Lappeenranta University of Technology
School of Engineering Science
Computational Engineering and Technical Physics
Technomathematics

Eddymurphy Ugochukwu Akwiwu

DISCRETE MODELING AND ANALYSIS OF CONTACT NETWORKS IN EPIDEMIC MODELS

Master's Thesis

Examiners: Professor Heikki Haario
Professor Matti Heiliö

Supervisor: Professor Heikki Haario

ABSTRACT

Lappeenranta University of Technology
School of Engineering Science
Computational Engineering and Technical Physics
Technomathematics

Eddymurphy Ugochukwu Akwiwu

DISCRETE MODELING AND ANALYSIS OF CONTACT NETWORKS IN EPIDEMIC MODELS

Master's Thesis

2018

45 pages, 8 figures.

Examiners: Professor Heikki Haario
Professor Matti Heiliö

Keywords: Contact network; SEIR; Discrete modeling; Calendar method; Graph theory

The use of compartmental models has been a dominant framework for modeling infectious diseases. These models assume that the underlying population is fully mixed. Studies have shown that in reality, the rate at which an infectious disease disease propagates in a population also depends on the pattern of contacts individuals make. Thus an alternative way of modeling which captures the heterogeneous nature of the population is the use of contact network models. Several contact network models have been proposed; however in this work we propose a simple and straightforward method of creating a contact network, where schedules in form of a calendar is created for each individual in the population. The contact pattern in the network model is limited or/and randomized in order to explore its effect on how the disease spreads. The results are compared to the mass-assumption model which shows an agreement with the randomized-mixing assumption of the traditional compartmental model, while the limitation in the contact pattern in the network model shows the effect of restricting movement in the spread of the disease.

PREFACE

I desire to thank the Almighty God, in whose air I take and have my being. For his mercies, love, provision and sustenance, for bringing me this far. I know, this height would have been unattainable devoid of the providence which can only come from the most high. I stand in awe and grateful for everything Lord! I also want to thank my astonishing supervisor, Professor Heikki Haario for your brand tutelage and advice, I want you to know that you helped to form me, and I am appreciative to you. I am also exceedingly thankful and indebted to Vladimir Shemyakin, for being considerate and helpful throughout this journey.

I appreciate African Institute for Mathematical Sciences for according me the chance to gain scholarship at Lappeenranta University of Technology and I am so enchanted that I didn't disappoint. To my well-read teachers at Lappeenranta University of Technology, with a heart filled with gratitude, I say thank you all, for your listening ears, corrections, advice and support. To my close friends , thanks for all the striking reminiscences we shared.

Finally, I convey my very thoughtful appreciation to my inimitable parents Mr. & Mrs. Anthony Akwiwu, my dearest siblings Anselm, Henry and Queen, for providing me with consistent support and unremitting back-up all through this episode of study at Lappeenranta University of Technology. I want you all to know that even in thorny moments, your prayers reserved me.

Lappeenranta, June 25, 2018

Eddymurphy Ugochukwu Akwiwu

CONTENTS

List of Figures

ABBREVIATIONS AND SYMBOLS

S	Susceptible individuals
E	Exposed individuals
I	Infected individuals
R	Recovered individuals
I	Infected individuals
pgf	Probability generating function
$\boldsymbol{A}$	Adjacency matrix
E-R	Erdös Rényi Model Networks
β	Transmission rate
α	Progression rate
γ	Recovery rate
k	Degree
P_k	Degree distribution
$\bar{k}$ or $\langle k \rangle$	Average degree
T	Average transmissiblity
T_c	Epidemic threshold
R_0	Basic reproduction number

1 Introduction and Background

1.1 Introduction

Epidemiology is the study (scientific, systematic, data-driven) of the distribution (frequency, pattern) and determinants (causes, risk factors that precipitate disease) of health-related states and events in a specified population and its application in controlling health-related problems [2]. Although an epidemiologist and a clinician deal with the existence and control of diseases, their views on "the patient" differ greatly. While an epidemiologist is concerned about the general well-being of individuals in a community or population, the clinician is concerned about the well-being of a single individual [2].

According to [3] and [4], the continuous problems created by infectious diseases gave rise to the need for rigorous study which entails the following:

- description of complex data to facilitate the dissemination of results.
- demonstration of general laws governing the dynamics of epidemics.
- estimation of the parameters that cannot be measured directly.
- prediction or forecasting of future disease burdens.

1.2 Modelling Methodologies in Mathematical Epidemiology

Answering specific questions to a problem forms the basis of model building. The type of models built depends on the availability and quality of data to be used, the background of the modelers, and the epidemiology of the disease to be modeled [5]. [5] stated that determining how complexity a model should be, is an art as well as a science. This is because including extra parameters or state variables may increase the complexity of the model without necessarily enhancing the quality of the model response. Also, ignoring some factors that are clearly not far from reality in the study of such disease could give rise to results that are misinforming [6].

In order to ensure that the epidemiological model resembles the system it represents , a crucial step in developing such model is model verification and validation [7]. In his com-

prehensive review, [7] defines model verification as the process of assessing the model's logic, simulation code, and that it works as expected.

Model validation ensures that such model resembles reality. There are no standard procedures to follow when validating models, however [7] provides the following guidelines:

- models should make biological sense.
- models should imitate real life.
- models should be fit for the use they are designed for.
- models should not be overly sensitive to the influence of uncertain parameters.

[1] state that mathematical modelling of infectious disease dynamics can be broadly classified into three categories as shown in Figure 1. For the purpose of this work, we shall focus on mathematical-mechanistic models, specifically complex network models to forecast the evolution of epidemic spread.

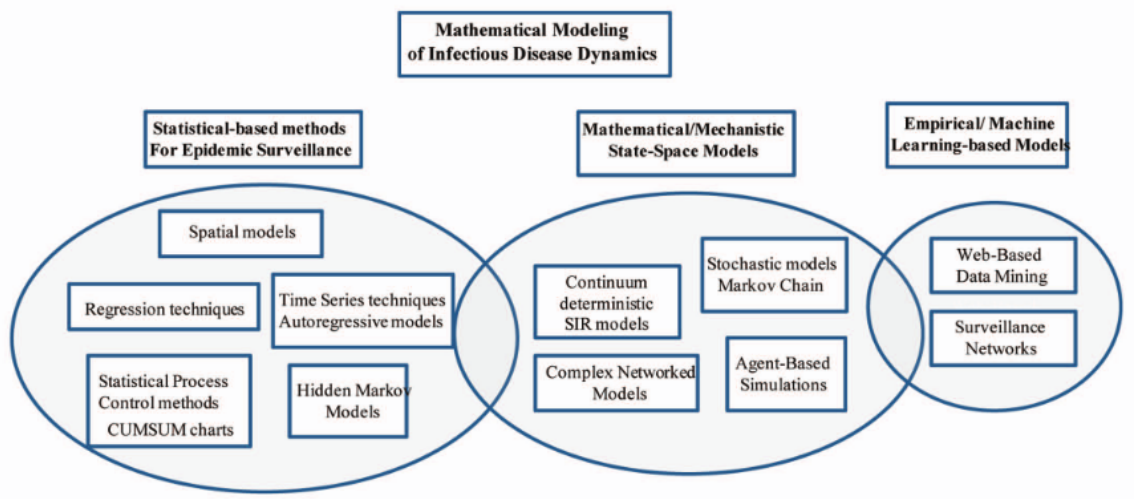


Figure 1. An overview of mathematical models for infectious diseases culled from [1]. It is broadly grouped into three categories: (i) statistical methods such as regression methods, time series analysis, spatial models e.t.c (ii) mathematical state-space models such as stochastic models, complex network models e.t.c (iii) machine learning-base models

1.3 Continuous Deterministic Models

The starting point in modeling an infectious disease is using continuous deterministic models also known as deterministic compartmental models [8] where individuals are grouped into small number of compartments based on their disease status [9, 10, 11, 12]. Deterministic in this context means that the model predictions are based entirely on the following: initial conditions, underlying system of equations, and the unknown parameter values [13], while continuous implies that effective number of contacts per unit time is continually changing [11]. Choosing which of compartments to include in a model is based on the attribute and nature of the disease under study and the reason for which the disease is being modeled [14]. The population under study is divided into disease classes. These classes according to [15] include:

- **S**usceptible: These are individuals who are not infected but can get infected.
- **E**xposed: These are infected individuals that are not yet capable of transmitting the disease to others.
- **I**nfected: These are infected individuals who may infect susceptible individuals given contact.
- **R**emoved/**R**ecovered: These are individuals who are immune to the disease or otherwise isolated from possible infection.

1.3.1 Susceptible-Exposed-Infectious-Recovered (SEIR) Model

Some diseases that are infectious like influenza are best modeled using an SEIR model [14]. SEIR models are preferred over other compartmental models since they take care of the exposed class. In this type of model, individuals move to the next compartment after a waiting time period. By waiting time period we mean either the latency period, infectious phase or recovering period [16]. Hence throughout this work, unless otherwise stated, we will refer to waiting time as either of the three periods or phase mentioned above. This type of compartmental model is used for our analysis in this thesis. The flowchart describing this type of model is shown in Figure 2 and the model's system of equations

is:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta\frac{SI}{N}, \\
\frac{dE}{dt} &= \beta\frac{SI}{N} - \alpha E, \\
\frac{dI}{dt} &= \alpha E - \gamma I, \\
\frac{dR}{dt} &= \gamma I,
\end{aligned} \tag{1}$$

where β is the transmission rate, α is the progression rate with $\frac{1}{\alpha}$ being the mean latent period and γ is the recovery rate [17].

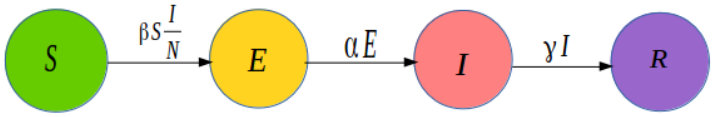


Figure 2. Flowchart of a simple SEIR model

1.4 Limitations of Continuous Deterministic Models

Continuous deterministic models have been long be useful in the field of mathematics of infectious disease, however they do not properly account for some certain realistic aspects in the spread of infectious disease [18]. The propagation of an infectious disease in a population is largely due to the network formed when there is an encounter between a susceptible and an infectious person [19]. Some of the limitations of this type of model are based of the following model assumptions:

(i) **Exponentially Distributed Waiting Times**
The model assumes that the time taken for a host to move to the next health state is exponentially distributed [20]. Exponentially distributed in the sense that the probability that such host moves to the next compartment during a given time interval is proportional to the duration of the given time period. According to [21], this model assumes that a host becomes infectious upon having contact with an infectious person and that the chances of an individual recovering from infection does not depend on the waiting time between the period of infection and recovery. According to [20] these two assumptions are unrealistic.

