12-2011

Early detection and control of potential pandemics.

Shengpeng Jin
University of Louisville

EARLY DETECTION AND CONTROL OF POTENTIAL PANDEMICS

By

Shengpeng Jin

A Thesis

Submitted to the Faculty of the

Graduate School of the University of Louisville

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

Department of Industrial Engineering

University of Louisville

Louisville, Kentucky

December, 2011

EARLY DETECTION AND CONTROL OF POTENTIAL PANDEMICS

By

Shengpeng Jin

A Thesis Approved on

October 20, 2011

By the following Thesis Committee:

_____

Suraj Alexander (Thesis Director)

_____

_____

Gerald W. Evans

_____

Weibin Zang

ACKNOWLEDGEMENTS

ABSTRACT

EARLY DETECTION AND CONTROL OF POTENTIAL PANDEMICS

SHENGPENG JIN

OCT 20th, 2011

Early information is crucial for policy makers and public health officials responsible for

protecting the public from the virulent spread of contagious diseases. Current indicators

of the spread of contagious outbreaks lag behind the actual spread of the epidemic,

leaving no time for a planned response. The studies of Christakis et al. in 2010 have

shown that social networks can provide more timely information for prediction. Our

focus, however, is on the effective control of the spread of contagious outbreaks in their

early stages. We do this by defining a more effective way to chart the spread of

contagious outbreaks, in a spatio-temporal sense, so that effective control actions can be

taken. In this paper, we use information from "sensors", such as, First Watch and EARS

(Early Aberration Response Systems) and "central" individuals in social networks for

early "spatio- temporal" prediction of virulent contagious outbreaks as a means to

allocate resources to "nip a potential pandemic in the bud." Specifically we combine

the research of Christakis et. al on social networks and that of Hongbo Yu on

"spatio-temporal" prediction of human activities to chart the spread of a virulent disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1   INTRODUCTION

1.1 Background

In 2009, the H1N1 influenza emerged out of Mexico and rapidly spread around the globe. Similarly in 2003 the respiratory illness SARS (Severe Acute Respiratory Syndrome) occurred in Guangdong province of China and lead to the death of many people all over the world. Therefore, public health agencies scrambled to understand and to control the spread of disease and prevent pandemics (Dimitrov et al. INFORMS Tutorial 2010). As we know, many infectious diseases are spread through populations via physical contacts among individuals. The patterns of these contacts tend to be highly heterogeneous. The traditional mathematical model used to understand the dynamics of epidemics is the compartmental SIR model, which assumes that the population groups are fully mixed and every individual has an equal chance of spreading the disease to another individual. However this is not the case in the real world (Meyers et al. 2005).

Our objective is to utilize the methods of social network epidemiology to predict the spatio-temporal spread of virulent contagious outbreaks as a means to allocate resources and "nip a potential pandemic in the bud". The careful collection of information from a sample of central individuals within a human social network could be used to detect contagious outbreaks before they spread among the population.

1.2 Problem Statement

It is well established that random immunization requires immunizing a very large fraction of the population in order to arrest diseases that spread through contacts between infected and susceptible individuals (Cohen et al. 2003). Mathematical modeling has long been an important tool for understanding and controlling the spread of infectious diseases. However, the current models espoused in the literature do not properly address some important aspects of disease spread even though it has proven to be quite useful in understanding epidemic dynamics.

In recent years, the study of social networks and in particular the spread of disease through these networks has attracted considerable attention in the academic community (Newman, 2002). Therefore, we will use the social network epidemiology to predict the effect of various control policies for a mildly contagious disease. It is well known that individuals near the center of a social network are likely to get infected sooner during the course of an outbreak than those at the periphery on average (Christakis et al. 2010). Our intent is to use social network epidemiology to define effective control policies to arrest the spread of a contagious disease. Unfortunately, it is typically very difficult to map a whole network to identify central individuals who might be monitored for infection (Christley et al. 2005). Therefore, an alternative strategy which does not require ascertainment of the global social network structure has been proposed by Christakis et al, namely, i.e. to simply monitor the friends of randomly selected individuals.

The current social network methods used to monitor disease spread seldom takes both the space and time factors and the travel pattern of individuals into account. Therefore, we will embellish the social network models with spatio-temporal information for more effective control of disease.

Finally, in order to make our approach more effective, we will utilize information from sources such as " First watch" and EARS" to identify individuals that are symptomatic and are diagnosed to have infections that could result in pandemics. We could also identify central individuals in the community who have been shown to have many connections and monitor them, for early signs of infections. We refer to these individuals as sensors. These sensor methods would be introduced in order to provide significant additional time to react to the epidemics especially at its early stages. The output of the disease spread prediction model introduced in this thesis could provide relatively important information to a decision support system that activates mitigatory actions for pandemic control, such as resource allocation, patient allocation, ambulance distribution and so forth.

1.3 Thesis Organization

Chapter 2 provides a literature review associated with the research of Christakis et al on social network and Hongbo Yu's research on spatio-temporal representation of travel patterns and interactions of individuals in GIS. Chapter 3 describes current approaches to disease prediction and proposed methods as well as some extension on the pandemic control. Chapter 4 discusses more advanced ways to predict the disease spread, such as risk analysis of infection, sensors to the disease prediction and their relationships with the other parts of a decision support system. Chapter 5 draws a conclusion of the thesis.

CHAPTER 2   LITERATURE REVIEW

2.1 Spatio-temporal GIS Design

Human activities are performed within a spatial and temporal context. GIS has been used

for representing human activity data, such as that obtained from travel and diary records

for the exploration of their spatio-temporal characteristics (Shaw and Wang, 2000). The

individual travel activities with their spatial, temporal and event attributes could be

organized by using a path-based representation of trips in a relational GIS environment.

A person's daily activities include physical and virtual activities. Four types of

communication modes have been suggested in the literature according to their spatial and

temporal requirements (Janelle, 2004). See Table 1.

(1) Conventional face-to-face meetings require participants to be at the same location

during the same time period. This communication mode requiring coincidence in both

space and time is classified as Synchronous Presence (SP).

(2) Post-it notes or bulletin boards must have people visit the same location, but these

visits can be at different times, to complete the information exchange. This type of

communications requires coincidence in space, but not in time, is called Asynchronous

Presence (AP).

(3) With the use of information and communication technologies (ICT)s, people are no longer required to be present at the same physical location for communications. Synchronous Telepresence (ST) only requires coincidence in time (e.g., two friends at different locations doing instant messages over the Internet).

(4) Finally, Asynchronous Telepresence (AT) is free from coincidence requirements in either space or time. E-mail between people belongs to this type of communications.

This classification system can be used to describe different types of human interactions based on their spatial and temporal requirements. The SP and AP types of human interactions are carried out in physical space and they are also what we are interested in because only physical activities could lead to the disease spread in real world. Therefore, in this thesis only SP and AP would be taken into consideration.

Spatial Temporal	Physical presence	Telepresence
Synchronous	SP Face to face (F2F)	ST Telephone Chat rooms Teleconferencing
Asynchronous	AP Post-it® notes Traditional hospital charts	AT Mail E-mail Web pages

Table 1. Communication modes based on spatial and temporal constraints

2.2 Space-time Path

Hagerstrand(1970) proposed a theoretical framework to study the constraints that affect an individual's presence in space and time and to portray individual activities in a space-time context, which is known as Time Geography. One fundamental concept is suggested under the time geographic framework to depict the capability of an individual to conduct activities in space and time which is space-time path.

A space-time path is the container of all activities performed by a person, since all activities take place at certain locations and time periods and each of them occupies a portion of the space-time path (Hagerstrand, 1970). It depicts the sequence of an individual's activities at various locations over a time period.

A space-time path offers a proper continuous representation of such a trajectory. Both physical activities and virtual activities performed by individuals leave traces in the physical space and time, which become contents of space-time paths. An individual's trajectory may pass through a location in the 2D space multiple times. When a space-time path is used to store the trajectory, every point on the space-time path possesses unique coordinates of (x, y, t) since a person only can be at a single physical location at any given time (Hongbo, 2008).
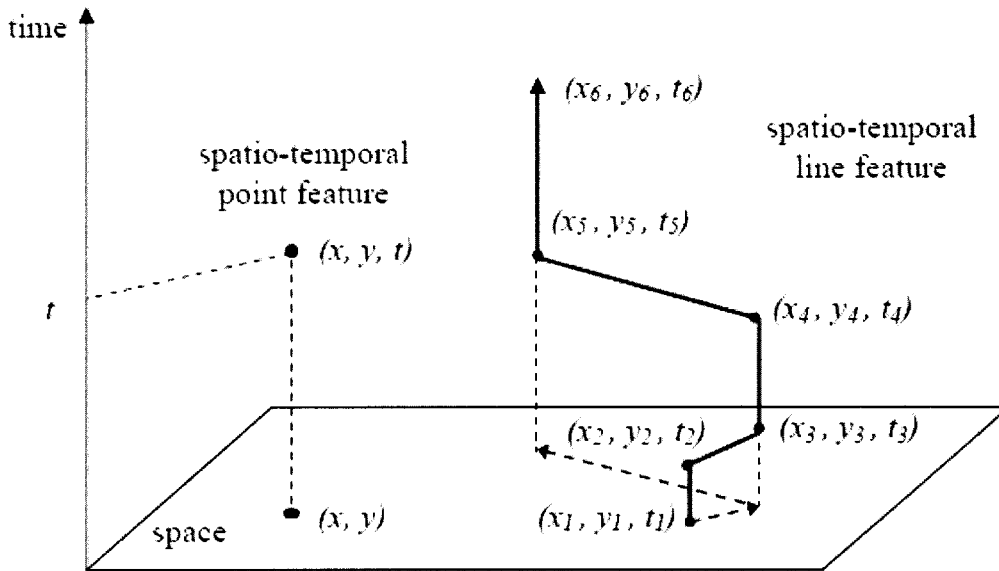
Figure 1. Spatio-temporal features in a 3D GIS framework

