50 percent, they switched their status with ordinary individuals. We therefore did not have to run simulations with a percentage above 50 percent.

Experimental results are shown Figures 5.5.7 and 5.5.9; their respective cumulative sums based on time sequences are shown in Figures 5.5.8 and 5.5.10. In Figures 5.5.7 and 5.5.9, the higher the ratio of heterogeneous individuals in a population, the earlier the occurrence of the first peak and the higher its maximum point. According to the other two figures, the six cumulative sum curves diverged from the very beginning, with statistically significant differences among all six. The six epidemic curves had completely different movements and fluctuations, showing that the percentage of heterogeneous individuals exerted a significant influence on simulation results and processes. A comparison of the 0 and 1 percent curves shows that even a tiny percentage of heterogeneous individuals are enough to make a difference. These results underscore the importance of being precise when setting individual attributes for a simulation model.

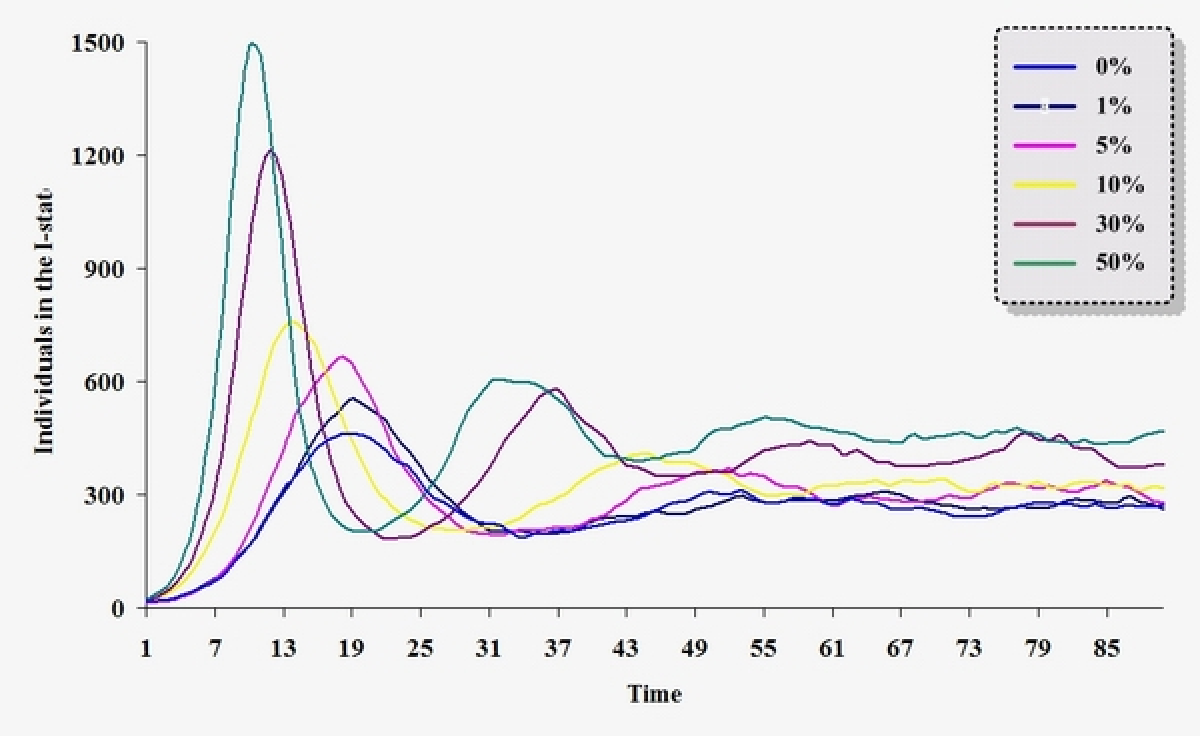


Figure 5.5.7: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

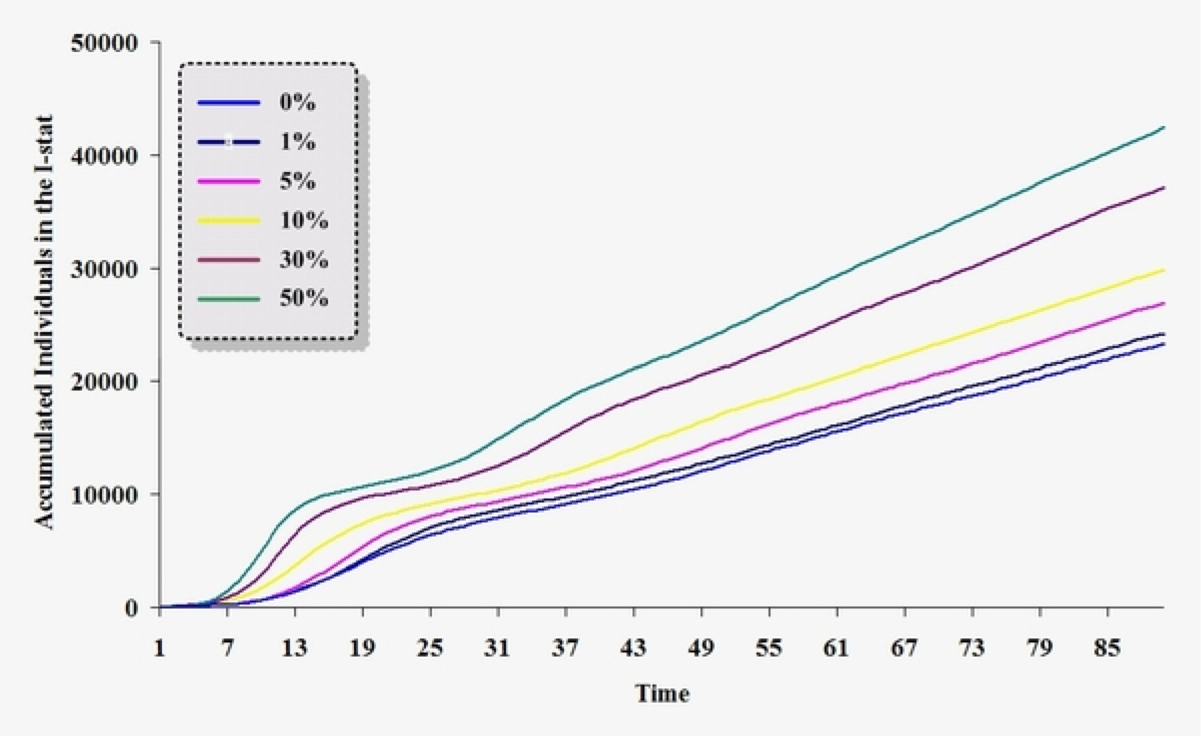


Figure 5.58: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

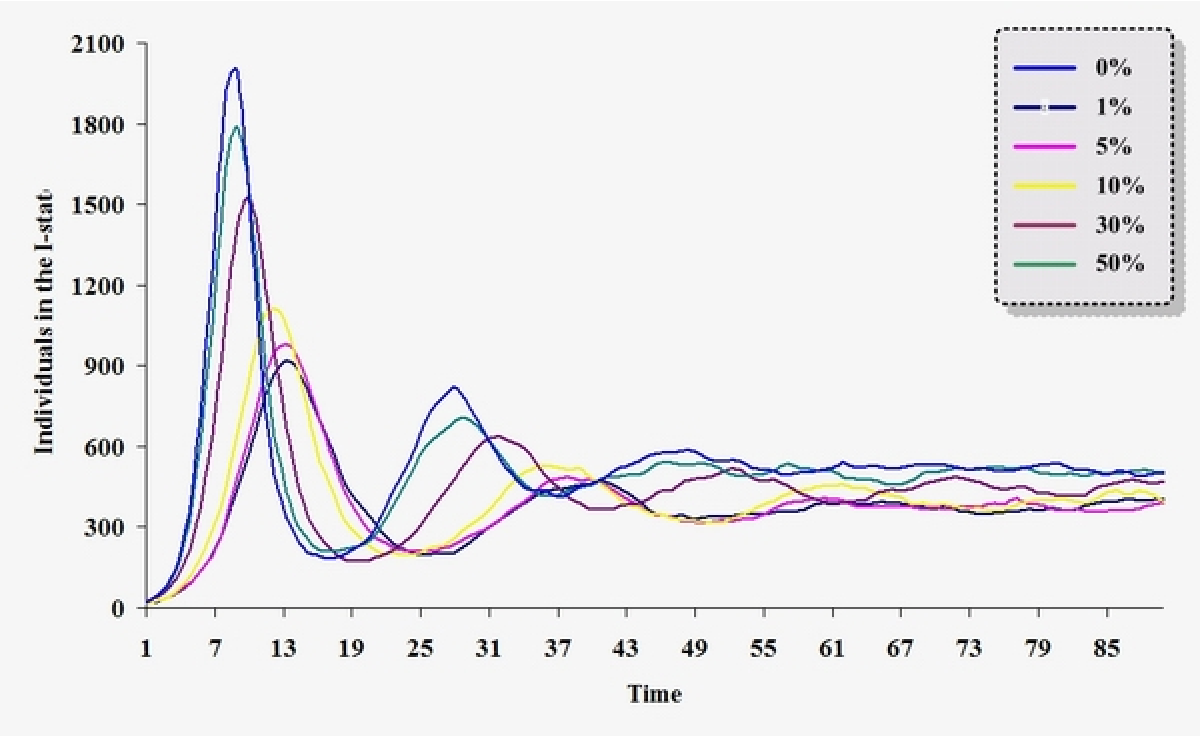


Figure 5.5.9: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

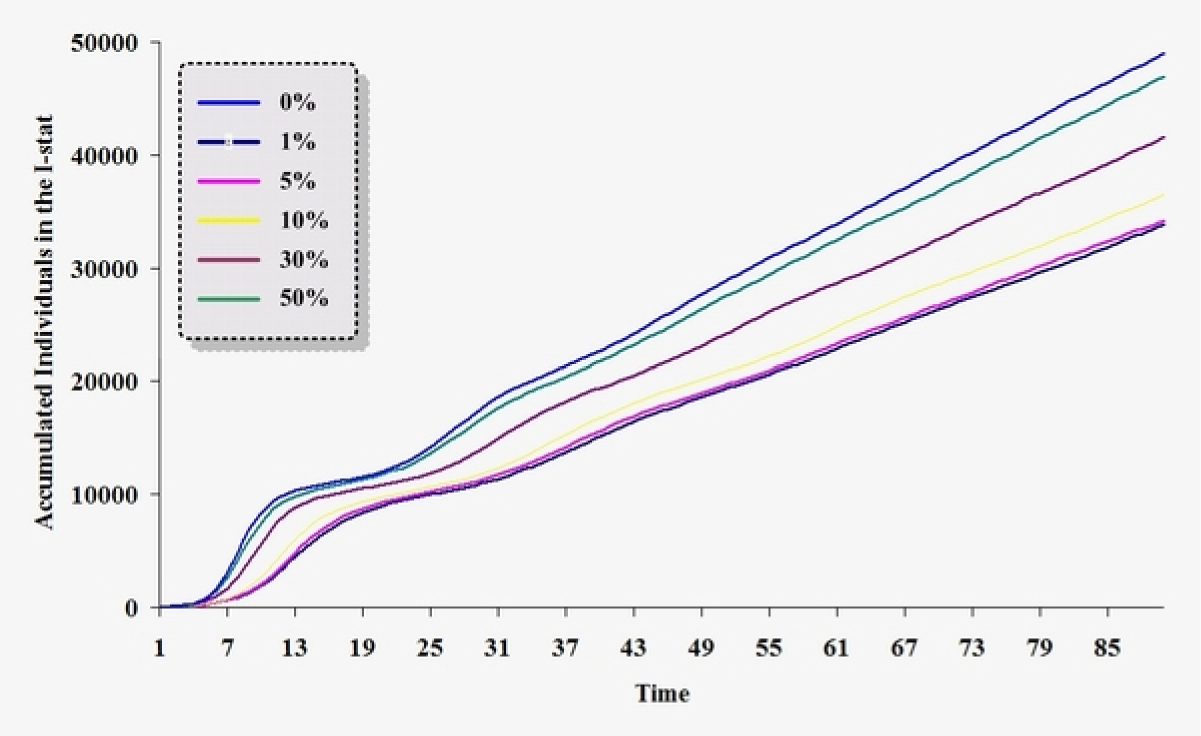


Figure 5.5.10: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

Experiment 3: Scattered Pattern of Heterogeneous Individuals

Our goal for the third experiment was to identify the effect on simulations of the same number of heterogeneous individuals under different settings—for example, when they are centrally distributed in a specific region or evenly distributed throughout a virtual environment.