(ii) **Homogeneous and Well-Mixed Population**

The model assumes that the population is homogeneous in nature and completely mixed [16,18,19,20,21,22]. This means that every host has equal but small chances of contacting and spreading the disease to other hosts in the group of population; thus the *mass-action assumption* [10,19,22,23]. According to [22], this assumption was made specifically to enable a modeler to write the model as differential equations. But in reality, this is not always the case since the spread of an infectious diseases differ in age groups, latent periods according to an individuals immune system or behavior, rates of transmission depending on the nature of contact since the rate of spread of some infectious diseases is higher during contacts at home than contacts at work for example [18]. Also for a population with large number of individuals, each person associates only with a few people with whom they can possibly transmit disease to [24].

(iii) **Large Population Size**

According to [21], the form of this model which is derived as a system of differential equation, for example in Equation (1) is determined by the assumption that the population size is very large or infinite in size strictly speaking. However if the underlying population size is small, modeling using this form of model as differential equations becomes problematic since stochastic effects which are crucial in the analysis are not achieved [21, 25].

1.5 Objectives of the Study

The nature of contacts made by hosts in a given population is an crucial factor in determining how the infectious disease spreads in such a population [26]. Furthermore, the existence of individuals with contacts or connections larger than the average number of encounters, the number of times and duration of these contacts made and most importantly, the topological structure of the contact network formed within the population all have major effects on how the disease propagates in the entire population [26, 27, 28]. Several statistical models incorporating contact networks have been proposed in previous studies in infectious disease modeling [11, 12, 13, 29, 30]. However in this work a different method of generating contact network is being proposed where schedules in form of a calendar is created for each individual. A disease is likely to be transmitted with some probability when an infectious person makes an encounter with a susceptible person as a result of having a common meeting point from their different schedules. Therefore our objectives in this work are:

- to introduce and understand the concept of contact network in modeling infectious diseases.
- to explore the impact and limitations of different network structures using our proposed method on the epidemic.

1.6 Structure of the Thesis

The rest of this thesis is organized as follows: Chapter 2 presents two discrete modeling approaches, concepts from graph theory which forms the basis of contact network, and our proposed calendar method of generating contact networks which forms the main methodology of this thesis. Next is Chapter 3 which entails the implementation of our proposed method using numerical simulations and the analysis of the results obtained. Finally Chapter 4 concludes the thesis by discussing its findings and outlines directions for future research.

2 CONTACT NETWORK MODELS

Our main focus in this chapter is contact network models which forms the main research methodology of this thesis. However, we shall take a look at a known discrete modeling algorithm, the Gillespie algorithm and another discrete modeling approach informally called as Euler's method. Since observations collected for individuals in different compartments are often discrete-time points (such as daily, weekly, monthly or annual incidence rates depending on the disease), discrete-time models are more suitable [8, 31]. Also since the change in health status of an individual is relatively proportional to the waiting times, a finite difference algorithm; that is discrete-time modeling would be appropriate [8].

2.1 Discrete Modeling: Gillespie Algorithm

An easy and efficient method of simulating stochastic models or Markovian processes was proposed by [32] where state variables changes are simulated by what now is called the Gillespie algorithm. One assumption of this algorithm which makes it particularly efficient is that all possible transitions from one compartment to another occur independently at each time step with given probability which depends on the number of individual in each compartment [33]. One disadvantage of this algorithm is that it is difficult to apply when the underlying process is non-Markovian with inter-event times which are not exponentially distributed [34]. According to [33, 34, 35], the idea behind Gillespie's algorithm is given as follows:

- The rates of all likely events happening are computed together with the total rate for a change to occur. Epidemiologically speaking, this implies finding the rate at which a susceptible host is likely to be infected and the recovery rates of each infected host.
- Using this computed rate of change, sample the waiting time before the next event happens as an exponential random variable with the calculated total rate of change as the rate of the distribution.
- Sample which event occurs, where the probability of the event happening is proportional to the rates of each event.
- Update and repeat the process.

2.2 Discrete Modeling: Euler's Method

An alternative approach to the event-based approach described in Section 2.1 is using the Euler's discretization method; a time-based approach [36]. This method is an improvement to the Gillespie algorithm in cases of slow steps in finding which host gets infected [34].

One of the assumptions of compartmental models is that the system is memoryless [8]. That is, there is no difference between individuals that have spent significant amount of time periods in a particular compartment compared to those individuals that are just included in any of the compartments. However with this method, instead of calculating the time for the next event to happen selected from all possible events like the case of Gillespie algorithm, we can calculate when an infected host would transmit the infectious disease to others and when the time of recovery would be; all depending on the waiting times to move to the next compartment [34]. This implies that the number of days spent by an individual in a particular compartment can be tracked, the progression of the underlying infectious disease through each persons in the population can be properly monitored [37].

Often, using this discrete-time modeling is more faster than the Gillespie algorithm [34] and hence is used in implementing our proposed calendar method of generating contact networks. Using this method analytically involves transforming the system of differential equations to difference equations. For more detailed example, see [36].

2.3 Contact Network Structure: Concepts from Graph Theory

In a contact network model, the interactions between individuals are studied. Unlike in compartmental models where a susceptible host can come in contact with any host in the population, the contact network model assumes that each host has a defined number of persons whom they come in contact with. This type of model describes the natural way in which members of a population interact, which is true in reality [29].

The idea of contact network models can be related to the concepts in graph theory. Each individual or group of individuals in a host population represents the *node* or *vertex* while the contacts or interactions among them is seen as the *edges* or *links* as shown in Figure 3 [10, 18, 29, 38].

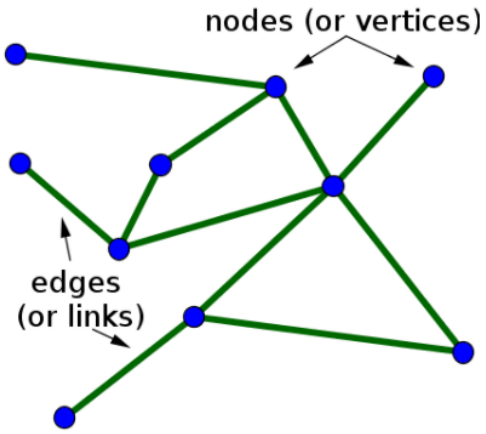


Figure 3. A simple network showing the nodes and edges connecting them.

An important feature of contact network models is the *degree distribution* P_k. The number of connections (edges) a particular node has is called the *degree* k of the node which signifies the number encounters which could result to a possible transmission of infection as shown in Figure 4 [10, 18, 34]. Clearly from Figure 4, the higher the number of links a vertex has, the higher chances that it is a neighbor of a vertex that is already infected.

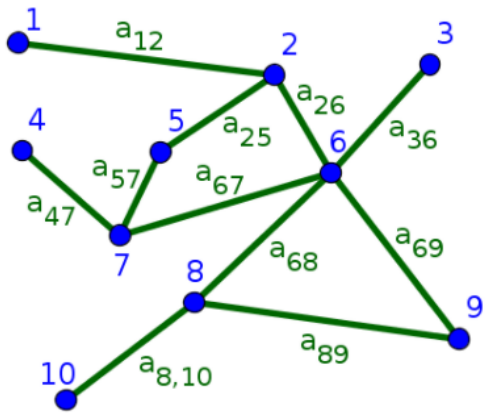


Figure 4. Here the degree for nodes 1 to 10 are $1, 3, 1, 1, 2, 5, 3, 3, 2$ and 1 receptively.

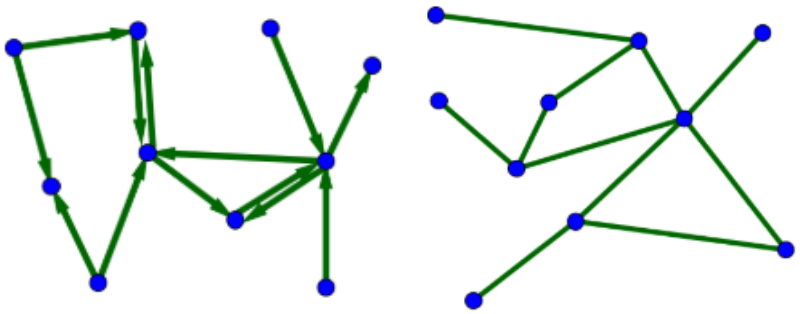
Hence the distribution of the number of such encounters within a given population is known as P_k [12, 18, 34]. According to [10, 39] P_k refers to the set of probabilities that a randomly chosen node has degree k. Also by counting the number of vertices each degree has, we can define P_k as the fraction of vertices in the network with degree k [40]. According to [28], plotting the degree distribution is one important way of showing the useful characteristics of a contact network (see Figure 9). P_k follows a Poisson law, implying that the number of vertices having degree k has an exponential decay between k

and the average degree $\bar{k}$ or $\langle\ k\rangle$ given by $\sum_{k=1}^{\infty} kP_k$ [39]. The variance of the degree of distribution can be used to determine if a network is homogeneous or heterogeneous [29]. If the variance calculated as $\sigma^2 = \sum(k - \bar{k})^2 P_k$ is zero then every person has the same number of contacts in which case we have a homogeneous network–same as deterministic compartmental models described in Section 1.3; else the network is said to be heterogeneous [29].

It is imperative to mention here that an encounter between any two nodes; an infected host and a susceptible host does not necessarily imply that there is transmission of infectious disease. Hence a network includes both edges that result to transmission of infectious disease and those edges that do not. At any particular time period, the chances of a node changing its health state is assumed to depend on its present health condition and that of its immediate neighbors [34]. For instance, a node which is susceptible has chances of getting infected if the surrounding number of nodes are infectious [34]. However, the rate of recovery of an infected node does not depend on its neighbor but on the infected node itself [34].

The structure of a network to be used depends both on the structure of population under study and the nature of the infectious disease itself [29]. Even within the same population, the type of network to be used would be different for diseases contaminated through unprotected sex like HIV or mere contact [29]. In addition to the latter, there might be difference between infections that occur as a result of lengthen close encounter in order for disease to spread and ones in which only a short encounter is enough to transmit the disease [29].

A network can either be directed or undirected as shown in Figure 5 depending on the mode of transmission.



(a) A directed network. (b) An undirected network.

