# NETWORK PLANNING OF WALK-IN CLINICS ON ROADSIDES IN AFRICA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

SINE TAYMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

AUGUST 2013

Approval of the thesis:

**NETWORK PLANNING OF WALK-IN CLINICS ON ROADSIDES IN AFRICA**

submitted by **SINE TAYMAZ** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Murat Köksalan _____
Head of Department, **Industrial Engineering**

Assist. Prof. Dr. Cem İyigün _____
Supervisor, **Industrial Engineering Department, METU**

Assoc. Prof. Dr. Zeynep Pelin Bayındır _____
Co-supervisor, **Industrial Engineering Department, METU**

**Examining Committee Members:**

Assoc. Prof. Dr. Canan Sepil
Industrial Engineering Department, METU _____

Assist. Prof. Dr. Cem İyigün
Industrial Engineering Department, METU _____

Assoc. Prof. Dr. Zeynep Pelin Bayındır
Industrial Engineering Department, METU _____

Assoc. Prof. Dr. Nico P. Dellaert
Industrial Engineering Department, TU/e _____

Assist. Prof. Dr. Özgen Karaer
Industrial Engineering Department, METU _____

**Date:** _____

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last Name: SINE TAYMAZ

Signature :

# ABSTRACT

NETWORK PLANNING OF WALK-IN CLINICS ON ROADSIDES IN AFRICA

Taymaz, Sine
M.S., Department of Industrial Engineering
Supervisor : Assist. Prof. Dr. Cem İyigün
Co-Supervisor : Assoc. Prof. Dr. Zeynep Pelin Bayındır

August 2013, 114 pages

This study discusses the problem of finding the optimal location of "walk-in clinics" specialized in healthcare along the transportation lines that would enable maximum coverage along the roads for the mobile populations and their related local communities. As the mobile populations are flowing on the routes unremittingly, the problem differs from other location problems. Every member of the mobile population would require a specialized service for their diseases and needs to access these services in a continuous manner along the roads, without any disruption. Therefore, the location of clinics should be adapted regarding these requirements and maximum continuum of care should be ensured for the demand populations. Additionally, as a results of the uncertain nature of the mobile demand, the risk associated with the lack of continuum of care provided to the population is an important component in the problem. While ensuring maximum level of continuum of care, the risk involved in the transportation lines which appears as a variation in mobile demand should not be overlooked.

Problem has been solved with the idea emerging from flow interception and coverage problems. Aims of maximizing the intercepted flow and coverage of roads are considered as the objectives of the model. The problem has been developed as Mixed Integer Program and it is

shown that model is capable of handling the different requirements resulting from the demand of mobile and static populations. The mathematical formulation is extended for the stochastic case, relaxing the assumption that demand is known and certain. Risk-averse measures are included in the mathematical formulation with the application of Conditional-Value-at-Risk risk measure. It is observed that with a stochastic model, when uncertainties are present in the network, with the help of the risk-averse measure, the risk on the network is kept under control and the amount of demand that is subject to risk is decreased.

Keywords: humanitarian logistics, coverage, facility location, continuum of care, coherent risk measures, CVaR, mixed integer programming

# ÖZ

AFRİKA'DAKİ ULAŞIM HATLARI ÜZERİNDE KLİNİK AĞI PLANLAMASI

Taymaz, Sine
Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Yöneticisi : Yrd. Doç. Dr. Cem İyigün
Ortak Tez Yöneticisi : Doç. Dr. Zeynep Pelin Bayındır

Ağustos 2013 , 114 sayfa

Bu çalışmanın amacı gezici populasyonlar ve onların yakınlarında yer alan sabit populasyonlar için maksimum kaplama alanı sağlayacak, yol kenarları üzerine kurulmuş bir klinik ağı planlamaktır. Gezici populasyonların yollar üzerinde sürekli olarak hareket ediyor olması, çalışmada incelenen problemi literatüredeki diğer tesis yerleştirme problemlerinden farklı kılmaktadır. Gezici populasyondaki her bir birey, taşıdığı hastalıklar için özelleşmiş bir servis talep etmektedir ve bu servislere üzerinde hareket ettiği yol boyunca kesintisiz bir şekilde ulaşabilmesi gerekmektedir. Dolayısıyla, kliniklerin yerleştirildiği noktalar bu özelleşmiş servis ihtiyaçlarını ele alınalarak planlanmalıdır ve hastalara bütüncül bir bakım uygulanmalıdır. Ek olarak, gezici populasyonların içinde bulunduğu belirsiz ortam, ihtiyaç duyduklarları servislerin tam olarak yerine getirilmesine engel teşkil etmektedir ve bu belirsizlik problem içerisinde önemli rol oynamaktadır. Servis düzeyini en üst seviyeye hedeflerken, gezici populasyonların yol üzerlerindeki talep miktarındaki varyasyonlar unutulmamalıdır.

Bu çalışmada ele alınan problem akış kesişimi ve kapsama problemlerinden ortaya çıkan fikirlerle çözülmüştür. Oluşturlan modelin amaç fonksiyonu kesilen akış miktarını ve kapsanan

insan sayısını enbüyüklemektir. Model Karma Tamsayılı Programlama olarak geliştirilmiştir ve oluşturulan model gezici ve sabit populasyonların farklı ihtiyaçlarına karşılık verebilmektedir. İkinci aşamada, oluşuturulan matematiksel model, talebin sabit ve kesin olduğu varsayımını esneterek, rassal durum icin adapte edilmiştir. Bu formulasyon için çekinceden kaçınan yöntemler kullanılmıştır ve bağdaşık çekince ölçülerinden Koşullu Çekince Değeri kavramı uygulanmıştır. Rassal model sayesinde, planlama ağında belirsizlikler mevcut olsa bile, çekince kontrol altında tutulmuştur ve çekince altında olan talep sayısı azaltılmıştır.

Anahtar Kelimeler: insani lojistik, kapsama, tesis yerlestirme, bağdaşık çekince ölçüsü, koşullu çekince değeri, karma tamsayılı programlama

To my grandparents Nafiye Gune and Ali Haydar

# ACKNOWLEDGMENTS

I have completed this thesis with the valuable support and motivation from many people. Initially, I would like to express my deepest gratitude to my mentors Dr. Cem İyigün and Dr. Pelin Bayındır, Middle East Technical University, and Dr. Nico Dellaert, Eindhoven University of Technology, for continuously guiding me throughout the thesis. Their support, advices and motivation were very valuable for me and it has been a great opportunity to work on this thesis together with them.

My sincere thanks also go to the examining committee members, Dr. Canan Sepil and Dr. Özgen Karaer for generously giving their time to review the thesis and offering valuable comments.

I am also grateful to Eindhoven University of Technology and The Scientific and Technological Research Council of Turkey (TÜBİTAK) for their financial support.

I would also like to express my deepest thanks to North Star who have supported my thesis by offering a valuable internship period in their organization. Special thanks must go to Richard van Dijk who have supported my research all the time by patiently answering my questions and putting full time observation over my study.

I would also like yo thank to my dear friends Gülcan Yayla, Esra Erişen, Ceren Kocaman and Okan Celep for their continuous motivation and support throughout the thesis. Even at the hardest times, they were always there to support me by putting a smile on my face and showing their trust in me. Furthermore, I would also like to thank to my friends in Eindhoven, Gizem Sezen and Mark van Logtestijn, for spending great time there.

Last but not least, I would like to thank to my mother and father for their endless love and support during my whole life. I am also grateful to them for their valuable insights and endless interest on my study, which almost put them in the position of an additional mentor of my thesis. With their motivation, I have learned to get through the agonizing times and be positive in every step.

# TABLE OF CONTENTS

# LIST OF TABLES

TABLES

# LIST OF FIGURES

FIGURES

# CHAPTER 1

# INTRODUCTION

Every year, millions of African people are dying as a result of the high burden of disease that is settled upon the African societies. Even though the diseases are preventable and treatable, the presence of challenges such as weak health systems, inefficient planning of resources, limited access to healthcare services, poverty and frequent occurrence of nature and human made disasters, put strong obstacles for tackling the problem. Existence of such challenges shows that there are still many fields that require intervention and actions need to be taken in order to increase the life expectancy and reduce the deaths associated with diseases.

There are many institutions and organizations that are devoted to tackle the problem of health in Africa. For example World Health Organization, which is the authority for coordinating health within United Nations, has been working on health related issues in Africa with the mission of achieving highest level of health for all people. Other branches under United Nations as United Nations Development Program has also been focusing with the help of Millennium Development Goals where 3 out of 8 goals are related with child mortality, maternal health, HIV/AIDS, malaria and other diseases. In addition to institutions like United Nations, there exist other organizations, whether working together with institutions or not, that are also committed to the problem.