Figure 2 (Hongbo, 2008) shows an example of typical activities of one individual.
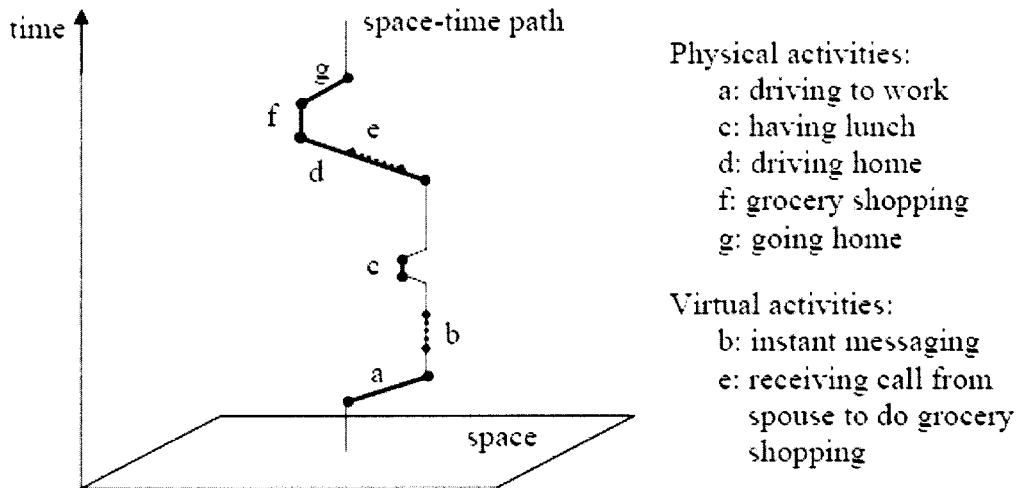
Figure 2. Locate individual activities on a space-time path

There exist three basic relationships of space-time paths between different individuals.
(Hongbo, 2008) See Figure 3.

7

(1) Co-location in time represents activities in different space–time paths that interact with each other within a common time window.

(2) Co-location in space occurs when activities in different space–time paths occupy the same location in different time windows.

(3) Co-existence describes the cases when activities take place at the same location and within a common time window.
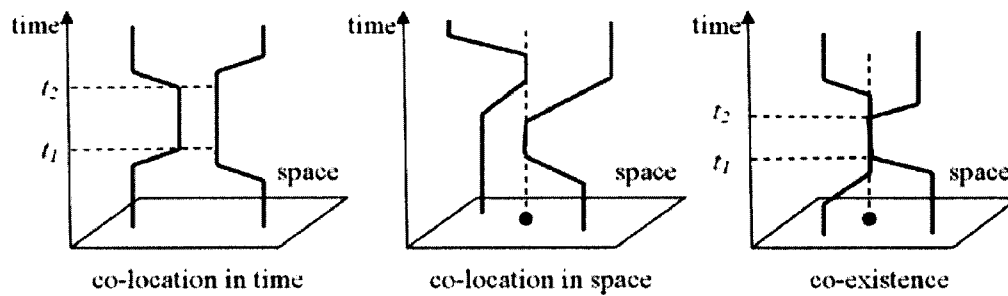
Figure 3. Space-time path relationships

2.3 Social Network

The careful collection of information from a sample of central individuals within human social networks could be used to mitigate the spread of contagious outbreaks before they happen in the population-at-large (Christakis et al. 2009). The social network itself will be an important conduit for the spread of an outbreak. However, mapping a whole network to identify particular individuals from whom to collect information is impractical, especially for large networks.

However, some other ways might be used to deal with this situation. Intuitively, it could enhance the population-level efficacy of a prophylactic intervention to vaccinate the central individuals in networks (Manhart and Holmes, 2005). Dr. Christakis et al monitored the spread of flu at Harvard College from September 1 to December 31, 2009 to evaluate the effectiveness of using nominated friends of randomly selected students as the central individuals in the social network.

They enrolled a total number of 744 undergraduate students from Harvard College as the randomly selected group and tracked whether they had the flu during that period of time. Another "nominated" group consisted of those who were named as a friend by at least once by a member of the random group. And the demographic information, such as, whether they were infected and vaccination status were collected by a completed brief questionnaire for each subject. The hypothesis is that the set of the nominated friends get infected earlier than the set of randomly chosen individuals (Christakis, 2010).
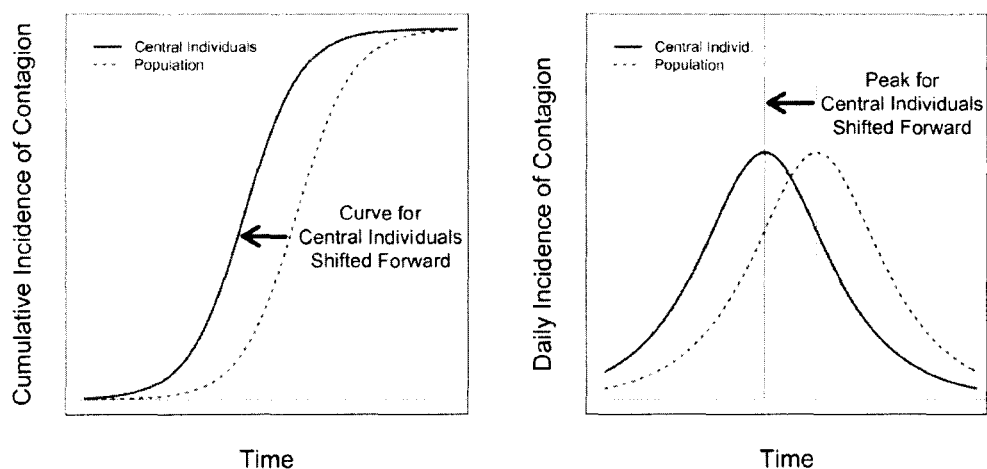
Figure 4. Theoretical differences in contagion between two groups

As hypothesized, the cumulative incidence curves for the friend group and the random

group diverge and then converge and the friends curve for flu diagnosed by medical staff

is shifted 13.9 days forward in time (95% C.I. 9.9–16.6), thus providing early detection.

The friend group showed a significant lead time prior to the estimated peak which could

be an effective technique for detecting outbreaks at early stages of an epidemic. See

Figure 5 (Christakis, 2010).
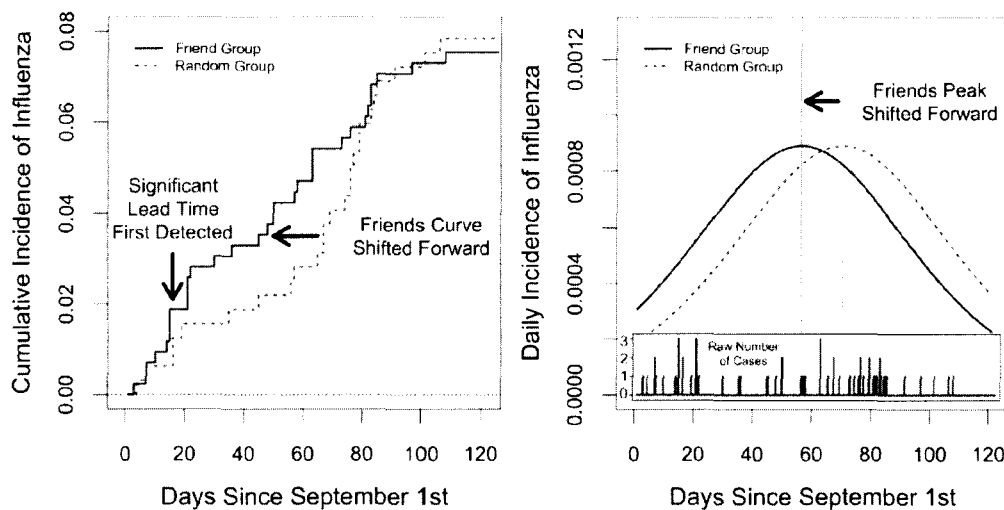
Figure 5. Empirical differences in contagion between two groups

For many contagious diseases, early knowledge of when – or whether – an epidemic is

unfolding is crucial to policy makers and public health officials responsible for defined

populations, whether small or large. In fact, with respect to flu, models assessing the

impact of prophylactic vaccination in a metropolis such as New York City suggest that

vaccinating even one third of the population would save lives and shorten the course of

the epidemic, but only if implemented a month earlier than usual (Khazen et al, 2009).

Also in case of influenza it takes time to develop the vaccine. In addition, resource planning requires early knowledge of the pandemic spread, and when it is expected to peak, etc.

In fact, this method could be used to monitor targeted populations of any size, in real time. For example, a health service at a university (or other institution) could empanel a sample of subjects who are nominated as friends and who agree to be passively monitored for their health care use.

There are mainly two steps by taking the social network epidemiology as an analytical framework to capture the disease transmission. The first step in this modeling approach is to build a realistic network model of contact pattern at an appropriate temporal and spatial scale. The second step is to predict the disease spread through the social network, based on the feature of both the disease and the network structure.

2.4 Travel Pattern

As we know, travelers are a rich source of information for infectious disease specialists. On returning from their journey, travelers can provide a representative sample of the diseases that abound in the places they have visited (Ross, 2006). There are a wide variety of travelers ranging from tourists and business people to immigrants, refugees

and foreign-born citizens who have visited friends and relatives in their home countries. Moreover, a traveller who returns home with an unusual disease could be the first clue to a new outbreak (Ross, 2006).

Therefore, a disease tracking system could be built for sharing information among their networks of travel medicine clinics or hospitals such that doctors could record travel histories and symptoms of their patients and their diagnoses in standardized electronic forms and submit these to the system which in fact, is also a central database. Then the system regularly examines the data to detect the symptoms which might indicate a new outbreak and warrants a warning to the clinics or hospitals.

2.5 Individual Centrality in Contact Network

The centrality in a social network is a parameter used to measure how central an individual is in a social network (Freeman, 1979). The concept of centrality was formally defined by Freeman. Specifically, Freeman identified three primary centrality measures: degree, closeness and betweenness.

(1) Degree centrality measures an individual's direct connectedness with other individuals;

(2) Closeness centrality provides a more global network prospective than degree centrality. Specifically, closeness centrality is a measure that indicates the degree to which an individual is near all the other individuals in the network not just those adjacent to them;

(3) Betweenness centrality is a measure of the strategic location of an individual along a potential communication path.