Figure 5. Types of networks

A network is said to be *undirected* (see Figure 5b) if the transmission of the infectious disease can occur in either direction between two nodes in contact (meaning that transmission is symmetric) [18]. However, there are cases where an individual gets infected but the converse is not true [18]. An example is transmission through blood transfusion where infection happens only in one direction. The type of network describing such a situation is called a *directed network* as seen in Figure 5a [11]. The connections within a population can be described using the *adjacency matrix*, **A** [11]. Therefore we define $\boldsymbol{A}_{ij} = 1$ if there is a connection with possible transmission of infection from person i to person j else $\boldsymbol{A}_{ij} = 0$ [11]. The adjacency matrix $\boldsymbol{A}$ gives a summary of the possible connections within a network [11]. For example in Figure 6, the adjacency matrix $\boldsymbol{A}$ with a total number of 100 entries out of which 13 are ones (13 edges in the graph) , is given as

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

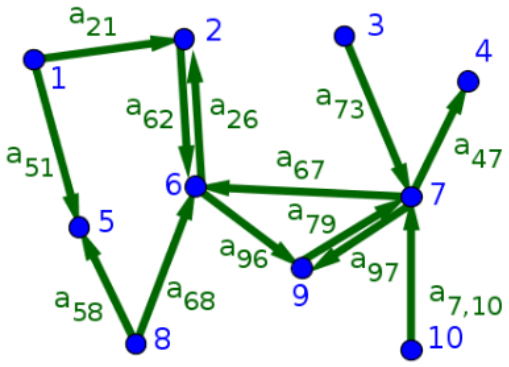


Figure 6. A directed graph with 13 edges and 10 nodes.

Epidemiologically, a network is said to be connected if the disease outbreak is able to spread through the entire population from any starting point by means of contact with infected persons [11, 29].

In the next section we shall briefly discuss methods for gathering data to describe a contact network since such data include details about contact among hosts in a population which can be a laborious task.

2.4 Methods for Sampling Contact Network Data

Gathering information in order to determine or build a complete random network could be a herculean task since it involves knowing every individual in a population and their interaction between others whom they contact [11, 38]. Moreover, issues surrounding individuals privacy in the case of sexual contacts, misreporting and error during sampling makes the entire data collection process more tedious [38]. According to [38], methods for collecting data in building contact network in the absence of an infectious disease can generally be grouped as either direct or indirect.

Direct methods such as diary-based studies or device-based studies which are primarily used for humans only, involves gathering information on the nature of the contact behavior of hosts in a population [38]. In dairy-based studies, individuals from a selected population sample record their contacts as they occur or immediately after they occur [11, 38]. Thus the idea of individuals not being able to recall or recount their contacts is avoided [11]. According to [11], this technique allows a larger number of persons in the population under study to be sampled extensively.

However this technique of data collection comes with a challenge. The collection of information as regards to the contact behavior of individuals relies solely on the individuals in question [11]; hence individuals in the population may have different definition of what a close contact means [11]. According to [38], sexual intercourse has been defined to be close contact in a sexually-transmitted disease like gonorrhea while an in-person two-way conversation exceeding two words has been defined to be close contact for infectious diseases spread through respiratory droplets or fluid.

Device-based study techniques of data collection depend on the usage of electronic recording devices which is used to measure the closeness between individuals most especially for infectious diseases spread through respiratory droplets or fluid [38]. Mobile phones and animal ear tags are examples of materials used in device-based studies [38].

Other techniques of data collection which are indirect involve using information obtained from general human or animal behaviors [38]. Examples include data collected on number of persons attending a school, census data and so on. Although this technique is less direct on humans, they serve as a guide in knowing the behavior of individuals in a population [38].

2.5 Idealized Contact Networks

Several forms of well-studied families of computer-simulated contact networks have been used extensively in the field of epidemiology to study disease spread and their implications for epidemic spread. They are either defined in terms of how connections are formed or how hosts are distributed either geographically or socially [11]. Here we briefly discuss some of these types of networks as follows:

(i) **Small-world Networks**
Small-world networks described in [41] are networks defined by high level of clustering (also known as transitivity, cliquishness or mutuality) and short path lengths [11] as shown in Figure 7. By clustering we mean that the neighbors or friends of any given vertex are likely to be neighbors or friends of each other [11, 12, 18], and by short path lengths, we mean that large number of randomly selected node can be reached in small number of steps [11, 28]. [10] referred to this type of network as *six degrees of separation*. For instance, six degrees of separation of two persons simply means that any two persons located anywhere in the universal can be linked through a chain of almost six acquaintances [10].

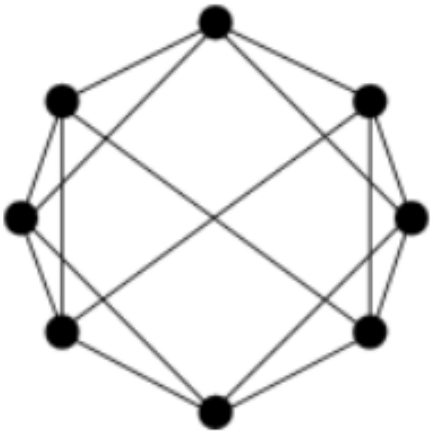


Figure 7. An Example of a small-world network with 8 nodes showing high level of clustering and short path length.

According to [11, 28], disease spread in this type of network with small region implies that it takes a short while for everyone in the population to be infected since the spread is rapid and the chances of containing the disease is small.

(ii) **Scale-free Networks**

Scale-free networks are characterized by few nodes with highly connected nodes and many nodes with small number of contacts as seen in Figure 8 [11, 12].

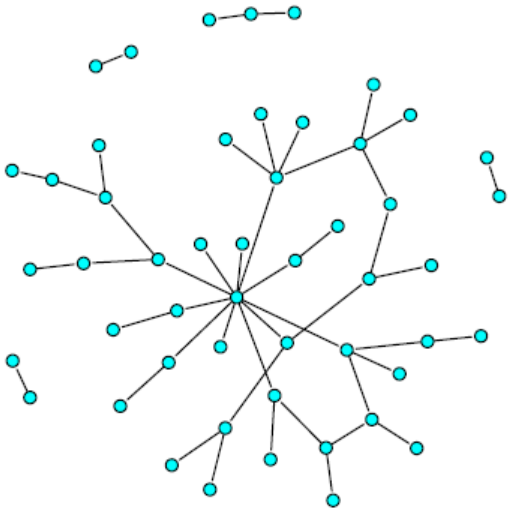


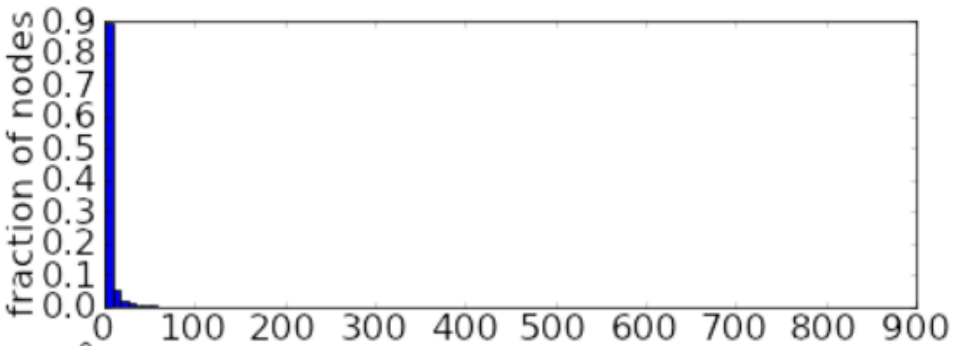
Figure 8. An Example of a scale-free network with hubs having high level of connectivity.

Nodes with highly connected nodes are called *hubs*. In epidemiological terms, they are called *super-spreaders* [10, 11]. This type of network is assumed to follow

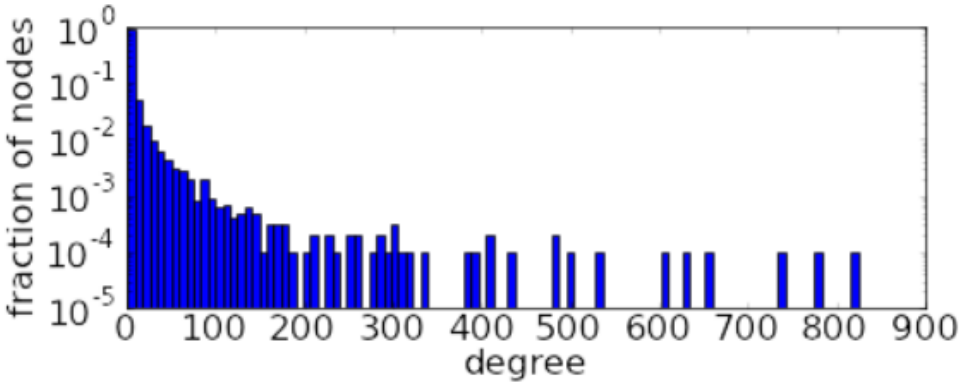
power-law degree distribution given as

$$P_k \sim k^{-x},$$

where x is some exponent in the interval $2 < x < 3$ as shown in Figure 9 [42].



(a) Power-law degree distribution on a linear scale.



(b) Power-law degree distribution on a log scale.

Figure 9. Histograms of a power law degree distribution with exponent of -2 and maximum degree of 1000, resulting in an average degree of around 7 of a network with $10,000$ nodes.

Here as the degree k increases, the degree distribution P_k decays slowly which implies that chances of finding a node with a very large degree increases.

Scale-free networks provide a means of modeling a population with high level of heterogeneity in their connectivity distribution by incorporating individuals which are highly connected to other individuals [11, 29]. This type of network can be constructed sequentially by adding new individuals one by one in a preferential manner to individuals who are already connected to large number of persons in the population [11, 29]. The connection of a new individual to an already existing individual is made at random with a probability which is proportional to the number of contacts of that existing individual [29,30]. This implies that new individuals join

the most popular people in the population and so the *rich get richer* leading to hubs with very high degree [29].

According to [10, 11, 12, 29], sexually transmitted diseases are best modeled using scale-free networks. Having too many connections like the hubs in the network has major effects on the population according to [11]. Firstly an individual with too many contacts is at greater risk of being infected and thus once infectious, he/she become infectious to many others after a certain time period [11].

(iii) **Spatial Networks**

These types of network are flexible when compared to other forms of contact networks [11]. Building a spatial network involves positioning individuals within a given area or volume who are then linked together with a certain probability which depends on their distance in form of a connection kernel [11]. In order to generate different forms of contact networks ranging from contact networks having scale-free properties to small-world structures, the distribution of the nodes or the connection kernel can be changed [11, 23]. The degree distribution of spatial networks, which is often approximated by a Poisson distribution, generally shows high degree of heterogeneity [11].

