Agent Based Decision Support System for Healthcare Applications

Nikhil Anand

A Thesis Submitted to

Indian Institute of Technology Hyderabad

In Partial Fulfillment of the Requirements for

The Degree of Master of Technology

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Computer Science & Engineering

June 2019

Declaration

I declare that this written submission represents my ideas in my own words, and where ideas or words of others have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

N Anand

(Signature)

Nikhil Anand

(Name)

CS17MTECH11021

(Roll No.)

Approval Sheet

This Thesis entitled Agent Based Decision Support System for Healthcare Applications by Nikhil Anand is approved for the degree of Master of Technology from IIT Hyderabad.

(Dr. Antony Franklin) Examiner
Dept. of Computer Science & Engineering
IITH

(Dr. Ramakrishna Upadrasta) Examiner
Dept. of Computer Science & Engineering
IITH

M.V.Panduranga

(Dr. M. V. Panduranga Rao) Adviser
Dept. of Computer Science & Engineering
IITH

N.B. Vineeth

(Dr. Vineeth N. Balasubramanian) Chairman
Dept. of Computer Science & Engineering
IITH

Acknowledgements

Firstly, I would like to thank my thesis advisor Dr. M. V. Panduranga Rao, Associate Professor at the Department of Computer Science & Engineering, Indian Institute of Technology, Hyderabad. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing. This work would not have been possible without his constant support and sincere advice throughout the work.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Vineeth N. Balasubramanian, Dr. Ramakrishna Upadrasta, Dr. Antony Franklin and Dr. Kotaro Kataoka, for their encouragement, insightful comments, and hard questions. I would also like to thank all other faculty members of the Department of Computer Science & Engineering at Indian Institute of Technology, Hyderabad for their continuous review and feedback of the work and the motivation to keep pursuing the research interest.

I thank my fellow labmates: Yenda Ramesh and Shiraj Arora, Research Scholars, Department of Computer Science & Engineering, Indian Institute of Technology, Hyderabad for all the help and support, for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had while working together. I would also like to thank all my friends who have always motivated me to work hard and have been my constant support in all the situations.

Last but not the least, I would like to thank my family: my parents and my sister, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This work would not have been possible without them. Thank you.

Nikhil Anand
CS17MTECH11021
Department of Computer Science & Engineering,
Indian Institute of Technology, Hyderabad.

Abstract

A healthy nation is a wealthy nation. Healthcare, being a universal area, concerns itself with the entire humanity. Healthcare, as an industry, will continue to grow and thrive as long as humans exist hence forming a large part of any countrys economy. In this context, the two most important domains are epidemics and healthcare coverage. The role of public healthcare facilities in both these domains is extensive. In this work, we introduce the application modules of an indigenously built tool used to simulate dynamic movement in a multi-agent based environment having complex social interactions. This work uses the tool to show it can be used in healthcare to perform probabilistic analysis effectively.

During an epidemic outbreak, strategic deployment of healthcare resources, which are usually limited in number, is very critical. In this work, dynamic health care units are used to represent those dynamic resources which can travel among cities and cure a fraction of the infected population. We have developed a tool based on agent based approach that effectively models the entire environment. We also propose a few strategies for their movement and study their effect in controlling the epidemic.

Another application of this tool is provided in the domain of universal healthcare schemes. Universal healthcare means equality in access to medical facilities irrespective of any basis. With various countries having a different organizational structure and existing levels of medical facilities, this work models two of Indias well-known healthcare schemes and analyzes their performance differences. Developed as an application module over the same tool, the models can be optimized or customized to evaluate the schemes over various parameters. This analysis also provides an estimate of how successful the scheme is likely to be.

Contents

Chapter 1

Introduction

Mankind has been plagued with epidemics for centuries. Since ancient times to the modern day, the world has been periodically faced with epidemics at different scales. It is probably the only thing which has affected people everywhere irrespective of time, location or any other factor. The above fact in itself makes it one of the most critical problems and hence the need to address it. However, even after centuries of technological advancement, there does not exist an efficient generic system which can predict the cause and spread of infection in a specific population.

Formally, epidemiology is defined by the World Health Organization (WHO) as "the is the study of the distribution and determinants of health-related states or events (including disease) and the application of this study to the control of diseases and other health problems." [1] Epidemiology is literally the "study of what is upon people" (Greek words *epi* = on or upon, *demos* = people & *logos* = the study of). Epidemiology focuses on the outbreak of a disease that affects many people at the same time, & may spread through one or more communities. However, It does not include diseases that exist permanently in a region like malaria in Africa.

The Greek physician Hippocrates, known as the father of medicine, coined the terms "epidemic" and "endemic" and explained the difference between them. He, however, wanted to discover the logic behind sickness. He is said to be the first person who had examined the relationships between the occurrence of a disease and the environmental factors that influence its progression in a population [2]. Epidemiological information forms the cornerstone of public health and can be used to plan ahead of time to design precautionary and control measures in case of an epidemic. It can also be used to shape policy decisions, evaluate strategies to curb the spread of infection and as a guide to the management of patients who are already infected.

1.1 Motivation

The relevance of computational epidemiology today is more than ever. The world we live in today can be described as a connected world. Today distances hold no bar in transportation and communication between people. Despite the advantages, this has brought, this has become a critical factor in the transmission of infection between different populations and even countries. Pathogens usually spread through contacts in social networks when susceptible individuals come in contact with infected people. These contacts can be viewed as dynamic links created by the everyday movement of people.

Earlier, the networks in which diseases were transmitted were restricted due to several social and economic factors. This led to slower geographic dissemination of diseases. Today, regarding infectious diseases that have emerged over recent years, pathogens can and have appeared unexpectedly, spread rapidly and has had devastating effects impacting millions of people. In this regard, individual behaviours become the centre of disease dynamics and controls. Factors like social distancing and early detection go a far way in controlling the spread and effects of the epidemic. Monitoring human social contact networks is equally difficult and critical due to their dynamic nature. The behavioural adjustments amid an epidemic modify social contact networks and affect the epidemic's advancement. An epidemic today cannot be on measured in terms of the spread of infection. As described by Andrea Apolloni in his work on Computational Epidemiology in a Connected World [3], "An epidemic is not simply a diffusion over a network, but a coevolving system of multiple networks, dynamic processes (spread of disease, fear, and so on), and individual behavioural adaptation." Epidemiology today is not just about predicting the spread of the pathogen in a network, but more about the anticipation of the above said things and having a model or methodology that can adapt to it. The blend of prediction, the expectation of complex social communications developing alongside limited and postponed data makes computational the epidemiology among the most challenging issues in social and computational science.

1.2 Types of Epidemiological Studies

Estimating disease recurrence in a population requires the stipulation of specific criteria. Epidemiological studies can range from observational to experimental designs and can be categorized in the following categories:

- Observational Study - In these studies, nature is permitted to follow through to its logical end while epidemiologists observe from outside. They can be further classified as:
  - Descriptive: A descriptive study tries to answer the 'Ws' of an epidemic. They try to describe an epidemic based on parameters like - 'who', 'what', 'where' and 'when'.
  - Analytic - An analytic study of epidemic tries to answer how the epidemic spread. The aim of this study is to examine the known associations further or hypothesize relationships.
- Experimental Study - The epidemiologists control all the factors in an experimental study and try to figure out the effect of various factors in the spread of infection. This type of studies is usually done in clinical or community trials and other interventions.

The term 'epidemiologic triad' is the combination of a host, an agent and the environment brought together to cause an epidemic outbreak. Identification of causal relationships between the exposures and outcomes is an important aspect of epidemiology. However, something of even more importance is to remember causal inference, i.e. the fact that correlation does not imply causation. Correlation, or if nothing else relationship between two factors, is an essential yet not adequate criteria for the derivation that one variable causes the other.

1.3 Evolution of Epidemic Modeling

Epidemic modeling began in the year 1976 when Edward Jenners research on smallpox virus led to the development of vaccines. Daniel Bernoulli at that time, developed mathematical models to demonstrate the benefits of inoculation from a mathematical perspective. The term 'epidemiology' was coined in 1802 by the Spanish physician Villalba in his work Epidemiologìa Española [4] to describe the study of epidemics. However, it was John Snow, father of (modern) epidemiology, who was the first person to link the London cholera epidemic in 1854 to a particular water source. The incident formally led to the development of the field of epidemiology. Today, cholera is an endemic which usually occurs in poverty-stricken countries. In 1897, Ronald Ross and George Macdonald [5] developed a mathematical model of mosquito-borne pathogen transmission. Anderson McKendrick studied with Ross on antimalarial operations and pioneered many discoveries in stochastic processes. The Kermack and McKendrick SIR model [6] of diffusion of infection in a population formed the basis of epidemiological research for years to come. The SIR model proposed by them to represent virus transmission used a set of nonlinear ordinary differential equations (ODEs) that associate a transition rate to the mobility of agents between compartments. Later, developments and variations to this came up in the form of SIER, SIS, SIERS, etc. models. All the above models are categorized as compartment models and suffer from many drawbacks with the most limiting ones being these models were deterministic and homogenous. Simply stated the models considered the complete population as having similar characteristics and gave the same results on multiple runs irrespective of the location or behaviour of people. To overcome the drawbacks of compartment models, Agent Based Models (ABMs) were developed.

An Agent Based Model (ABM) can be viewed as a bottom-up approach to recreate a real life dynamic system consisting of multiple entities, each with a different set of attributes and behaviours. ABM s are stochastic models that treat every person as an individual in the form of an agent with a unique behaviour based on a set of rules and try to simulate the expected path an infection can follow. The concept behind following a bottom-up approach to discover the emergent phenomena created by individual actions of the agents, which usually is not intuitional or easy to predict. ABMs in recent times have come up as an alternative to the traditional mathematical mass action models used for epidemiology. It has increasingly gained acceptance in recent years due to the promising results achieved in various fields of social science. The use of ABMs in every field provides us with numerous advantages over traditional mass action or mathematical models in every field. However, with the advantages, ABMs require a lot of simulation and computation power in order to emulate real world scenarios along with sophisticated models. However, with the progress in the field of high performance computing, epidemiologists want to harness the computation power to solve the epidemiological problems.

1.4 Classification of Computational Epidemiology Models

Mathematical models for epidemiology can be classified under different aspects. Models may be classified on the approach followed by them or the time at which are used. Some of the classification categories are described below.

1.4.1 Classification based on approach

Formally, based on approach, epidemiological models can be classified as follows:

- Mass Action Compartment Models - These are differential rate based models which follow a set of mathematical equations to depict the diffusion of infection in a network. These consist of deterministic and stochastic ordinary differential equation based models. The examples include Bernoulli [7], Ross, Kermack and McKendrick models for deterministic ODEs and Bartlett [8], Bailey [9] and Brauer models for stochastic ODEs.
- Agent Based Models - These are interaction based models which are based over a network of people. These can depict complex interactions in networks. Examples include RandomNets (Barabasi, Meyer) and Realistic Social Networks (NDSSL, Salathe).

1.4.2 Use-time classification of models

Epidemiological models in real time can be classified based on their time of use as follows:

- Modeling before an epidemic - Modeling before an epidemic is usually done as precautionary or preventive measures. The targets of this model include determining the (non) medical intervention required, the feasibility of containment, the optimal size of the stockpile and best use of supplies once an epidemic begins.
- Modeling during an epidemic - Modeling in real time during an epidemic is done to contain the spread of the epidemic and form health measures to be adopted based on the current situation. The objective of these models is quantifying transmission parameters, interpreting real time epidemiological trends, measuring antigenic shift and assessing the impact of an intervention.