One part of the health related problems is related with the mobile populations that are situated in Africa. Mobile populations are groups of people, which are not necessarily located in a defined place but rather move through transportation lines and change positions. The main group of people under the mobile population can be given as truck drivers. Africa has many hubs, harbors and airports making large number of goods to flow over the continent every day requiring extensive amount of truck drivers who are working in hard conditions, sepa-

rated from home almost for every day of the month. Another group of people under mobile population is the sex-workers, who are also transporting occasionally with the truck drivers.

These two groups of people, truck drivers and sex-workers, interact at the "hot-spots" along the transportation network, where truck drivers spend long waiting hours. With the combination of being away from home, working under hard conditions and other factors, at the hot-spots the sex activity peaks. This situation leads to spreading of communicable diseases especially high impact diseases as HIV/AIDS, malaria, tuberculosis and sexually transmitted infections. As the population moves along the network, they act as vectors of transmission, carrying the diseases along with themselves and spreading in other regions of Africa. Furthermore, the local populations along the transportation lines are also affected by this situation, with the transmission of diseases.

Even though the mobile populations are having higher prevalence rates compared to other populations in the region and requiring healthcare services extensively, it is not very easy for them to access the healthcare service they need. The truck drivers work on strict time windows making it difficult to access the hospitals that are far away from the transportation lines which increases the travel time and requires them to deviate from their route. Furthermore, the road conditions to access the hospital are not always suitable for trucks to enter and parking is also a problem. The concept of Roadside Wellness Centers (RWC), which are small clinics that are situated along transportation lines and hot-spots, developed by North Star Alliance was an idea put in practice to deal with the issue of health problems related with mobility. By putting RWCs along the transportation lines and hot-spots, North Star aims to cover the continent by offering a continuous service that mobile populations can access.

In this thesis, the idea of the RWC and their operations is taken as a starting point and a model that would optimally plan the transportation network with the establishment of walk-in clinics is studied. Planning of the walk-in clinics along the transportation lines, however, is not a very straightforward task. By establishing the walk-in clinics, the initial aim is to give benefit to people as maximum as possible. In this context, what is defined as benefit has to be determined carefully and planning should be based on this definition.

One of the measures for benefitting the people can be given as capturing as many patients as possible in the walk-in clinics. For this purpose, the hot-spots having high number of waiting duration and population can be considered as a suitable candidates. Another aim is

to provide continuous care for the mobile population as they move along the transportation lines. Certain services for certain diseases are not one time and would require couple of visits to clinics in given time frames. In order to achieve adherence to protocols and complete the treatment of the patients, the walk-in clinics should be available when required. Every service of a disease has different protocols and therefore, the locations of the walk-in clinics require adjustments given these specifications. Accordingly, while planning the network, the questions of where to establish the walk-in clinics and which services to offer in these clinics that would ensure continuous care and capture maximum number of patient visits are asked. With the developed model, effects of specialized conditions on the planning of walk-in clinic locations is investigated in order to evaluate whether the model responds to the initial requirements.

An additional question that is studied in this thesis is how to handle the uncertainties that are included in the problem itself. Since the model is related with healthcare and human lives, risk is an important component and the model requires adapting the planning of walk-in clinics accordingly with the development of an appropriate stochastic model. The effects of risk on the optimal solution given the conditions of high and low risk along the network are another part of the study that is included in the thesis. It is also analyzed how the risk factor affects the performance of the model and benefits the target population.

The thesis starts with a literature review based on two subjects where in Section 2.1, the network location problems and in Section 2.2 stochastic programming about risk are overviewed. In Chapter 3, the problem definition is made by describing the motivation behind the study, the environment and the problem that is studied. After explaining the problem, in Chapter 4 problem formulations are described in detail which initially starts with a deterministic model and followed by a stochastic model in Chapter 5. The results of the both formulations are analyzed in computational studies given in Chapter 6. The thesis is concluded in Chapter 7.

# CHAPTER 2

# LITERATURE REVIEW

The literature review is given in two parts. In Section 2.1, literature review about network location problems is described. This is further divided as the flow interception problems and the coverage problems in Section 2.1.1 and 2.1.2, respectively. In the second part of the literature review, the focus is on stochastic programming. In Section 2.2.1 methods to include risk in the stochastic optimization is explained. A description about coherent risk measures is given in Section 2.2.2. Finally, out of the coherent risk measures, Value-at-Risk and Conditional-Value-at-Risk is discussed in Section 2.2.3.

## 2.1 Network Location Problems

### 2.1.1 Flow Interception Problems

Network location problems are represented with a set of arcs and nodes and the objective is to find the optimal location of a facility that is chosen from a discrete or continuous set of possible locations. The network locations problems can vary with respect to the different objective functions, different shape of facilities, number of types of facilities and so forth [30]. In these studies, initially it was assumed that the demand was occurring at static points where the customers would be willing to travel to the facility within certain limits. In 1990s, the interest has been also put in flows, where the facilities were no longer a static point to attract demand flows, but rather serve as interception points for the demand occurring on the flows. These types of problems are referred as flow interception problems which is discussed further in this section.

Table 2.1: Notation used for flow interception problem

| | |
|---|---|
| Sets: | |
| $Q$ | Set of all origin and destination pairs, $q = \{1,2,\ldots,Q\}$ |
| $J$ | Set of all potential facility sites, $j = \{1,2,\ldots,J\}$ |
| $j \in q$ | Set of all nodes on the shortest path between O-D pair |
| Decision Variables: | |
| $x_q$ | Equals 1 is flow $q$ is intercepted by a facility, 0 otherwise |
| $y_j$ | Equals 1 if a facility is opened at location $j$, 0 otherwise |
| Parameters: | |
| $f_q$ | Volume of flow between origin and destination on flow $q$ |
| $p$ | Upper bound on number of locations to be located |

In [34], the basic formulation for the generalized flow-interception location-allocation model is given. The notation for the problem is summarized in Table 2.1 and the problem formulation can be given as follows:

$$\text{Max} \quad \sum_{q \in Q} f_q x_q \tag{2.1}$$

$$\text{s.t.} \quad x_q \leq \sum_{j \in q} y_j \quad \forall q \in Q, \tag{2.2}$$

$$\sum_{j \in J} y_j = p, \tag{2.3}$$

$$x_q \in \{0,1\} \quad \forall q \in Q, \tag{2.4}$$

$$y_j \in \{0,1\} \quad \forall j \in J. \tag{2.5}$$

In this optimization problem, with the objective function (2.1), the volume of flow that is intercepted is maximized. A flow can be categorized as intercepted only if a facility is opened along the flow, which is ensured by constraint (2.2). The number of facilities that can be opened on the network is limited by an upper bound, $p$, as written in (2.3). The last two constraints, (2.4) and (2.5), are for ensuring the decision variables are binary.

One of the oldest articles for flow interception problems is [7], where an environment for placing discretionary service facilities (services that are consumed in an optional way, meaning that the customer does not embark a tour only to receive the service) is considered. In this setting, customers are on their pre-planned trip and if they pass a discretionary service along the path, they have potential to stop in the service center. Example for such services can be given as gasoline stations or automatic teller machines. The objective of the problem is to

maximize the flow of potential customer that is intercepted by the services by deciding on the location of services. When calculating the customer visits, multiple passing of a customer is not counted separately; i.e., each customer is captured only once. In the model traffic flows are assumed to be known.

Initially, the system is defined as a transportation network of arcs and nodes. Before setting up the mathematical model, it is proved that the optimal set of locations will exist on the nodes of network as it has the maximum amount of flow that can be intercepted. The problem is solved by two approaches, which are greedy heuristic and exact branch-and-bound algorithm. The greedy heuristic provides near optimal solution, which can also be used as initial solution for the branch-and-bound algorithm.

This initial model has provided a basis for the extensions and with changes in objectives and constraints, special cases of the flow interception problem have been studied. As an extension to [7], in [5], location of units on network where multi counting is also considered is discussed by Berman et al. This asserts that, the consumption of discretionary service by consumers along the path does not only depend on the existence of facility but also on the number of facilities located. The consumption by consumers is assumed to be a non-decreasing concave function.

Two different problem definitions are used in this paper. First one aims to minimize the number of facilities required to ensure the maximum level of consumption and the other one is maximizing consumption given a restriction on number of facilities to be opened. For the first problem, a polynomial algorithm on a tree is formulated. For the latter, it is proven that the problem is NP-hard even on a tree therefore, a greedy heuristic with its worst-case performance analysis is presented. In case of restricting the problem to having at most one facility in the nodes, both problems become NP-hard. Even though the authors did not attempt to solve them through the computer, they expect that computer code that combines heuristic and integer programming can be successful for solving the problem.

In [7], the location of discretionary facilities are planned to be on the pre-planned route of the customers, believing they would not change their route. This assumption of planning the location exactly on the pre-planned route is relaxed in [6] by considering the fact that the customer may also deviate from the path for visiting certain facilities.