(iv) **Configuration Networks**

In this type of network, a set of contacts that each person wishes to have is linked to such a person and then these contacts are randomized and assigned to other persons, depending on the number of contacts they wish to be connected [28]. This simple form of creating network of contacts has an advantage that no short forms of loops are created within the network since it is basically created from individuals that are randomly connected [28].

(v) **Erdös Rèyni (E-R) Model Networks**

In this type of network, an edge connecting two given nodes is randomly created with probability $p = \frac{2M}{N(N-1)}$ independent of all other $\frac{N(N-1)}{2}$ possible edges, where N and M denote the needed number of nodes and edges respectively of the actual network to be simulated [27, 29, 30]. In order to produce a given number of simulated networks with the number of edges on the increase, the probability p can be varied [30].

More details of other forms of contact networks such as lattices and exponential random graph networks can be found in [11, 12, 13, 29, 30].

2.6 Predicting Disease Spread in Contact Network Models

One of the steps in contact network modeling according to [12] is to mathematically predict the disease dynamics in the population using the attributes of the disease and the topological properties of the network.

Suppose that there is a disease outbreak, first from a random vertex in a contact network; where such a vertex is called *patient zero* in epidemiological terms [12]. The disease spreads through the network in the same manner as the compartmental models with the exception that the structure of the contact network replaces the random-mixing contact patterns of the compartmental models, which assumes a Poisson degree distribution [12, 18]. The initially infected nodes would remain exposed after a given waiting time period and then become infectious, during which time there is a tendency to transmit the disease to its neighbor [12, 18]. This process continues for all infected nodes. This spread of disease across the network looks similar to bond percolation from the branch of statistical physics which describes, for example, how a liquid flows through a porous media [43]. Generally speaking, percolation theory can be used to describe the connectivity in random contact networks, and hence the theory can be well applied to forecast the size or number of nodes reached along the links in a contact network during the infection disease spread [18]. This idea was first suggested in [44] and [22]. It has recently been extended into a more flexible approach in modeling infectious diseases by [19, 45, 46].

According to [12], how an infectious disease spread a contact network relies on its structure and the infection rate. Following the works of [22], to model the rate of infection, every contact or edge in a contact network is given a per unit time probability of infectious disease transmission p_{ij}; that is the chances that if a node i is infected, it will transmit the infectious disease to node j during its period of infection. Therefore if we assume discrete time steps (say in an interval of one hour), if node i is infectious for h_i time steps, then the probability that a susceptible node j will be infected when there is an encounter as shown in [12, 22] is given as

$$T_{ij} = 1 - (1 - p_{ij})^{h_i} . \tag{2}$$

However for continuous time, the probability $1 - T_{ij}$ that node j would not be infected by

node i as shown in [12,22] is given as

$$1 - T_{ij} = \lim_{\delta t \to 0} (1 - p_{ij}\delta t)^{h_i/\delta t}, \tag{3}$$

$$1 - T_{ij} = e^{-p_{ij}h_i} \tag{4}$$

$$T_{ij} = 1 - e^{-p_{ij}h_i}. \tag{5}$$

Suppose we assume that p_{ij} given in Equations (2) and (5) is independent and identically distributed (iid) random variable chosen from a certain probability distribution $\varphi(p)$, then T_{ij} is also an iid random variable [12,22]. Furthermore, the quantity p_{ij} gives a summary of the main features of infectious disease transmission in addition to the likelihood that a disease would be transmitted if there is contact between a susceptible node and an infectious node [12,18]. Therefore the chances that a disease would spread is conditioned entirely on the mean probability of transmission (better known as *average transmissiblity*) between two hosts, given by

$$T = \langle T_{ij} \rangle = 1 - \int_0^\infty \phi(p)dp, \tag{6}$$

where $\phi(p) = 1 - \varphi(p)(1-p)^h$ and $\phi(p) = 1 - \varphi(p)e^{-ph}$ for discrete and continuous time, respectively [12].

2.7 Predicting The Fate of an Outbreak in Contact Network Models

The concept of probability generating functions (pgf) from probability theory can be used to predict the fate of an infectious disease outbreak and also give useful summaries about the structure of the underlying contact network [12, 18, 19]. The pgf for the degree distribution P_k with degree k of a contact network is defined as [12, 19, 22]

$$G_0(x) = \sum_{k=1}^{\infty} P_k x^k. \tag{7}$$

From Equation (7), the derivative at $x = 1$, $G_0'(1)$ gives the average degree $\langle k \rangle$ mentioned in Section 2.3.

Suppose we choose an edge randomly and trace it to one of its nodes. The number of edges different from the randomly chosen edge but still connected to that particular node is termed *excess degree* of the node or vertex [12,19]. Specifically, the chances of arriving

at a vertex with degree k (thus with excess degree of $k-1$) is directly proportional to k [12, 19]. Therefore we can define $\frac{kP_k}{\langle k \rangle}$ as the probability that a randomly chosen edge with a vertex at its end has excess degree $k-1$ [12, 19]. With this, our pgf for an excess of a vertex or node as [12, 19]

$$G_1(x) = \frac{\sum_{k=1}^{\infty} kP_k x^{k-1}}{\sum_{k=1}^{\infty} kP_k}, \tag{8}$$

and upon finding the derivative of Equation (8) at $x = 1$ we obtain our average excess degree as

$$\langle k_e \rangle = \frac{\sum_{k=1}^{\infty} k(k-1)P_k}{\sum_{k=1}^{\infty} kP_k} = \frac{\langle k^2 \rangle}{\langle k \rangle} - 1. \tag{9}$$

Usually an infectious disease will spread through some but not all of the links or edges in a contact network when an infectious disease is introduced into the network according to the average transmissiblity given in Equation (6) [12]. Those edges that were infected as a result of the spread of the infectious disease are termed *occupied* [12]. Therefore the pgf for the occupied is given as [12, 18]

$$G_0(x;T) = G_0(1 + (x-1)T) \tag{10}$$

Similarly, the pgf for the excess occupied is given as [12, 18]

$$G_1(x;T) = G_1(1 + (x-1)T) \tag{11}$$

Generally speaking, the percolation theory gives a description of the average behavior of clusters connected together in a random graph which can be applied to forecast the size of infected group of nodes [12]. For a fixed network, the critical transmissiblity which is the minimum transmissiblity (T) required for an infectious disease outbreak to be called small

outbreak or large-scale epidemic is given by the *epidemic threshold* T_c as [12, 18, 19, 22]

$$T_c = \frac{1}{G_1'(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \tag{12}$$

According to [12, 19], the well-known traditional *basic reproduction number* R_0 is related to the average transmissiblity T in Equation (6) by

$$R_0 = TG_1'(1) = \frac{T}{T_c} = T\left(\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}\right). \tag{13}$$

Thus from Equation (13), when $R_0 = 1$, then $T = T_c$. Similar to the interpretation of R_0, recall from Equation (12) that T_c is the critical transmissiblity value above which the underlying population is liable to large scale epidemic (although this is not a guaranteed for the epidemic to happen) and below which will result to a small outbreak [12, 18, 19, 22].

Another quantity that is of epidemiological importance is the average size $\langle s \rangle$ of an outbreak beginning from a vertex chosen at random, given by [12, 18, 19, 22]

$$\langle s \rangle = 1 + \frac{G_0'(1)}{1 - TG_1'(1)} = 1 + \frac{T\langle s \rangle}{1 - T\left(\frac{\langle k^2 \rangle}{\langle k \rangle - 1}\right)}. \tag{14}$$

From Equation (14), $\langle s \rangle$ is valid if $TG_1'(1) < 1$ [12, 18, 19, 22].

Predicting the probability and size of a full-blown epidemic is feasible when the transmissiblity of an infectious disease is above the epidemic threshold T_c. Thus Equation (14) becomes no longer applicable in determining the size of the resulting full-scale epidemic [12, 18, 19, 22]. Therefore the probability of a large-scale epidemic S is given by

$$S = 1 - \sum_{k=1}^{\infty} P_k (1 + (u-1)T)^k, \tag{15}$$

where the sum term gives the likelihood that an outbreak would erupt from a single infection instead of a full-scale epidemic and u is the probability that an individual connected at the end of the edge is not infected with the disease in question [12, 19, 22]. This probability is obtained numerically by finding the solution (that is, the root) to the equation

$$u = \frac{\sum_{k=1}^{\infty} k P_k (1 + (u-1)T)^{k-1}}{\sum_{k=1}^{\infty} k P_k}. \tag{16}$$

Other relevant epidemiological quantities which has been extended by [12, 19] from the results proposed [22] have been derived. One of such is the probability that an initially infected person known as *patient zero* with degree k being the number of k contacts will spark an epidemic is given by [12, 19]

$$\varepsilon_k = 1 - (1 - T + Tu)^k, \tag{17}$$

while the probability that a group of infected nodes with N cases will spark an epidemic on a large-scale is given as [12, 19]

$$1 - \prod_{i=1}^{N}(1 - \varepsilon_{k_i}), \tag{18}$$

where k_i is the degree of each host i. Equation (18) is simply the probability that an epidemic will not be sparked by N hosts that are infected with the disease [12, 19].

Also, the probability that none of contacts of a specific infected host will result to an epidemic is given by [12, 19]

$$\frac{\sum_{k=1}^{\infty} kP_k(1 - T + Tu)^{k-1}}{\sum_{k=1}^{\infty} kP_k}, \tag{19}$$

while the probability that an outbreak of size N would result to an epidemic is give as [12, 19]

$$1 - \left(\frac{\sum_{k=1}^{\infty} kP_k(1 - T + Tu)^{k-1}}{\sum_{k=1}^{\infty} kP_k} \right)^N \tag{20}$$

Finally according to [12, 19] the probability that a random chosen host with degree k will be infected due to the full-blown epidemic is given as

$$v_k = \varepsilon_k = 1 - (1 - T + Tu)^k. \tag{21}$$

Equation (21) is simply the risk of a host being infected as a result of an outbreak which is a function of the host's number of contacts k.

2.8 Network Formulation: Calendar Method

So far we have mentioned some well-studied types of contact networks that are designed to mimic real contact networks in the study of infectious diseases. They all focus on some of the characteristics of a real contact network such as the level of connectivity, path lengths, clustering, heterogeneity, and so on. However in this work, we propose a simple and straightforward method of describing the network of contacts for a given population implemented using MATLAB. In this method, schedules in form of calendar is created for each person in the given population. A disease is likely to be transmitted with probability (see Equation (22)) when an infectious person comes in encounter with a susceptible person. This can happen depending on the probability of infection at each encounter and duration of contact as a result of having one or more common meeting points (such as train station, school or shopping mall) from their different schedules within the population or between sub-populations. The network generated is limited and/or randomized in order to vary the structure of the network.