1.5 Epidemic Quantities of Interest

The following section lists the entities and quantities related to an epidemic which are useful and of interest to us. Ideally, an epidemiological model should be able to give us information about these terms. These quantities combine to form the results and are the basis of health policy drafted in case of an epidemic.

- Susceptible Population (S) - The group of people who at current instant dont have the disease/pathogen but may catch it from their network or environment.
- Exposed Population (E) - The group of people who are the carriers the pathogen but havent developed symptoms themselves.
- Infected Population (I) - The group of people who are currently infected by the pathogen. All infected people are also the carriers of pathogen and therefore exposed as well.
- Recovered Population (R) - The group of people who were infected earlier but are now cured. People in this category are assumed to develop immunity against the infection, meaning that they cannot be infected again. If this is not the case, then they are assumed to lie in a susceptible group. Depending on the infection, there may or may not be a group of recovered population.

- Infection Rate (β) - Infection rate for infection is defined as the number of interactions made by an infected person per time unit.
- Recovery Rate (γ) - Recovery rate of infection can be defined as the number of infected people who are likely to recover per unit time. The recovery rate is inverse of infectious period/time taken to recovery.
- Basic reproduction number R_0 - It is the expected number of secondary cases produced by a single primary infection in a completely susceptible population. For an infectious disease with an average infectious period of $\frac{1}{\gamma}$ and transmission rate β, the value of $R_0 = \frac{\beta}{\gamma}$. For an infection to turn into an epidemic, the infection rate should be greater than the recovery rate. This essentially transforms $\beta > \gamma$ to $R_0 > 1$.
- Epicurve - Graphical display of the numbers of incident cases in an outbreak or epidemic, plotted over time is known as epicurve.
- Peak of an epidemic - The highest number of infections at any given instance is known as the peak of the infection.
- Time to peak - The time unit at the peak of the epidemic is said to be the time to peak.
- Total number of infections - The cumulative of all infection population is known as the total number of infections.
- Duration of an epidemic - The time when the number of daily infections falls below a daily threshold, and the epidemic is assumed to be in control is known as the duration of the epidemic.

Chapter 2

Computational Epidemiology

Customarily, the study of disease transmission has been founded on information gathered by general health agencies through health personnel in medical clinics, specialists' workplaces, and out in the field. Epidemiological studies began with a desire to predict the about the next epidemic but till date have only been able to study the past. Now with the increase in processing power of computer systems, researchers want to use it to create network systems to predict various factors associated with epidemics. In his book Veterinary Epidemiology [10], Michael Thrusfield defined computational epidemiology as "the application of computer science to epidemiological studies". Similarly, Habtemariam, in his work on population models [11] in 1988, explained the need and effectiveness of computer based interaction models to understand the behaviour of complex biological systems. These models suffer from the drawback of an oversimplification of complex interactions but still provide an approach to conducting experiments as compared to classical laboratory experiments which may neither be practical nor feasible.

Computational epidemiology deals with the development of computational and mathematical methods, tools and techniques to support epidemiology using various aspects of computer science. The use of computational models for broader policy questions in epidemiology has opened up a distinct view to solve the age old problems in this field. However, the use of newer technologies poses a new set of problems to create a computation system. This field relates to the study of uses, applications and challenges posed by the use of computation systems in the field of epidemiology. Computational Epidemiology is a collective term including but not restricted to the following:

- Discrete Event Simulation (for Discrete & Continuous time)
- Numerical Computation (Scientific Computation)
- Web crawling & surveillance (Information Retrieval)
- Data Mining & Machine Learning (Statistics)
- Algorithms & optimization applied to scientific concerns
- High Performance Computing (Parallel & Distributed techniques)

Despite the works done in statistical epidemiology, some of the current areas of application of computational epidemiology are:

- Spread of epidemics in complex networks
- Surveillance of diseases in social media
- Opinion mining in social media (behave like viral epidemics)

2.1 Mass Action Compartment Model

Mass action compartment models, as previously explained, are traditional ordinary rate based models which split the population into compartments that represent the different stages of a disease. A set of nonlinear ordinary differential equations (ODEs) associates a transition rate to the mobility of agents between compartments. The disease dynamics are characterized by a parameter R_0, the basic reproduction number. Traditional compartment models are the SIR (Susceptible-Infected-Recovered) model and its variations (SI, SIR, SIS, SEIR, etc.). The SIR model was proposed by Kermack & McKendrick in 1927 which formed the basis of mathematical modeling in epidemiology. Mass action compartment models follow some simplifying assumptions to describe the system under consideration.

- The transmission of infection is immediate on contact without any idle period.
- All the individuals in a compartment are equally susceptible or infected.
- The infection is transmitted by contact (direct or indirect) between a susceptible and an infected individual.
- All the people in a compartment are homogeneous (having same behaviour & characteristics).
- SIR is a deterministic model since the population measure is sufficiently extensive to deal with the changes in the spread of the infection
- The population under experimentation has a fixed size, i.e. no births, deaths or immigration is considered to have taken place.

2.1.1 SIR Model - Kermack & McKendrick (1927)

The SIR model proposed by Kermack & McKendrick in 1927 [6] divides the total population into three compartments, namely the susceptible, infected and recovered compartments. To begin the SIR model, we initially identify the independent and dependent variables. The independent variable, in this case, is time t. There are two categories of the dependent variable in the SIR model. The first category of variables is the population dependent on time. We express the population as a function of time. The number of susceptible individuals is:

$$S = S(t) \tag{2.1}$$

The number of infected individuals is given by:

$$I = I(t) \tag{2.2}$$

Similarly, the number of recovered individuals is:

$$R = R(t) \tag{2.3}$$

The second category of variables represents the fraction of the total population in each compartment. For a total population count of N, we have the susceptible fraction of the population as:

$$s(t) = \frac{S(t)}{N} \tag{2.4}$$

The infected fraction of the population is denoted by:

$$i(t) = \frac{I(t)}{N} \tag{2.5}$$

Similarly, the recovered fraction of the population is:

$$r(t) = \frac{R(t)}{N} \tag{2.6}$$

However, both the sets of dependent variables are proportional to each other and will give us the same information about the progress of the epidemic. Also, at each time instant,

$$s(t) + i(t) + r(t) = 1 \tag{2.7}$$

The rate of changes for the dependent variables is defined as follows:

- Assume that each infected individual has a fixed number of β contacts per day. Although all these contacts may not necessarily be with the population in the susceptible compartment. Assuming a homogeneous mixing of the population, the fraction of these contacts with susceptible people is $s(t)$. Therefore, on average, there are $\beta s(t)$ new infections spread by each infected individual. Another simplifying assumption is that due to the large population size, we can ignore the fact of encountering the same susceptible individual multiple times in a single day.
- We also assume that a fixed fraction of γ of the infected population will recover on any given day. This essentially means that the time to recover is $\frac{1}{\gamma}$ days.

On differentiating the dependent variables, we get the following SIR equations. The susceptible equation depicts the rate of flow at which the susceptible population is getting infected.

$$\frac{ds}{dt} = -\beta s(t)i(t) \tag{2.8}$$

The recovered equation depicts the rate at which the infected people are recovering from the infection.

$$\frac{dr}{dt} = \gamma i(t) \tag{2.9}$$

The infected equation depicts the rate of change of population in the infected compartment. It is a combination of the incoming flow of susceptible population and outgoing flow of recovered population. To derive the infected equation, we should remember from Eq. 2.7:

$$s(t) + i(t) + r(t) = 1$$

On differentiating this we get,

$$\frac{ds}{dt} + \frac{di}{dt} + \frac{dr}{dt} = 0$$

From the above derived equations we have,

$$-\beta s(t)i(t) + \frac{di}{dt} + \gamma i(t) = 0$$

$$\frac{di}{dt} = \beta s(t)i(t) - \gamma i(t) \tag{2.10}$$

The SIR model is completed by giving an initial condition for the differential equations. This is usually the parametric values for the total population of N and an initial number of people in the SIR compartments. The values of β and γ are based on the infection and can be adjusted as required based on the data. Other similar variants of Kermack & McKendrick SIR model such SIS, SEIR or SEIS work on similar assumptions and follow the above methodology.

2.2 Evaluation of Mass Action Models

The mass action models are quite general. They are used to simplify mathematical models for epidemiology. Considered as the workhorse for computational epidemiology, these models are considerably quick to build and put to use. These can be easily modified to consider different compartments or more recently variation in population size due to births, deaths and immigration. These models are easy to solve, using well defined ODEs in mathematics. However, these models follow what can be best explained as a black box approach and do not focus on factors like individual behaviour or the mechanism of transfer of the disease. These models generalize the population as well and pay no individual focus towards the local healthcare factors where the epidemic outbreak is considered to have happened. Moreover, these dont even take into consideration the structure of the disease, which in real life, play a pivotal role in the diffusion of infection. Depending on the disease, the mode of transmission of infection could be directly in case of viral or bacterial agents or indirectly due to vectors or carriers such as malaria by mosquitoes. Thus modeling for communicable disease requires careful consideration. The table below summarizes the advantages and disadvantages of compartment models.

2.3 Interaction Based Modeling

Interaction based approach and ABMs derive their roots from Complex Adaptive Systems (CAS) [12]. CAS was designed as a simple approach to understanding how individual behaviours affect the whole system. CAS is a subset of nonlinear dynamical systems having multiple interdisciplinary uses. CAS can provide insights about various factors in natural and social sciences consisting of heterogeneous agents, system state transition and emergent phenomena. The CAS is a system with a complex

Advantages	Disadvantages
Compartment models are considered as workhorse models for epidemiology	They assume the complete population to be homogeneous
These are quick to build and implement	They are unable to capture complex human interactions
Compartment models are easy to extend	They suffer from behavioural generalization
ODEs are well developed and defined in mathematics	They follow a black box approach and hence are difficult to refine

Table 2.1: Advantages & Disadvantages of Compartment Models

interaction of dynamic agents along with their relationships. CAS is considered to be adaptive and self-dependent. They are viewed as a complex naturally visible accumulation of generally comparable and in part associated miniaturized scale structures framed to adjust to the changing conditions and increment their survivability as a full-scale structure. With many of the conceptual, implementation details and features being similar to ABMs, they are considered as the predecessor of what we know today as ABM. However, with the focus being on high-level properties, CAS differs from ABM due to the former having self-similarity and self-organization features.

2.3.1 Simulation Modeling

Interaction Based Modeling is a specific kind of simulation modeling technique which can be used to find solutions to real world problems. A simulation model is a set of rules that defines how the system will change its state in the future. Simulation modeling gives us a way to optimize our solutions for the real world before actually implementing them. The need for simulation modeling arises from the basic fact that it is not possible to have an analytical model for all the problems in the world. In epidemiological applications, real life experimentation is not an option as we cannot infect some population to test our strategies. We need to settle on choices based on limited data. However, unlike in the real world, we can set up a scale model and run the simulations to test our hypothesis before implementing them in the real world. This saves us a lot of money and effort that goes into implementing an intervention or a policy at a large scale only to see it fail. Modeling a real world problem can be divided into three stages - an abstraction of a real problem into a scale model, analysis and optimization of the model, and mapping the solution back in the real world. Some of the areas of application of simulation modeling are:

- High level abstraction areas - are typically the areas which deal in aggregates, global causal dependencies and feedback dynamics. These models are known as macro level models with fewer details and more focus on strategic implementation. Examples of these models are population dynamics, marketplace & competition, ecosystem, health economics, R&D project management, etc.

- Medium level abstraction areas - mainly deal with component or module aggregates of the system. These models are known as meso level models with medium details and focus on the tactical level. Some of the examples in this category are supply chain management, call centre, asset management, electrical power grid, etc.