Three sub-problems are examined in this study. The first one is a delta coverage problem where a customer passes through a pre-planned route of the trip and if a service is not encountered, the customer can deviate a maximum distance of *delta* (given that delta is a predefined maximum distance value) from the path, which in the end increases the length of the tour by two deltas. In case the facility is located more than delta distance away, consumers do not prefer to visit the facility. The other one is maximizing market size problem where the customer becomes less inclined to visit a facility if the distance from the pre-planned trip is increased. Therefore, the probability of visiting a facility is a decreasing convex function of the distance traveled. For both problems, objective is to locate a given number of centers to maximize the total number of visiting customers to the service center. The final type of problem is minimizing inconvenience problem where customers will deviate from the path to receive the service, making them not discretionary anymore and the objective becomes minimization of the distances traveled by the customers. Delta coverage problem can essentially be solved by the same algorithm given in [7] and the minimizing inconvenience problem is $m$-median problem. For maximizing market share problem, a specific greedy algorithm and a branch-and-bound based solution is given as an extension to approach [7].

In [12] by Dandan et al., the one type facility interception problem is extended with multiple types of flows that have mutual positive influence on each other such as consumption of certain facility increases with the co-location of an other . An example for such multi-type flow is given as the gasoline stations and the convenience stores. The flow is defined separately for each type of service, including the combination of multiple services. In the paper, flow interception problem is solved for multi-type flows that involve multi-purpose customer flows and their mutual positive influences. Every service is assigned with a certain profit and the overall objective of the model is to maximize the mutual influence with the placement of locations. Since the formulation is NP-hard and there is no polynomial time algorithm, heuristic solutions are made. As greedy algorithm heuristics are known to be effective in these types of problems, variant of greedy heuristic is implemented and near optimal solution is obtained.

Zeng further investigates the flow interception by introducing the pick-up problem and show how preferences can have influence on the locations [35]. The model is divided under four scenarios as “video”, “hamburger”, “coffee” and “pizza” where each of the items needs to be obtained by the customer at a certain part of the path. Out of these scenarios, “video” does not require a specific location on the flow and can be obtained at any point situated on

the flow. Each of the locations would return the same benefit from acquiring for "video" scenario. However, for the "coffee" scenario, the situation is different, as travelers would require having it at the beginning of their trips. Therefore, the "coffee" facilities are associated with more benefit for those that are located closer to the origin. The objective changes to maximizing benefit, where benefit of intercepting one unit of flow at a specific node is added in the objective function calculations.

### 2.1.2 Coverage Problems

A coverage problem can be described as within a network of facilities as customers, every customer has a certain radius to travel for the facility. The aim is to maximize the gain from locating the services with a certain upper bound on the number of facilities or covers the whole customer demand to optimize a certain performance measure [19].

A recent review of coverage problems literature is given by Farahani et al. [15]. The survey includes a comprehensive summary of coverage models starting from the year 1993. In the article, coverage problems are categorized into two parts as *Set Covering Problem* (SCP) and the *Maximal Covering Location Problem* (MCLP) where in the first topic, coverage is required and in the latter, coverage level is maximized. Even though SCP is also examined, the rest of the focus of this literature review about coverage problems will be on MCLP as it is decided to be more relevant to the scope of the problem. This article is used as a guideline for analyzing the important coverage problems in general. Afterwards, those that are considered to be important and relevant are further studied.

In the article by Church and ReVelle, a very basic model for coverage problem is developed by defining a critical distance from the demand locations [10]. This groups the demand locations as either covered or not by looking at their distance from the facilities and the required coverage radius. The model tries to maximize the number of covered demand given the demand amount in each demand locations and limit of number of facilities to be opened.

Table 2.2: Notation used for coverage problem

| | |
|---|---|
| Sets: | |
| $I$ | Set of demand nodes |
| $J$ | Set of facility nodes |
| $N_i$ | Set of facilities that can cover node $i$ |
| | |
| Decision Variables: | |
| $x_j$ | Equals 1 if a facility is located at site $j$, 0 otherwise |
| $y_i$ | Equals 1 if the demand node $i$ is covered, 0 otherwise |
| | |
| Parameters: | |
| $a_i$ | Population at node $i$ |
| $p$ | Upper bound on number of facilities to be located |

The notation for the problem is given in Table 2.2 and formulation for the model is given as:

$$\text{Max} \quad \sum_{i \in I} a_i y_i, \tag{2.6}$$

$$\text{s.t.} \quad \sum_{j \in N_i} x_i \geq y_i \quad \forall i \in I, \tag{2.7}$$

$$\sum_{j \in J} x_j = p, \tag{2.8}$$

$$x_j \in \{0,1\} \quad \forall j \in J, \tag{2.9}$$

$$y_i \in \{0,1\} \quad \forall i \in I. \tag{2.10}$$

In the problem, the objective is to maximize the amount of demand covered with (2.6). A demand node is considered as covered if a facility is opened out of the set of facilities that can cover the demand node (2.7). The number of facilities that can be opened is limited (2.8) and the decision variables belong to set of binary variables as given in (2.9) and (2.10).

The model is extended in various studies by examining further factors such as including coverage levels, partial coverage possibilities and stochasticity. One of the extensions of the model is written by Current and Schilling in which a tour must be completed which would visit certain number of facilities on the network [11]. The objective is to minimize the tour length and maximize the access to tours on the node that are not directly on it. An example to this type of problems can be given as the mail delivery service. The model works by defining a covering radius and groups the demand points as either covered or not, similar to the one developed in [10]. The main differentiation is the inclusion of directed arcs in the network and this is represented by a binary variable to see if there exists an arc from location $i$ to $j$.

Regarding the constraints of the problem, it is made sure that solution contains a single tour as it was aimed.

Another interesting expansion to the coverage models is implemented by Alexandris and Giannikos by applying the Geographical Information System (GIS) in the partial coverage model [2]. The main differentiation in this model is, the demand points are not considered as "points" but rather as "spatial objects". In the linear model, the set of demand *points* is changed with set of demand *areas* and their coverage by the facilities that are located within a distance is defined accordingly. Similar to previously discussed models; objective function tries to maximize the covered demand.

Additional extension to the coverage model comes with the building of hierarchical locations models, which are based on partial coverage. This is an important approach to the problem as from this point, demand locations are not divided as either "covered" or "uncovered". In these types of models, unlike the previously discussed versions of the problems, if a facility is located further than the required a threshold distance from the demand, the demand can still be partially covered. This situation is modeled by a coverage function, which returns the ratio of coverage of demand location changes with respect to the relation between critical coverage radius and distance between facility and demand location. The model also aims to maximize the coverage of all facilities.

The paper by Balcik and Beamon is also helpful for examining the coverage problems for relief operations [4]. In this paper, facility location and the stock positioning decisions for responding quick-onset disasters are examined. There are certain complexities associated with this problem due to unpredictability of demand, in very large amounts, which requires action in short times, involves high stakes and lacks resources. In the paper, a maximal covering model is built where different items require different coverage. It also considers budget and capacity constraints and allows stepwise partial coverage of customers. Also, inventory decisions are integrated to the facility location decisions.

The relief items are distinguished in accordance with their criticalities and response times. The lower and upper limits are set on the response times and coverage benefits are coupled with these values. For the modeling, sets of scenarios are generated and candidates for distribution centers are defined. The objective function maximizes the total expected demand covered by the established distribution centers for the sum of all given scenarios. The model

is used for solving an example case about the earthquake-caused disasters. Due to having a small-scaled problem, the problem was solved optimally without using heuristics. However in future, for larger problems, development of heuristic models might be needed.

Coverage problems are also helpful when having multiple facilities along the path is beneficial or necessary. In flow interception problems, once a flow passes through a node, it is directly assumed to be covered. However, it might be necessary to stop at several points in order to categorize a point as covered and successfully contain the entire path. This is discussed in the paper [18] by Kuby and Lim for the location of fueling stations on the path. The vehicles might need more than a single refueling operation and therefore, it is necessary to have multiple facilities to successfully cover the path. The paper includes the term "vehicle's range" which denotes the distance that can be travelled by the vehicle along the route to denote where it would require refueling. The combination of facilities that can fulfill the refueling criteria are determined and incorporated in the flow coverage model. The model forces to have at least one eligible combination of facilities to be open rather than the single facilities. The model is solved using the Model modeling language and compared with traditional models. For flow capturing models, the optimal solutions were proved to be at the nodes of the network, however, with this model considering the total round trip relative to vehicle range and spacing of nodes, suboptimal solutions might result with the allocation of facilities only at the nodes and middling locations can be more beneficial.

### 2.1.3 Further Literature Review

Not directly related to flow interception or coverage problems, there are some other literature review materials that deserve examination due to the characteristics of the problem. One of them is the article by Oppong which discusses the rainy seasons' effect on location-allocation problems [25]. In this paper, the general location-allocation models are criticized, as they are not focused on the characteristics of the rural areas that affect the locations. The paper particularly focuses on the issues of rainy seasons and argues that effects of rainy seasons are not observed in $p$-median or coverage models, where the limitation of access to certain routes can be problematic in real life. The study takes place in Ghana, for the establishment of primary health care services. For locating the services, the seasonality issue is included in the model and the effects of extending the model is observed.