The probability P_{inf} of an individual getting infected at some other points of encounters is derived by first calculating the probability of an individual not being infected during an encounter, that is $1 - p_e$ which is replaced by $(1 - p_e)^h$ for several hours of encounter, and then subtracting this value from one:

$$P_{inf} = 1 - (1 - p_e)^h, \tag{22}$$

where p_e is probability of getting infected for each encounter and h is the hours of contact.

The following assumption are made in order to simplify the contact network generated:

(i) The underlying population size is small so as to obtain more accurate information concerning the daily activities of the individuals in the population and also due to the computational cost.

(ii) The rate of transmission of the disease when there is contact between at least two persons is the same and equal for all.

(iii) Each individual spends at least one hour in a place(s) for contact that may lead to transmission of infection. That is, at least one hour of contact is enough for a susceptible person to get infected from an infectious person.

3 NUMERICAL SIMULATIONS & RESULTS

In this chapter we present the results obtained from the numerical simulation of our proposed method of generating contact network discussed in Section 2.8 using MATLAB. Since one of our aim in this work is to explore only the effects of the contact network structure, certain quantities such as infection probability for each encounter, a total number of 10 contact places (such as schools, shopping mall, bus station, homes and so on), and time steps are all kept constant or otherwise mentioned.

First we verify our contact network generated by first calibrating the SEIR model with and without network to obtain Figure 10.

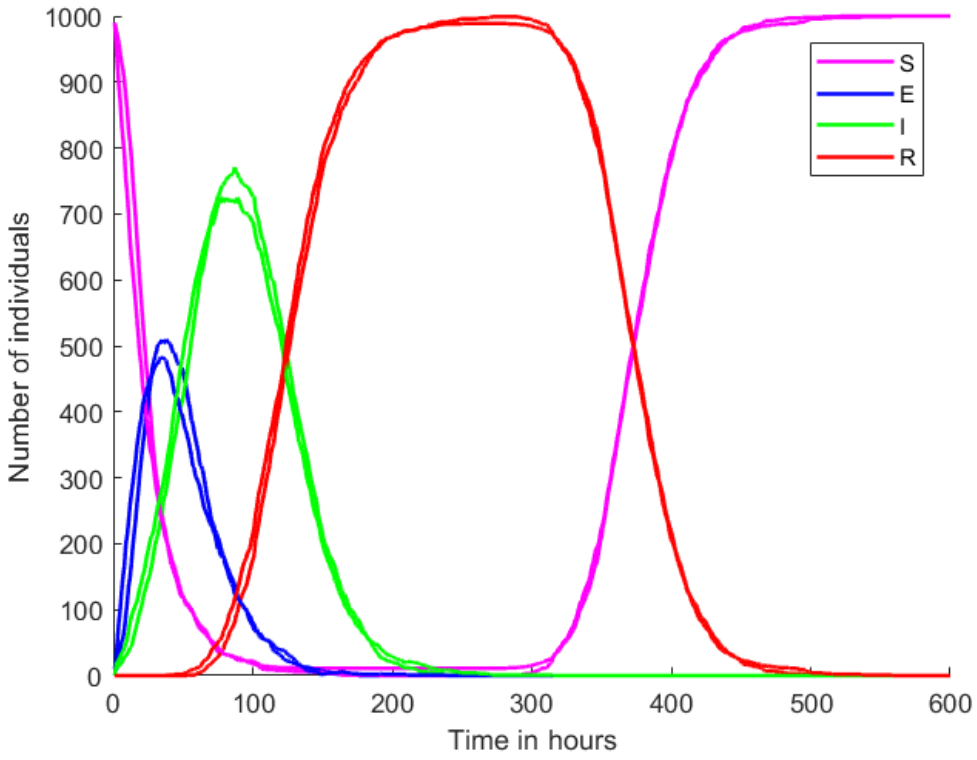


Figure 10. Comparison of the SEIR model with and without a network using discrete-time modeling approach, with initial conditions: $S(0) = 990$, $E(0) = 0$, $I(0) = 10$, $R(0) = 0$ for 600 hours.

Figure 10 shows the comparison of our SEIR model with and without network structure for a small population of 1000 individuals. 10 infectious persons where randomly chosen from the total population as the initial number and we carried out a simulation of the

epidemic for 600 hours (25 days). The probability of infection for each encounter or interaction after calibration of the model was found to be 0.0038 while the rate of infection was found to be 0.9 per hour . For the non-network model, the probability of infection was drawn from an exponential distribution with a constant rate of infection. For both cases, the waiting times before moving to the exposed, infectious and recovery health states are all drawn from normal distributions with mean values (in hours) as $28, 100,$ and 350 respectively and constant standard deviation of 26 for all. The plots for each compartment in both models (with and without network) are almost the same because for our network model, the contact pattern was randomized for every individual in the population which is one of the assumptions of the standard epidemic SEIR model mentioned in Section 1.4.

Suppose we restrict the interactions within the population by spitting the entire population of size 1000 into three subpopulations I, II, III of sizes 300, 400, and 300 respectively. With 10 initial number of infected persons chosen randomly from subpopulation I only, the situation changes as seen in Figure 11.

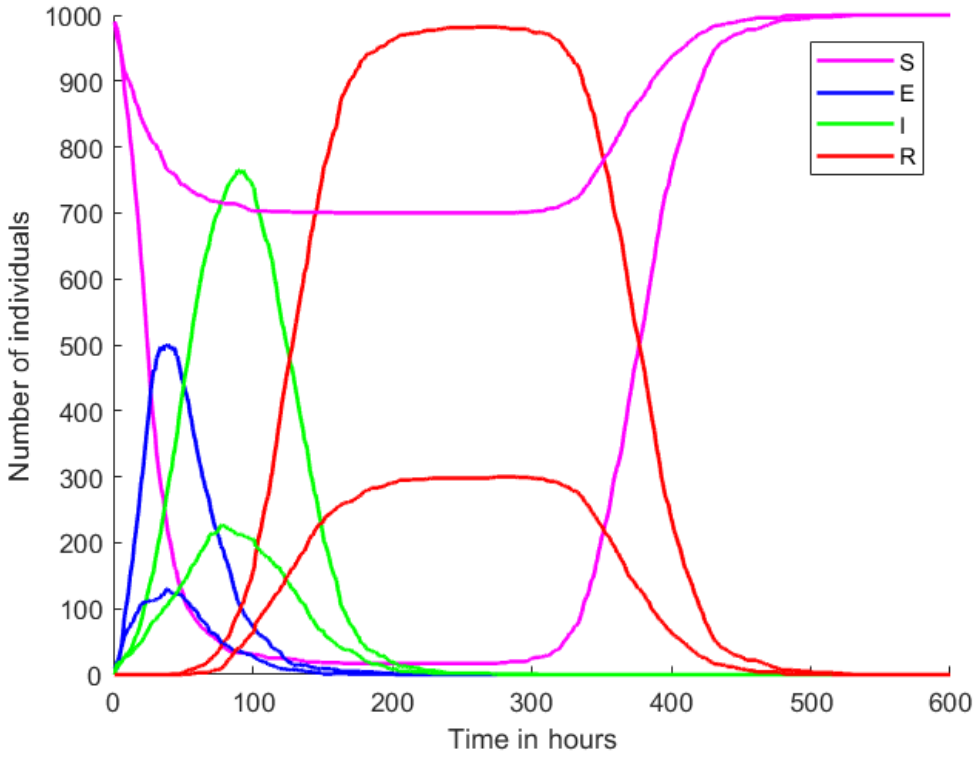


Figure 11. Contact network model with three subpopulations having no common meeting place using discrete-time modeling approach with initial number of infected persons chosen randomly from the first subpopulation only; initial conditions: $S(0) = 990$, $E(0) = 0$, $I(0) = 10$, $R(0) = 0$ for 600 hours.

Similar situation is also observed in Figure 12 when the initial number of infected persons are chosen randomly from subpopulation I and II only.

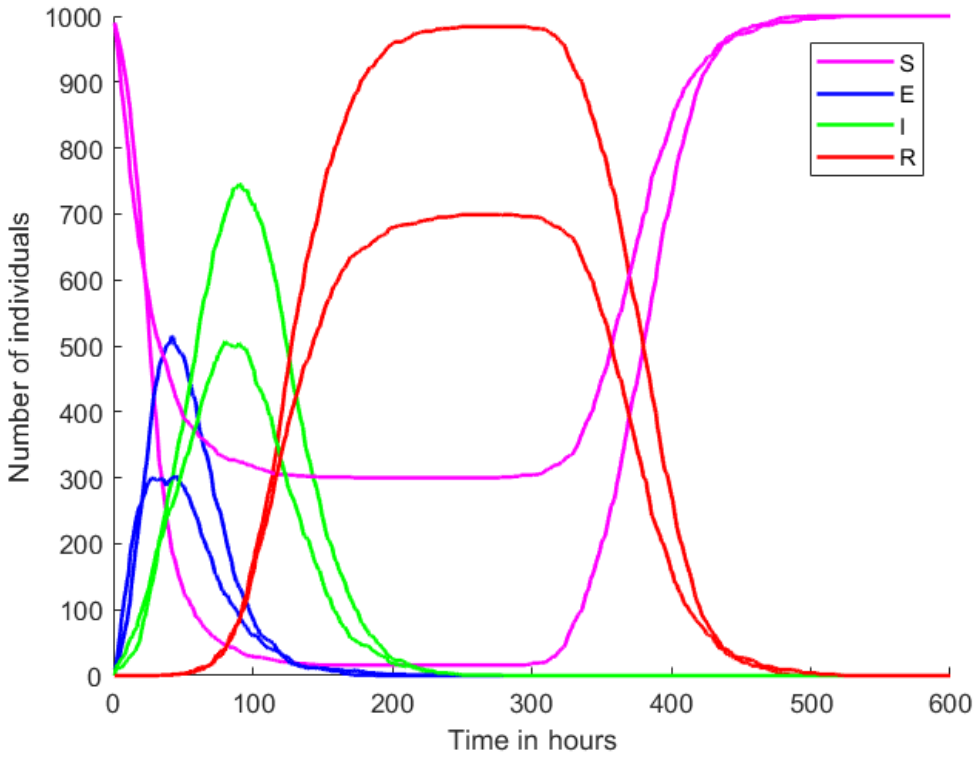


Figure 12. Contact network model with three subpopulations having no common meeting place using discrete-time modeling approach with initial number of infected persons chosen randomly from the first two subpopulations; initial conditions: $S(0) = 990$, $E(0) = 0$, $I(0) = 10$, $R(0) = 0$ for 600 hours.