The study of the centrality in a social network could determine the most central individuals who play a critical role in the disease spread. Obviously, the rate of spread of a potential pandemic could be mitigated to some extent by monitoring and immunizing those with higher centrality values.

2.6 Compartmental SIR Model

The compartmental SIR model, which is relatively simple and widely used, is a traditional approach to model infectious disease dynamics.

Consider a population of N individuals and the following simple discrete-time, discrete-state epidemic model. Each individual begins in one of the three possible states:

13

(1) susceptible, meaning that the individual has never had the disease and is susceptible to being infected;

(2) infected, meaning that the individual currently has the disease and can infect other people; and

(3) resistant, meaning that the individual does not have the disease, cannot infect others, and cannot be infected.

The model simulates the progression of the disease through the three states. Individuals are first susceptible, then infected, and then become resistant by acquiring immunity to the disease (Anderson et al, 1979; Bernoulli and Blower, 2004).

The model then evolves in discrete time steps, with all individuals simultaneously acting as follows in each time step:

(1) Each susceptible individual draws a uniformly random person from the population. If the person drawn is infected, then the susceptible individual changes his state to infected with probability β.

(2) Each infected individual changes his state to resistant with probability γ.

(3) Each resistant individual remains resistant.

The parameter β captures the ability of the disease to be transmitted from one person to another; the parameter γ is related to length of the period for which an individual can transmit the disease, called the infectious period. The population in this model is a homogeneously mixed population which interacts in such a uniformly random and independent way between time steps. In this model, there is a very important parameter R_0, called the basic reproduction number, which is the expected number of new infections created by an infected individual under the most favorable conditions for transmission. For the compartmental SIR model, $R_0 = \beta/\gamma$ and generally the disease can become an epidemic only if $R_0 > 1$.

Here is a mass-action SIR compartmental model, where X(t), Y (t),and Z(t) denote the number of susceptible, infected, and resistant individuals in the population at time t. and X(t) + Y (t) + Z(t) = N.

$$\frac{dX(t)}{dt} = -\beta \cdot X(t) \cdot \frac{Y(t)}{N}, \tag{1}$$

$$\frac{dY(t)}{dt} = \beta \cdot X(t) \cdot \frac{Y(t)}{N} - \gamma \cdot Y(t). \tag{2}$$

$$\frac{dZ(t)}{dt} = \gamma \cdot Y(t). \tag{3}$$

The model can also be extended to a more complex disease spread model with a more complex population structures. For example, a natural birth/death rate or a latent period of disease could be included in the model. For more information on the SIR model and its extension, see Dimitrov and Meyers, INFORMS Tutorial 2010.

Figure 6 provides an example of typical epidemic curves defined by the SIR model (Dimitrov and Meyers, INFORMS Tutorial 2010).
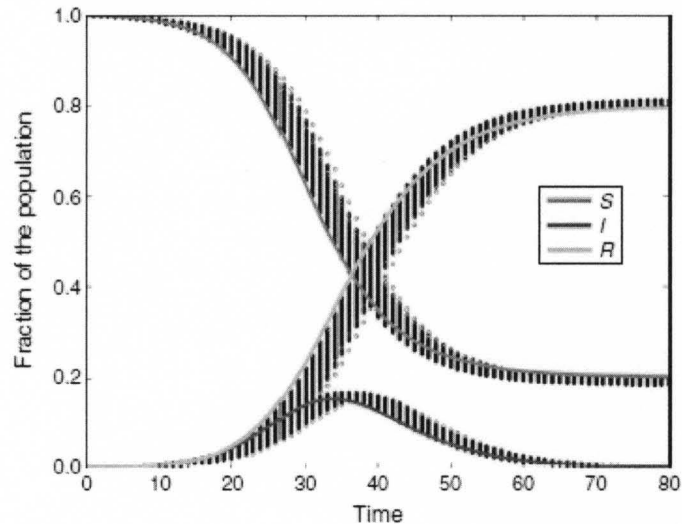
Figure 6. A typical curve of the compartmental SIR model

2.7 The Comparison of three Disease Spread Models

Although the compartmental SIR models have proven to be quite useful in modeling epidemics, they do not properly model some important aspects of disease. Moreover, the compartmental SIR models assume a fully mixed, homogeneous population in which each individual has the same amount of contacts as every other individual. Thus, simple SIR models do not accurately model the increased rate of contact in the hospitals and the decreased rate of contact of quarantined individuals. If the population at large had as many contacts as the population within a hospital, possibly the estimates of R_0 would have been more accurate, and SARS (Severe acute respiratory syndrome, see http://en.wikipedia.org/wiki/SARS) would have infected many more people.

Incorporating realistic contact patterns of the population is just one possible way to increase the fidelity of epidemic models. Diseases often spread at different rates based on age and the type of contact; they also have varying incubation periods in different age groups. For example, contacts at home tend to be more intimate than contacts at work. Hence, an infected person's family members are more susceptible to the disease. Also, disease spread is affected by both geographic location and seasonality.

Therefore, researchers have attempted using high-fidelity agent-based simulation models, where each individual is tracked as they move from home to work and back. Such models involve complex parameterization and often require extensive computation that deems the models intractable and of limited usefulness. The social network modeling approach utilized in this research provides acceptable fidelity and tractable formulations. This is illustrated in the figure below (Dimitrov and Meyers, INFORMS Tutorial 2010).
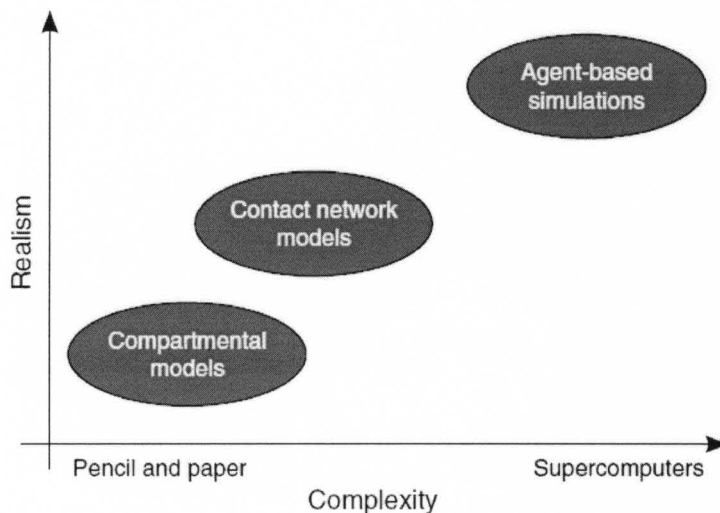
Figure 7. Complexity of epidemiological models

The Compartmental SIR models are easy to analyze but miss important, realistic details, such as heterogeneous patterns and types of contacts. Agent-based simulations are able to model reality with a great amount of detail, but are difficult to parameterize and analyze, and require large amounts of computation. Social network models capture disease transmission with a higher fidelity than compartmental models yet remain analytically tractable.

CHAPTER 3   BASIC METHODS OF DISEASE SPREAD

3.1 Basic Concepts of A Social Network

A social network is a social structure made up of individuals (or organizations) represented as "nodes", that are tied (connected) by one or more specific types of interdependency, such as friendship, kinship, common interest, financial exchange, dislike, sexual relationships, or relationships of beliefs, knowledge or prestige (Horton, 2006).

Social network analysis views social relationships in terms of network theory consisting of nodes and ties (also called edges, links, arcs or connections). Nodes are the individuals within the networks, and ties are the relationships between the individuals. The resulting graph-based structures are often very complex (Social network, wikipedia). There can be many kinds of ties between the nodes. Research in a number of academic fields has shown that social networks operate on many levels, from families up to the level of nations, and play a critical role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals (see http://en.wikipedia.org/wiki/Social_network).

Moreover, social network analysis has also been used in epidemiology to help understand

how patterns of human contact aid or inhibit the spread of diseases such as HIV in a

population (Parker, 2002). The evolution of social networks can sometimes be modeled

by the use of agent based models, providing insight into the interplay between

communication rules, rumor spreading and social structure.

3.2 The Traditional Method of the Centrality

Consider a social network modeled as a direct graph, G(V, E). Let V = (v1, v2,..., vn)

denote the set of nodes in the network, and let E = (e1, e2, ..., em) denote the number of

edges between the nodes (Hamill et al. 2006). Specifically, consider the modified

network from Hamill with n = 11 nodes with the social relationships (or paths of

communication) depicted in the figure below (Schneider et al 2011).
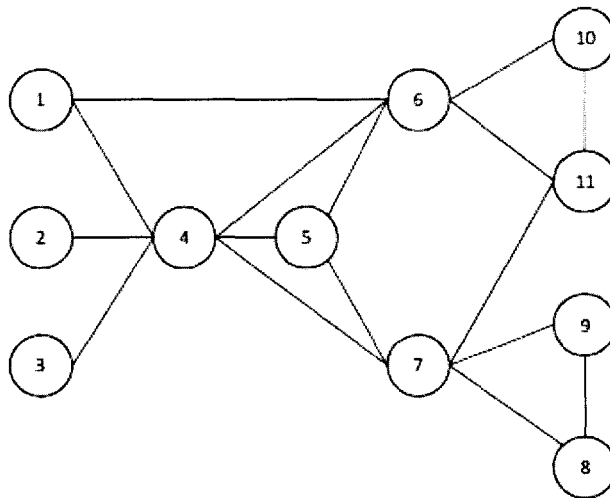
Figure 8. Sample network

One of the most important metrics in the social network analysis (SNA) is the centrality

of an individual (*each node in the network*). There are three main centrality measures:

degree, closeness and betweenness (Freeman, 1979). The traditional way to calculate

these three centrality measures is introduced below:

Degree centrality $C_D(v)$: measures an individual's direct connectedness with other

individuals. The degree of a node (or vertex) is the number of edges connected to it. Let

deg(v) denote the degree of an individual v in the network which have n individuals, and

then the degree centrality is given by

$$C_D(v) = \frac{\deg(v)}{n-1} \tag{4}$$

Closeness centrality $C_C(v)$: a measure that indicates the degree to which an individual is

close to all the other individuals in the network not just those adjacent to them. It

provides a more global network prospective than degree centrality. Let $d_G(v, c)$ denote

the length of shortest path connecting individual v with individual c, so that the closeness

centrality of an individual v is given by

$$C_C(v) = \frac{1}{\sum_{c \in V} d_G(v,c)} \tag{5}$$

Betweenness centrality $C_B(v)$: a measure of the strategic location of an individual along a

potential communication path. Let σ_{bc} denote the number of shortest paths from

individual *b* to individual *c*, and let $\sigma_{bc}(v)$ denote the number of shortest paths from

individual b to individual c that contain individual v. The betweenness centrality of an

individual v is given by

$$C_B(v) = \sum_{b \neq v \neq c \in V} \frac{\sigma_{bc}(v)}{\sigma_{bc}} \qquad (6)$$

The centrality measures for the network in Figure 8 are shown in the Table 2 (Schneider et al 2011). From Table 2, we can see that individual 4 has the largest centrality values for each measure. Individuals 6 and 7 exhibit the second highest degree centrality and meanwhile individual 7 exhibits the second highest closeness and betweenness centrality as well. All these results could indicate which individuals are the most "central" people in a social network to some extent. However, these metrics treat each individual in the network identically and assume a perfect contact chain. In reality, certain individuals within the network may be more persuasive and the contact between individuals in the network may not be that perfect since their centrality values are so close such that we could not tell the differences among individuals based on their centralities. In these instances, the centrality metrics may not adequately quantify the criticality of individuals within the network (Schneider et al 2011).

Individual	Degree	Closeness	Betweenness	Betweenness*
1	0.20	0.0500	0	0
2	0.10	0.0435	0	0
3	0.10	0.0435	0	0
4	**0.60**	**0.0714**	**43**	**1**
5	0.30	0.0588	2	0.0465
6	**0.50**	0.0588	17	0.3953
7	**0.50**	**0.0667**	**36**	**0.8372**
8	0.20	0.0435	0	0
9	0.20	0.0435	0	0
10	0.20	0.0455	0	0
11	0.30	0.0526	8	0.7273

Table 2. Traditional Centrality Measures of the Sample Network

3.3 A Proposed Centrality Measure

Consider a social network with n nodes with the connection probabilities between every two nodes given. The connection probability is the likelihood of contact between two individuals and it could be obtained by recent frequency of physical contact between the two individuals. The connection probability could reflect the possibility for one individual to get an infection from another individual to some extent, although the chance of getting infected also depends on other factors, such as the feature of the disease, the feature of the population and the social network structure.

There are several steps to calculate this centrality value of a specific node i:
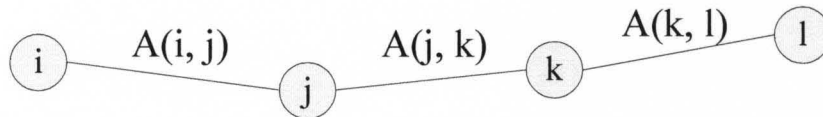
Figure 9. Sample arc in a network

1.   For each node, find the first-nearest-layer (FNL) nodes which have direct connection with the original node i. In this case, node j is the only FNL nodes and the impact of FNL nodes (FNL values) could be obtained by the formula below.

$$\text{FNL value} = \sum_{j}^{n} A(i,j) \tag{7}$$

Where A(i, j) is the degree of connectedness between node i and j.

23

2. For each FNL node, generate a second loop to their FNL nodes, which would be the second-nearest-layer (SNL) nodes of the initial node i. In this illustrative example, node k is the SNL node and their impact (SNL value) towards the initial node i could be obtained by the formula below:

$$\text{SNL value} = \sum_{k}^{n} \{A(i,j) \times A(j,k)\} \quad (8)$$

Where A(i, j) is the degree of connectedness between node i and j; A(j, k) is the degree of connectedness between node j and k.

3. Similarly, for each SNL node, generate a loop to their FNL nodes, which are the third-nearest-layer (TNL) nodes of the initial node i and their indirect impact values (TNL value) towards the initial node i could be calculated by the formula below:

$$\text{TNL value} = \sum_{l}^{n} \{A(i,j) \times A(j,k) \times A(k,l)\} \quad (9)$$

Where A(i, j) is the degree of connectedness between node i and j; A(j, k) is the degree of connectedness between node j and k; A(k, l) is the degree of connectedness between node k and l.

4. Finally, the centrality of node i would be the summation of all the impacts of FNL, SNL and TNL.

3.4 Utilizing the Proposed Measure to Identify Central Nodes in a Sample Social Network

Consider a sample non-dynamic social network model which contains 30 individuals represented in Figure 10. The number beside each arc indicates the connection probability between two adjacent nodes. Using this sample we illustrate the utilization of the proposed centrality measure to identify "central" nodes in the network.
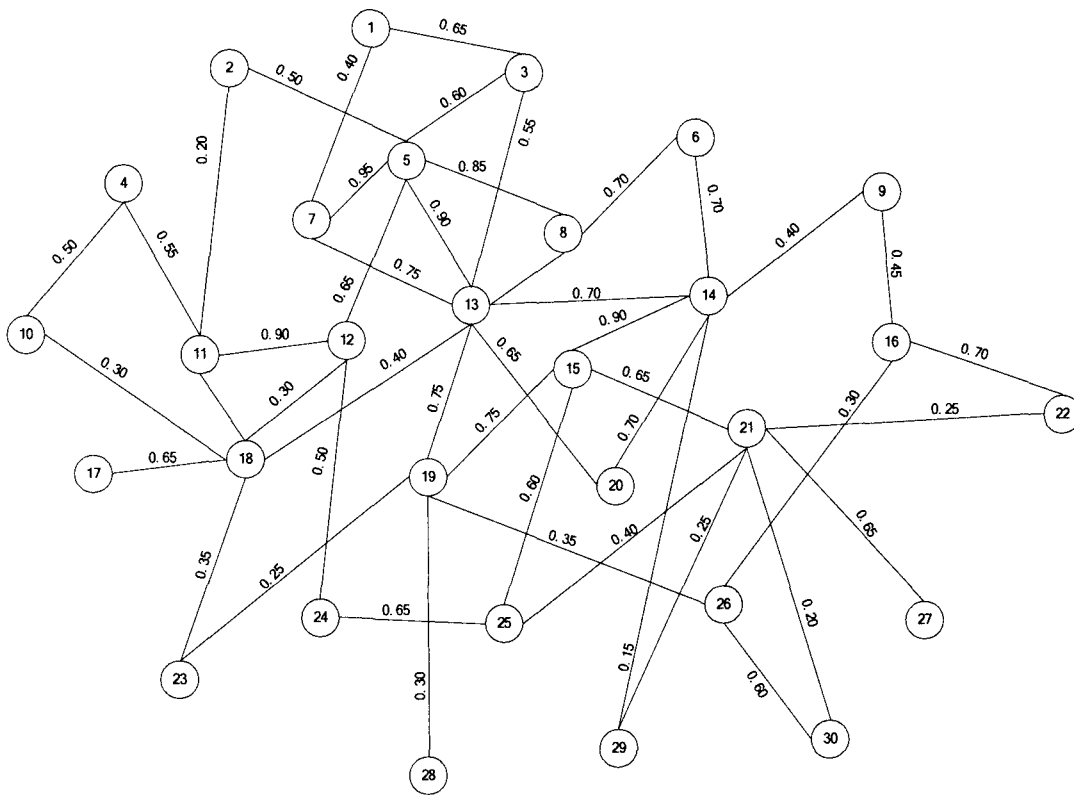
Figure 10. A 30-size sample population network

The centrality values of all nodes based on the proposed measure are shown in Table 3.

Node	1	2	3	4	5	6
Value	4.5590	6.7520	7.3490	4.9158	17.7090	9.0615
Node	7	8	9	10	11	12
Value	15.3947	16.3542	4.8150	3.3313	10.3566	12.0382
Node	13	14	15	16	17	18
Value	16.7642	14.3414	12.1193	3.6458	5.4600	10.7714
Node	19	20	21	22	23	24
Value	12.4678	11.7649	6.2289	2.5223	4.0230	7.0408
Node	25	26	27	28	29	30
Value	7.2977	4.5477	2.8990	2.2320	2.4004	2.5375

Table 3. The centrality values of a 30-size population network

From the result table, we can see that nodes 5, 7, 8, 12, 13, 14, 15 and 19 seem more central than other individuals. The Matlab program developed to automate the calculation is shown in the Appendix.
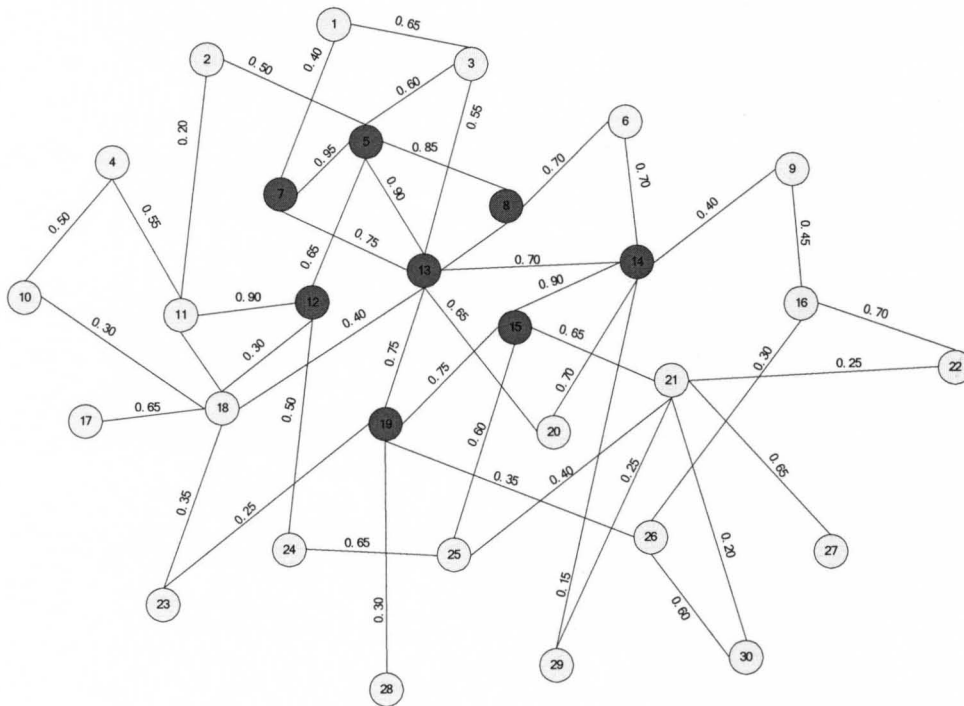
Figure 11. The more central individuals in the sample network

Visually examining, the identified central nodes (red nodes) and their probabilities in the network provide face validity to the suggested measure of centrality.