We used a radius parameter *r* to represent the ratio of heterogeneous individuals' distribution scope in an environment relative to the entire simulation environment (Fig. 5.5.11). An *r* of 0 meant that all heterogeneous individuals were in close proximity to each other, an *r* of 1 meant that they were evenly distributed throughout a simulation environment, and an *r* between 0 and 1 meant that they were evenly distributed throughout a specific region within the radius parameter. When the number of heterogeneous individuals exceeded the size of the *r* region, the radius parameter automatically adjusted itself to the smallest possible ratio to contain all of them. We maintained a 1% level of heterogeneous individuals in the total population and ran six simulations with radius parameter *r* values of 0, 0.2, 0.4, 0.6, 0.8 and 1.

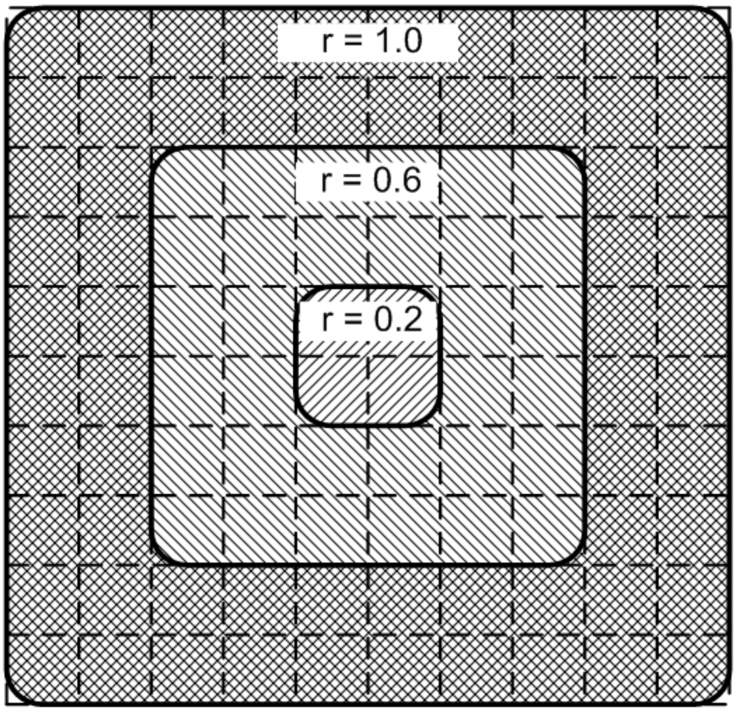


Figure 5.5.11: Schematic diagram of regions covered by different radius parameters (*r*).

Experimental results are shown in Figures 5.5.12 and 5.5.14; their respective cumulative sums based on time sequences are shown in Figures 5.5.13 and 5.5.15. Figures 5.5.12 and 5.5.14 show that with the exception of the maximum point of the first peak in the sixth epidemic curve, the time points were very consistent (no statistically significant differences). Figures 5.5.13 and 5.5.15 indicate the near-overlapping of the six cumulative sum curves, indicating that the six epidemic curves in Figures 5.5.12 and 5.5.14 were very consistent in terms of movement and fluctuation. In other words, the random dynamic process did not affect development. According to these results, it is not particularly important to determine the pattern of scattered (concentrated or distributed) heterogeneous individuals in simulations that match or come close to matching real-world situations, since satisfactory results can be produced from a random distribution approach.