The paper discusses two ways to include the effects of rainy seasons in location-allocation model. One method is to edit geographic distances in the model relative to difficulty of traveling. Roads can be categorized in classes and each can be assigned with a different weight, which depends on maximum speed limit, frequency of traffic flow along the road and year round accessibility. However, according to the authors, as it is difficult to justify different weights, this approach is not considered as feasible. Another approach suggested is that locating facilities only on the places linked by year-round drivable roads. The authors divide the model in three scenarios to evaluate the possible location models. One of them is the dry season scenario where no rain is expected and all of the roads are accessible. The other one is an impaired scenario where the facilities not on all-season roads are assigned to the nearest facility on all-season roads. The final scenario locates facilities only on all-season roads. As a result it is stated that not considering the accessibility problems for location allocation models in third world countries would result in “elegant but useless" solutions. This situation can be improved by consideration of simple modification, making the results of the models be applicable in real life.

In addition to the article by Oppong, Rahman and Smith also considers the special conditions of the developing nations for the location-allocation models in health services [26]. The main motive behind the article asserts that, the conditions in developing nations may not be very appropriate for the application of sophisticated models. The paper does not focus on development of a new model, however focuses on the assumptions and lacking points of the existing models that needs to be further considered especially for health care in developing nations. One of the issues is the capacity constraint where in most of the cases facilities are assumed to be incapacitated. Another point they call attention to is the multi criteria nature of the problems. Referring to Erkut and Neuman, they claim that models can be useful for suggesting candidate sites [14]. However, final decision should be made with the application of multi-objective decision-making tools.

Another useful article that would bring a different approach to the problem is by Abdel-Latif [1]. The paper combines two different research methods in the field of location allocation planning which are the mathematical programs, similar to the ones discusses previously, and the GIS Spatial Analysis techniques. The paper focuses on the selection of site for a school in Egypt where many criteria are involved. As a first step of the application, candidate locations are determined. The information about different criteria is collected via surveys and the

information resulting is mapped with ArcGIS Builder to identify candidate sites. After this step, a mathematical model is built using the $p$-median problem with network distances. With Environmental Systems Research Institute models in ArcInfo GIS different algorithms for optimization model is run and satisfactory results are achieved with the $p$-median algorithm.

A final study which is important to discuss is the thesis [13]. The problem discussed in this study is very relevant to the problem examined in this thesis and guided the problem definition. De Vries describes a model for the problem of investing optimally in the network of medical centers which are located in the African highways. The model tries to increase the number of truck drivers who can access the health care services as well as ensuring their continuum of care. For this purpose two objectives which are maximizing patient visits and minimizing expected traveling time the to the next medical centers are applied. The model is extended further which plans the allocation of budget to investment decisions such as opening or closing new centers and hiring or firing staff members. The model is formulated with a MIP model and with the computation studies it is seen that with the application of the investment model and systematic approach, the benefits achieved in the network are increased. The model developed by De Vries guided the mathematical formulations that are established in this thesis as they were both focused on the same real-life situation, yet showed changes as both of the models are approaching the problem from different perspectives and taking into account different requirements.

## 2.2 Stochastic Programming

As a second part of the literature review, literature about stochastic programming and risk related concepts are provided. As briefly introduced in the introduction of the thesis, the demand that is included in the problem is not known and also shows variation. Therefore, the randomness in the problem is the demand along the transportation lines. Since the problem is related with healthcare and demand is an important component, risk is also included in the problem context. The review about stochastic programming is consequently focused on the risk-averse measures.

### 2.2.1 Risk Aversion in Stochastic Programming

In stochastic programming, the randomness is handled using the expected values in many cases. For instance, as discussed by Miller and Ruszcynski, in the context of two stage stochastic programming, assuming the first stage objective function coefficients as deterministic and the second stage objective function coefficients as known after the first stage, builds the model as risk neutral [20]. However, especially in the context of humanitarian logistics where the life of people is a concern, risk neutral programming may not be the best approach.

Including the risk in the problem formulation can be done in various ways. One of the well known method is to use the expected utility theory, with the utility function over set of decisions. The risk aversion can be included in the problem such as:

$$F : \mathbb{R}^n \times \Omega \to \mathbb{R} \text{ and } X \subset \mathbb{R} \text{ is a feasible set,}$$

$$\min_{x \in X} \mathbf{E}[u(F(x, \omega))]$$

where $u$ is the disutility function and $\omega$ is the state of nature which is an element of the set $\Omega$ . However, choice of the utility function to be optimized is not always easily determined and different utility functions may lead to different interpretations. Therefore, while working with the expected utility theory, careful determination of utility function is required to achieve meaningful results.

Another approach is the mean risk model where the objective function is represented as the summation of expected outcome and the measure of uncertainty multiplied with the price of uncertainty involved. A generic mean risk model can be expressed as,

$$\min_{x \in X} \mathbf{E}[Z_x] + c\mathbf{D}[Z_x]$$

where $\mathbf{E}[Z_x]$ is the expected outcome, $\mathbf{D}[Z_x]$ is the measure the uncertainty of outcome and $c$ is the price of the risk.

Another important approach is the use of risk measures. In a given probability space, the risk measure is a function which assigns random variables to a risk value which results as holding the set of decisions [3]. The objective is to minimize the risk measure associated with the decision variables. In this approach, the risk measure function should be defined carefully to achieve significant results in the problem. As a subsection of risk measures, Artzner et

al. defines the “coherent risk measures" in which the risk measures satisfy certain properties [3]. In the following paragraphs, the coherent risk measures and its properties are explained extensively.

Given that $\Omega$ is the set of states of the nature and $G$ is the set of all risk, meaning that set of real valued functions on $\Omega$, a risk measure, $\rho$, is defined as mapping from $G$ to $\mathbb{R}$:

$$\rho : G \to \mathbb{R}.$$

A risk measure satisfying the axioms of *translational invariance*, *subadditivity*, *monotonicity* and *positive homogeneity* is named to be coherent. In the following paragraph these axioms are further elaborated as discussed by Artzner et al [3].

- Translational invariance: For all $X \in G$, $\alpha \in \mathbb{R}$ and $r$ is a strictly positive price, $\rho(X + \alpha \cdot r) = \rho(X) - \alpha$
  This property asserts that risk of the $X + \alpha \cdot r$ is less than the risk of $X$ as $\alpha$ behaves like insurance to the value risked. It can be also further noted that, if $\alpha$ is as much as $\rho(X)$, then $\rho(X + \rho(X) \cdot r) = 0$.

- Subadditivity: For all $X_1$ and $X_2 \in G$, $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$
  This property states that the “a merger does not create extra risk” which is a result of the diversification [3]. If this property was failed to satisfy, for example a firm requiring extra capital would be willing to separate into two incorporated affiliates.

- Monotonicity: For all $X$ and $Y \in G$ with $X \leq Y$, $\rho(Y) \leq \rho(X)$
  With this axiom, it is understood that out of two options where one has more future value, $Y$, it should always have a lower risk. This means that less has to be added to option $Y$ compared to option $X$ to make it more acceptable, where the amount added can be interpreted as risk measure.

- Positive homogeneity: For all $\lambda \geq 0$ and all $X \in G$, $\rho(\lambda X) = \lambda \rho(X)$
  This axiom would in brief imply that if amount invested in portfolio is multiplied by a certain amount, it also means that an associated risk is also multiplied with that amount. The position held directly influences the risk.

If a risk measure satisfies the axioms given above, it is said to be a “coherent risk measure”. The reason that coherency is a desired property in the class of risk measures is that the ad-

vantages that is provides for the optimization problems. In [28] by Rockafellar, comparison between coherent and non- coherent approaches are made and the implications over optimization problem are explained. In Section 2.2.2, the relation between optimization and coherent risk measures will be clarified.