3.5 Extension of Network Analysis

In fact, each node in the network could be colored according to their centrality values in order to classify different importance of the nodes like the social network below (Christakis, 2010).
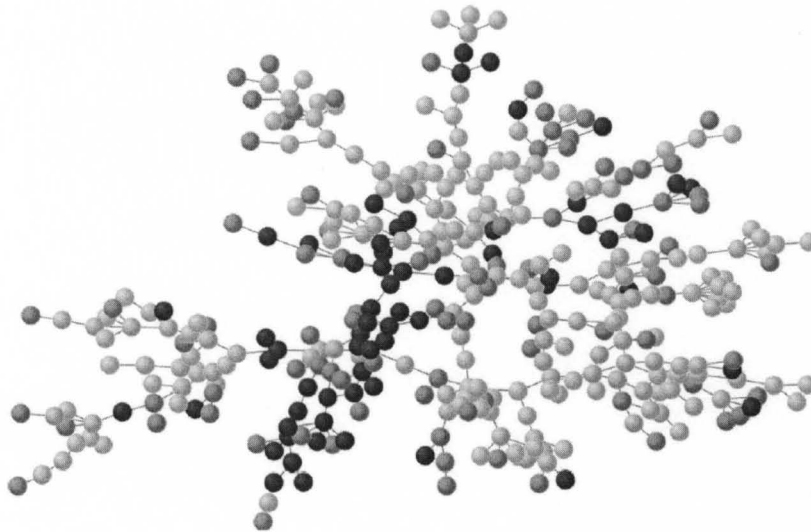
Figure 12. The different-colored network

The social network visualization can support such intervention in numerous ways.

First of all, they can be used to identify clusters of connected individuals with similar disease susceptibility and health-relevant attributes. The clusters could be targeted for collective interventions. Secondly, they can be used to identify and target individuals for

public heath interventions. Thirdly, the knowledge of the overall network structure may

be crucial to the design of public health intervention strategies.

CHAPTER 4   ASSESSMENT OF DISEASE SPREAD BASED ON SPTIO-TEMPORAL INFORMATION

Although there are a variety of different tools to predict and detect the disease spread including the SIR model, contact network models and mathematical model, we suggest that it would be more effective and beneficial to predict the disease spread considering the intensity of exposure to the disease of susceptible individuals. This would involve the consideration of both space and time factors. Since diseases caused by either viruses or bacteria involve some type of contact, either direct (shaking hands) or through the atmosphere (e.g. coughing or sneezing) between the susceptible and infected individuals.

In this thesis we use the concept of space-time paths to embellish the prediction of disease spread. As we know, the space-time path records the main activities and tracks during a certain period of time for an individual. Suppose there are two individuals and one is susceptible and the other is infected. By considering the space-time paths of the individuals, we could estimate exposure intensity and assess the likelihood for the susceptible individual to get infected.

In this case, both space and time would be taken into account which is unique in this thesis as well as other parameters, such as the characteristics of the disease. The length of time for overlapping parts of the two-individual space-time paths could be a descriptive method to evaluate the chance for one individual to get infected from the other. More detail is on the procedure is provided in the next section.

4.1 Risk Analysis of Infection

Consider two individuals A & B. A is infected and B is susceptible. The locations of A & B at any point in time t, specified by geographic coordinates $(X_a(t), Y_a(t))$ and $(X_b(t), Y_b(t))$ respectively.

Therefore, their relatively proximity at any time t, can be represented by the Euclidean distance $D_{ab}(t)$ as indicated below:

$$D_{ab}(t) = \sqrt{(X_a(t) - X_b(t))^2 + (Y_a(t) - Y_b(t))^2} \qquad (10)$$

Consider the start and end times that A & B are in the "infection range" at location i (i = 1, 2, ..., n) as Ts_i and Te_i.

In fact, for each location i, D_{ab} could be regarded as a constant.

30

Therefore, $D_{ab} = D_i = \sqrt{(X_{ia} - X_{ib})^2 + (Y_{ia} - Y_{ib}))^2}$

Then such a metric parameter could be obtained, which to some extent can describe how likely it is for B to get infected from A (from the start time Ts to the ending time Te)

$$F_{ab} = \sum_i^n \int_{T_{s_i}}^{T_{e_i}} \frac{\gamma(t - T_{s_i})^p}{D_{i_{ab}}{}^{q+\varepsilon}} dt \qquad (11)$$

Where p is the time parameter and q is the distance parameter. p should be between 0 and 1; q should be greater than 1. In fact, both p and q could be fixed by a large amount of experimental data and this would be the future research.

Additionally, ε is a very small positive number which prevents the denominator from being zero.

γ is the disease parameter which describes the spread rate of the disease and is related to multiple factors, such as temperature, seasons, and features of the disease.

However, in the real world it is unlikely that individuals could specify the locations where they had been in terms of exact geometric coordinates. It is conceivable however, that they could identify landmarks, such as buildings, museums, libraries, restaurants, etc.

where they had been. Therefore, the formula above could be approximately rewritten as follows:

$$F_{ab} = \sum_i^n \int_{T_{s_i}}^{T_{e_i}} \frac{\gamma(t-T_{s_i})^p}{R_i^q} \, dt \qquad (12)$$

Where R_i represents radius of a circle that could encompass the areas of possible movement at common location i between two individuals.

F_{ab} is called infection index in this thesis and in standardized form, it would be a value between 0 and 1, which could represent the likelihood of B being infected by A.

The following example illustrates this concept:

Suppose there are two individuals and their activity information during two days are in Table 4.

	Start Time	End Time	Activity & Place		Start Time	Arrival Time	Activity & Place
			Individual 1				**Individual 2**
	08:40	09:00	Walked to IE Dept		08:15	08:57	Drived to IE Dept
	10:02	10:08	Walked to UofL Book Store		10:05	10:12	Walked to UofL Library
	10:41	10:47	Walked to UofL Library		10:55	10:58	Walked to Classroom A
1st Day	12:24	12:40	Walked to Restaurant A for Lunch	1st Day	12:15	12:25	Walked to SAC for Lunch
	13:12	13:30	Returned to IE Dept		13:10	13:22	Returned to IE Dept
	17:10	17:32	Returned Home		16:30	16:47	Walked to SAC for Sports
					19:08	19:35	Drove Home
	09:00	09:18	Walked to UofL Library		09:50	10:20	Drove to Company A for Co-op
2nd Day	12:30	12:47	SAC for Lunch	2nd Day	11:25	12:10	Left Company A for Restaurant A
	13:25	13:41	UofL Library		12:40	12:53	Drove Back to Company A
	16:40	17:02	Returned Home		15:14	15:38	Drove Home

	Common Place	Time	Duration	Radius
	IE Dept	9:00-10:02	62 mins	25
1st Day	UofL Library	10:47-10:55	8 mins	35
	IE Dept	13:30-16:30	180 mins	25

Table 4. The Activity information of two individuals during two days

And by using ArcGIS software, we can generate the space-time paths of their activity information as shown in Figure 13. The red line represents the first individual's activity information and the blue one for the other individual. On the space-time path, there are several overlapping parts, which represents synchronous activities, where they are both are in close proximity. This, as noted previously, would factor into risk of getting infected.
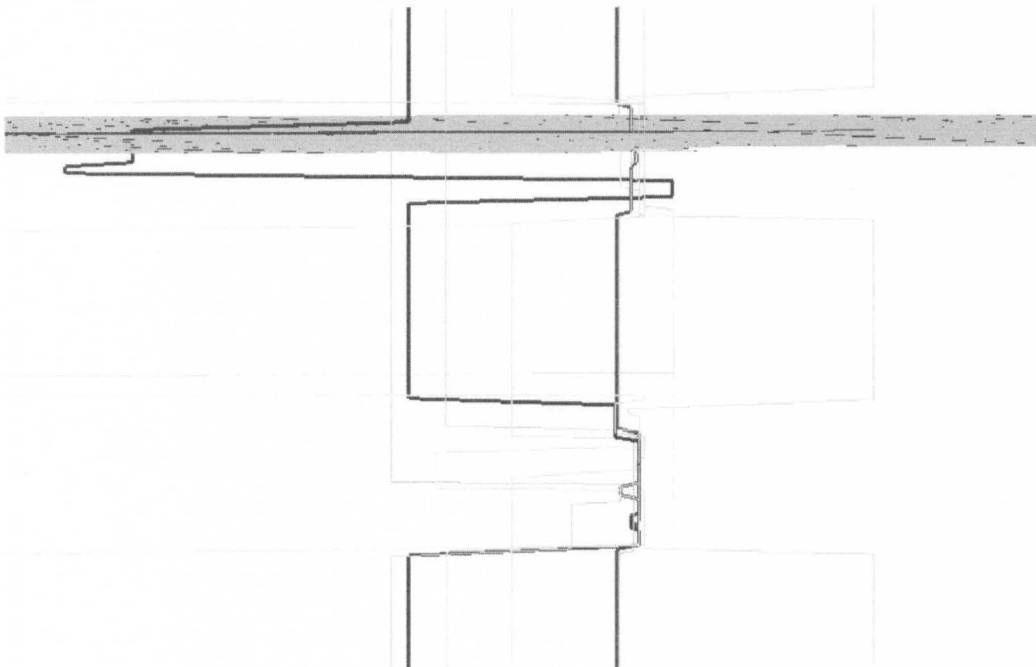
Figure 13. The Space-time Paths of two individuals

Given different parameters such as γ, p and q, F_{ab} could have different values which are displayed in Table 5 and their data analyses are shown in Figure 14.