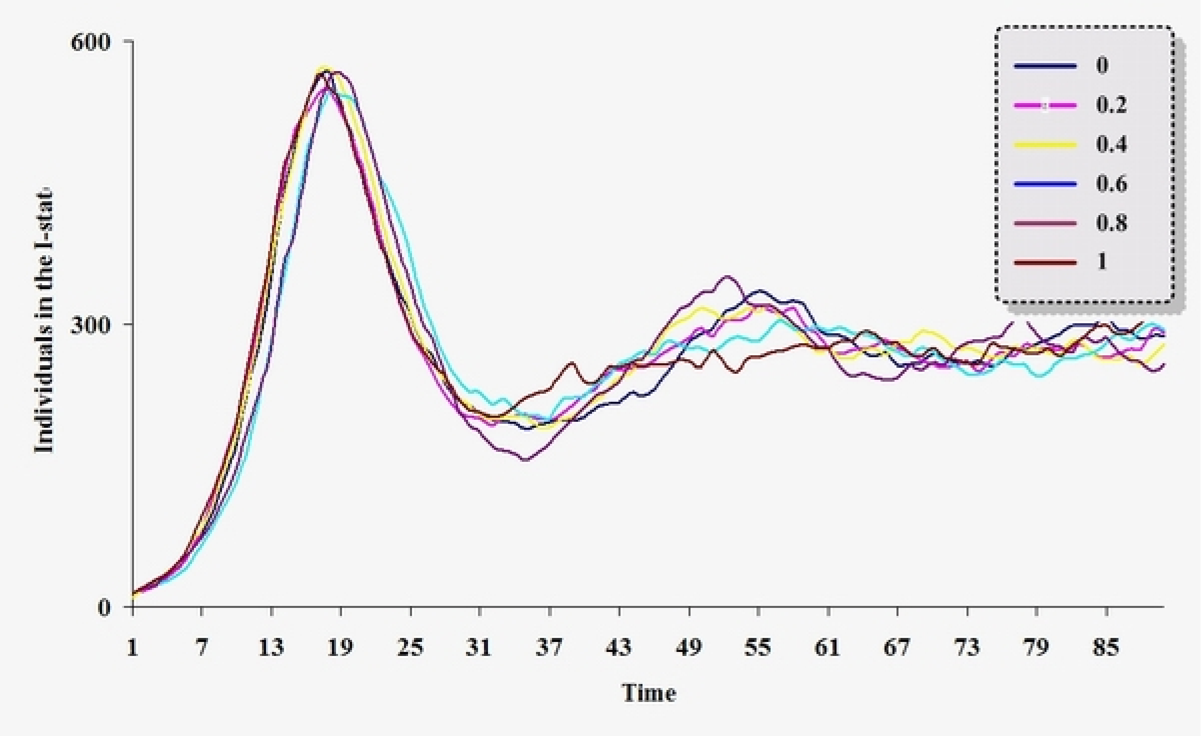


Figure 5.5.12: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

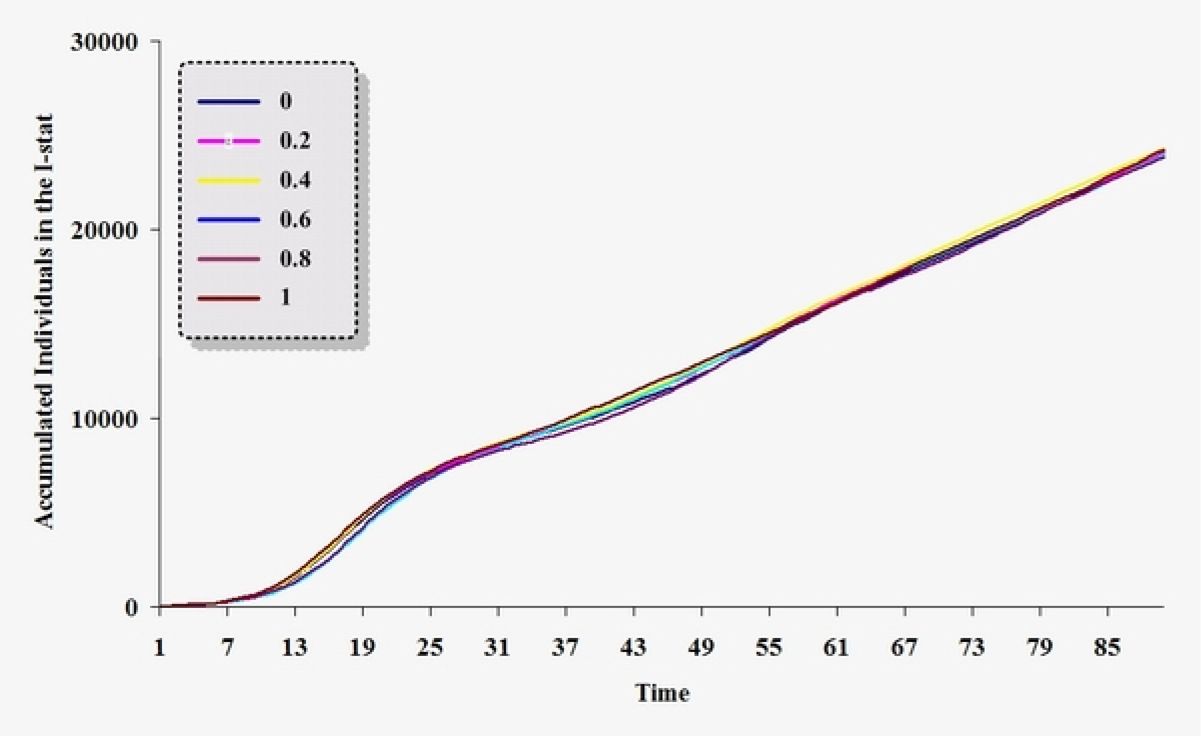


Figure 5.5.13: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

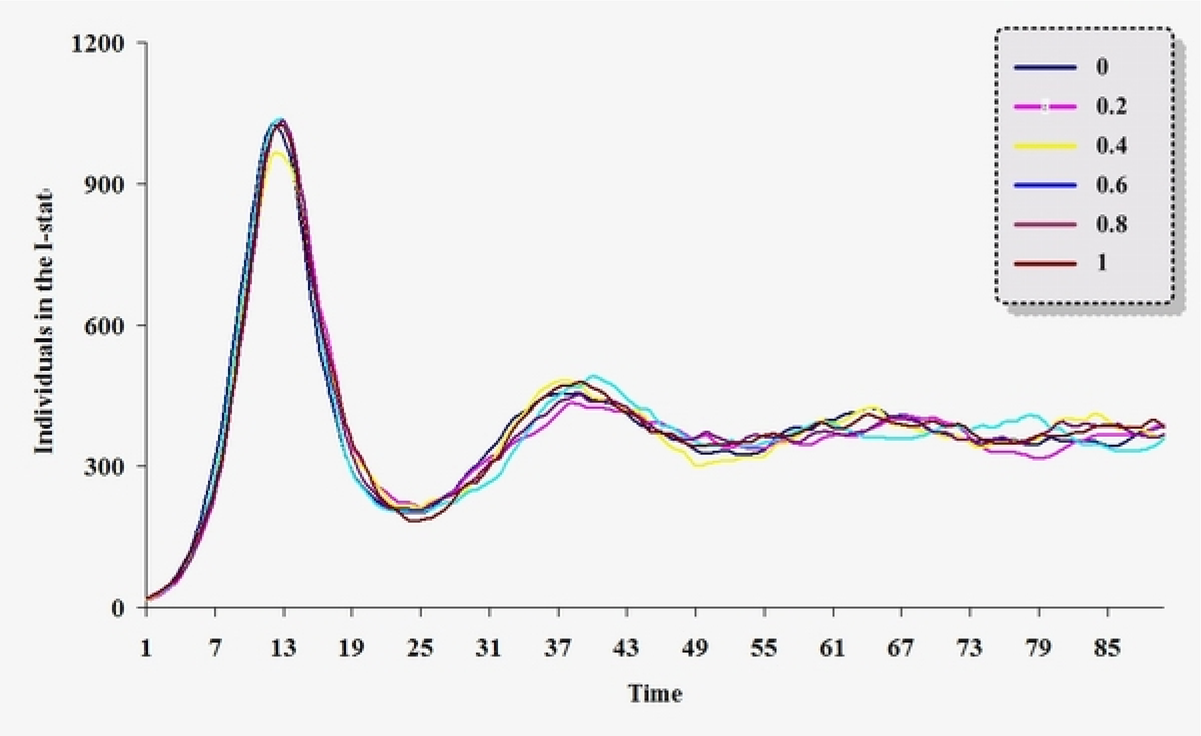


Figure 5.5.14: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the resulting number of I (infectious state) individuals produced by our proposed model.