- Low level abstraction areas - usually deal with individual objects, exact sizes, distances and timings. These models are known as micro level models with more details and focus on the operational level. Traffic micro models, pedestrian movement, computer hardware, automotive control system, etc. are examples of this category.

2.4 Agent Based Modeling

Agent-based modeling is a powerful simulation modeling technique that has been used and gained acceptance in an increasing number of areas, including applications to real world business problems. Steven Banks [13] in his work on agent based models had written that revolutionary tools are distinguished from other tools not because of their technological importance but because of the social impact they have. In today's world, when we see a sea of new tools and technologies every day, only some of these can have a measurable impact on society. In this regard, Agent Based Modeling has set itself apart from the others. The most important reason for the popularity of ABMs is the restrictions imposed by traditional methods. As seen earlier, that list is long with the most important ones being linearity, homogeneity and generalization. To overcome the disadvantages of compartment models, epidemiologists have switched to interaction based approach for mathematical modeling. Interaction based approach, popularly known as Agent Based Model, is designed to capture the behaviour of each unique individual (agent) and its inherent fuzziness by representing every person as an agent. An agent based model can be defined as a system of autonomous decision making entities. Each agent in an ABM has a unique behaviour based on a set of rules. Agents execute various behaviours appropriate for the system they represent. For an application like epidemiological modeling, it is necessary to have a model which can capture the inherent nature complex nature of interactions between different agents. The model can depict the behavioural and mobility patterns of the agents based on its attributes and environmental responses. However, ABM simulations, especially for a large number of agents, requires a lot of memory and can be computationally expensive.

2.4.1 Components of an Agent Based Model

At the core level, an ABM consists of agents, relationships between them and a set of rules followed by them. Based on the agent attributes and rules, even the simplest agent based model can exhibit complex behaviour pattern. In more sophisticated systems, the agents may incorporate networks, evolutionary algorithms or other learning techniques. ABM can be viewed as a decentralized, bottom-up modeling approach. The method is said to be bottom up as it uses as building blocks at the lowest level to build up a system using the individual behaviours which might simulate the operation of a complex organizational system at a higher level. An ABM even though capable of capturing complex structures and dynamics is easier to maintain. Such a system can be modelled even in the absence of knowledge about the global interdependencies between entities as the model itself will find those based on interactions.

An agent in ABM can be formally defined as "an autonomous decision making entity". Agents are the independent atomic entity. It is a discrete individual with a set of attributes and characteristics. The behaviour of agents is governed by certain rules. The agents have particular decision making capability based on the current state. The agents of a system are self-contained and live in a specific environment where they interact with other agents. The purpose of an agent is to achieve a specific goal, i.e. agents are always goal-directed. In high level ABMs, the agents may also have some memory or intelligence. Agents can contain both base-level rules for behaviour and the intelligence to change them as well.

2.4.2 Advantages of Agent Based Models

ABMs hold an edge over traditional mathematical modeling techniques. The most trivial reason behind the success of ABMs is that they overcome a lot of disadvantages of compartment models. However, the ABMs don't stop at this but give some added advantages as well, which make it useful in many applications. Some of the advantages are:

- Emergent phenomena - ABMs can successfully capture the emergent phenomena of a group. Emergent phenomena can be defined as "a groups collective behaviour" (which can be often counterintuitive). The emergent formula is the result of interaction between agents in the same or different modules. An emergent formula might show a behaviour which could be quite distinct from individual behaviour. The characteristic of the emergent formula is that it is usually challenging to understand and predict. In ABMs, a small change in a simple individual rule can change the dynamics of the complete system and intuition is an inferior guide in case of emergent behaviour beyond a specified complexity.

- Natural description - ABMs provide a natural description of the model as compared to traditional models. Due to the bottom-up approach, each entity has to follow a certain set of rules which define his behaviour irrespective of any other factors or its impact on the system. This is close to a real life organizational structure where each person has a dedicated job and can explain it easily. However, it becomes quite tricky if we want to observe his behaviour in context to the impact it has on the organizational dynamics. The bottom-up approach allows us to capture individual behaviour, and the emergent phenomena combine it to find its effect on the organization. ABM looks at an organization from the perspective of individuals and their contribution in an organization rather than from an organization's perspective.

- Flexibility - The flexibility of an ABM can be seen in numerous situations. ABMs can be best represented as an object oriented approach for modeling. This enables the user to modify the attributes and the rules of an agent. ABM provides a natural way of tuning the model and complexity of agents. Another situation is being able to change the levels of abstraction and aggregation.

2.5 Areas of Applications

Agent based modeling technique has a lot of application areas. They can particularly be used in areas where we can exploit one or more advantages mentioned above. ABMs can be useful in the following cases:

- ABMs can be used in cases where we have non-linear individual behaviours, or the behaviour can be expressed as a cascade of if-else-then rules. It can also be used where agents have some memory, show path dependence or have temporal correlations.

- ABMs can be used for models where we have heterogeneous interactions in a network.

- ABMs are also useful in scenarios where the average for a system will not work since averages tend to smoothen sharp changes which could have otherwise been amplified and have had a more substantial impact.

- ABMs can be applied in circumstances where singular conduct is complex and can't be adequately characterized to total change rates.

- In situations where activities are a progressively natural method for depicting the framework, for example, a hierarchical structure, ABMs can be useful to demonstrate the framework.

- ABMs are useful in situations where stochasticity is involved in agent behaviour that may lead to change in the state of the system.

ABMs today are useful in various social, economic and political science problems. With the increase in complexity to analyze these areas, ABMs are gaining both applications and acceptance. Based on the above use cases, some of the real life applications of ABMs are:

- Network Flows - A flow network can be depicted as a directed graph with each edge having a flow and capacity(maximum flow). A flow network can be used to solve multiple problems like maximizing the flow, minimizing the cost, finding the source of flow, etc. Real world examples of flow networks are evacuation from a building in case of fire, traffic flow in an area, data flow in an internet network or even customer flow management. Crowd stampedes during a fire evacuation or at times without any significant reason other than to reach the destination in time have caused a loss of life and infrastructure. Panic is a common reaction observed during such situations which make it very difficult to model it computationally or control it in real life. Agents in a network flow tend to have a variety of responses which can be captured with an ABM. The result of collective panic or any interventions at that time is the emergent behaviour of the system. Another similar example is the traffic flow with a focus on solving problems such as traffic light timings, accident minimization and congestion minimization and recovery due to bad driving habits. Industrial applications related to this are have been developed such as $TranSims$ [14] at the Los Alamos National Laboratory (LANL) in the United States and are under progress as a part of the Smart Cities project in India.

- Strategic Positioning - Situations or conditions in which a strategy is required for placement or position of an item or an agent concerning the environment can use ABMs for strategic positioning. It can be used in places like stock markets, shopping marts and strategic economic simulations. The stock market is one of the most fluctuating fields in the world with the largest environment comprising the whole wide world. Any action of an individual or any organization in any part of the world can trigger a one-way movement in the market which usually gets caught on very soon. ABMs are an acceptable technique to try and simulate what the possible outcomes in particular scenarios are. A live application [15] has been developed for the National Association of Security Dealers Automated Quotation (NASDAQ), the American stock exchange to analyze the effects of changes in its trading policies. Another example of strategic placement is of products at a shopping mart. The idea behind the problem is that the shopkeeper wants the customers to take longest possible paths so that they buy more products, but the customers would want the shortest path so that they can quickly finish their

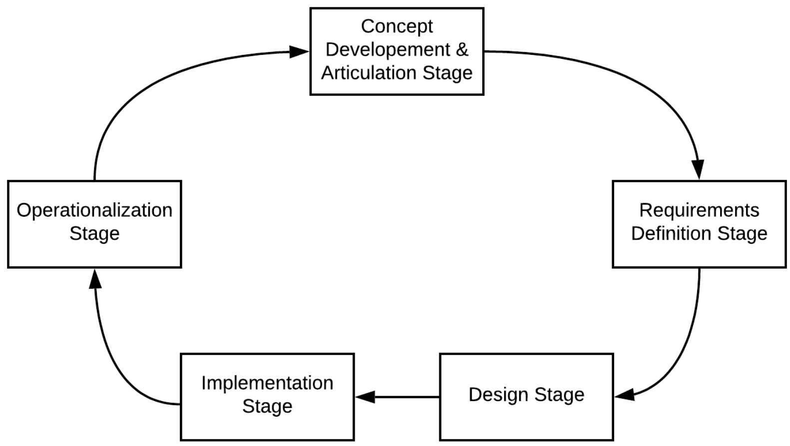

Figure 2.1: Modeling Lifecycle of an ABM

shopping. This can be considered a high-end example of minimax problem. ABMs such as *SimStore* [16] have been developed which simulate such a problem.

- Organizations - Corporations and organizations are a direct application area for ABMs. Modeling operational risk and organizational design are two popular applications of ABMs in organizations. It is very trivial to understand how a bottom-up technique will be able to capture the emergent formula in an organization. Operational risk is the probability of having a loss due to employee errors, system failures or any other external activity such as fraud. An ABM can be used to estimate the loss in such cases and how it will have an effect on the system. Moreover, we can use the ABM to design interventions as well as test them using an ABM. Similarly, organizational design and induction of resources in a particular project or organization and its effect can be modelled using an ABM.

- Diffusion - Diffusion in a network captures the underlying mechanism of how events propagate in a complex network. Events like the spread of the virus in a computer network, the spread of infection in an epidemic, viral market or transmission of fake news in social media. ABMs are used to find the hubs and transmission rates in these cases. ABM can be applied in these cases where people are influenced or affected by their contact network, be it online or offline. ABM can be created for an application like the spread of fake news in a Twitter network or real life. Similarly, ABMs have already been created for epidemiology and computer virus networks. The ABMs can be used to decide on public policy based on the result. The policies decided should strike a balance between the regular working of a system vs control measures.

2.6 Modeling Lifecycle of an Agent Based Model

Similar to a software development model, the modeling of an ABM is done in multiple cyclic phases. The modeling life cycle can be broadly divided into five stages described below and represented in Figure 2.1.

- Stage 1 - The modeling begins with concept development and articulation stage. The target of this stage is having an idea of what is expected from the model.
- Stage 2 - The requirements definition stage follows next. The target of this stage is to make the model goal specific.
- Stage 3 - The design stage follows next with the objective of giving the model a definitive structure and function.
- Stage 4 - The implementation stage refers to the development, coding and testing of the model.
- Stage 5 - Once the implementation is finished, the model is deployed for use in the operationalization stage.

Chapter 3

Framework

Resource allocation and intervention management are the two highest priority tasks which need to be done during an epidemic. Strategic deployment of healthcare resources during an epidemic is critical to prevent its rapid spread. In this work, we propose a tool that considers a complex network setting and some limited healthcare resources and then performs a statistical simulation to find the performance of various strategies.

3.1 Overview of the tool

To solve the above-given problem description, we have developed *DyNeMoC* as a generic tool which is customizable to simulate various problems pertaining to dynamic movement of agents on graphs. Examples of these problems are computational epidemiology (discussed later), variants of cops and robbers problem and diffusion of news/infection in a network. DyNeMoC has a multi-layered architecture which is depicted in figure 3.1. The tool is divided into two main parts: the core components and application based components. The core components consist of the agent module and the city module. It essentially holds information about the agents and the city. For the agent module, these include the agent id, current location and list of neighbours of each agent. For the city module, it includes the city id and agent id of the people currently residing in the city. For the simulation in various scenarios, an application-based module is provided to contain the information required for the specific application. In the case of epidemiological modeling, the application specific information consists of the original location (home city), health status (SIR), infection timestamp, scan status and neighbour list for the current timestamp. This work, *HStrat*, primarily focuses on the application of the tool in epidemiology.