	F_{ab}		
	q\p	0.3	0.6	0.9
$\gamma = 0.2$	1.5	1.3263	4.8136	18.4013
	2	0.2649	0.0622	3.6794
	2.5	0.0529	0.1924	0.7358
	q\p	0.3	0.6	0.9
$\gamma = 0.4$	1.5	2.6526	9.6272	36.8026
	2	0.5298	1.9244	7.3588
	2.5	0.1058	0.3848	1.4716
	q\p	0.3	0.6	0.9
$\gamma = 0.6$	1.5	3.9789	14.4408	55.2039
	2	0.7947	0.1866	11.0382
	2.5	0.1587	0.5772	2.2074
	q\p	0.3	0.6	0.9
$\gamma = 0.8$	1.5	5.3052	19.2544	73.6052
	2	1.0596	0.2488	14.7176
	2.5	0.2116	0.7696	2.9432
	q\p	0.3	0.6	0.9
$\gamma = 1.0$	1.5	6.6315	24.068	92.0065
	2	1.3245	0.311	18.397
	2.5	0.2645	0.962	3.679

Table 5. Different F_{ab} values with different parameters
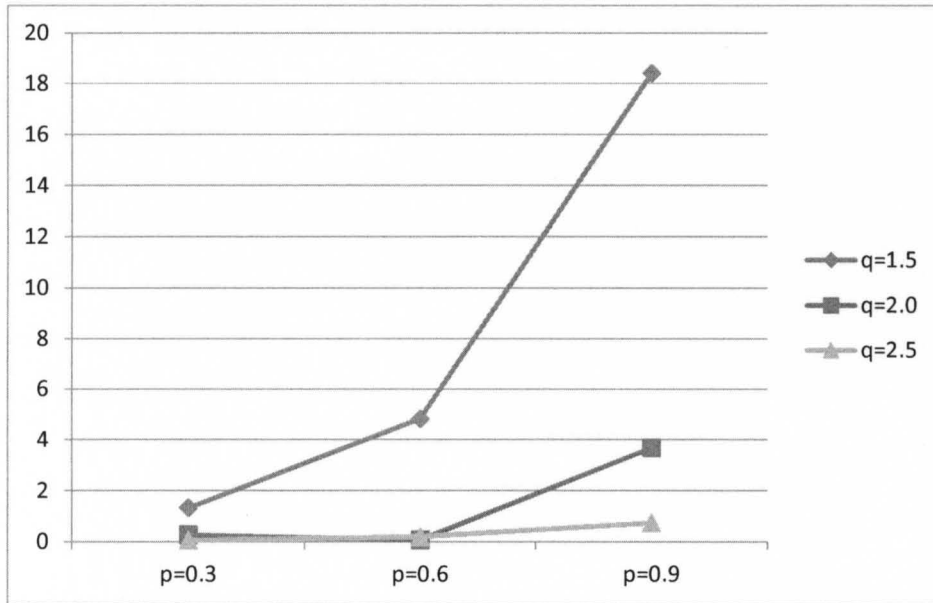
Figure 14. The Curve of different F_{ab} values with different parameters (γ= 0.2)

So from both Table 5 and Figure 14, we can see that for the same value of γ, as p increases, F_{ab} would be increased accordingly; Conversely, F_{ab} would decrease when q increases. For the same p and q, the risk would increase if γ increases. Intuitively this makes sense, since higher the infectivity of the disease, higher risk of getting infected.

4.2 Possible Process of Getting information for Spatio Temporal Analysis

The whole process is shown in Figure 15.
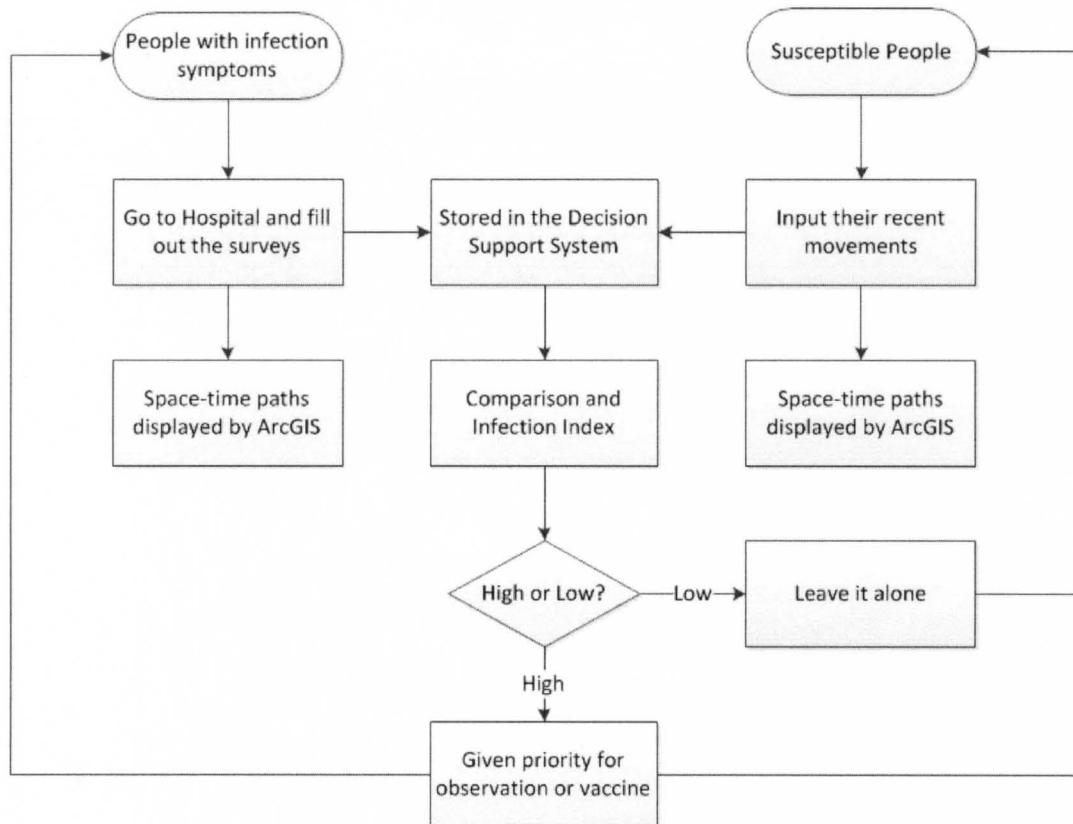
Figure 15. Possible Process of Getting information for Spatio Temporal Analysis

(1) The people who seem to have the infection symptoms go to the hospitals spontaneously and then they have to fill out the surveys provided by the hospital. The surveys contain some questions about the information of where they have been to recently and when.

(2) Then the geographic information of these infected individuals could be obtained by space-time paths, which would be displayed in ArcGIS environment.

(3) For those who care more about their risk of getting infected and mostly susceptible or not detected to be infected yet, they can input their information about their recent

tracks into the decision support system. Then they will get a feedback or report about their infection risk based on the infection index.

(4) If the risk is high, they would be given priority for observation or vaccines. Otherwise, they could do the routine activities as normal but still need to keep an eye on the spatio-temporal trend of the disease spread.

4.3 Sensors to the Proactive Intervention

As we know, it is generally costly, time-consuming, and often impossible to map a whole social network, especially for a large population to identify central individuals to serve as sensors for early detection of possible virulent disease outbreaks and to track their spread. However, there are possibly other options for identifying suitable "sensors".

For example, the "Louisville Connectors" could be considered as one set of sensors for the Louisville metropolitan area. The Louisville connectors are a diverse group of 128 individuals ranging in age from 28 to 71, from 5,500 nominations submitted by people through Louisville and Southern Indiana and they could be considered as "central individuals" to monitor. The hypothesis is that members of this group are more likely to get infections early. In case of an infection in this group, their social networks and spatio-temporal interactions would be considered immediately, and mitigatory actions to stop a possible epidemic would be taken.

Also individuals identified by sources, such as, hospital laboratories, and "First Watch" with symptoms of virulent infections can answer questionnaires, which would provide information to construct their social networks and spatio-temporal interactions.

Additionally, The Early Aberration Reporting System (EARS, see www.bt.cdc.gov/surveillance/ears/) of the Centers for Disease Control and Prevention (CDC) allows the analysis of public health surveillance data using available aberration detection methods. The primary purpose of EARS is to provide national, state, and local health departments with several alternative aberration detection methods. EARS helps assist local and state health officials to focus limited resources on appropriate activities during epidemiological investigations of important public health events. Finally, EARS allows end users to select validated aberration detection methods and modify sensitivity and specificity thresholds to values considered to be of public health importance by local and state health departments.

Also, we can explore the use of readily available or extant data, such as that available using social network websites such as Facebook or MySpace or mobile phone networks or perhaps overlapping lists of memberships in some organizations or clubs or lists of users of public services (Christakis, 2010). Another proposal to use Google Trends to monitor online searches for information about the flu suggests that this approach could

offer a better indicator, providing evidence of an outbreak at least a week before published CDC reports (Carneiro and Mylonakis, 2009).

Finally, the survey methodology or self-report could also be considered as an effective senor. Such sources of data would allow us to visualize potential direction of disease spread in large networks composed of hundreds of thousands of people and to intervene proactively.

4.4 Comprehensive Application

In a similar fashion, infection spread within a hospital or treatment facility could be mitigated by tracking spatio-temporal contact with infected patients with infection stay long in hospital. Beside, most hospitals do not track which equipment used on which patient.

One way to accomplish this would be through the use of RFID tags that track the movement of patients, hospital staff and equipment and their spatio-temporal interactions with an infected person. This would enable the defining a risk level for all hospital residents who are susceptible to infection, which would allow efficient prophylactic measures to be appropriately scheduled and adopted.

4.5 Future Research

Future research should aim to extend and validate the application of infection index and the likelihood of infection to large networks. The infection index itself could be expanded to include more detail, such as the influence gender, age, population, climate, etc. Finally, feedback of values of infection index and the likelihood of infection to the general public based on interaction with a web based tool would be a valuable contribution to society.