### 2.2.2 Coherent Risk Measures and Optimization

A generic optimization problem, where the parameters are set to be deterministic can be summarized in the following form,

$$\textit{minimize } c_o(x) \text{ over all } x \in S \text{ satisfying } c_i(x) \leq 0 \text{ for } i = 1,2,\ldots m$$

where $S$ is the subset of $\mathbb{R}^n$ composed of vectors $x = x_1,\ldots x_n$ and each $c_i$ is a function from $S$ to $\mathbb{R}$ [28]. With the introduction of uncertainty in the problem, it is not possible to represent the coefficients solely by $x$, as the new parameter $\omega$, which belongs to set $\Omega$ representing future states, is included in the problem. The set of $\Omega$ can be regarded as mathematical structure of a probability space with probability measure $P$, for the comparison of future states of $\Omega$. The new coefficients take the form $c_i(x,\omega)$ such that:

$$\underline{c_i}(x) : \omega \rightarrow c_i(x,\omega) \text{ for } \ i = 1,2\ldots m.$$

In [28], measure of risk is defined as representing "overall cost" which maps between random variable to a single value. A risk measure applied to random variable $X$ will always stand for the cost values, meaning positive outcomes of $X(\omega)$ of $X$ are disliked and tried to be minimized (keeping $\leq 0$). The random variable $X$ is a function from $\Omega$ to $\mathbb{R}$ that are the part of the linear space $\mathscr{L}^2$ which is introduced relative to probability space $P$ on $\omega$. When the risk is required to be quantified, each value of $X$ which belong to set $\mathscr{L}^2$ is assigned to $\mathscr{R}(X)$. It is assumed that the value will belong to the set of $(-\infty,\infty]$. Regarding these discussions, the deterministic optimization formulation can be rewritten including risk and stochastic formulations as,

$$\textit{minimize } \overline{c_0}(x) \text{ over all } x \in S \text{ satisfying } \overline{c_i}(x) \leq 0 \ for \ i = 1,2,\ldots m$$

where $\underline{c_i}(x)$ is replaced by function $\overline{c_i}(x) = \mathscr{R}_i(c_i(x))$ and $\mathscr{R}_i$ quantifies risk by $\mathscr{L}^2 \rightarrow (-\infty,\infty]$ [28].

The risk measure developed for the problem above is an important aspect of the formulation and definition of the function should be carefully made. Within this respect, the coherent risk measures act as strong candidates to be included. Coherent risk measures hold a strong position in optimization due to several properties resulting from their properties. These are summarized in the following theorem by Rockafellar.

**Theorem 1** *(Theorem 1 in [28]) Assuming each* $\overline{c_i}(x) = \mathscr{R}_i(c_i(x))$ *for* $i = 1,2,\ldots m$, $\mathscr{R}_i$ *is a coherent risk measure,*

*(a) Convexity is preserved. For the initial deterministic formulation where uncertainty was not included is convex programming, when uncertainty is added, the convexity of the problem formulation is maintained and the advantage remains. This is a result of when* $c_i(x,\omega)$ *is convex with respect to x for each* $\omega$*, accordingly the function* $\overline{c_i}(x) = \mathscr{R}_i(c_i(x))$ *is convex. This property is a result of the subadditivity and monotonicity properties of the coherent risk measures.*

*(b) Certainty is preserved. If* $\underline{c_i}$ *(x) is a constant random variable for each x such that* $c_i(w,\omega) = c_i(x)$ *without depending on* $\omega$*, then* $\overline{c_i}(x) = c_i(x)$*. Accordingly the features of the problem without uncertainty are not distorted by the composition techniques.*

*(c) It is insensitive to scaling. Regarding the positive homogeneity property, the risk included problem formulation remains the same when the denominated values of* $c_i(x,\omega)$ *are rescaled.*

Regarding the above theorem and the advantages offered with the coherent risk measure, for a development of a risk averse problem, the coherent risk measures can be considered to be a good alternative for the optimization procedures. It would both allow risk averse settings in the formulation and at the same time provide optimal solution by preserving the convexity of the linear programs.

There are several measures of risk that are considered to be coherent by Rockafellar which are guessing the future, worst case analysis and relying on expectations.

Another important coherent risk measure is the Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR) measures. In the following chapter the two measures are explained in details.

### 2.2.3 Value-at-Risk and Conditional Value-at-Risk

For definition purposes, suppose $x$ is the decision vector chosen from the set of $\mathbb{R}^n$, $y$ is the random vector in $\mathbb{R}^m$ and $f(x,y)$ is the loss associated with $x$ and $y$. The probability distribution for $y$ is denoted as $p(y)$ assuming it has density. However, this is not a requirement for the definition as it will be seen in the upcoming steps.

Given a threshold $\alpha$, the probability for loss function to not exceed the threshold is stated as:

$$\psi(x,\alpha) = \int_{f(x,y)\leq\alpha} p(y)dy.$$

Regarding this definition, Rockafellar and Uryasev define the $\beta$-VaR and $\beta$-CVaR values respectively with any associated probability level $\beta$ in (0,1) as follows [29]:

$$\alpha_\beta(x) = \min\{\alpha \in \mathbb{R} : \psi(x,\alpha) \geq \beta\}, \tag{2.11}$$

$$\phi_\beta(x) = (1-\beta)^-1 \int_{f(x,y)\geq\alpha_\beta(x)} f(x,y)p(y)dy. \tag{2.12}$$

VaR can also be referred as the $\alpha$- quantile of the loss function. It can be further explained as the threshold value in which loss associated with the set of decisions variables will exceed this threshold value with the given probability level. CVaR is defined as the conditional expectation of the loss in the conditional distribution of its upper $\alpha$ tail.

The relation between VaR and CVaR values and their characterization is defined with a function $F_\beta$ on $X \times \mathbb{R}$:

$$F_\beta(x,\alpha) = \alpha + (1-\beta)^-1 \int_{y\in\mathbf{R}^m} [f(x,y)-\alpha]^+ p(y)dy \tag{2.13}$$

where $[t]^+ = t$ when $t > 0$ and $[t]^+ = 0$ when $t \leq 0$. Given the function $F_\beta$, the following theorem is stated by Rockafellar and Uryasev.

**Theorem 2** *(Theorem 1 in [29]) As a function of $\alpha$, $F_\beta(x,\alpha)$ is convex and continuously differentiable. The $\beta$-CVaR of the loss can be determined by,*

$$\phi_\beta(x) = \min_{\alpha\in\mathbf{R}} F_\beta(x)$$

*and $\beta$-VaR associated with loss value can be determined by,*

$$\alpha_\beta(x) = left\ endpoint\ of A_\beta(x)$$

*where $A_\beta(x) = argmin_{\alpha \in \mathbf{R}} F_\beta(x,\alpha)$, the non-empty, closed, bounded set consisting of values for $\alpha$ for which the minimum is attained. Consequently, it is always stated as,*

$$\alpha_\beta(x) \in argmin_{\alpha \in \mathbf{R}} F_\beta(x,\alpha) \text{ and } \phi_\beta(x) = F_\beta(x,\alpha_\beta(x)).$$

From the theorem, it can be seen that, without explicit calculation of VaR value, CVaR can be calculated. This definition of CVaR can further be approximated by sampling the distribution of $y$. This corresponds to:

$$\tilde{F}_\beta(x,\alpha) = \alpha + \frac{1}{1-\beta}\sum_{k=1}^{q} \pi_k [f(x,y_k) - \alpha]^+. \tag{2.14}$$

Both VaR and CVaR values can be used for optimization where risk is involved. Out of two concepts, CVaR has been defined to be more preferable. Artzner et al. argue that VaR does not have the property of sub-additivity and convexity which is not very desirable [3]. VaR is not a coherent risk measure, unless it is based on the standard deviation of normal distribution. VaR can also show difficulties in optimization procedures. As explained in Section 2.2.2, coherent risk measures have properties that are inline with optimization methods and CVaR as a coherent risk measure, satisfies these properties.

Furthermore, in the article [9], the VaR risk measure (which is referred as $\alpha$-reliable Minimax regret), is discussed to be not very powerful as it focuses on the only $\alpha$-reliable scenario set and ignores the magnitude of regrets at the tail . This can result in overwhelmingly high regret values at the tail. It is also noted that computationally, as $\alpha$-reliable Minimax model is non-smooth, non-convex and multi extreme function which makes the model difficult to solve. Consequently, CVaR risk measure (referred as $\alpha$-reliable Mean-excess regret), is more advantageous as it makes up the drawbacks of VaR risk measure.

The following theorem by Rockafellar and Uryasev explains how minimization of CVaR values can be achieved.

**Theorem 3** *(Theorem 2 in [29]) Minimizing the $\beta$-CVaR of the loss associated with $x \in X$, in is equivalent to minimizing the function $F_\beta(x,\alpha)$ over all $(x,\alpha) \in X \times \mathbb{R}$, in the sense that,*

$$\min_{x \in X} \phi_\beta(x) = \min_{(x,\alpha) \in X \times \mathbf{R}} F_\beta(x,\alpha) \tag{2.15}$$

*where the pair $(x^*,\alpha^*)$ achieves the second minimum if and only of $x^*$ achieves the first minimum and $\alpha^* \in A_\beta(x^*)$. Therefore, for the circumstances where the interval $A_\beta(x^*)$ reduces to*

*single point, minimizing $F(x,\alpha)$ over $(x,\alpha) \in X \times \mathbb{R}$ produces a pair $(x,\alpha^*)$, not necessarily unique, such that $x^*$ minimizes $\beta$-CVaR and $\alpha^*$ gives the corresponding $\beta$-VaR.*

This theorem supports the optimization approach associated with the calculation of $\beta$-CVaR values with respect to the function $F_\beta(x,\alpha)$. The minimization of the function $F_\beta$ is in the category of stochastic optimization and one way to deal with this formulation can be the approximation as described in (2.14).