3.2 Structure of HStrat

The HStrat simulator is designed to capture dynamic movement phenomenon across multiple cities. Building on the structure of DyNeMoC, HStrat has a configuration file containing the experimental parameters which are taken as input by the tool. The structure of HStrat is explained below.

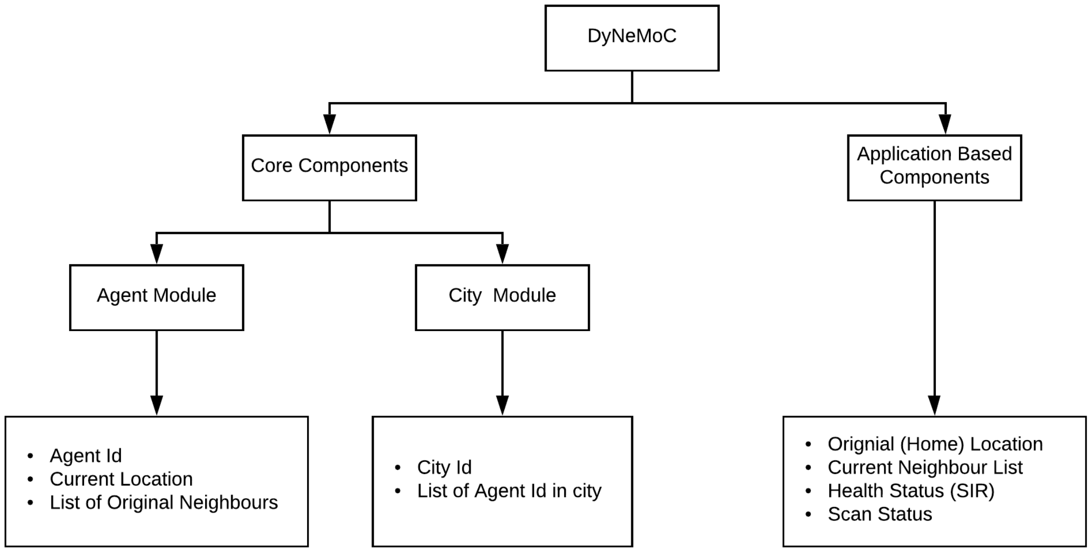

Figure 3.1: Architecture of DyNeMoC (for Epidemiological Application)

3.2.1 Network of cities

The model has a graphical representation of M cities, where M is input by the user. Implemented as a part of a core module, the vertices of an undirected graph are used to represent the different cities. The edges represent the path or connections between these cities. The tool, by default, considers all the cities to be interconnected as all major cities have a direct path to each other. However, the user is provided with the ability to modify the intercity connection as per his experiment. The total population of N people is distributed among these M cities based on some ratio. By default, the tool takes the value of M as 50 and distributes the population in the ratio of population in the top 50 cities in India.

3.2.2 Social Contact Network

The people within the cities form an intracity connection. This network implemented as a core component forms the social contact network of people. The model can be provided with an option of either having a random connection or following a BA Topology [17]. BA Topology is used by default since it best represents random scale-free dynamic networks. The BA network of every agent is updated at each time step based on his movement.

3.2.3 Mobility Model

As described in the problem statement, we assume that a person has a home city. However, dynamic movement is possible in the tool. Implemented in the application based component, the mobility model consists of movement policy of people. A custom transition probability matrix can be provided by the user as input. By default, the model a high probability (around 0.8) of a person staying at his home location and the remaining is divided equally among the remaining locations. When a person makes a transition, his current location is updated. Based on the current location, the person is assigned a new set of neighbours based on probability. However, when he returns to his home, he joins his original contact network with the same neighbours.

3.2.4 Transmission Model

The transmission model defines at what the rate the infection is transmitted among the population. Implemented in the application based component, the probability of infection depends on the number of infected neighbours and total numbers of neighbours in a persons contact network. The model allows the infection model to be modified or completely replaced with a rate based infection model. The tool by default has three states an agent can have, namely, susceptible, infected and recovered. The user has the choice to modify the states and replace the infection model with one of his choices, such as SEIR, SIRS, etc. The current model calculates the infection probability as:

$$P(i) = \frac{\text{No. of infected neighbours}}{\text{Total no. of neighbours}} \tag{3.1}$$

3.2.5 Recovery Model

The tool implements an application based recovery model which allows the infected people to recover naturally with some probability. However, since the chances of this happening in an epidemic scenario are less, the predefined value of recovery rate, γ, is very less with more focus on dynamic health care resources. The recovery model follows an exponential rate of increase of recovery probability with the increase in the number of days. The recovery probability is calculated as:

$$P(r) = 1 - (1 - (\frac{1}{t_r})^{d_i}) \tag{3.2}$$

where d_i is no. of days since infection & t_r is the average recovery period or $\frac{1}{\text{recovery rate}}$.

3.2.6 Healthcare Model

This tool focus on optimal resource allocation in times of epidemic. These resources could represent various things, including healthcare professionals, medicines, vaccinations, interventions and other various resources. We have implemented, what we call, dynamic Health Care Units (HCUs) which represent temporary medical facilities capable of treating people infected due to the epidemic. We assume that these HCUs are capable of testing as well as recovering the people. Once an infected person is given an intervention, we assume him to be cured of the infection as he cannot spread the infection. Each of these HCUs can test and cure a limited number of persons each day. This limit K, can be input by the user to represent the capacity of the HCUs. Once an HCU has treated the people, it can either stay in the same city for another day or move to a new city. In any case, each city is allocated only one HCU. The decision of the HCU staying or moving is another strategic decision which is taken care of by the tool. The model has a limited number of HCUs (usually lesser than the number of cities) as we want a resource allocation strategy which in real life will be limited and costly as well.

3.3 Implementation Details

The above tool developed from scratch in Java to make it platform independent and python has been used for generating input. An object oriented approach has been selected for the implementation as it is best suited to depict an agent based approach. As described above, the tool is highly customizable for the users to implement their scenarios and test the strategies. The tool has an Agent class which implements the agent module and a city class to implement city module. Objects of these classes can be created based on user input. These two classes form our graph structure and population for the model. The attributes of objects are then initialized based on some rules. The user has the ability to add new attributes and specify the initialization rules for them. The tool has other classes which implement the modules described above viz. Infection model, recovery model and HCU model. Each of these models can be customized as well. These modules update the attributes of the agents and provide us with the results of the simulation. Parallel programming using multithreading approach has been used to improve the performance of the tool. It exploits the fact the people in different cities are unrelated and independent and hence can be updated simultaneously. Figure 3.2 depicts how multithreading has implemented for parallel execution. Once the simulation is completed, the results are stored in the hard disk.

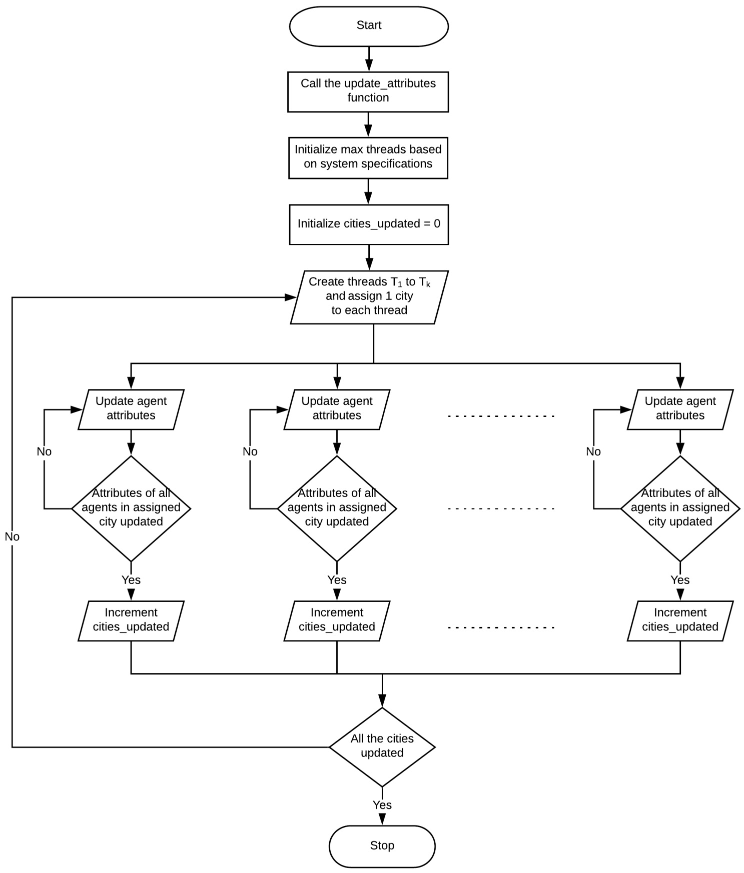

Figure 3.2: Multithreading as implemented in DyNeMoC

3.3.1 Simulation Process

This subsection describes a sample simulation run and also serves as a manual on how to run the tool.

- The tool takes a configuration file as input which contains all the input parameters such as the number of agents and their attribute values, movement thresholds, infected rate, recovery rate, file paths, etc.
- After deciding the input parameters, the graph input is generated using a python script (graph generation.py).
- After generating the graph structure, agents are generated based on input parameters using a shell script (Generate Input.sh).
- Once input is generated, the simulation is begun with a shell script (Simulation.sh).

The user part is completed here, and the tool begins the simulation. The simulation is done as follows:

- The system takes the input parameters and initializes all the agents. This includes changing the state of some random agents to infected based on input.
- The tool then runs the infection module and finds the probability of every agent being infected. Based on the probability, the next state of the agent is decided.
- After the infection module, the recovery module is called to find the agents who will recover naturally from the disease. Based on their recovery probability, their next state is set to recovered.
- The HCU module is called to sample some population and cure the infected people from the sample. All the infected people get cured by HCU.
- The movement module is called to find the next location of all the agents based on the movement probability.
- All the updates in the above steps are done in parallel for high performance.
- After all the simulations, the transitions take place and simulation begins for the next day.

Chapter 4

Resource Allocation

4.1 Problem Statement

We consider a real life setting of people residing in multiple cities interconnected with each other. All the people belong to exactly one of these cities which can be said to be his home city. Every person will have a social contact network consisting of people from his town. It is assumed that the person will interact with other people in his contact network. Despite having a home city, there is a chance that people travel outside their city for multiple reasons. This leads to having a model capable of depicting these dynamic movements as well. During an epidemic, some people randomly get infected by some pathogen, which causes a communicable disease. Now when an infected person interacts with other susceptible persons in his contact network, there is a probability he will spread the infection. This will lead to an epidemic situation. In order to control the epidemic, some healthcare resources are deployed with the capability of treating the infected people, which will slow down the spread of the infection. The objective of this work is to implement multiple deployment strategies and find out their performance under various constraints. This work has been published at the Sixth Social Networking Workshop, 11^{th} International Conference on Communication Systems & Networks (COMSNETS), 2019 [18].

4.2 Strategies for Resource Allocation

As described in subsection 3.2.6 above, the HCUs have a strategic decision to make about their movement. Their movement can be seen as a resource allocation problem that we want to solve with our tool. This section describes the staying algorithm followed by them, along with the allocation strategies.

4.2.1 Staying Algorithm

The HCUs have an upper limit of K people they can test and treat in a single day. However, based on the level of infection spread in the city, there will be a possibility that the city needs more medical supplies. In such a scenario, the HCU will decide to stay back for another day. The decision of whether an HCU should stay back or move to a new location is decided by an algorithm which tries to find the optimum threshold of the fraction of people who should be infected out of K tested people indicating that there is still a high level of spread of the infection. This means that if the fraction of infected people out of the tested sample is less than a certain threshold Th_m, then the HCU will infer that the infection level is not high and move to a new city. The algorithm followed is given below.