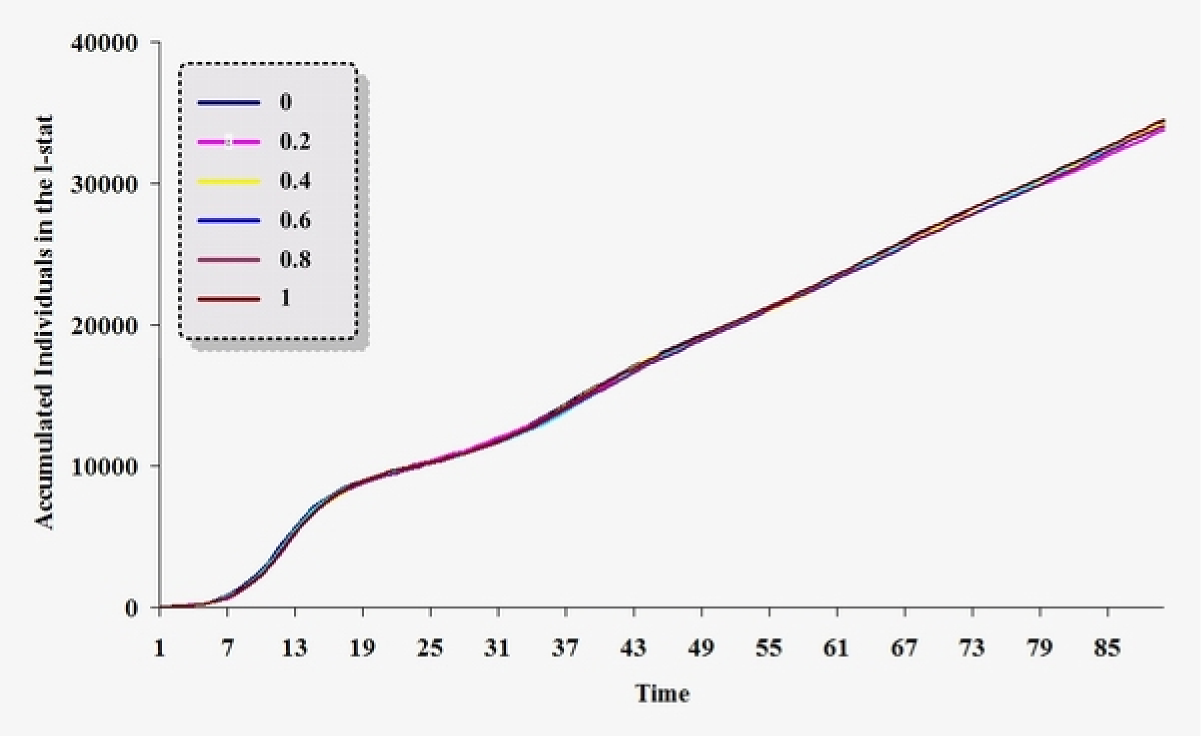


Figure 5.5.15: A comparison of three experimental results using different methods and vertex degree information in a simulation where one individual interacts with three neighboring individuals per time step. Each time grid shows the accumulated number of I (infectious state) individuals produced by our proposed model.

Chapter 6. Conclusions

In this chapter, we proposed a novel small-world model consisting of cellular automata with mirror identities representing daily-contact social networks for running epidemiological simulations. We established the mirror identity concept to integrate long-distance movement and geographic mobility into the model, which can be used to simulate the transmission dynamics of infectious diseases among social networks and to investigate the efficacies of various public health policies and epidemic prevention strategies—alone and in combination. The model successfully exhibits epidemiological behaviors in the form of daily interactions among heterogeneous individuals, and expresses such present-day small-world properties as high degrees of clustering, low degrees of separation, and long-distance movement.

According to the results of simulations that we ran based on data collected during the 2002-2003 SARS outbreaks in Singapore, Taipei, and Toronto, we suggest that this model can be applied to different infection scenarios and used to simulate the development of epidemics with considerable accuracy. A comparison of simulation and real-world data indicate that our model can be used to test epidemic report systems and to identify the best public health policy suites for specific scenarios. The simulation results also indicate considerable flexibility in the model—that is, we believe it can be applied to a wide range of contagious diseases (e.g., influenza, enteroviruses, and HIV/AIDS) that have well-defined epidemic parameters.

From this investigation of the influences of local information on communication problems involving small world networks, we found that the influence of vertex degree is not significant, that the influence of percentage of heterogeneous individuals is significant, and that the influence of the pattern of how heterogeneous individuals are scattered is not significant. We believe these results will be helpful for determining simulation parameters. Our immediate research plans are to study the influences of other types of local information (e.g., edge-related) and the influences of local information on SF, RN, and other small world models to determine if they are similar to those associated with the SWN model.

BIBLIOGRAPHY

ABELSON R P and Carroll J D (1965) Computer Simulation of Individual Belief Systems. *American Behavioral Scientist*, 8, pp. 24-30.

AHMED E and Agiza H N (1998) On Modeling Epidemics. Including Latency, Incubation and Variable Susceptibility. *Physica A*, 253, pp. 347-352.

AHMED E and Elgazzar A S (2001) On Some Applications of Cellular Automata. *Physica A*, 296, pp. 529-538.

ALBERT R and Barabási A-L (2002) Statistical Mechanics of Complex Networks. *Review of Modern Physics*, 74, pp. 47-91.

ALBERT R, Jeong H, and Barabási A-L (1999) Diameter of the World Wide Web. *Nature*, 401, pp. 130-131.

ANAND K, Ziebuhr J, Wadhwani P, Mesters J R, and Hilgenfeld R (2003) Coronavirus Main Proteinase (3CL^{pro}) Structure: Basis for Design of Anti-SARS Drugs. *Science*, 300(5626), pp. 1763-1767.

ANDERSON R M and May R M (1982) Directly Transmitted Infectious Diseases: Control by Vaccination. *Science* 215(4536), pp. 1053-1060.

AXELROD R (1997) Advancing the Art of Simulation in the Social Sciences. *Complexity*, 3(2), pp. 16-22.

BARABÁSI A-L and Albert R (1999) Emergence of Scaling in Random Networks. *Science*, 286(5439), pp. 509-512.

BARABÁSI A-L, Albert R, and Jeong H (1999) Mean-field Theory for Scale-free

Random Networks. *Physica A*, 272, pp. 173-187.

BENYOUSSEF A, Hafidallah N E, Elkenz A, Ez-Zahraouy H, and Loulidi M (2003) Dynamics of HIV Infection on 2D Cellular Automata. *Physica A*, 322, pp. 506-520.