In addition to optimization of $\beta$-CVaR values in the objective function, the risk measure can also be subject to constraints regarding the problem context as it may enforce requirements of different CVaR levels with different confidence intervals. This condition may yield constraints in the form of

$$\phi_\beta(x) \leq \omega \tag{2.16}$$

where $\phi_\beta(x)$ is the $\beta$-CVaR risk measure on the loss function and $\omega$ is the upper limit on the risk.

In the paper [17], the theorem below has been proposed for the use of CVaR as constraint in the linear programs, which is useful for further calculations.

**Theorem 4** *(Theorem 4 in [17]) The two problems below,*

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} -R(\boldsymbol{x}), \quad \phi_\beta(\boldsymbol{x}) \leq \omega, \quad \boldsymbol{x} \in \boldsymbol{X}$$

*and*

$$\min_{(\alpha,\boldsymbol{x}) \in \boldsymbol{X} \times \mathbb{R}} -R(\boldsymbol{x}), \quad F_\beta(\boldsymbol{x},\alpha) \leq \omega, \quad \boldsymbol{x} \in \boldsymbol{X}$$

*are equivalent from the aspect that both of the problems' objective functions achieve the same minimum.*

Referring to the above theorem and (2.14) about the approximation of function $F$ by $\tilde{F}$, the constraints can be re-written in the form,

$$\alpha + (1-\beta)^{-1} \sum_{k=1}^{q} \pi_k z_k \leq \omega, \tag{2.17}$$

$$z_k \geq \sum_{l=1}^{n} f(x,y_k) - \alpha, \quad z_k \geq 0, \quad k = 1,\ldots,q. \tag{2.18}$$

The literature review explained in this section is a foundation for the stochastic model that that will be explained in Chapter 5 of the solution approach. The initial deterministic model is converted to a stochastic model where the risk is included in the problem context regarding the definitions and theorems explained in this section.

# CHAPTER 3

# PROBLEM DEFINITION

In this section, the problem studied will be thoroughly explained. The motivation behind the problem is explained in Section 3.1. The environment of the problem given in motivation is described in Section 3.2. The definition for the problem is given in Section 3.3. Finally, problem scope is described in Section 3.4.

## 3.1 North Star Alliance

The studied problem is emerged from a real life problem which was researched extensively within a non-governmental organization named North Star Alliance. North Star Alliance, which will be hereafter referred as North Star, is established as a result of the public-private partnership between United Nations World Food Program (WFP) and TNT Express. The primary motive for building partnership was to combat hunger in Africa with the provision of relief food to the demanded regions. However, during the project, the target of dealing with hunger problem in Africa made it possible for both parties to realize a further problem. As a part of distribution of relief food to communities in hunger, the problem of having insufficient number of truck drivers to deliver the relief food to the communities, revealed the relation between mobility and health problems.

Even though the initial purpose of the partnership did not include the health and the safety of the truck drivers, both of the organizations were aware that their operations were based on transportation and it would be impossible to overlook the problems arising as a result of the relation between mobility and healthcare. In order to tackle this issue, they have initiated a separate pilot program by opening a walk-in clinic in Mwanza border crossing in Malawi,

which provided primary health care that can be accessed by truck drivers along the roadside, aiming to make health care service accessible for truck drivers. After observing success of the program, it was understood that having similar clinics located along major transport corridors across Africa would have significant effect on health, lessening the health problems resulting from the mobility condition. Consequently, North Star was founded in 2006 with the aim of establishing similar clinics that cover the continent and make life safer for mobile populations and their communities [24].

North Star developed the concept of "Roadside Wellness Center" (RWC) to extend the local health care infrastructure and provide service for the mobile populations. An RWC is a converted container, which can be either mobile or not, that is placed on critical locations along the roadsides where the mobile population congregate. With this placement, mobile populations can have easier access and get the primary help for basic health care services. RWCs also act as a “funnel” which captures the mobile populations, increase the awareness and refer to other better equipped and larger local healthcare services when required.

RWCs employ at a minimum two persons where one of them is the counselor and the other is the nurse. Counselors are responsible for conducting behavior change communication sessions with the patients. Groups of people that are involved in trainings are also attracted to the sessions by the out reach workers that are also a part of the RWC employees who go in the local region to attract people for the education sessions. Nurses are responsible for health related operations where they diagnose and treat the patients. They are also required to keep extensive amount of documentation both for the national health authorities and for North Star to keep track of the operations and patients.

RWCs are established based on set of standards that conform to WHO requirements. To maintain a high quality service, North Star also developed standard operating procedures. Furthermore, North Star signs cooperation agreements with the local and national health authorities in the operating countries to make sure that RWCs’ activities are in line with the requirements and strengthen the national health priorities. Currently, North Star has 29 operating RWCs in Africa, located mainly in the Southern and Eastern regions.

The mission of North Star is to provide mobile workers and the communities near the transportation corridors with sustainable access to high quality health and safety services [22]. North Star has three offices, where the head office is based in Utrecht, the Netherlands, and

two other offices in South Africa and Kenya employing approximately 115 number of personnel. North Star has core partners that support its activities and it also continues its operations with the donations from private companies who would like to invest in clinics in African regions.

## 3.2 The Environment

When operations of North Star are investigated, several factors are identified as playing an important role. These factors are particularly worth discussing as they provide a foundation for the problem studied. These factors are explained in the following sections.

### 3.2.1 Demand Population

Mobile populations have a special position compared to the rest of the population within the scope of healthcare and well-being. Separation from home and partners, hard working conditions, lack of access to care and treatment services make the mobile populations vulnerable and increases the risk of exposure to numerous diseases including the high risk diseases as HIV and tuberculosis. Mobile populations are not only under risk of getting infected, but also play an important role for the transmission of these diseases making them acting as vectors of transmission.

This susceptibility of mobile populations is further compounded with their inability to access high quality healthcare services because of the demand and nature of the work they are subject to. The services they require are not usually designed in accordance with their needs, making it nearly impossible to receive the required care.

The main target group for North Star within the mobile population is the truck drivers. Due to hard working conditions, which generally require days of traveling with strict time pressure on traveling times, access to even basic health services can be very difficult. These conditions make truck drivers more vulnerable to many diseases. Other barriers, such as not accessing health centers due to unavailability of parking conditions for the trucks, lack of security and opening hours for health care facilities pose obstacles for them to access the services whenever required.

Even though North Star's primary focus is the truck drivers, there are two other target populations that require attention in order to tackle with the problem. The second target group of North Star in addition to truck drivers is the sex-workers that are generally located in the hot-stops where the truck drivers spend long waiting hours. They can also be mobile with the truck drivers, traveling within the trucks across borders, making them also a part of the mobile population. Sex-workers are especially vulnerable to infectious diseases, making them in the need of having access to health services. However, as they are occasionally subject to extensive abuse and violence, it is hard to receive the healthcare needed with regards to social stigma. Additionally, the separation of truck drivers from regular partners usually result in getting engaged in sexual activities in the hot-spot locations with the sex-workers. This situation makes the sex-workers act as a hidden link in transmission chain of diseases and an important part of the target population.

The final target group is the local communities that are situated near the roads and hotspots. Similar to sex-workers, local communities can also act as transmission link for the communicable diseases and they need to be considered as well.

Within the scope of the target patient groups, North Star operations are limited within the continent Africa currently. The primary reason for focusing on Africa is that, it is one of the regions that has the highest burden of disease. According to World Health Organization's (WHO) Global Burden Disease Report, when the measure DALY (disability-adjusted life year, representing the loss of the equivalent of one year of full health) values are examined, Africa has at least twice as much as values compared to any other region. This situation can also be examined in the Figure 3.1.

From the communicable diseases, Africa is also the most heavily affected by HIV worldwide, accounting for the 71% of the deaths related with HIV/AIDS all over the world according the Foundation for AIDS Research [31]. This situation is compounded by exposure to other communicable diseases, which are mainly malaria, tuberculosis and sexually transmitted infections (STIs). This persistent burden of communicable diseases which puts Africa in a special position, makes North Star interested in getting in action to tackle the issue within African region.

Another categorization of demand population for the purpose of inclusion in the problem formulation can be made by broadly categorizing demand population as "Mobile" and "Static".