CHAPTER 5   CONCLUSION

In recent year, the infectious diseases have emerged from some part of the world and rapidly spread around the globe. In the last decade alone disease, such as SARS in 2003 and H1N1 in 2009, have spread globally, and have received needed attention from the public as well as the public health agencies. It is essential to understand, predict and control the spread of the disease. However, as discussed previously, the mathematical techniques used to understand, forecast and control the spread of infectious disease is not effective and sometimes lags the actual spread of disease, and hence is of limited value for proactive actions to mitigate the spread.

The approach that has been developed and suggested in this thesis is built on existing methods from diverse fields such as contact network modeling, graph theory, space-time path development and risk analysis. Some "sensors", which in fact are the sample group of individuals, or web-based tools, or even survey methodology, and comprehensive application of the approach, are also introduced in this thesis to mitigate the disease spread.

REFERENCE

Aschwanden C. Spatial Simulation Model for Infectious Viral Diseases with Focus on SARS and the Common Flu. in Proceedings of the 37th Hawaii International Conference on System Sciences, 1-5, 2004.

Bansal S., Grenfell B., and Meyers L.A. When individual behavior matters: Homogeneous and network models in epidemiology. Journal of the Royal Society Interface 4:879–891, 2007.

Bansal S., Pourbohloul B., Hupert N., Grenfell B., and Meyers L.A. The shifting demographic landscape of influenza. PLoS One, in press, 2010.

Bansal S., Pourbohloul B., and Meyers L. (2006) A comparative analysis of influenza vaccination programs. PLoS Med 3: e387.

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

Bernard L., and Kruger T. Integration of GIS and spatio-temporal simulation models: interoperable components for different simulation strategies. Transaction in GIS. 4(3). June, 2000.

Boccaletti S., Latora V., Moreno Y., Chavez M., and Hwang D.U. Complex networks: Structure and dynamics. Phys. Rep., vol. 424, p.175, 2006.

Carneiro H.A., Mylonakis E (2009) Google trends: a web-based tool for real-timesurveillance of disease outbreaks. Clinical Infect Dis 49: 1557–1564.

Cassa C.A., Iancu K., Olson K.L., Mandl K.D. A software tool for creating simulated outbreaks to benchmark surveillance systems. BMC Medical Informatics and Decision Making 2005, 5:22.

Christakis N.A., Fowler J.H. (2010). Social Network Sensors for Early Detection of Contagious Outbreaks. PLoS ONE 5(9): e12948. doi:10.1371/journal.pone.0012948.

Christakis N.A., Fowler J.H. The spread of obesity in a large social network over 32 years. N Engl J Med 2007;357: 370–379.

Christakis, N.A., and Fowler J.H. 2009. Social network visualization in epidemiology. Norsk Epidemiologi 19(1): 5-16.

Cohen R., Havlin S., and ben Avraham D. Efficient immunization strategies for computer networks and populations. Physical Review Letters, 91:247901, 2003.

Craft M.E., Volz E., Packer C., and Meyers L.A. (2009). Distinguishing epidemic waves from disease spillover in a wildlife population. Proc. R. Soc. Lond., B, Biol. Sci., 276, 1777–1785.

Davey V.J., Glass R.J., Min H.J., Beyeler W.E., and Glass L.M. (2008). Effective, robust design of community mitigation for pandemic influenza: A systematic examination of proposed US guidance.

DellValle S.Y., Hyman J.M., Hethcote H.W., and Eubank S.G. Mixing patterns between age groups in social networks. Social Networks. 2007, 29:539-554.

Dimitrov N., Goll S., Meyers L.A., Pourbohloul B., Hupert N. Optimizing tactics for use of the U.S. antiviral strategic national stockpile for pandemic (H1N1) influenza, 2009. PLoS Currents: Influenza, 2009; RRN1127.

Dimitrov N.B. and Meyers L.A. Mathematical approaches to infectious disease prediction and control. Tutorials in Operations Research, INFORMS 2010, 1-25, 2009.

Eubank S., Guclu H., Kumar A., Marathe M.V., Srinivasan A., Toroczkai Z., and Wang N. Modeling disease outbreaks in realistic urban social networks, Nature 429 (2004). pp.180-184.

Farnsworth M.L., and Ward M.P. Identifying spatio-temporal patterns of transboundary disease spread: examples using avian influenza H5N1 outbreak, Veterinary Research 40 (2009), p. 20.

Freeman L.C. (1979). Centrality in Social Networks: Conceptual Clarification, Social Networks, 1, 215-239.

Hamill J.T., 2006, Analysis of Layered Social Networks, Dissertation. Air Force Institute of Technology.

Keeling M.J., Danon L., Vernon M.C., and House T.A. Individual identity and movement networks for disease metapopulations. Proceedings of the National Academy of Sciences of the United States of America, vol. 107, no. 19, pp. 8866–8870, 2010.

Khazen N., Hutton D.W., Garber A.M., Hupert N., Owens D.K. (2009). Effectiveness and cost-effectiveness of vaccination against pandemic influenza (H1N1). Ann Intern Med 151: 829–839.

Kwan M.P. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. Transportation Research Part C 2000, 8:185-203.

Leskovec J., Krause A., Guestrin C., Faloutsos C., VanBriesen J., and Glance N. Cost-effective outbreak detection in networks, in Proc. of ACM KDD, 2007.

Lin J., Jane C., and Yuan J. (1995). On Reliability Evaluation of a Capacitated-Flow Network in Terms of Minimal Pathsets, Networks, 25, 131-138.

MacNab Y.C., and Dean C.B., 2002. Spatial-temporal modelling of rates for the construction of disease maps. Stat. Med. 21, 347–358.

Mangili, A. and Gendreau, M.A. (2005) Transmission of infectious diseases during commercial air travel, Lancet, 365, 989–996.

Mark D., Egenhofer M., Bian L., Rogerson P., and Vena J. Spatio-temporal GIS Analysis for Environmental Health, in: 2nd International Workshop on Geography and Medicine (GEOMED'), Paris, France, 1999.

Meyers L.A., Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. Bulletin of the American Mathematical Society 44:63–86, 2007.

Meyers L.A., Newman M.E.J., and Pourbohloul B. Predicting epidemics on directed contact networks. Journal of Theoretical Biology 240:400–418, 2006.

Meyers L.A., Newman M.E.J., Martin M., and Schrag S. Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. Emerging Infectious Diseases 9:204–210, 2003.

Meyers L.A., Pourbohloul B., Newman M.E.J., Skowronski D.M., and Brunham R.C. Network theory and SARS: Predicting outbreak diversity. Journal of Theoretical Biology 232:71-81, 2005.

Mokbel M.F., Ghanem T.M., and Aref W.G. Spatio-temporal Access Methods. IEEE Data Engineering Bulletin, 26(2), 2003.

Mossong J., Hens N., Jit M., Beutels P., Auranen K., Mikolajczyk R., Massari M., Salmaso S., Tomba G.S., Wallinga J., Heijne J., Sadkowska-Todys M, Rosinska M., and Edmunds W.J. Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS Med 5 (2008), p. e74.

Newman M.E.J. The spread of epidemic disease on networks. Phys. Rev. E, 66(016128), 2002. cond-mat/0205009.

O'Malley A.J., Christakis N.A. (2011). Longitudinal Analysis of Large Social Networks: Estimating the Effect of Health Traits on Changes in Friendship Ties. Statistics in Medicine 30: 950–964.

Ostfeld, R.S., Glass, G.E. and Keesing, F. (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. Trends in Ecology and Evolution, 20, 328–336.

Perez, L., and Dragicevic, S. An agent-based approach for modeling dynamics of contagious disease spread. International Journal of Health Geographics, 8:50, 2009.

Pourbohloul B., Meyers L.A., Skowronski D.M., Krajden M., Patrick D.M., and Brunham R.C. Modeling control strategies of respiratory pathogens, Emerg. Infect. Dis. 11 (2005), pp. 1249–1256.

Rosenquist, J.N., Fowler, J.H. and Christakis, N.A. (2011). Social network determinants of depression. Molecular Psychiatry. 16(3): 273-281.

Rossion B., Delvenne J.F., Debatisse D., Goffaux V., Bruyer R., Crommelinck M., and Guérit J.M. Spatio-temporal localization of the face inversion effect: an event-related potentials study, Biol. Psychol. 50 (1999), pp. 173–189.

Salathe M., Kazandjieva M., Lee J.W., Levis P., Feldman M.W., and Jones J.H. A high-resolution human contact network for infectious disease transmission, Proc. Natl. Acad. Sci. U.S.A. 107 (2010), pp. 22020–22025.

Schneider K., Rainwater C., and Pohl E.A. Investigating Actor Importance in a Multi-State Social Network, Proceedings of the 2011 Industrial Engineering Research Conference.

Shaw S-L, Bombom L., and Yu H. 2008 A space-time GIS approach to exploring large individual-based spatiotemporal datasets. Transactions in GIS 12: 425–41.

Shaw S. L., and Yu H. A GIS-based Time-geographic Approach of Studying Individual Activities and Interactions in A Hybrid Physicalvirtual Space. Journal of Transport Geography. 2009, vol.17, pp.141-149.

Soubeyrand S., Held L., Hohle M., and Sache I. Modelling the spread in space and time of an airborne plant disease, Journal of the Royal Statistical Society C 57 (2008), 253-272.

Sparks R., Carter C., Graham P.L., Muscatello D., Churches T., Kaldor J., Turner R., Zheng W., and Ryan L. Understanding sources of variation in syndromic surveillance for early warning of natural or intentional disease outbreaks, IIE Transactions 42 (9) (2010), pp. 613–631.

Tao Y., Papadias D., and Sun J. The TPR*-Tree: An Optimized Spatio-temporal Access Method for Predictive Queries. In VLDB, 2003.

Theophilides C.N., Ahearn S.C., Grady S., and Merlino M. Identifying West Nile virus risk areas: the dynamic continuous-area-space-time system, Am. J. Epidemiol. 157 (2003), pp. 843–854.