BOCCARA N and Cheong K (1992) Automata Network SIR Models for the Spread of Infectious Disease in Populations of Moving Individuals. *Journal of Physics A*, 25, pp. 2447-2461.

BOCCARA N and Cheong K (1993) Critical-behavior of a Probabilistic-automata Network SIS Model for the Spread of an Infectious-disease in a Population of Moving Individuals. *Journal of Physics A*, 26, pp. 3707-3717.

BOCCARA N, Goles E, Martinez S, and Picco P (1993) *Cellular Automata and Cooperative Phenomena*. Kluwer Academic Publisher, Boston.

BOCCARA N, Cheong K, and Oram M (1994) A Probabilistic Automata Network Epidemic Model with Births and Deaths Exhibiting Cyclic Behavior. *Journal of Physics A*, 27, pp. 1585-1597.

BYRNE D (1997) Simulation – A Way Forward? *Sociological research Online*, 2(2). URL:http://www.socresonline.org.uk/socresonline/2/2/4.html.

CDC (CENTERS FOR DISEASE CONTROL) (2003a) *SARS: Frequently asked questions*, http://www.cdc.gov/ncidod/sars/faq.htm.

CDC (CENTERS FOR DISEASE CONTROL) (2003b) Update: SARS --- Singapore, *MMWR*, 52(18), pp. 405-411.

CDC (CENTERS FOR DISEASE CONTROL) (2003c) Cluster of SARS Cases Among Protected Health-Care Workers-Toronto, Canada, *MMWR*, 52(19), pp.

433-436.

CDC (CENTERS FOR DISEASE CONTROL) (2003d) SARS --- Taiwan, *MMWR*, 52(20), pp. 461-466.

CDC (CENTERS FOR DISEASE CONTROL) (2003e) Update: SARS --- Toronto, Canada, *MMWR*, 52(23), pp. 547-550.

CDC (CENTERS FOR DISEASE CONTROL) (2003f) Update: SARS --- Worldwide and U.S., *MMWR*, 52(28), pp. 664-665.

CDC (CENTERS FOR DISEASE CONTROL) (2003g) Use of Quarantine to Prevent Transmission of SARS---Taiwan, *MMWR*, 52(29), pp. 680-683.

CDC (CENTERS FOR DISEASE CONTROL) (2003h) Revised U.S. Surveillance Case Definition for Severe Acute Respiratory Syndrome (SARS) and Update on SARS Cases --- United States and Worldwide, *MMWR*, 52(49), pp. 1202-1206.

CHOWELL G, Fenimore P W, Castillo-Garsow M A, and Castillo-Chavez (2003) SARS Outbreaks in Ontario, Hong Kong and Singapore: the Role of Diagnosis and Isolation as a Control Mechanism. *Journal of Theoretical Biology*, 224(1), pp. 1-8.

COLEMAN J S (1962) Analysis of Social Structures and Simulation of Social Processes with Electronic Computers. In H S Guetzkow (Ed.), *Simulation in social science: Readings*, pp. 61-79. Englewood Cliffs, NJ: Prentice Hall.

COMELLAS F and Sampels M (2002) Deterministic Small-world Networks. *Physica A*, 309, pp. 231-235.

COMELLAS F, Ozon J, and Peters J G (2000) Deterministic Small-World Communication Networks. *Information Processing Letters*, 76, pp. 83-90.

CONTE R and Gilbert N (1995) Introduction: Computer Simulation for Social Theory. In N Gilbert and R Conte (Eds.) *Artificial Societies: the Computer Simulation of Social Life*, London: UCL Press.

COWEN R A and Miller J H (1990) *Economic Life on a Lattice* (Santa Fe Institute Working Paper 90-010). Santa Fe, NM: Santa Fe Institute.

CRICK F (1988) *What Mad Pursuit: A Personal View of Scientific Discovery*. New York: Basic Books.

CUSACK T R and Stoll R J (1990) *Exploring Realpolitik: Probing International Relations Theory With Computer Simulation*. Boulder, CO: Lynne Rienner.

DONNELLY C A, Ghani A C, Leung G M, Hedley A J, Fraser C, and Riley S (2003) Epidemiological Determinants of Spread of Causal Agent of Severe Acute Respiratory Syndrome in Hong Kong. *Lancet*, 361(9371), pp. 1761-1766.

DYE C and Gay N (2003) EPIDEMIOLOGY: Modeling the SARS Epidemic. *Science*, 300(5627), pp. 1884-1885.

EDELETEIN-KESHET L (1988) *Mathematical Models in Biology*: Random House, New York.

EGUILUZ V M and Klemm K (2002) Epidemic Threshold in Structured Scale-free Networks. *Physical Review Letters*, 89(10), 108701.

ENSERINK M (2003) SARS IN CHINA: The Big Question Now: Will It Be Back? *Science*, 301(5631), pp. 299.

EPSTEIN J M and Axtell R (1996). *Growing Artificial Societies*. Washington, DC: Brookings Institution.

ERDÖS P and Renyi A (1959) On the Evolution of Random Graphs. *Publication of the Mathematical Institute of the Hungarian Acadamy of Science*, 5, pp. 17-60.

FALOUTSOS M, Faloutsos P, and Faloutsos C (1999) On Power-law Relationships of the Internet Topology. *ACM Computer Communications Review*, 29, pp. 251-263.

FARMER D, Toffoli T, and Wolfram S (Eds.) (1984) *Cellular Automata*. New York: North-Holland.

FIORINA P (1975) Formal Models in Political Science. *American Journal of Political Science*, 19, pp. 133-159.

FORRESTER J W (1969) *Urban Dynamics*. Cambridge, MA: MIT Press.

FORRESTER J W (1980) *Principles of Systems*, 2nd preliminary edn. MIT Press, Cambridge, MA (1st edn 1968).

FUENTES M A and Kuperman M N (1999) Cellular Automata and Epidemiological Models with Spatial Dependence. *Physica A*, 267, pp. 471-486.

GARSON G D (1994) Social Science Computer Simulation: Its History, Design, and Future. *Social Science Computer Review*, 12, pp. 55-82.

GILBERT N (1999) Simulation: A New Way of Doing Social Science. *American Behavioral Scientist*, 42(10), pp. 1845-1847.

GILBERT N and Troitzsch K G (1999) *Simulation for the Social Scientist*. Buckingham: Open University Press.