In both cases as seen in Figures 11 and 12, we observe that not all the 1000 susceptible persons of the entire population were infected with the disease. This implies that a restriction in the level of interactions within the entire population has the tendency to slow and reduce the spread of the infectious disease. Also when compared to the case where the entire population is randomized, we observe a decrease in the maximum number of exposed, infected and recovered persons as a result of no interaction between the subpopulations.

However, the situation changes when we have just one infected person in subpopulation I only, with subpopulations I and II, and subpopulations II and III having just one common meeting point each as shown in Figure 13.

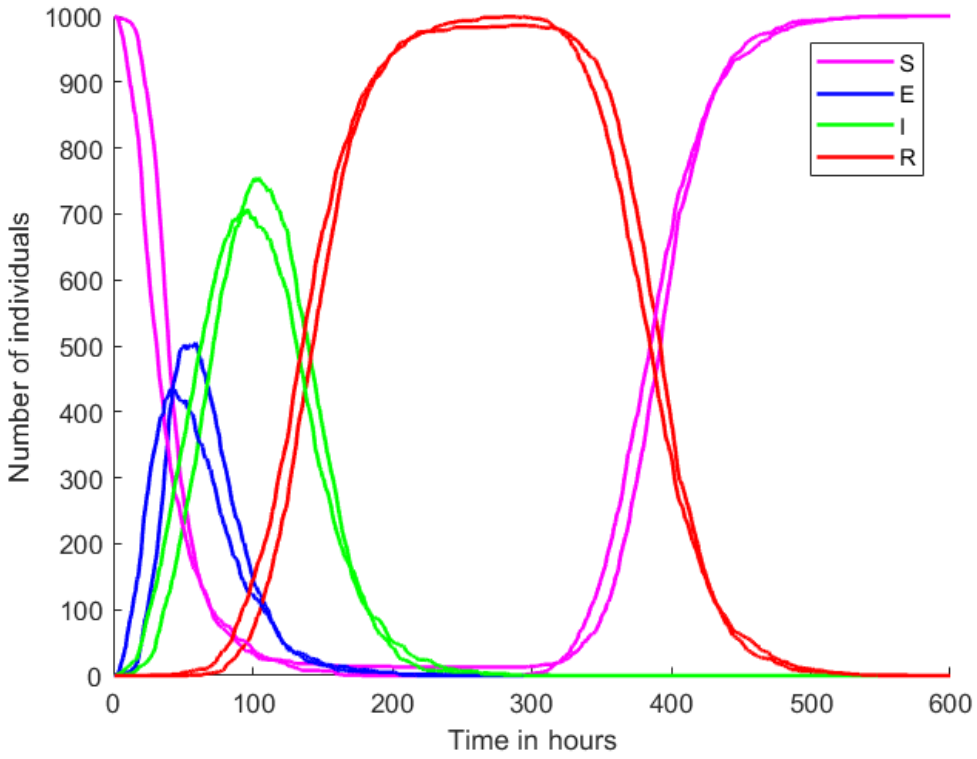


Figure 13. Contact network model with three subpopulations having just one common meeting place each using discrete-time modeling approach with initial number of infected persons chosen randomly from the subpopulation I only; initial conditions: $S(0) = 999$, $E(0) = 0$, $I(0) = 1$, $R(0) = 0$ for 600 hours.

Figure 13 shows almost similar result in Figure 10 probably because the rate of infection and probability of getting infected are relatively high. This implies that once a susceptible gets in contact with an infected person, the chances of such person getting the infectious disease is relatively high. Thus the disease is spread at a quicker rate.

Lets consider a different but realistic scenario in the level of contact pattern of persons in the population. Suppose that only a few initially uninfected persons (say politicians for example) in subpopulations I and II are allowed to meet at a particular meeting place (say a five-star hotel). With only 3 persons from subpopulation I as the initial number of infected, the situation changes when we simulate the process for several times as shown in Figure 14.

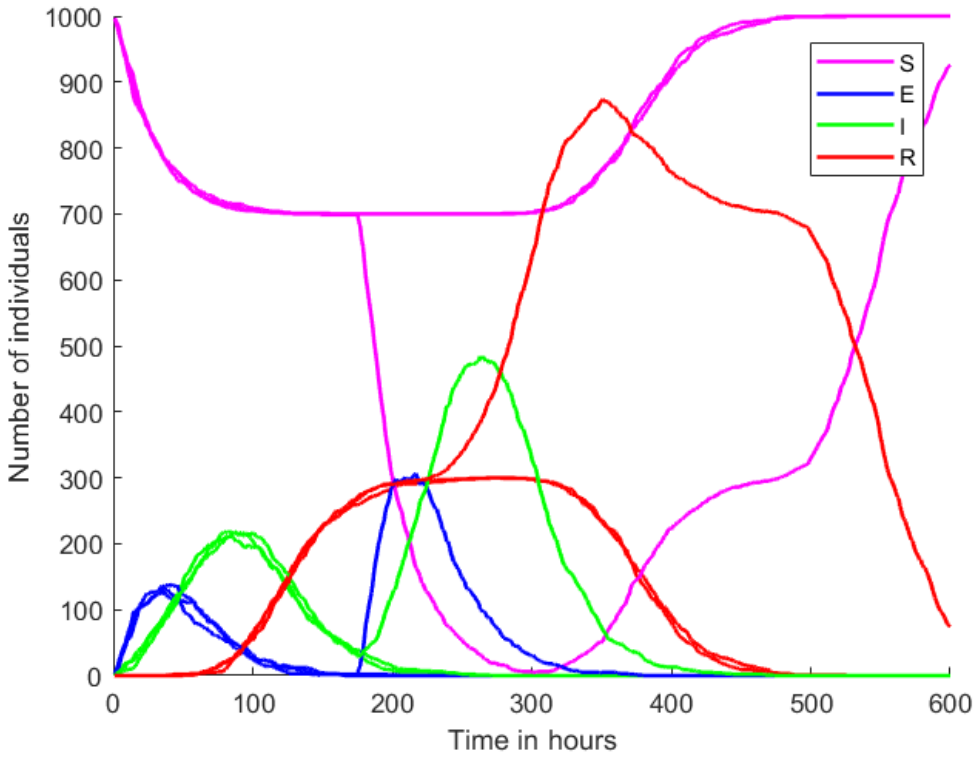


Figure 14. Contact network model with subpopulations I and II having just one common meeting place by few persons using discrete-time modeling approach with initial number of infected persons from the subpopulation I only; initial conditions: $S(0) = 997$, $E(0) = 0$, $I(0) = 3$, $R(0) = 0$ for 600 hours. The three trajectories for each compartment represents different realizations from the simulation.

Figure 14 shows the effect of the contact network when the contact pattern or movement of persons in a particular population or subpopulation is limited. At the beginning of the epidemic, about 300 persons get infected and after a while nothing happens and then suddenly everyone gets infected eventually. Thus, when compared to Figure 10, we observe that for each of the three runs, the rate at which the infectious disease spreads is limited by restriction in the contact network.

Figure 14 shows a limitation in the rate at which the disease spreads when compared to the situation where everybody has equal chances of spreading the disease as seen in Figure 10. However having more few persons (who might eventually get infected in the long run) to meet at a common place has the tendency to propagate the spread of the disease as shown in Figure 15.

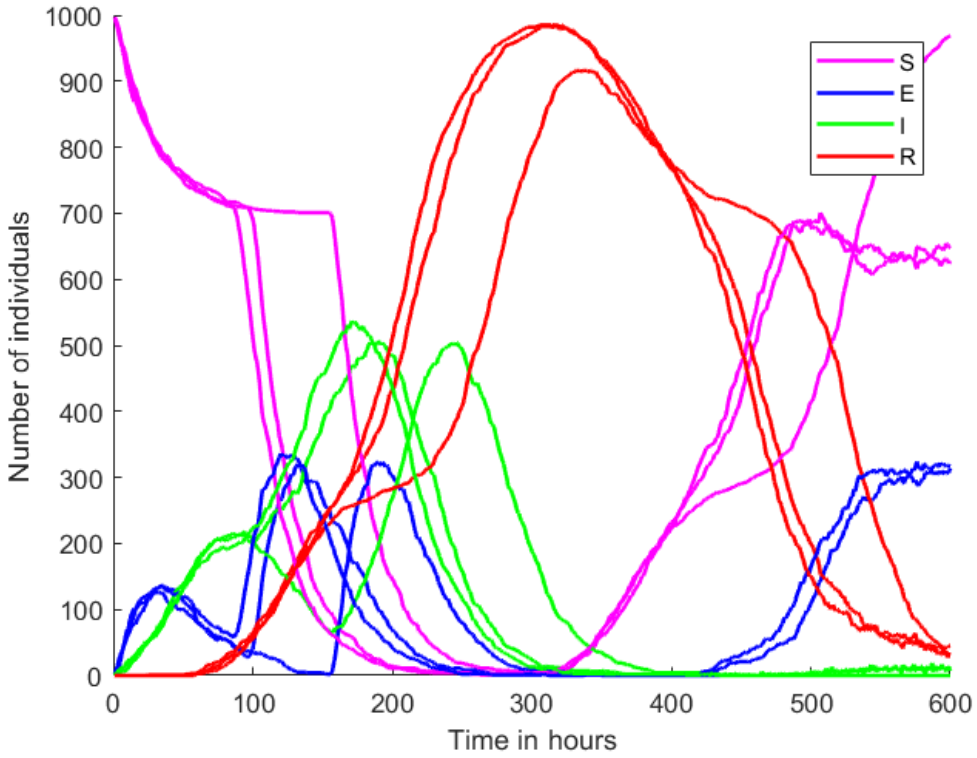


Figure 15. Contact network model with subpopulations I, II and III having just one common meeting place by few more persons using discrete-time modeling approach with initial number of infected persons from the subpopulation I only; initial conditions: $S(0) = 997$, $E(0) = 0$, $I(0) = 3$, $R(0) = 0$ for 600 hours. The three trajectories for each compartment represents different realizations from the simulation.

Comparing Figure 15 and Figure 14 shows that the susceptible individuals in the population are infected much faster in the former. Nevertheless in both cases, the spread of the disease is limited by contact network.

We further restrict the mobility of few persons who have a defined schedule for the entire period of time in the population. The effect of such situation is shown in Figure 16.