Volz E. SIR dynamics in structured populations with heterogeneous connectivity. Journal of Mathematical Biology 56:293–310, 2007.

Volz E., and Meyers L.A. Susceptible-infected-recovered epidemics in dynamic contact networks. Proceeding of the Royal Society B 274:2925–2933, 2007.

Volz E., and Meyers L.A. Epidemic thresholds in dynamic contact networks. Journal of the Royal Society Interface 6:233‑241, 2009.

Watkins R.E., Eagleson S., Beckett S., Garner G., Veenendaal B., Wright G., Plant A.J. Using GIS to create synthetic disease outbreaks. BMC Med Inform Decis Mak 2007, 7:4.

Yu H. (2006). Spatio-temporal GIS design for exploring interactions of human activities. Cartography and Geographic Information Science 33: 3–19.

Yuan M. (1996). Temporal GIS and spatio-temporal modeling. In: International Conference/Workshop Integrating GIS and Environmental Modeling, January.

APPENDIX

1. Matlab codes for the individual centrality model

```
function V = getCentrality

A = [0, 0, 0.65, 0, 0, 0, 0.40, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0.50, 0, 0, 0, 0, 0, 0.3, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0.65, 0, 0, 0, 0.60, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0.50, 0.55, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0.50, 0.60, 0, 0, 0, 0, 0.85, 0, 0, 0, 0.65, 0.90, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0, 0, 0, 0.70, 0, 0, 0, 0, 0, 0.70, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0.40, 0, 0, 0, 0.95, 0, 0, 0, 0, 0, 0, 0, 0.75, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0.85, 0.70, 0, 0, 0, 0, 0, 0, 0.90, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.40, 0, 0.45, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0.50, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0.30, 0, 0.55, 0, 0, 0, 0, 0, 0, 0, 0.90, 0, 0, 0, 0, 0, 0.75,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0.65, 0, 0, 0, 0, 0, 0.90, 0, 0, 0, 0, 0, 0, 0.70,
0, 0, 0, 0, 0, 0.50, 0, 0, 0, 0, 0, 0;
     0, 0, 0.55, 0, 0.90, 0, 0.75, 0.90, 0, 0, 0, 0, 0, 0.70, 0, 0,
0, 0.40, 0.75, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0, 0.70, 0, 0, 0.40, 0, 0, 0, 0.70, 0, 0.90, 0, 0,
0, 0, 0.70, 0, 0, 0, 0, 0, 0, 0, 0.15, 0;
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.90, 0, 0, 0, 0, 0.75,
0, 0.65, 0, 0, 0, 0.60, 0, 0, 0, 0, 0;
     0, 0, 0, 0, 0, 0, 0, 0, 0.45, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0.70, 0, 0, 0, 0.30, 0, 0, 0, 0;
```

      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.65, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0.75, 0.70, 0.40, 0, 0, 0, 0.65,
0, 0, 0, 0, 0, 0.35, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.75, 0, 0.75, 0, 0, 0, 0,
0, 0, 0, 0.25, 0, 0, 0.35, 0, 0.30, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.65, 0.70, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.65, 0, 0, 0, 0, 0,
0, 0.25, 0, 0, 0.40, 0, 0.65, 0, 0.25, 0.20;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.70, 0, 0, 0, 0,
0.25, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.35, 0.25,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.50, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0.65, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.60, 0, 0, 0, 0, 0,
0.40, 0, 0, 0.65, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0, 0, 0.35,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.60;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.65,
0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.15, 0, 0, 0, 0, 0, 0,
0.25, 0, 0, 0, 0, 0, 0, 0, 0, 0;
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.20,
0, 0, 0, 0, 0.60, 0, 0, 0, 0]

    n = size(A);
    M = zeros(n);   % an anrry which will indicate which 3-layer nodes
are related to a specific node
    M1 = zeros(n);   % an anrry which will indicate which nodes have the
first relationship with specific node
    M2 = zeros(n);   % an anrry which will indicate which nodes have the
second relationship with specific node
    M3 = zeros(n);   % an anrry which will indicate which nodes have the
third relationship with specific node

    MV = zeros(n);
    MV1 = zeros(n);
    MV2 = zeros(n);
    MV3 = zeros(n);

```
V = zeros(1,n); % record the value of each node



flag1 = 0;
flag2 = 0;
flag3 = 0;



% for i = 1:n
%    for j = 1:n
        for k=1:n
            for m = k:n


                if A(k,m) > 0 & flag1 == 0
                    flag1 = flag1 + 1;
                    V(1, k) = A(k, m);
                elseif A(k, m) > 0.0 & flag1 > 0
                        flag1 = flag1 + 1;
                        V(1, k) = (1 - V(1, k)) * (1 - A(k, m));
                        V(1, k) = 1 - V(1, k);
                end


            end
        end
%    end
% end

%--- The codes below is to find 3-layer nodes related to each node.
for i = 1:n
    for j = 1 : n

        if A(i, j) ~=0 & i ~= j
            M1(i, j) = 1;
            MV1(i, j) = A(i, j);
            for k = 1: n
                if A(j, k) ~=0 & j ~=k & i ~= k
                    M2(i, k) = 1;
                    MV2(i, k) = A(i, j)* A(j, k);
                    for m = 1:n
                        if A(k, m)~=0 & k~= m & i~= m & j~=m
                            M3(i, m) = 1;
```

49

```
                                       MV3(i, m) = A(i, j)* A(j, k)* A(k, m);
                            end
                        end
                    end
                  end

                 end
            end
        end

        MV = MV1 + MV2 + MV3;

        for i = 1: n
            for j = 1: n
                V(i) = V(i) + MV(i, j);
            end
        end

%------------------------------------------
```

2. Matlab codes for the risk analysis model

```
function getInfect

syms x;

A = [ 540, 602, 1 ;
      608, 641, 2 ;
      647, 744, 3 ;
      760, 792, 4 ;
      810, 1030, 1];

B = [ 537, 605, 1 ;
      612, 655, 3 ;
      658, 735, 6 ;
      745, 790, 5 ;
      802, 990, 1 ;
      1007, 1148, 5];

R = [1, 25;
     2, 5;
     3, 35;
     4, 8;
     5, 20;
     6, 3;
     7, 100]

s = size(A, 1);    % get the number of rows in Matrix A
t = size(B, 1);    % get the number of rows in Matrix B

num = 0 ; % define the number of rows in the result Matrix C

for i = 1:s
  for j = 1:t

      if A(i, 3) == B(j, 3)
         if B(j, 1) <= A(i, 1)
            if B(j, 2) <= A(i, 2) & B(j, 2)>=A(i, 1)
               num = num + 1;
               C(num, 1) = A(i, 1);
               C(num, 2) = B(j, 2);
               C(num, 3) = A(i, 3);
               C(num, 4) = R(A(i, 3), 2);
```

```
                elseif B(j, 2) >= A(i, 2)
                    num = num + 1;
                    C(num, 1) = A(i, 1);
                    C(num, 2) = A(i, 2);
                    C(num, 3) = A(i, 3);
                    C(num, 4) = R(A(i, 3), 2);
                end
            elseif B(j, 1) >= A(i, 1) & B(j, 1) <= A(i, 2)
                    if B(j, 2) >= A(i, 2)
                    num = num + 1;
                    C(num, 1) = B(j, 1);
                    C(num, 2) = A(i, 2);
                    C(num, 3) = A(i, 3);
                    C(num, 4) = R(A(i, 3), 2);

                elseif B(j, 2) <= A(i, 2)
                    num = num + 1;
                    C(num, 1) = B(j, 1);
                    C(num, 2) = B(j, 2);
                    C(num, 3) = A(i, 3);
                    C(num, 4) = R(A(i, 3), 2);
                    end
                end

            end
        end

    end

    m = size(C, 1);
    m1 = size(C, 1);
    m2 = size(C, 1);
    Eps = 0.00001;   % Epsilon is a very small positive number
    Gam = 0.2;       % Gamma is a medical parameter about the feature
of the disease
    f = 0.0;
    f1 = 0.0;
    f2 = 0.0;
    p = 0.30;
    q = 1.5;
    q1 = 2;
    q2 = 2.5;

    for k = 1: m
```

```
        fx = Gam * (x - C(k, 1))^p/ (C(k, 4)^q + Eps);
        f = f + int(fx, x, C(k, 1), C(k, 2));
    end


    g = 0.00;
    g = double(f)


    for k = 1: m1
        fx1 = Gam * (x - C(k, 1))^p/ (C(k, 4)^q1 + Eps);
        f1 = f1 + int(fx1, x, C(k, 1), C(k, 2));
    end


    g1 = 0.00;
    g1 = double(f1)


    for k = 1: m2
        fx2 = Gam * (x - C(k, 1))^p/ (C(k, 4)^q2 + Eps);
        f2 = f2 + int(fx2, x, C(k, 1), C(k, 2));
    end


    g2 = 0.00;
g2 = double(f2)
```

CURRICULUM VITAE

Name: Shengpeng Jin

Address: Department of Industrial Engineering
                  J.B. Speed School of Engineering
                  University of Louisville
                  Louisville, KY 40292

Education: Bachelor of Science, Automation
                  Dalian University of Technology
                  2002-2006
                  Master of Science, Control Theory and Control Engineering
                  Dalian University of Technology
                  2006-2009

Awards & Honors:
                  Graduate Fellowship, University of Louisville
                  2009-2011
                  University Outstanding Student Scholarship, Dalian University of
                  Technology
                  2002-2005

Work Experience:
                  Teaching Assistant
                  Industrial Engineering, University of Louisville, 2011-Represent

Professional Societies and Memberships:
                  Institute for Operations Research and the Management Sciences
                  Institute of Industrial Engineers
                  Vice President of INFORMS Student Chapter, University of
                  Louisville

Presentations: Shengpeng Jin, Suraj M. Alexander, "Early Detection and Control
                  of Potential Pandemics", INFORMS 2010 Annual Conference,
                  November 7-10, 2010, Austin, Texas.

54

Shengpeng Jin, Suraj M. Alexander, "Early Detection and Control of Potential Pandemics", IIE 2011 Annual Conference, May 21-25, 2011, Reno, Nevada.