GOLDBERG D E (1989) *Genetic Algorithms in Search, Optimization, and Machine learning*. Reading, MA: Addison-Wesley.

GUAN Y, Zheng B J, He Y Q, Liu X L, Zhuang Z X, Cheung C L, Luo S W, Li P H, Zhang L J, Guan Y J, Butt K M, Wong K L, Chan K W, Lim W, Shortridge K F, Yuen K Y, Peiris J S M and Poon L L M (2003) Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China. *Science*, 302(5643), pp. 276-278.

GUETZKOW H S and Bowes A (1957) The Development of Organizations in a Laboratory. *Management Science*, 3, pp. 380-402.

GUTOWITZ H (Ed.) (1991) *Cellular Automata: Theory and Experiment*. Cambridge, MA: MIT Press.

HAKEN H (1983) *Advanced Synergetics*. Berlin: Springer-Verlag.

HALFPENNY P (1997) Situating Simulation in Sociology. *Sociological Research Online*, 2(3). URL:http://www.socresonline.org.uk/socresonline/2/3/9.html.

HANNEMAN R and Patrick S (1997) On the Uses of Computer-assisted Simulation Modeling in the Social *Sciences*. *Sociological Research Online*, 2(2). URL:http://www.socresonline.org.uk/socresonline/2/2/5.html.

HANNON B and Ruth M (1994) *Dynamic Modeling*. New York: Springer-Verlag.

HASTIE R (1988) A Computer Simulation Model of Person Memory. *Journal of Experimental Social Psychology*, 24, pp. 423-447.

HERMANN C F and Hermann M G (1967) An Attempt to Simulate the Outbreak of World War I. *American Political Science Review*, 61, pp. 400-416.

HOLLAND J H (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

HOLLAND J H and Miller J H (1991) Artificial Adaptive Agents and Economic

Theory. *American Economic Review*, 81, pp. 365-370.

HOOVER S V and Perry R F (1990) *Simulation: A Problem-Solving Approach*. Reading, MA: Addison-Wesley.

HUANG C Y, Sun C T, Hsieh J L, and Lin H (2004a) Simulating SARS: Small-World Epidemiological Modeling and Public Health Policy Assessments. *Journal of Artificial Societies and Social Simulation*, ?.

HUANG C Y, Sun C T, and Lin H C (2004b) Influence of Local Information on Social Simulations under the Small-World Network Models. *Journal of Artificial Societies and Social Simulation*, ?.

HUCKFELDT R R, Kohfeld C W, and Likens T W (1982) *Dynamic Modeling: An Introduction*. Sage University paper series on Quantitative Applications in the Social Sciences, 07-027. Newbury Park, CA: Sage.

ILGEN D R and Hulin C L (Eds.) (2000) *Computational Modeling of Behavior Organizations: The Third Scientific Discipline*. Washington DC: American Psychological Association.

JOHNSON P E (1999) Simulation Modeling in Political Science. *American Behavioral Scientist*, 42(10), pp. 1509-1530.

JONES B, Radcliff B, Taber C and Timpone R (1995) Condorcet Winners and the Paradox of Voting: Probability Calculations for Weak Preference Orders. *American Political Science Review*, 89, pp. 137-144.

KEELING M J (1999) The Effects of Local Spatial Structure on Epidemiological Invasions. *Proceedings of the Royal Society of London*, 266, pp. 859-867.

KERMACK W O and McKendrick A G (1927) Contributions to the Mathematical Theory of Epidemics. In *Proceedings of the Royal Society of London*, pp. 700-721.

KING G, Keohane R O, and Verba S (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

KHEIR N A (Ed.). (1988) *Systems Modeling and Computer Simulation*. New York: Marcel Dekker.

KLECZKOWSKI A and Grenfell B T (1999) Mean-field-type Equations for Spread of Epidemics: the 'Small World' Model. *Physica A*, 274, pp. 355-360.

KOOPMAN J (2004) Modeling Infection Transmission. *Annual Review of Public Health*, 25, pp. 303-326.

KUPERMAN M and Abramson G (2001) Small World Effect in an Epidemiological Model. *Physical Review Letters*, 86(13), pp. 2909-2912.

LAVE C A and March J G (1993) *An Introduction to Models in the Social Sciences*. New York: University Press of America.

LEIK R K and Meeker B F (1995) Computer Simulation for Exploring Theories: Models of Interpersonal Cooperation and Competition, *Sociological Perspectives*, 38(4), pp. 463-482.

LIPSITCH M, Cohen T, Cooper B, Robins J M, Ma S, James L, Gopalakrishna G, Chew S K, Tan C C, Samore M H, Fisman D, and Murray M (2003) Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. *Science*, 300(5627), pp. 1966-1970.

MARRA M A, Jones S J M, Astell C R, Holt R A, Brooks-Wilson A, Butterfield Y S

N, Khattra J, Asano J K, Barber S A, Chan S Y, Cloutier A, Coughlin S M, Freeman D, Girn N, Griffith O L, Leach S R, Mayo M, Mcdonald H, Montgomery S B, Pandoh P K, Petrescu A S, Robertson A G, Schein J E, Siddiqui A, Smailus D E, Stott J M, Yang G S, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth T F, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples G A, Tyler S, Vogrig R, Ward D, Watson B, Brunham R C, Krajden M, Petric M, Skowronski D M, Upton C, and Roper R L (2003) The Genome Sequence of the SARS-Associated Coronavirus. *Science*, 300(5624), pp. 1399-1404.

MARTINS M L, Ceotto G, Alves S G, Bufon C C B, Silva J M, and Laranjeira F F (2001) A Cellular Automata Model for Citrus Variegated Chlorosis. *Physica A*, 295, pp. 42-48.

MCPHEE W N (1963) *Formal Theories of Mass Behavior*. New York: Free Press.

MEADOWS D H, Meadows D L, Randers J, and Behrens W W (1972) *The Limits to Growth*. New York: Universe Books.

MEDINA A, Matta I, and Byers J (2000) On the Origin of Power-laws in Internet Topologies. *ACM Computer Communications Review*, 30(2), pp. 18-28.

MILGRAM S (1967) The Small World Problem. *Psychology Today*, 2, pp. 60-67.

MIRAMONTES O and Luque B (2002) Dynamical Small-world Behavior in an Epidemical Model of Mobile Individuals. *Physica D*, 168-169, pp. 379-385.