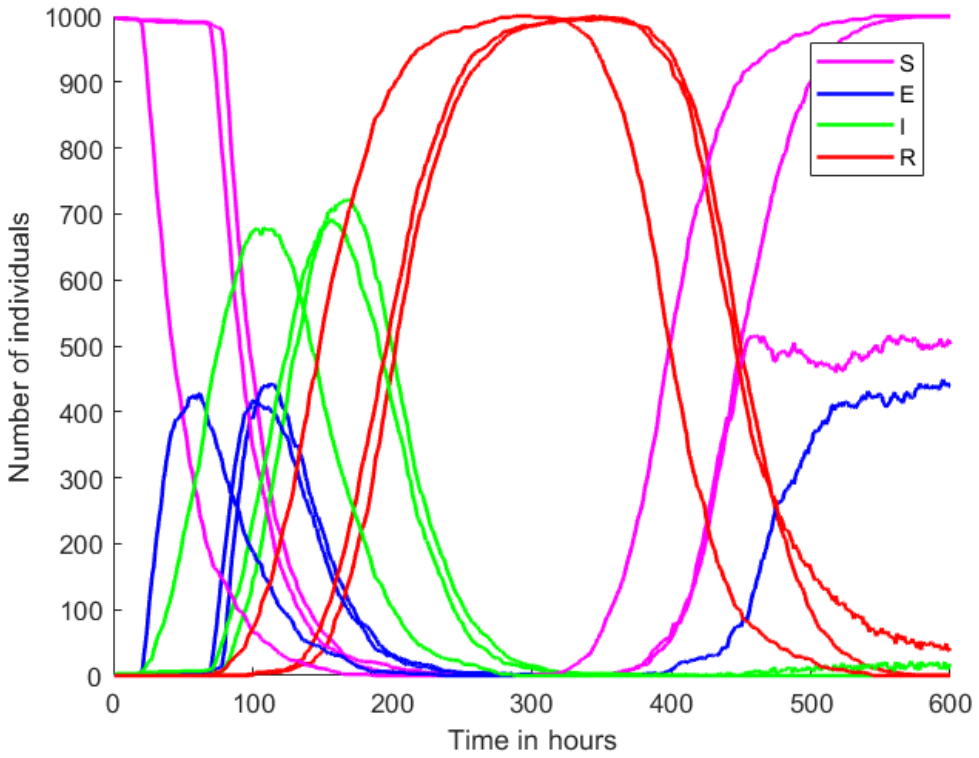


Figure 16. Contact network model with few persons having a defined schedule using discrete-time modeling approach with initial number of infected persons chosen such persons; initial conditions: $S(0) = 997$, $E(0) = 0$, $I(0) = 3$, $R(0) = 0$ for 600 hours. The three trajectories for each compartment represents different realizations from the simulation.

Figure 16 shows that the infectious disease does not immediately start spreading throughout the entire population. Individuals in the population start getting exposed and infected to the infectious disease after a certain time period. This situation for example, can be linked to school kids in kindergarten who go to school from home, and after a while in school, they return back home. Here we assume that 3 of the kids in the kindergarten where initially infected with the disease.

Suppose we have a similar scenario as Figure 16 but with two different subgroups of persons in the population. The effect of such situation is shown in Figure 17.

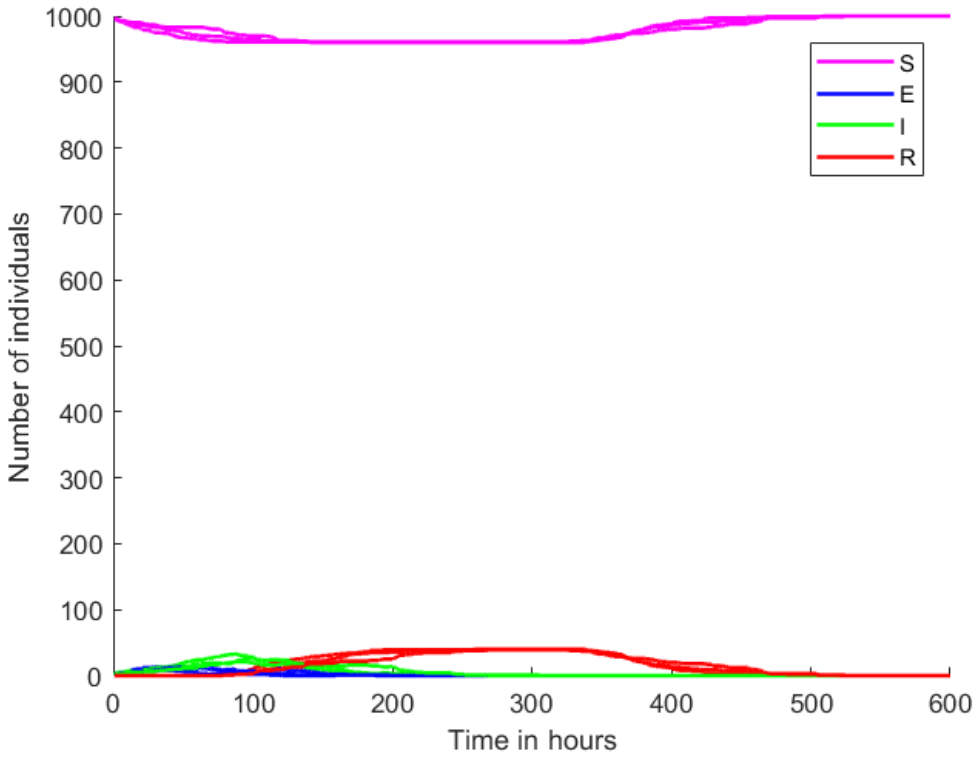


Figure 17. Contact network model with few persons in two different subgroups of the population having a defined schedule using discrete-time modeling approach with initial number of infected persons chosen from the first group; initial conditions: $S(0) = 997$, $E(0) = 0$, $I(0) = 3$, $R(0) = 0$ for 600 hours. The three trajectories for each compartment represents different realizations from the simulation.

Figure 17 shows that only a few number of persons in the population get exposed and infected in the entire population. This kind of scenario can be used to describe a situation where we have two different boarding schools in a population. Besides the teachers interacting with their respectively students, the teachers in both schools are only allowed to interact with themselves probably as a result of having a general meeting. Here we assume that 3 of the students in one of the boarding schools where initially infected with the disease.

So far we have considered cases where individuals in the population meet at every time step (that is, on a regular basis). However, suppose we now consider a case where the contact network is limited by the frequency of contacts between persons in the population. An example of such scenario is when the contact network described in Figure 14 is further restricted by the frequency of contacts between few persons meeting at a particular place as shown in Figure 18.

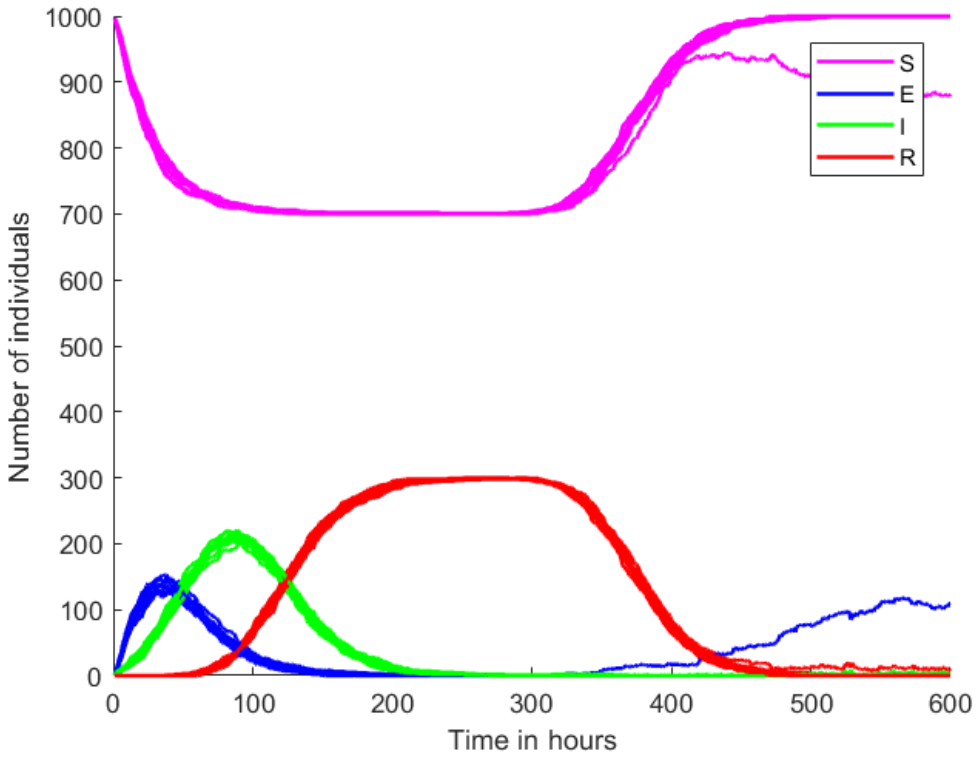


Figure 18. Contact network model with subpopulations I and II having just one common meeting place by few persons using discrete-time modeling approach with initial number of infected persons from the subpopulation I only; initial conditions: $S(0) = 997$, $E(0) = 0$, $I(0) = 3$, $R(0) = 0$ for 600 hours. The 10 trajectories for each compartment represents different realizations from the simulation when the contact network is limited by the frequency of contact.

Comparing Figures 18 and 14 shows that in Figure 18, all the 10 realizations from the simulation remains the same. That is, only a few number of persons always get exposed and infected. Here we consider a case where few persons in the population meet at a common place three times a week.

4 CONCLUSION AND FUTURE WORK

This last chapter concludes the thesis by discussing its findings and outlines directions for future research.

4.1 Summary and Conclusions

In this work, we have proposed a simple and straightforward method of generating contact network applied in modeling infectious diseases. The waiting time modeling approach together with the proposed network method gives us a way of exploring the effects of contact heterogeneity on the dynamics of infectious diseases.

The spread of an infectious disease through the contact network is similar (that is, changes in health states of individuals) to modeling with standard compartmental models; except that the Poisson distribution of contacts and the probability of getting infected in standard compartmental models which is chosen from an exponential distribution is replaced by the structure of the contact network and the probability described in Equation (22) receptively.