Algorithm 1: Staying Algorithm for HCU

Sample K people without replacement;
if *no. of infected people* $\geq Th_m$ **then**
| stay for another day;
else
| move to a different city;
end

4.2.2 HCU Allocation Strategies

HCUs being dynamic agents have the capacity of being allocated to a new city. However, the movement is crucial as this has a huge impact on the performance of the model in terms of the spread of an epidemic. Three strategies have been described below, which have been implemented in the tool. The user, like earlier, has the option of implementing and testing his own strategy.

- Random Allocation - The most primitive and basic is the random strategy where HCUs are allocated randomly to any city based on equal probabilities. However, the probability in this as seen traditionally is not $\frac{1}{m}$ as the current city and cities that have already been allocated an HCU are removed from consideration.

- Population Based Allocation - This strategy allocates HCUs among the cities based on weighted probability. The probability is in the ratio of the initial population of the cities. This means that a city with a higher population has a higher chance of being allocated an HCU. However, since a weighted probability is used and a top-K approach, the smaller cities also have a chance of being allocated an HCU. The idea behind this strategy is to exploit the consensus that a more populated city tend to have more infection. Following this approach, the probability for a city being allocated an HCU is:

$$p(city_i) = \frac{\text{No. of infected people in } city_i}{\text{Total no. of infected people}} \tag{4.1}$$

- Infection Based Allocation - The infection based allocation follows a different approach as compared to the above strategies. In the above strategies, we allocate HCUs based on the assumption that we are unaware of the overall level of spread of infection. It can also be seen as having no information about the number of resources needed in a city. However, if we do have some ground information, then we can use it to have a smarter allocation strategy. This strategy separates the testing capability of the HCU and provides each city with a local testing facility. This enables each city to test patients parallelly at each timestep irrespective of the location of HCU. However, these testing facilities cannot cure the patients as only HCUs have those resources. Instead, the testing information is passed to a central control where the allocation of HCUs is done. The allocation is done based on the weighted probability of the need for resources of a particular city with respect to the total requirement of resources (or

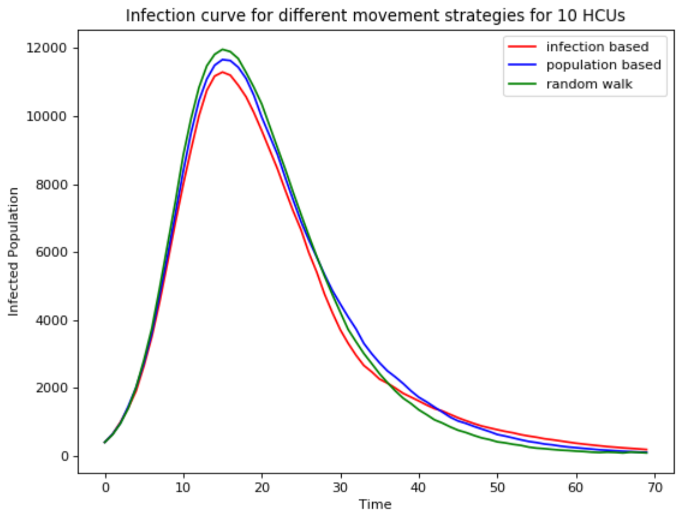

Figure 4.1: Infection curve for different HCU movement strategies for 10 HCUs

the number of infected people). Thus, a city will have the following probability of an HCU being allocated to it:

$$p(city_i) = \frac{\text{No. of infected people in } city_i}{\text{Total no. of infected people}} \tag{4.2}$$

4.3 Experiments & Results

4.3.1 Experimental Setup

To set up the experimental environment, we have generated a graphical structure of 50 cities with a total population of 20000 people. Agents are allowed to move among the cities with different probabilities. Experiments were conducted with the probability of people staying at home city, local probability, $P(home)$, as 0.80, 0.85, 0.90 and 0.95. For the HCUs, the capacity K is set to 100 and the movement thresholds Th_m were set to 0.2, 0.4, 0.6 and 0.8. The effect of varying the number of HCUs was observed with HCUs in the range between 0 & 10. The recovery rate γ, in all these experiments was set to 0.005. In the next section, we report simulation results in terms of epidemic curves while varying other parameters and allocation strategies.

4.3.2 Results

In this section, we describe the experiment conducted and the observations from the different results. We ran the experiments with above parameters having the initial infected population as 2% of the total population.

- **Experiment 1:** In this experiment, we assess the performance of the various allocation or movement techniques followed by the HCUs. The techniques under consideration are Random allocation, Population based allocation and infection based allocation techniques.

Observation: The presentation of 10 HCUs following distinctive movement procedures is assessed with estimations of $P(home)$ and Th_m as 0.95 and 0.8 separately. In Figure 4.1, we can see that the infection based movement methodology plays out the best since because of the communicate correspondence, it can achieve most infected cities and clear greatest disease. This is trailed by the population based movement technique, which can be best credited to the way that cities with more population are probably going to have more elevated amounts of diseases. The random movement technique plays out the most unfortunate as it doesn't pursue a fixed pattern and consequently neglects to clear an extensive piece of disease.

- **Experiment 2:** In this experiment, we analyze the effect of different HCU movement thresholds Th_m while keeping all the other parameters constant. The thresholds are used to decide the cutoff level at which the HCUs should move to a new city.
  Observation: The performance for Th_m of 0.2, 0.4, 0.6 and 0.8 with $P(home)$ as 0.95 for 10 HCUs following random allocation is assessed. From Figure 4.2, we can see that a higher value of Th_m results in a lower crest in the infection curve. This shows for the parameters considered in the experiment, HCUs should move to work in a different city if just a small fraction of the inspected populace tests positive for the disease.

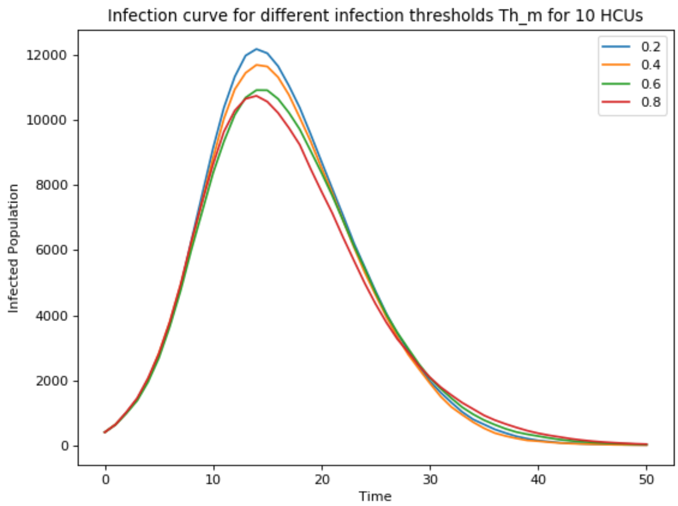

Figure 4.2: Infection curve for different movement thresholds, Th_m, for 10 HCUs

- **Experiment 3:** In this experiment, we assess the effect of movement of people for different local probabilities $P(home)$, on the execution of HCUs. This experiment analyzes how the movement of people among various affects the spread of infection.
  Observation: The presentation of HCUs for estimations of $P(home)$ as 0.80, 0.85, 0.90 and 0.95 with Th_m as 0.2 for 10 HCUs following random movement technique is assessed. From Figure 4.3, it tends to be seen that if there should arise an occurrence of an epidemic, it is better if the general population have lesser movement as it hinders the spread of infection.

- **Experiment 4:** In this experiment, we vary the number of HCUs and study its effect on controlling the experiment. Though an increasing number of HCUs should perform better,

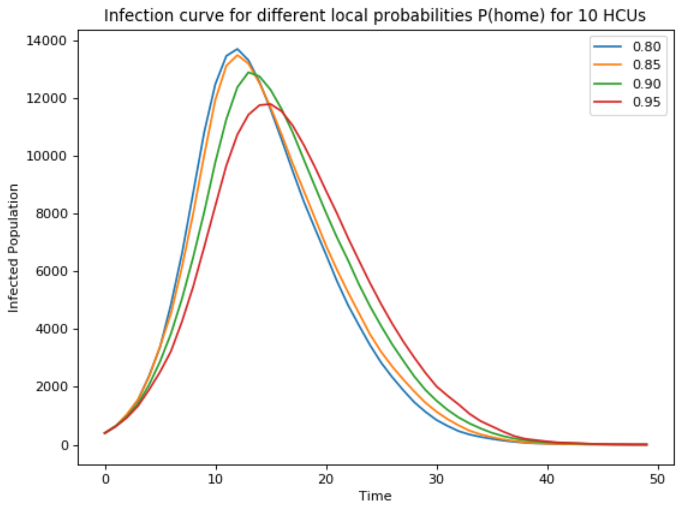

Figure 4.3: Infection curve for various local probabilities, $P(home)$, for 10 HCUs

there will be a cost associated with it, so a cost-effect analysis is required.
Observation: The presentation of Infection based movement strategy for no. of HCUs as 2, 4, 6, 8 & 10 is assessed with estimations of $P(home)$ and Th_m as 0.95 and 0.8 separately. In Figure 4.4, we can see that there is a diminishing in infection with the expansion in no. of HCUs. The bend with 0 HCU gives us the benchmark correlation for increment in the execution of HCUs. On closer perception, we can likewise observe that there is a postponement in the time to peak for infections as the HCUs increment in number.

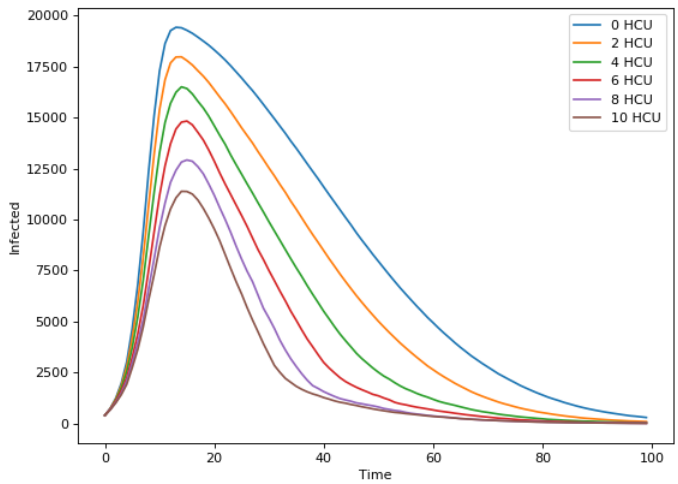

Figure 4.4: Infection curve for Infection based movement strategy for various HCUs

Chapter 5

Universal Healthcare

Healthcare can be defined as "the maintenance or improvement of health via the prevention, diagnosis, and treatment of disease, illness, injury, and other physical and mental impairments in people" [19]. It includes taking preventive measures or necessary medical procedures to improve an individuals health. Healthcare covers procedures ranging from administering basic medicine to advanced surgery. These administrations are normally offered through a health care framework made up of medical clinics and doctors. Healthcare is a quickly developing field, and a portion of the main areas of concentration in the healthcare business are given below:

- Clinical Healthcare - This incorporates the medical specialists, medical attendants, and collaborators who assess, diagnose and treat patients, and regularly give safeguard care to enable patients to keep up excellent health. A couple of instances of clinical specializations incorporate cardiology, dentistry, emergency operations, paediatrics, and so on.
- Recovery Therapist - These include medical professionals who help patients during recovery and rehabilitation from some physical or mental damage. These include rehabilitation counselors, physiatrists, psychologists, prosthetists, etc.
- Healthcare Administration - These are non-clinical roles which aim at management services in the healthcare area. These include hospital-, medical resource- or nursing management,
- Public Health - This area focuses on larger groups of the general population rather than a few individuals and tries to design universal health policies for the general public. These include epidemiologists, health policy managers, biostatistics, etc.