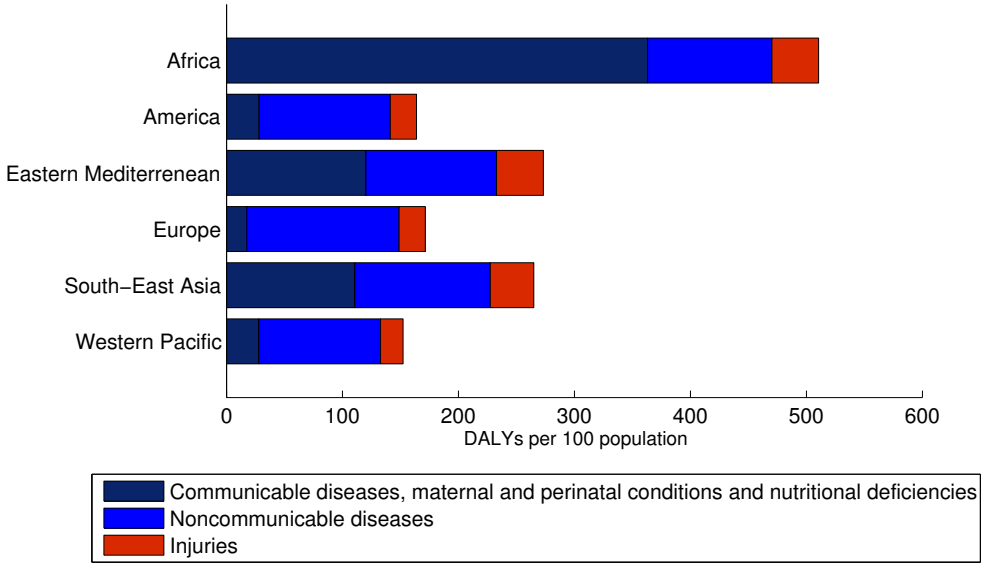


Figure 3.1: The disease burden in Africa

In mobile demand group, the patients which travel throughout the time are considered. The majority of this group is the truck drivers as they are the ones spending their most of the time on the roads. Another part of mobile demand is the sex-workers which are not as dominant as truck drivers in terms of numbers. The amount of mobile demand can be defined as the average daily number of people on a certain transportation line defined over network to understand how transportation lines are congested. In the static demand group, local populations are the major element. They are the communities near the transportation lines and their demand to the RWCs would depend on the patient visits occurring to the local healthcare facilities nearby. The amount of local population visiting the RWCs is affected by factors as distance to the walk-in clinics, availability of local health care infrastructure and number of inhabitants in the local communities. Sex-workers also belong to the group static demand.

A further remark on demand should be made about its stochastic nature. The value of demand for each of the target patient group is not estimated straightforwardly. There is a huge amount of uncertainty involved in the nature of the problem and this should be considered through the problem formulation. Some of these uncertainties arise as a result of the unstable situation in Africa, in which field research and data gathering activities are not very effectively handled. Finding accurate estimates for the traffic flows can be problematic as traffic counts are not completed properly.

### 3.2.2 Health Strategy

North Star's health strategy is grouped in two categories as *key* and *working* principles. With the three key principles, North Star prioritizes and justifies what they do and with the ten working principles, they explain and justify how they deliver the services and how to improve their quality and effectiveness [23].

When the key principles are analyzed further, it is seen that North Star prioritizes the actions based on the health needs of the society which is determined by looking at the impact of the diseases on the society. This impact on society can be observed with the concept of "disease burden" which was discussed in Section 3.2.1. Another idea included in key principles is that primary focus is given to mobility related diseases. For instance, even though maternal health is a serious health issue, it is not the main health concern addressed by North Star. However, such health issues are included in the secondary focus area, which is an adapted service package depending on the requirements of the region where the RWCs are built. The final idea behind the key principles is to offer feasible and affordable services. For individual cases, the evaluation of both whether the health condition can be effectively prevented, diagnosed, treated and/or cured and whether an effective contribution can be made is determined. Accordingly, if a given solution is not feasible, such as requiring complex diagnostic methods, the patients are referred to other health services, which are capable of carrying necessary services. In summary, with the key principles, North Star focuses it health care services on health conditions that:

- Play a high impact on society
- Are related mobility
- Can be prevented/diagnosed/treated in an effective and affordable way

Among the ten working principles, even though each of them is required for operations of North Star, principle stating "providing continuum of care" is critical for the study. The continuum of care definition is further divided as "vertical" and "horizontal" continuity.

- Vertical Continuum of Care: In terms of "vertical" continuity, the continuity of care is ensured by aligning the facilities with local district health systems and by establishing referral systems to other preferred local health care providers.

- Horizontal Continuum of Care: The "horizontal" continuity is concerned with the mobility of the truck drivers along the transport corridors. The care is ensured by promoting the follow-up service of activities for drivers along the roads even if the services are situated in different districts, regions or even countries. This is supported by individual tracking of the truck drivers with the identification system developed within North Star, which enables every RWC to access the health related historical data of the patients. North Star wants to be accessible whenever a truck driver requires health service and provide them with high quality and effective health services.

Currently North Star implements a standard service package in RWCs where the services offered for every disease is divided into categories as screening, diagnosis/test, treatment and care and referral. As the organization grows and acquires more supporting donors, the services are planned to be expanded and improved.

### 3.2.3 Diseases and Services

One of the most important parts of the problem is the diseases and services that are served in the RWCs. Starting with the diseases, diseases can be divided in two as communicable and non-communicable diseases. Communicable diseases are the ones which can spread from one person to another, making it play a threatening role in the society. WHO tracks the infectious diseases, sounds alarm when required and plays a protective role from the consequence of epidemics [32]. Important communicable diseases are cholera, malaria, measles, meningitis and tuberculosis. From the perspective of mobile populations several communicable diseases are relevant more compared to others which are HIV, malaria, tuberculosis and STIs. Especially for the case of STIs and HIV, the prevalence rates are found to be more drastic for mobile populations compared to other communitites. It was reported that 60% of truck drivers reported having STIs in previous six months time and in total, truck drivers and sex-workers have a total of 56% HIV prevalence rate in South African region [27]. It is not only the Southern African countries, that have high prevalence rates for HIV and STIs but also other regions in Africa are affected. In a study, out of 331 truck drivers from East and Central Africa, 18% has been diagnosed HIV positive and around 4% for sexually transmitted diseases. In specific, the prevalence rate for Central African truck drivers was more with value of 34% [8]. The non-communicable diseases are named as chronic diseases and they do not spread from person to

person [33]. They usually evolve slowly and are of long duration. WHO categorizes them as cardiovascular diseases, cancers, chronic respiratory diseases and diabetes. These are also diseases that can be encountered in mobile populations, due to hard working conditions for truck drivers make them more vulnerable to risks imposed by these diseases and also for sex-workers. The non-communicable diseases are mainly treated to make drivers feel connected and familiar to the clinics.

In order to achieve the desired effectiveness from walk-in clinics, services to be offered inside should be determined carefully as well. According to North Star's health service package, which is developed together with professional medical institutions, the services offered in the RWC are listed as:

- *Screening:* The process in which symptoms from the patients are recorded and medical history is obtained. The social and economical background of the patient is understood and for some cases physical examination is conducted.

- *Diagnosis and Tests:* It involves simple tests as weight and body mass index. Depending on the disease more advanced tests as rapid HIV test and urine tests can be done.

- *Care and Treatment:* This involves treatment or symptomatic treatment of the diseases. Basic care is given including counseling about the disease and adherence to treatment. Patients are advised periodic follow up (their treatments) for treated cases.

- *Referral:* Not all diagnosis, tests, treatment possibilities and medications are available in the RWCs. In case further examination is required for diagnosis, the disease cannot be controlled and is severe, the patient is advised to refer to the nearest hospital.

Each disease has its own different actions listed for every service category resulting in different types of intervention. The progression speed and the requirements change from disease to disease and therefore separate planning for every disease and service combination is essential. For instance, the treatment follow-ups with HIV and malaria are not the same and a network specializing on HIV treatment would be different compared to a network specialized on malaria treatment. This idea should be kept in mind through problem formulation.

### 3.2.4 Local Healthcare Infrastructure

One other important point as a part of the problem is the local health care infrastructure. As described with "Continuum of Care" definition in Section 3.2.2 and Section 3.2.3, referral is a key point for operation of the walk-in clinics. It is important to examine the local health care infrastructure and identify the blank spots on the map to include the RWCs. RWCs should not only be present where local healthcare infrastructure is lacking but also it should designed in line with them too to work in cooperation. If there is a possibility of referral service nearby the RWCs when required, the patient can be easily referred to bigger hospitals and it would be more likely that the patient will not be lost. It is a problem with the network that once a patient requires further examination in more equipped healthcare providers, they tend to postpone the time for examination or even completely ignore it if they are not directly referred. In the case where the referral possibilities are present in close distance, it would be more effective for achieving the referral aim.

### 3.2.5 Delays

Through the transportation lines, not everything goes along very smoothly and trucks often do not reach the final destination in very short time frames. There are certain obstacles that lengthen the travel time considerably causing undesirable results. These can be named as, including but not limited to, border crossings, check points and weightbridges. Trucks have to check through the borders where they wait for long hours or days in the queue and spent a lot of time with document procedures. There are also other check points, such as the police checks, which result in delays. Weightbridges are the points where the trucks are weighed and checked if they comply with the standards. These delay locations at the same time present a good opportunity for the location of walk-in clinics as truck drivers accumulate at these points and spend long waiting hours.