We further highlighted the differences between the use of standard random-mixing compartmental model and contact network model generated using our proposed calendar method in the spread of an infectious disease by carrying out a simulation for 600 hours (25) days. Randomizing the network fully gives similar result to that of the mass-assumption model where everyone has equal chances of getting infected. Several results showing the effect of varying the contact network structure where obtained. It was found that the spread of an infectious disease through a contact network depends on two factors: the topological structure of the underlying network under study, and the probability of getting infected when there is an encounter. This argument is in agreement with the findings made by [12]. It was indeed shown that restricting the movement of group of persons in a population tends to slow and reduce the rate at which an infectious disease propagates throughout the entire population. In addition, we argue that our proposed method generates a network which is dynamic in nature; where the connections within a defined subgroup of the population changes over time. This agrees with our intuitive perception of human interactions breaking and forming.

Finally our proposed method of creating contact network is more appropriate for populations that are small in size, such as community, group of people attending a conference

(see [26] for more details) , and so on, where accurate information concerning the schedules of individuals needed for building the network is feasible. A summary of the contact network settings is shown in Table A1.1.

4.2 Future Work

This thesis has demonstrated the potential of building contact networks, yet many opportunities for extending the scope of this thesis remain. The assumptions made using the proposed method and the findings in this work provide a strong foundation for future work in determining which particular type of disease is best modeled using our proposed method, and thus comparing the results with other forms of contact network models used for such particular infectious disease.

While calibrating the proposed contact network, it was observed that parameter values such as the rate of infection in no-network models and the probability of getting infected for each encounter in network models has significant effect on the dynamics of the disease. Thus, an interesting direction for future work would be to investigate the relationship between these two parameter values and also carry out a sensitivity analysis of these two parameters.

Another area for future work is to improve on the proposed method by making it more realistic. Firstly, by using real data and thus carrying out statistical inference in estimating the model parameters; and secondly by relaxing or/and improving on some of the assumptions made . Further investigation could be made on how the frequency of contacts affects the spread of the disease like the situation described in Figure 18.

Finally, future research could investigate how to implement various forms of intervention strategies described in [12, 18], which are seen as control strategies in eliminating the threat of a large-scale epidemic outbreak.

REFERENCES

[1] Constantinos I Siettos and Lucia Russo. Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4):295–306, 2013.

[2] Centers for Disease Control, Prevention, et al. Principles of epidemiology in public health practice: an introduction to applied epidemiology and biostatistics, 2006.

[3] H Kipruto, J Mung'atu, K Ogila, A Adem, S Mwalili, E Kibuchi, E Masini, R Kiplimo, and JR Ong'ang'o. An application of deterministic and stochastic processes to model evolving epidemiology of tuberculosis in kenya. 2013.

[4] A-J Valleron. Roles of mathematical modelling in epidemiology. *Comptes Rendus de l'Academie des Sciences Series III Sciences de la Vie*, 323(5):429–433, 2000.

[5] Caroline Dubé, Graeme Garner, Mark Stevenson, Robert Sanson, Conrad Estrada, and Preben Willeberg. The use of epidemiological models for the management of animal diseases. In *OIE Conference*, pages 13–23, 2007.

[6] M Graeme Garner, Caroline Dubé, Mark A Stevenson, Robert L Sanson, Conrad Estrada, and John Griffin. Evaluating alternative approaches to managing animal disease outbreaks–the role of modelling in policy formulation. *Vet. ital*, 43(2):285–298, 2007.

[7] Nick Taylor. Review of the use of models in informing disease control policy development and adjustment. *DEFRA, UK*, 26, 2003.

[8] Wayne M Getz, Richard Salter, Oliver Muellerklein, Hyun S Yoon, and Krti Tallam. Modeling epidemics: A primer and numerus software implementation. *bioRxiv*, page 191601, 2017.

[9] MG Roberts and JAP Heesterbeek. *Mathematical models in epidemiology*. EOLSS, 2003.

[10] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge university press, 2016.

[11] Matt J Keeling and Ken TD Eames. Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307, 2005.

[12] Lauren Meyers. Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. *Bulletin of the American Mathematical Society*, 44(1):63–86, 2007.

[13] NATHAN GILL. Deterministic and stochastic models of infectious disease: Circular migrations and hiv transmission dynamics.

[14] Puntani Pongsumpun, Rujira Kongnuy, Diana Garcıa Lopez, I-Ming Tang, Marc A Dubois, Weeraphat Pon, Wattana Kanbua, Seree Supharatid, Nipaphat Charoenthai, Ramanathan Suryanarayanan, et al. Contact infection spread in an seir model: An analytical approach. *Month*, 6(01), 2013.

[15] Maia Martcheva. *Introduction to Mathematical Epidemiology*, volume 61. Springer, 2015.

[16] Liliana Perez and Suzana Dragicevic. An agent-based approach for modeling dynamics of contagious disease spread. *International journal of health geographics*, 8(1):50, 2009.

[17] Gerardo Chowell, Nick W Henartner, Carlos Castillo-Chavez, Paul W Fenimore, and JM Hyman. The basic reproductive number of ebola and the effects of public health measures: the cases of congo and uganda. *Journal of theoretical biology*, 229(1):119–126, 2004.

[18] Nedialko B Dimitrov and Lauren Ancel Meyers. Mathematical approaches to infectious disease prediction and control. *INFORMS tutorials in operations research*, 7:1–25, 2010.

[19] Lauren Ancel Meyers, Babak Pourbohloul, Mark EJ Newman, Danuta M Skowronski, and Robert C Brunham. Network theory and sars: predicting outbreak diversity. *Journal of theoretical biology*, 232(1):71–81, 2005.

[20] Alun L Lloyd. Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theoretical population biology*, 60(1):59–71, 2001.

[21] A Huppert and G Katriel. Mathematical modelling and prediction in infectious disease epidemiology. *Clinical Microbiology and Infection*, 19(11):999–1005, 2013.

[22] Mark EJ Newman. Spread of epidemic disease on networks. *Physical review E*, 66(1):016128, 2002.

[23] Matt Keeling. The implications of network structure for epidemic dynamics. *Theoretical population biology*, 67(1):1–8, 2005.

[24] N Sherborne, JC Miller, KB Blyuss, and IZ Kiss. Mean-field models for non-markovian epidemics on networks: from edge-based compartmental to pairwise models. *arXiv preprint arXiv:1611.04030*, 2016.

[25] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2011.

[26] Juliette Stehlé, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Vittoria Colizza, Lorenzo Isella, Corinne Régis, Jean-François Pinton, Nagham Khanafer, Wouter Van den Broeck, et al. Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. *BMC medicine*, 9(1):87, 2011.

[27] Rehan Ahmad and Kevin S Xu. Effects of contact network models on stochastic epidemic simulations. In *International Conference on Social Informatics*, pages 101–110. Springer, 2017.

[28] Leon Danon, Ashley P Ford, Thomas House, Chris P Jewell, Matt J Keeling, Gareth O Roberts, Joshua V Ross, and Matthew C Vernon. Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases*, 2011, 2011.

[29] Alun L Lloyd and Steve Valeika. Network models in epidemiology: an overview. In *Complex population dynamics: nonlinear modeling in ecology, epidemiology and genetics*, pages 189–214. World Scientific, 2007.

[30] Mark DF Shirley and Steve P Rushton. The impacts of network topology on disease spread. *Ecological Complexity*, 2(3):287–299, 2005.

[31] S Alonso-Quesada, M De la Sen, and A Ibeas. On the discretization and control of an seir epidemic model with a periodic impulsive vaccination. *Communications in Nonlinear Science and Numerical Simulation*, 42:247–274, 2017.

[32] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[33] Gilles Kratzer. *Simulation Based Inference in Epidemic Models*. PhD thesis, Zurich Universtity, 2015.

[34] István Z Kiss, Joel C Miller, and Péter L Simon. *Mathematics of Epidemics on Networks*. Springer, 2017.

[35] David F Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *The Journal of chemical physics*, 127(21):214107, 2007.

[36] Veronika Malakhovskaia et al. Modelling approaches for epidemiological models. 2017.

[37] Liliana Perez and Suzana Dragicevic. An agent-based approach for modeling dynamics of contagious disease spread. *International journal of health geographics*, 8(1):50, 2009.

[38] David Welch, Shweta Bansal, and David R Hunter. Statistical inference to advance network models in epidemiology. *Epidemics*, 3(1):38–45, 2011.

[39] Jean-Loup Guillaume and Matthieu Latapy. A realistic model for complex networks. *arXiv preprint cond-mat/0307095*, 2003.

[40] Frédéric Mertens, Johanne Saint-Charles, Marc Lucotte, and Donna Mergler. Emergence and robustness of a community discussion network on mercury contamination and health in the brazilian amazon. *Health Education & Behavior*, 35(4):509–521, 2008.

[41] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440, 1998.

[42] Zhaoyang Zhang, Honggang Wang, Chonggang Wang, and Hua Fang. Modeling epidemics spreading on social contact networks. *IEEE transactions on emerging topics in computing*, 3(3):410–419, 2015.

[43] Geoffrey Grimmett. What is percolation? In *Percolation*, pages 1–31. Springer, 1999.

[44] Peter Grassberger. On the critical behavior of the general epidemic process and dynamical percolation. *Mathematical Biosciences*, 63(2):157–172, 1983.

[45] Lauren Ancel Meyers, MEJ Newman, Michael Martin, and Stephanie Schrag. Applying network theory to epidemics: control measures for mycoplasma pneumoniae outbreaks. *Emerging infectious diseases*, 9(2):204, 2003.

[46] Lauren Ancel Meyers, MEJ Newman, and Babak Pourbohloul. Predicting epidemics on directed contact networks. *Journal of theoretical biology*, 240(3):400–418, 2006.

Appendix 1. Contact Network Settings

We now present a summary of the settings used in generating the contact network, and the results in this thesis using the proposed calendar method. It is important to state here that the specific number of common meeting points indicated in Table A1.1 , are the points considered with respect to the initial infected persons in the contact network.

Table A1.1. Summary of Different Network Settings

	Initial No. of Infected Persons	Sub. Population	Meeting Places	Common Meeting Points
Figure 10	10 (Randomly)	1	10	Randomized
Figure 11	10 (Sub. Population I only)	3	10	Randomized
Figure 12	10 (Sub. Population I & II only)	3	10	Randomized
Figure 13	1 (Sub. Population I only)	3	10	2 (1 in each)
Figure 14	3 (Sub. Population I only)	2	10	1 (Few Persons)
Figure 15	3 (Sub. Population I only)	2	10	1 (More Persons)
Figure 16	3	1	10	1
Figure 17	3 (Sub. Population I only)	2	10	1
Figure 18	3 (Sub. Population I only)	2	10	1 (Few Persons meeting three times only)