However, in general, when healthcare is talked about, it refers to clinical healthcare. The same convention is followed in this chapter from hereon. Even though healthcare may have different meanings based upon various environments and circumstances, healthcare systems usually have a primary healthcare center serving as the first contact point in the structure. Hospitals and medical facilities offering advanced care form the secondary and tertiary levels of treatment. Health care can be characterized as either public or private, which essentially means that the healthcare provider (like doctors and hospitals) can either be private or government. The various levels of healthcare are described below:

- Primary care - This alludes to the health experts who have the primary purpose as consultation for the patients. It alludes to crafted by health experts who go about as the principal purpose of discussion. Such an expert would typically be a general physician, such as a family doctor. Primary care is frequently utilized as the term for the health care benefits that assume a role in the neighbourhood network.

- Secondary care - Secondary Healthcare refers to the second tier of the healthcare system. In secondary healthcare, patients from primary health care are referred to specialists in higher hospitals for treatment. Secondary care incorporates intense care, fundamental treatment for a brief timeframe, sickness, damage, or other health condition. In India, the health centres for secondary health care include district hospitals and community health centre at the block level.

- Tertiary care - Tertiary Healthcare refers to the third level of the health framework, where particular consultative care is given more often than not on referral from primary and secondary medical care. Services like intensive care units and advanced diagnostic support form the critical features of tertiary health care. In India, tertiary care service is provided by advanced medical research institutes.

5.1 Healthcare in India

State governments are responsible for healthcare in India. State governments give healthcare administrations and health instruction, while the central government offers authoritative and specialized administrations. The health care framework in India is universal. Figure 5.1 gives an overview of the organization of the health system in India. That being stated, there is an extraordinary inconsistency in the quality and inclusion of medical treatment in India. However, there are numerous distinctions in quality among rural and urban regions. Rural areas frequently experience the ill effects of doctor deficiencies, and variations between states imply that occupants of some states, regularly have less access to satisfactory healthcare than inhabitants of other states. Rural areas in India have a deficiency of medical experts. About 74% of specialists are in urban regions that serve 28% of the population living in urban zones [20]. This is a noteworthy issue for rustic access to healthcare. The absence of medical resources makes people resort to private healthcare suppliers. There is also a lack of infrastructure for health benefits in rural territories, which makes healthcare resources scarce. Urban public hospitals have twice the number of beds as rural hospitals, which are deficient in provisions. The issue of healthcare access emerges in rural areas as well as in quickly developing small urban regions. Here, there are less accessible alternatives for healthcare administrations, and there are less sorted out government bodies. In this way, there is regularly an absence of responsibility and collaboration in healthcare divisions in little urban zones. Moreover, factors like different residential areas and socioeconomic status create even further inequalities in healthcare.

Similar to the rural-urban variations, healthcare provided by public and private sector varies a lot as well. The public health care system, which was initially created to provide a means to healthcare access regardless of socioeconomic status incorporates 18% of complete outpatient care and 44% of complete inpatient care [20]. Although India has a comprehensive health care framework, the

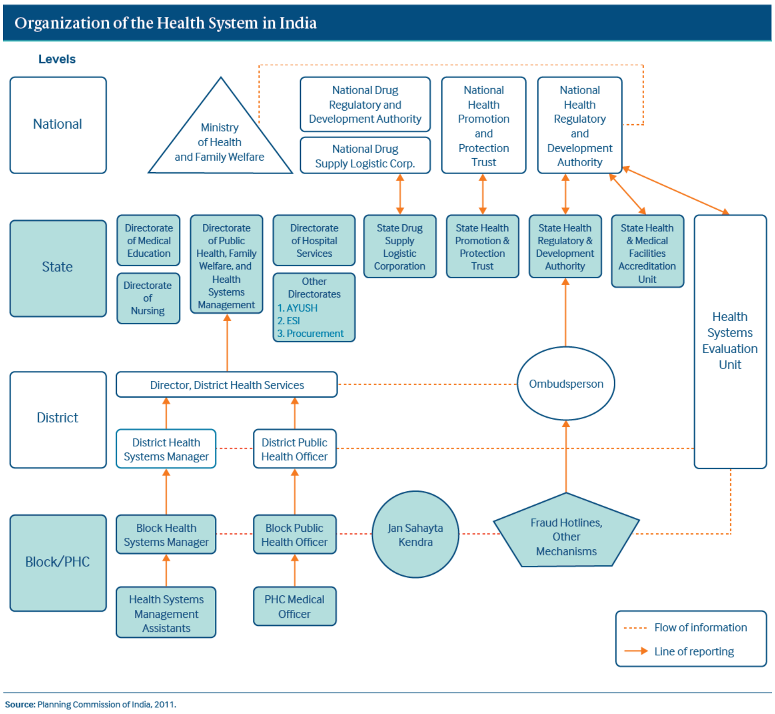

Figure 5.1: Healthcare in India

absence of inclusion and foundation in the public division implies that numerous Indians go to the private area for healthcare needs. On a basic level, government health administrations are accessible to all natives under the taxation based public framework. By and by, bottlenecks in getting to such services force families to look for private care, bringing about high personal expenses. Nonetheless, dependence on public and private healthcare areas differs altogether between states. A few reasons are cited for depending on the private as opposed to the public sector; the fundamental reason being low quality of care in the public division. As far as healthcare quality in the private segment is concerned, a report by Sanjay Basu et al. [21], showed that health care suppliers in the private area were bound to go through a longer duration of procedure with their patients and lead physical tests as a piece of the visit contrasted with those working in public healthcare. Due to several reasons, middle and upper-class people, in general, utilize public healthcare lesser than those with a lower standard of living. This alternative is more difficult for poor people as they cannot bear the high costs of private healthcare facilities. In any case, the higher personal expenses to cover the costs of private healthcare strain the financial resources of a family, making it challenging to maintain a basic standard of living. The higher expenses of the private division, be that as it may, are far from numerous Indians, producing interest for private health protection in India. The individuals who can stand to buy private insurance to help pay for the expenses of private health care. However, a large section of Indian citizens requires health coverage since personal expenses make up a significant part of medicinal costs in India.

5.2 Universal Healthcare Schemes

Universal healthcare aims at affordable healthcare to citizens of a state or country by providing financial protection. A healthcare scheme is built around delivering a specific benefits package to all individuals with the objective of improving healthcare outcomes by offering improved healthcare services at a lesser or even without a cost. The goal of universal healthcare is to create an environment in which every individual irrespective of socioeconomic status or any other factor has access to the best medical care available. The term universal healthcare, in any way, does not mean that coverage is provided for everything. Instead, it focuses on the fact that everyone has access to quality healthcare facilities. The crucial aspects of a scheme on which it can be evaluated are - the target demographic, medical facilities covered and the amount of medical cover provided by the plan. A universal healthcare scheme is described by the World Health Organization (WHO) as "a situation where all individuals and communities receive the health services they need without suffering financial hardship." [22] The Director-General of WHO described universal health coverage as the "single most powerful concept that public health has to offer since it unifies services and delivers them in a comprehensive and integrated way." [23] Some of the universal healthcare schemes which are currently implemented have been described in further subsections.

5.2.1 Patient Protection and Affordable Care Act (ObamaCare)

Popularly referred to as Obamacare or abbreviated as ACA, The Patient Protection and Affordable Care Act (PPACA) [24] [25] was enacted by President Barack Obama on March 23, 2010, as a national healthcare plan to provide affordable healthcare facilities in the United States of America (USA). This was a landmark scheme in the history of the USA as it did not just provide insurance but made a significant difference in the American healthcare industry as a whole. With the primary objective of providing affordable healthcare and reduce the health-related expenses of an ordinary person, Obamacare can be described as a combination of three major components, as described below.

- The major component of the ACA was providing the healthcare insurance itself. However, before this, a common practice which is followed majorly by all insurance companies was excluding people with pre-existing conditions. An alternative to this is provided by some companies to include the pre-existing conditions by paying the extra premium amount, which again is high. Due to this, people who most needed the insurance were the ones often without it leading to either lack of proper medical care or unaffordable expenses affecting their daily lives. The ACA allowed everyone, including those with pre-existing conditions, to get health cover at affordable prices and put a check to the rising costs of healthcare.

- However, since the people with pre-existing conditions were most likely to require medical care, the ACA needed to ensure that the insurance companies should be able to afford the costs. This was done by subjecting the people to one of the two conditions - that either everyone should have health insurance for at least nine months in a year or a tax would be levied on them. This increased the number of healthy patients taking health insurance as well and contributing to the premium costs rather than having a few individuals bear it.

- The above two components were concerned majorly between the public and the insurance companies. So the third component was the federal subsidies provided by the government to enable everyone to afford the required insurance.

5.2.2 Aarogyasri

With the aim of the government to achieve “Health for All”, Aarogyasri [26] was presented as a flagship healthcare scheme in 2014 by the then state government of Andhra Pradesh, India. However, post the splitting of the state into Telangana and Andhra Pradesh, this scheme was adopted by and adapted to the state of Telangana under the management of Aarogyasri Health Care Trust. For the state of Andhra Pradesh after the split, the scheme was revamped in 2016 to Dr NTR Vaidya Seva under the care of Dr NTR Vaidya Seva Trust. This healthcare insurance scheme provides an annual financial cover of up to ₹ 2 lakhs to families living below the poverty line (BPL) in the state. By paying the entire premium on behalf of all BPL families, the scheme aims to improve access to better medical facilities and treatment of ailments requiring hospitalization or surgery. The beneficiaries of the scheme are provided with an insurance card under the scheme which can be used for cashless transactions for all the procedures covered under this scheme. However, some of the high-end diseases are excluded under this scheme.

5.2.3 Pradhan Mantri Jan Arogya Yojana (PMJAY)

The PMJAY [27], popularly known as the Ayushman Bharat Yojana, is the national health protection scheme launched by the central government of India in 2018. Targeting medical care at all levels, the scheme carries forward the goal of India’s National Health Policy of 2017 of attaining universal health coverage. The PMJAY is a two-step scheme through which it aims for an increase in accessibility, availability and affordability of primary-, secondary- and tertiary-care health services in India. Figure 5.2 depicts the organization of how the PMJAY scheme will be implemented across India.

- Health and Wellness Centres - The first of the two initiatives of PMJAY is upgrading 1.5 lakh out of existing 1.8 lakh sub healthcare centers (SHCs) and primary healthcare centers (PHCs) to health and wellness centers by December, 2022. This would lead to an increase in the scope and range of services and facilities as compared to what is currently offered. This would make the healthcare services accessible within 30 minutes of walking distance by every community.

- Pradhan Mantri Rashtriya Swasthya Suraksha Mission (PM-RSSM) - The other component of PMJAY which has been dubbed as “the worlds largest government-funded healthcare (insurance) program”. Under the PM-RSSM, the government will provide a medical cover of up to ₹ 5 lakh for secondary and tertiary level hospitalization expenses. The current estimate of beneficiaries of this scheme is 535 million citizens, which is equivalent to 107.4 million families encompassing about 40% of the current population of India. This number goes beyond the traditional approach of targeting people living below the poverty line. Instead, with an all-inclusive vision, the government through census data has included the “vulnerable and deprived population”. Post the upgradation of SHCs and PHCs to HWCs; their benefits would be available for everyone without any cost. Further, the schemes and its benefits are applicable all over the country at all the registered private and public hospitals.