MOORE C and Newman M E J (2000) Epidemics and Percolation in Small-world Networks. *Physical Review E*, 61(5 Pt B), pp. 5678-5682.

MORENO Y, Pastor-Satorras R, and Vespignani A (2002) Epidemic Outbreaks in

Complex Heterogeneous Networks. *European Physical Journal B*, 26(4), pp. 521-529.

NEWMAN M E J (2000) Models of the Small World: A Review. *Journal of Statistical Physics*, 101, pp. 819-841.

NEWMAN M E J (2002) Spread of Epidemic Disease on Networks. *Physical Review E*, 66 (1 Pt 2). 016128.

NEWMAN M E J, Jensen I, and Ziff R M (2002) Percolation and Epidemics in a Two-dimensional Small World. *Physical Review E*, 65 (2 Pt 1), 021904.

NEWMAN M E J and Watts D J (1999) Renormalization Group Analysis of the Small-world Network Model. *Physics Letters A*, 263, pp. 341-346.

NG T W, Turinici G, and Danchin A (2003) A Double Epidemic Model for the SARS Propagation. *BMC Infectious Disease*, 3(19), pp. 1-16.

NISHIURA H, Patanarapelert K, Sriprom M, Sarakorn W, Sriyab S, and Ming T I (2004) Modelling Potential Responses to Severe Acute Respiratory Syndrome in Japan: the Role of Initial Attack Size, Precaution, and Quarantine. *Journal of Epidemiology and Community Health*, 58(3), pp. 186-191.

OSTROM T M (1998) Computer Simulation: The Third Symbol System. *Journal of Experimental Social Psychology*, 24, pp. 381-392.

PEIRIS J S, Lai S T, Poon L L, Guan Y, Yam L Y, Lim W, Nicholls J, Yee W K, Yan W W, Cheung M T, Cheng V C, Chan K H, Tsang D N, Yung R W, Ng T K, Yuen K Y, and SARS Study Group (2003) Coronavirus as a Possible Cause of Severe Acute Respiratory Syndrome. *Lancet*, 361(9366), pp. 1319-1325.

RAPOPORT A (1957) A Contribution to the Theory of Random and Biased Nets.

Bulletin of Mathematical Biophysics, 19, pp. 257-571.

RIETMAN E (1989) *Exploring the Geometry of Nature: Computer Modeling of Chaos, Fractals, Cellular Automata, and Neural Networks*. Blue Ridge Summit, PA: Windcrest.

RILEY S, Fraser C, Donnelly C A, Ghani A C, Abu-Raddad L J, Hedley A J, Leung G M, Ho L M, Lam T H, Thach T Q, Chau P, Chan K P, Lo S V, Leung P Y, Tsang T, Ho W, Lee K H, Lau E M, Ferguson N M, and Anderson R M (2003) Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. *Science*, 300(5627), pp. 1961-1966.

ROTA P A, Oberste M S, Monroe S S, Nix W A, Campagnoli R, Icenogle J P, Penaranda S, Bankamp B, Maher K, Cher M H, Tong S, Tamin A, Lowe L, Frace M, Derisi J L, Chen Q, Wang D, Erdman D D, Peter T C T, Burns C, Ksiazek T G, Rollin P E, Sanchez A, Liffick S, Holloway B, Limor J, Mccaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus A D M E, Drosten C, Pallansch M A, Anderson L J, and Bellini W J (2003) Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *Science*, 300(5624), pp. 1394-1399.

SEBASTIAN B and Hoffmann C (2003) *SARS Reference*. Flying Publisher.

SIRAKOULIS G C H, Karafyllidis I, and Thanailakis A (2000) A Cellular Automaton Model for the Effects of Population Movement and Vaccination on Epidemic Propagation. *Ecological Model*, 133, pp. 209-223.

STARFIELD A M, Smith K A, and Bleloch A L (1990) Analyzing Political Decision Making from an Information Processing Perspective: JESSE. *American Journal of Political Science*, 34, pp. 74-123.

SCHRODT P A (1995) *Patterns, Rules and Learning: Computational Models of International Behavior*. Ann Arbor: University of Michigan Press.

TABER C S and Timpone R (1994) Pictures in the PC: Computational Modeling and Political Psychology. In M X Delli-Carpini, L Huddy, and R Y Shapiro (Eds), *New Directions in Political Psychology*. Greenwich, CT: JAI.

TSIMRING L S and Huerta R (2003) Modeling of Contact Tracing in Social Networks. *Physica A*, 325, pp. 33-39.

TURING A M (1950) Computing Machinery and Intelligence. *Mind*, 59, pp. 433-460.

WANG X F and Chen G (2003) Complex Networks: Small-World, Scale-Free and beyond. *IEEE Circuits and Systems Magazine*, First Quarter, pp. 6-20.

WATTS D J and Strogatz S H (1998) Collective Dynamics of 'Small-world' Networks. *Nature*, 393(6684), pp. 440-442.

WATTS D J (1999) *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, Princeton, NJ.

WHICKER M L and Sigelman L (1991) *Computer Simulation Applications: An Introduction* (Applied Social Research Methods Series, Vol. 25). Newbury Park, CA: Sage.

WHO (WORLD HEALTH ORGANIZATION) (2003) *Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)*, http://www.who.int/csr/sars/en/WHOconsensus.pdf.

YACOUBI S E and Jai A E (2002) Cellular Automata Modelling and Spreadability. *Mathematical and Computer Modelling*, 36, pp. 1059-74.

ZANETTE D H (2003) Models of Social Processes on Small-World Networks. *AIP Conference Proceeding*, 658, pp. 187-203.

ZEKRI N and Clerc J P (2001) Statistical and Dynamical Study of Disease Propagation in a Small World Network. *Physical Review E*, 64(5), 056115.

Vita

Chung-Yuan Huang received his M.A. in Computer Science (2000) from National Chiao Tung University. He earned his Ph.D. in Computer Science (2005) from the National Chiao Tung University. From 2000 to 2005 he joined the faculty at China University of Technology. He joined the faculty at Yuanpei Institute of Science and Technology in 2005. He is currently serving as an assistant professor and chairperson in the university's Department of Computer Science and Information Engineering. His research interests include network-based epidemic simulations, the computational modeling of complex systems, distributed artificial intelligence, and the applications of learning classifier systems for evolutionary computations.