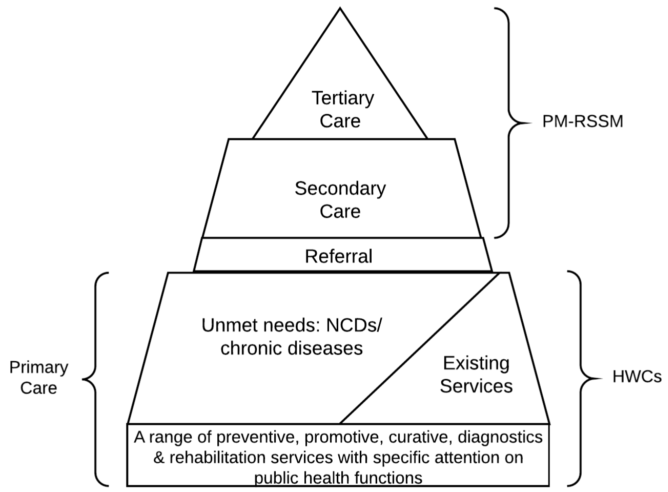

Figure 5.2: Pradhan Mantri Jan Arogya Yojana (PMJAY)

5.3 Problem Statement

As seen from the schemes described above in section 5.2, there can be several structural and implementation variations in the healthcare schemes. These variations can be attributed to various factors such as existing medical infrastructure, resources and facilities, their accessibility and affordability by the citizens, the funding mechanism for the schemes and the organizational structure of the system as a whole among other things. The performance of a scheme is the direct outcome of its design and implementation based on which it can be said to be successful or not. The objective of our work is to design different healthcare schemes and analyze the impact of structural differences on their performance. This is done by implementing two different models of healthcare schemes using our tool described earlier in chapter 3 of this work.

5.4 Modeling Healthcare Schemes

Taking a cue from real world healthcare schemes, we have modeled and implemented two variations of healthcare schemes. The models are then evaluated over different parameters and situations, and their performance is analyzed.

5.4.1 Direct Access Scheme (Model 1)

The first model is a direct access scheme, much like Aarogyasri, in which whenever a patient feels he is unwell, he can go directly to a hospital for treatment. After the patient is admitted to a hospital, he will be diagnosed for ailments and treated accordingly. Once the patient is treated, he is said to be recovered and sent home. However, similar to the real world, there is a probability that the treatment might be unsuccessful. In that case, the person is removed from the simulation. Another dimension to this model is the involvement of government agents which reimburse the treatment costs of the patients. Delay in reimbursement after some tolerant period would result in private in hospitals withdrawing their participation from the scheme and hence putting more load on the public healthcare system.

5.4.2 Referred Care Scheme (Model 2)

The other model is a referral based system where Health and Wellness Centers (HWCs) are established to provide primary care to the citizens. The HWCs will act as the first contact point in order to obtain the benefit of the scheme. Based on the initial diagnosis conducted by the HWC, the person will either be treated at HWC in case of minor treatments or will be referred to a hospital for secondary- or tertiary care. Similar to the previous model, the medical units here, i.e. the HWCs and the hospitals, will be able to treat accurately diagnose and treat the patient with some probability based on their efficiency.

5.4.3 Model design and implementation

The model is designed over the in-house developed tool, DyNeMoC. Using the same base structure, modules for healthcare schemes have been implemented to model the variations in their design. Both the models follow the same network of cities along with the intracity BA topology for social contact network. The transmission of infection, though, is done through a similar model, has a reduced infection rate since we want to analyze the performance during regular days and not during an epidemic. Since, we are concerned with the secondary - and tertiary level of care; the recovery model has been disabled by keeping the self-recovery probability as 0. The HCU model, which was used earlier, has now been replaced by a combination of public and private hospitals. However, the difference between the two models lies in the fact that the Referred Care Model has an additional HWC component. All the patients who require advanced care will be referred by the HWC. There is also a chance that due to an incorrect diagnosis, the HWC may wrongly refer a patient for advanced care or send him back. If the patient is incorrectly sent back, he remains in the infected stage and will have to come back to HWC again. Some of the assumptions taken are:

- Infection in Referred Care Model has two infected stages, i.e. early stage & advanced stage. HWCs can test all the patients but can treat only early stage patients.
- Private hospitals are more efficient, & more expensive.
- People choose or are referred to a hospital uniformly at random.
- Both the above schematic implementations have a cost associated with it. However, we are only concerned with running costs and not establishment costs.

5.4.4 Evaluation Parameters

The objective of these experiments is evaluating the effect that HWCs have on the efficiency of a scheme. Further, we also want to perform a cost vs efficiency analysis for the different models implemented. The efficiency of a medical unit is given by:

$$\text{Efficiency} = \frac{\text{No. of patients cured}}{\text{No. of potential patients visiting}} \tag{5.1}$$

The running cost of scheme is given as:

$$\text{Cost} = \sum_{0}^{N_i} \text{cost of individual treatment}; \tag{5.2}$$

where N_i is no. of infected people.

5.5 Experiments & Results

5.5.1 Experimental Setup

The experiments were performed for a population of 5 lakh people residing in 50 cities connected through a BA social contact network. The number of hospitals in a city is dependent on the population living in that city. For experiments, this value is set to 1 hospital for every 1000 people. Out of the total hospitals in a city, 35% are classified as public hospitals and remaining 65% as private hospitals. The capacity of public hospitals is in the range of 60 to 80 patients, whereas the capacity is 40 to 60 patients for private hospitals. The number and capacity of public and private hospitals are based on a combination of heuristics with publically available data [28]. The efficiency of public hospitals is assumed to be between 60% to 70%, whereas the same for private hospitals is assumed between 80%-90%. In the case of model 2, HWCs are set up in cities with a capacity of diagnosing 100 to 120 people daily with an efficiency of 65% to 75%. The number of HWCs per city is varied from 1 to 5 for every 1000 people to observe its impact on the performance of the scheme. The cost of treatment is assumed to be 1 unit in a public hospital and ten units in a private hospital. Treatment in an HWC is free of cost. All the above parameters being input parameters can be tweaked and set by the user. The model is simulated and averaged over a total of 100 runs.

5.5.2 Results for Direct Access Scheme (Model 1)

In this section, we will discuss the performance of the Direct Access Scheme (Model 1), which has no HWCs through a series of experiments. We will also observe the effect of delay in reimbursement on the scheme over three parameters - efficiency, cost and time.

- **Experiment 1:** In this experiment, we analyze the cost required to run the scheme with variation in the delay of reimbursement time.

  Observation: On increasing the delay in reimbursement, from Figure 5.3, we observe a reduction in the amount. This indicates that with an increase in delay, private hospitals are more likely to quit the scheme resulting in a larger population going to public hospitals for treatment. Since the cost of treatment at public hospitals is less, the overall cost required to finance the scheme reduces.

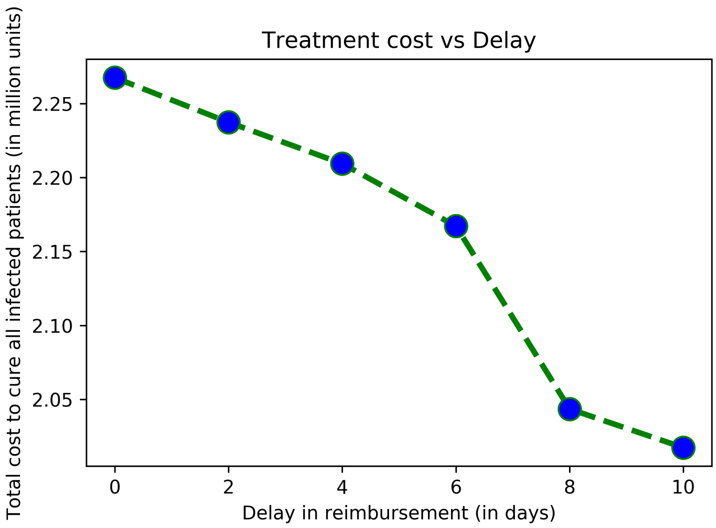

Figure 5.3: Treatment cost to cure all infected patients

- **Experiment 2:** In this experiment, we analyze the effect of delay in reimbursement over the time required to treat the population sample.

  Observation: The previous experiment displayed a reduction in cost as a more significant number of people visit public hospitals for treatment. However, with medical resources being constant, this would put a strain on them. From Figure 5.4, we can see that with an increase in delay and consequently, the number of patients, there is an increase in time taken to treat the patients. This longer recovery time is also because more people are expected to get infected within an increased period.

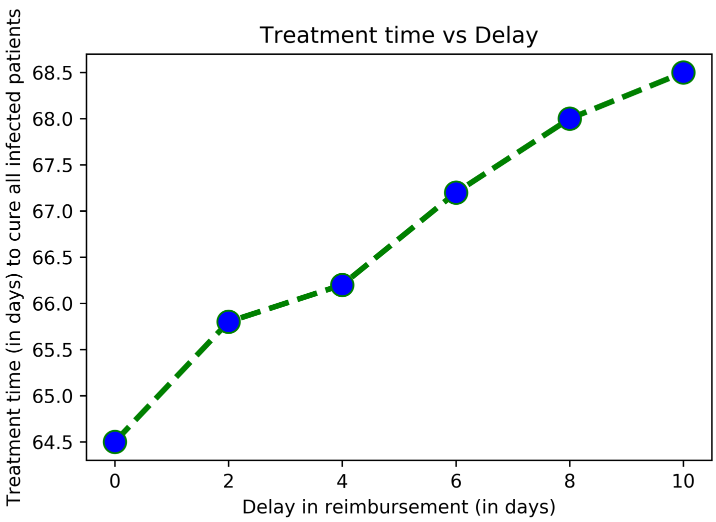

Figure 5.4: Treatment time to cure all infected patients

- **Experiment 3:** In this experiment, we analyze the performance in terms of efficiency of the scheme concerning a delay in reimbursement to the private hospitals.

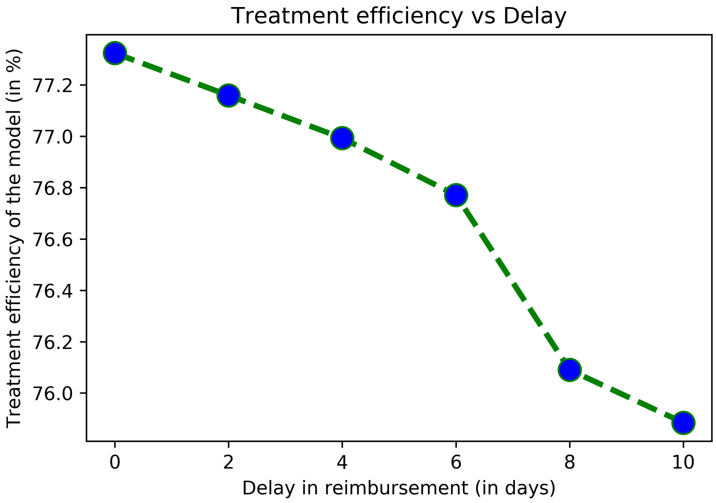

Figure 5.5: Treatment efficiency of the model

Observation: Since the efficiency of public hospitals is assumed to be lesser than that of private hospitals coupled with the fact that a reimbursement delay would result in more people approach public hospitals, the overall efficiency of the model will reduce. From Figure 5.5, we can observe a drop in the efficiency of the model, which becomes more apparent with the increase in the delay. This is also indicative of the fact that excessive pressure on a medical establishment will lead to lower efficiency.

5.5.3 Results for Referred Care Scheme (Model 2)

In this section, we will discuss the performance of the Referred Care Scheme (Model 2) having HWCs acting as primary healthcare centers. We will also observe the effect of variation of number of HWCs on the scheme over three parameters - efficiency, cost and time.

- **Experiment 1:** In this experiment, we analyze the cost required to run the scheme with an increase in the number of HWCs.

  Observation: From Figure 5.6, we can see that the scheme without using any HWCs (equivalent to Direct Access Scheme) is comparatively expensive. This difference in the cost can be attributed to the fact that advanced diagnosis is conducted unnecessarily for a lot of patients, especially in private hospitals. This leads to a spike in the cost of the scheme. However, with the introduction of HWCs, a lot of early-stage diseases are treated at primary centers saving the costs of advanced diagnosis and treatment. Also, this cost shows a decreasing trend as we can treat more number of people at HWCs at the initial stage itself, resulting in a lesser spread of infection and hence a lower cost.

- **Experiment 2:** In this experiment, we analyze the effect of the increase in the number of HWCs over the time required to treat the population sample.

  Observation: Figure 5.7 gives us an interesting insight into the time required to treat the population sample. The scheme without HWCs seems to perform better in terms of time taken

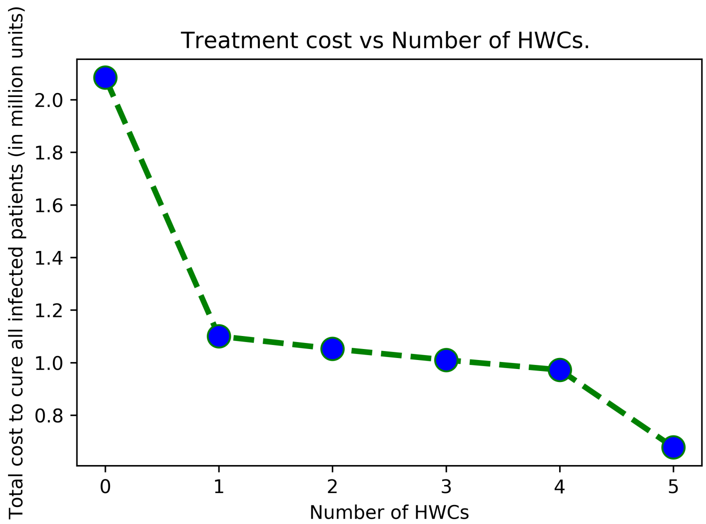

Figure 5.6: Treatment cost to cure all infected patients

as compared to 1 HWC. This is due to the bottleneck created by using an HWC. Since being the only referral point, people have to visit the HWCs necessarily. However, having a limited capacity, the HWCs cannot diagnose and treat more patients than its capacity resulting in people having to wait longer to get their turn. This will also result in more spreading of the infection and more patients. However, as the number of HWCs are increased, enough to treat the number of patients every day, the time taken improves and even betters the Direct Access Scheme.

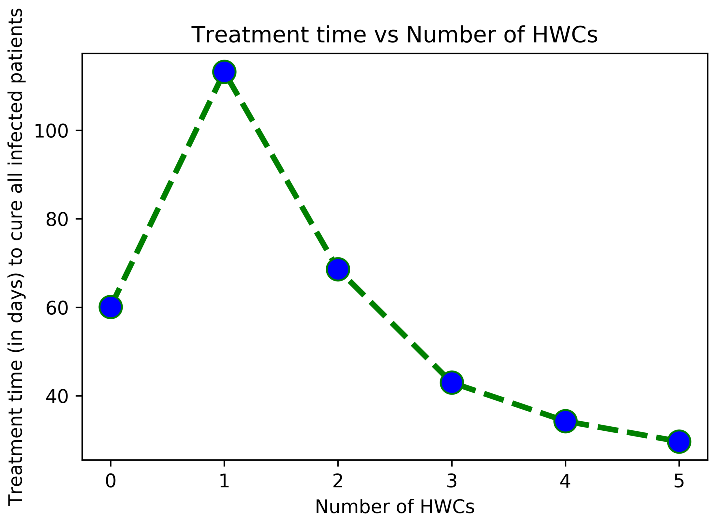

Figure 5.7: Treatment time to cure all infected patients

- **Experiment 3:** In this experiment, we analyze the performance in terms of efficiency of the scheme with respect to an increase in the number of HWCs.

  Observation: From Figure 5.8, we can observe that there is a decrease in efficiency when we

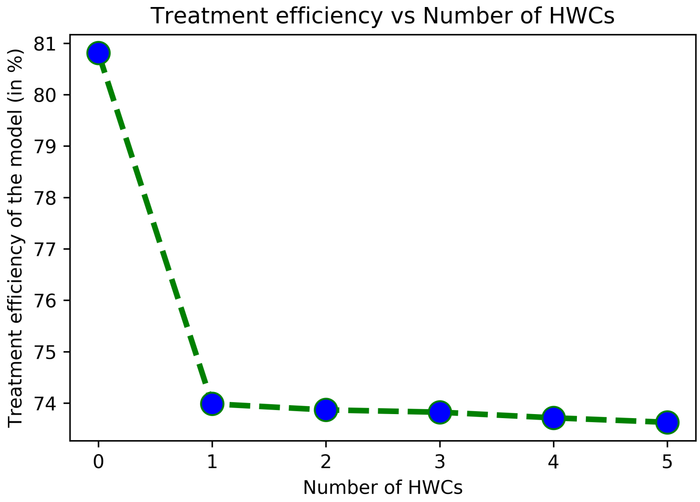

Figure 5.8: Treatment efficiency of the model

introduce HWCs in the model. The model is designed such that HWCs form a bottleneck for referral, and patients have to visit them necessarily. However, being only the primary healthcare centers, the efficiency of the HWCs is assumed to be low. Therefore, due to less efficient HWCs, there is a decrease in efficiency as it is dependent on all the successful treatments which are performed by HWCs, public- and private hospitals on all the potential patients visiting any medical unit at any time instant and for any number of times. Also, this decrease is not dependent on the number of HWCs since all of them have almost similar efficiency, and thus we see very little change with an increase in the number of HWCs.

Chapter 6

Conclusion

In this work, we have developed application modules for two critical healthcare applications over an in-house developed tool DyNeMoC. The first application was we studied was resource management and strategic intervention during epidemics. In our work, we designed the application module for the spread of infection and recovery from it. We then developed algorithms for resource allocation in the form of HCUs and tested various intervention strategies. The second application was designing the application module to simulate the various healthcare schemes. Taking the structure from schemes implemented or under planning in India, we designed two models with structural variations. We then tested the performance of both the schemes based on various parameters, which helped us understand their advantages and drawbacks under different scenarios.

6.1 Future Work

We believe that a significant direction to this work can be given with the inclusion of learning. The agents in both the multi-agent models can be given memory to remember and learn rules by interacting with its environment. By providing the agents with a learning capacity using a technique like Reinforcement Learning, it will be possible to simulate the spread of not only infection but human behaviour such as fear as well. This will help in making the model even closer to the real world. Using learning techniques, it will be interesting to see if the model can give better resource allocation strategies on the fly than what we have tested. Similarly, decision making in healthcare schemes can also be learning based rather than rule-based. A general learning system can be employed to evaluate the performance of schemes as well.

References

[1] Epidemiology, World Health Organization. Retrieved from `https://www.who.int/topics/epidemiology/en/`.

[2] Hippocrates. On Airs, Waters, and Places, The Internet Classics Archive, 400 B.C.E.

[3] A. Apolloni, V. A. Kumar, M. V. Marathe, and S. Swarup. Computational Epidemiology in a Connected World. *Computer* 42, (2009) 8386.

[4] J. d. Villalba. Epidemiologia española o Historia cronológica de las pestes, contagios, epidemias y epizootias que han acaecido en España : desde la venida de los cartagineses hasta el año 1801... Madrid : [s.n.], 1802.

[5] R. Ross. The Prevention of Malaria. *Public Health* 24, (1910) 287.

[6] W. O. Kermack and A. G. Mckendrick. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 115, (1927) 700721.

[7] D. Bernoulli and S. Blower. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Reviews in Medical Virology* 14, (2004) 275288.

[8] M. S. Bartlett. Deterministic and Stochastic Models for Recurrent Epidemics. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health. University of California Press, Berkeley, Calif., 1956 81–109.

[9] N. T.J Bailey. The Mathematical Theory of Infectious Diseases and Its Applications. *SERBIULA (sistema Librum 2.0)* 34.

[10] M. Thrusfield. Veterinary epidemiology. Wiley Blackwell., 1978.

[11] T. Habtemariam, A. Ghartey-Tagoe, E. Mamo, and V. Robnett. Epidemiologic modelling of diseases a case example using Schistosoma and Trypanosoma. *Mathematical and Computer Modelling* 11, (1988) 244249.

[12] C. Macal and M. North. Tutorial on agent-based modeling and simulation. *Agent-based Modeling and Simulation* .

[13] S. C. Bankes. Agent-based modeling: A revolution? *Proceedings of the National Academy of Sciences* 99, (2002) 71997200.

[14] R. Beckman, C. Barrett, K. B. Berkbigler, K. R. Burris, B. Bush, S. D. Hull, J. M. Hurford, P. Medvick, D. Kubicek, M. Marathe, J. D. Morgeson, K. Nagel, D. J. Roberts, L. L. Smith, M. J. Stein, P. Stretz, S. J. Sydoriak, K. Cervenka, and R. Donnelly. Transportation ANalysis SIMulation System (TRANSIMS) - The Dallas-Ft. Worth Case Study 2002.

[15] W. Brian Arthur, S. Durlauf, and D. Lane. The Economy as an Evolving Complex System II 1997.

[16] SimStore Layout Simulator. Retrieved from `https://www.simworld.co.uk/simstore-layout-simulator.html`.

[17] A.-L. Barabsi and R. Albert. Emergence of Scaling in Random Networks. *Science* .

[18] Y. Ramesh, N. Anand, and M. V. Panduranga Rao. Statistical Model Checking for Dynamical Processes on Networks: A Healthcare Application. In 2019 11th International Conference on Communication Systems Networks (COMSNETS). 2019 720–725.

[19] Health care. Retrieved from `https://en.wikipedia.org/wiki/Health_care`.

[20] J. Thayyil and M. C. Jeeja. Issues of Creating a new Cadre of Doctors for Rural India. *IInternational Journal of Medicine and Public Health* .

[21] S. Basu, J. Andrews, S. Kishore, R. Panjabi, and D. Stuckler. Comparative performance of private and public healthcare systems in low- and middle-income countries: a systematic review. *PLoS medicine* .

[22] Universal health coverage (UHC). Retrieved from `https://www.who.int/en/news-room/fact-sheets/detail/universal-health-coverage-(uhc)`.

[23] Matheson and Don. Will Universal Health Coverage (UHC) Lead to the Freedom to Lead Flourishing and Healthy Lives?; Comment on "Inequities in the Freedom to Lead a Flourishing and Healthy Life: Issues for Healthy Public Policy". *International Journal of Health Policy and Management* .

[24] J. M. Adkinson and K. C. Chung. The patient protection and Affordable Care Act: a primer for hand surgeons. *Hand clinics* .

[25] The Patient Protection and Affordable Care Act. Retrieved from `https://www.dpc.senate.gov/healthreformbill/healthbill04.pdf` 2010.

[26] Aarogyasri Scheme, Government of Telangana. Retrieved from `https://www.aarogyasri.telangana.gov.in/`.

[27] C. Lahariya. Ayushman Bharat Program and Universal Health Coverage in India. *Indian Pediatrics* 55, (2018) 495506.

[28] National Health Profile 2018, Central Bureau of Health Intelligence, Government of India. Retrieved from `https://cdn.downtoearth.org.in/pdf/NHP-2018.pdf`.