There can be other delays resulting from road and weather conditions along the transport lines. The roads are not fully paved and sometimes poor road conditions lower the speed and complicates the truck drivers job. Weather condition in combination with the poor road can also be another obstacle such as muddy or blocked roads resulting from heavy rains.

The delay durations along the transport line show that the travel times should not be only

considered as the kilometers difference between origin and destination locations but also the long waiting hours should be integrated and while planning the RWC network, it should not be disregarded.

### 3.2.6 Budget and Investments

Intuitively, the very best solution in terms of health service would be opening all of the walk-in clinics in every possible location, fully equipped to provide all services for all diseases. Unfortunately, this is constrained by the budget and investment availabilities in the region. In the case of North Star, the organization opens RWCs with the donations provided from Ministries of Health, global health partners and private companies. Additionally, the investment process for North Star is donor dependent. This means that, usually donors contact North Star specifying their requirements and donations to a certain extent. For instance, a donor can approach North Star stating that they would like to invest in a specific country or region and give funds for a certain number of RWCs. In these situations, North Star and the donor reach a common understanding and define the set of acceptable conditions and invest in definite number of RWCs.

The cost of an RWC include the (one off) establishment cost, maintenance cost and the employee cost over the duration of RWC's life span. It is also quite difficult to derive a cost table for the RWC as every country in which the RWC operate has different rules and regulations, changing the amount invested in RWC and making the cost data complex. Therefore, given an investment amount for establishing new RWCs, there is an upper limit on number of RWCs that can be opened regarding the specifications of the region.

## 3.3 Problem Formulation

The challenges behind North Star's planning of the RWCs and difficulty of the work can be understood from the conditions that they work through and the importance of the activities. The problem studied in the thesis has been greatly inspired from this mission of North Star and provided a basis for the problem definition.

As explained, the solution North Star provided to the problem of dealing with the issues

resulting from the detrimental relation between health care and mobility is the concept of RWCs. These walk-in clinics near the roads and hot-spots are playing a life changing role for the mobile populations and throughout time, they are becoming an inseparable part of the job routine for mobile populations. Having such a significant meaning for the mobile populations, the RWCs need to be planned carefully to meet the expectations and add value to the transportation network.

The problem studied in this study deals with the issue of planning of the walk-in clinics along the transportation lines by making them situated as beneficial as possible for the mobile populations. This planning includes three important decisions:

1 What should be the location of clinics that are decided to be opened?

2 Which diseases should be managed within the opened walk-in clinics?

3 Which services of the managed diseases should be offered within the opened walk-in clinics?

Therefore, the main objective of this thesis is to propose a model to provide benefit at the maximum level. Currently, North Star makes these decision by rule of thumb methods, situational analysis and directions provided by the donors. Nevertheless, it is certain that there should be an improved methodology and strategy behind this procedure.

It is not only the locations that matter for the problem, but also the diseases and services that are offered within the clinics. Currently in North Star, there is a standard health care service package that is offered in every open RWC, but from a general perspective, these services should be tailor made regarding different requirements of the target groups and areas in which the walk-in clinic is established. A more specific and detailed planning of the walk-in clinic network would be a better solution for this purpose.

To answer these questions above, criteria should be determined which would appropriately define the effectiveness of the walk-in clinics for the target population. However, this definition may not be straightforward and requires careful analysis of both the target group and the region. Firstly, as a general motivation of location planning problems, the walk-in clinics should be placed in a way which would capture as many patients as possible. It is important to locate the clinics near the hot-spots which have high volume of trucks waiting and at the same

time near local communities with a significant population. Therefore, one aim would be to maximize the patient visits that would be occurring at the walk-in clinics. Another important objective would be to maximize the "continuum of care", as described under in Section 3.2.2. The walk-in clinics should offer a continuous care along the roads for the mobile populations and be aligned in accordance with the local health care infrastructure. The continuum of care of the mobile population should also be maximized while answering the questions for planning the walk-in clinics. These two objectives do not have be necessarily conflicting however it should be also noted that, one solution that is optimal with one objective may not be optimal with the other. Therefore, through optimization process, this multi criteria objective should be considered with appropriate adjustments.

The problem of determining the optimal location of the walk-in clinics, resembles the network location problems in which a network is represented by set of arcs and nodes. In [21], a network is defined as $\mathscr{N} = (\mathscr{G}, l)$ where $\mathscr{G}$ is the underlying graph with $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where the node set $\mathscr{V} = \{v_1, \ldots, v_M\}$ and edge set $\mathscr{E} = \{e_1, \ldots, e_N\}$. Every edge of the network $e \in \mathscr{E}$ has an associated positive length through the function $l : \mathscr{E} \to \mathbb{R}_+$.

Within this respect, it would be required to make a network definition for the given problem definition. The arcs of the networks can be considered as the routes in which mobile population transport. These routes can be the major transport corridors in Africa. There are several major transport corridors in Africa which accommodate significant truck traffic. There are also other transport routes which are high-density roads within Africa. These roads on which mobile population travel should be determined with their origin and destination coordinates.

The set of nodes can be determined by the potential locations on which a walk-in clinic can be built. These nodes can be hot-spots on the lines of transportation where large number of mobile population is accumulated while traveling. Within this respect, there will be high number of demand in the hot-spot locations and the patient visit objective described in problem definition can be achieved. In addition to these hot-spots, other available locations which can be more appropriate for local population can be considered as potential walk-in clinic locations, and also as nodes of the network.

## 3.4 Scope of the Problem

As modeling the optimal locations for the walk-in clinics can be too complex after including all factors in the environment of North Star explained in Section 3.2, some aspects of the problem are narrowed or considered as out of scope. In this section, these aspects of the problem are explained.

Initially, the range of diseases that are handled in the walk-in clinics are large in number. Since the non-communicable diseases are simpler to handle, more affordable and usually present in the clinics with the aim of capturing patients, they do not require detailed planning. However, the focus is on the communicable diseases as it is more relevant with the mobility issues of the target population. The prevalence, treatment and care of communicable diseases are related with the mobility and can be used for measuring the continuum of care objective, consequently in the study, the non-communicable diseases part will not be addressed directly. In terms of communicable diseases, the four high impact diseases, HIV, tuberculosis, malaria and STIs, will be in scope for observing the continuum of care.

For the services to be handled, a new categorization is made. Screening is considered out of scope as it can be included in the walk-in clinics with no or relatively low costs and therefore does not require to be planned. For the care and treatment services, it would be more appropriate to divide in separate headings as each of the services are conducted with differing protocols. Treatment for malaria and care for malaria are handled under different approaches and a model for planning walk-in clinics should take it into account. Consequently, the services that are in scope can be listed in four categories as diagnosis, treatment, care and referral.

Another topic is about the delay that the truck drivers face during their trips. As described the delays can be caused as a result of formal procedures such as border crossings or regular checks or other cause of delay can be due to unplanned reasons such as road and weather conditions. However, predicting the weather conditions and collecting reliable information about road conditions and the amount of delay associated with it is very difficult. Including them in the calculations may mislead the model and may not add value. Therefore, the road and the weather conditions are considered as out of scope for the beginning. In case reliable information can be obtained, the ways to associate and integrate them to model for further developing can be considered.

A final point considered as out of scope is the budget and investment related topics. As described before in Section 3.2.6, it is not possible to come up with detailed estimates that can be used for approximating the costs associated with opening and running an walk-in clinic. If in future, North Star becomes a donor independent organization by raising money from different operations, which can also include selling of services, North Star would require different tools and financial analysis to optimize the investments. Nevertheless, at this stage of operations, investment and budget considerations are irrelevant and will not be included in the model. Given the requirements from donors, the agreements and field studies, the limit on opening clinics can be provided as "number of walk-in clinics to be opened" or even more specifically, "number of walk-in clinics which would offer the given service for a given disease".

# CHAPTER 4

# MATHEMATICAL FORMULATION - DETERMINISTIC MODEL

In this section, the mathematical formulation will be explained in detail. Initially in Section 4.1, the ideas that are developed for handling the calculation of continuum of care values will be given. Following these descriptions, in Section 4.2, the deterministic model will be developed which allows the optimal planning of walk-in clinic based on services and diseases.

## 4.1 Definition of Continuum of Care

The term “continuum of care" is a critical term for determining to what extent the mobile populations are served along the transportation lines. As explained in Section 3.2.2, the definition has two branches as *horizontal* and *vertical* where the horizontal continuum of care reflects the service given along the roads, both by being accessible whenever needed and following the patients along the road, and vertical continuum of care reflects the service beyond the roads, having referral availabilities near the walk-in clinics. In order to plan the walk-in clinics providing the maximum continuum of care, a mathematical formulation which would reflect the different service and disease requirements extensively is required. For this purpose, several approaches are developed. This section explains the continuum of care approaches both in horizontal and vertical groups.

### 4.1.1 Horizontal Continuum of Care

#### 4.1.1.1 Approach 1: Binary Coverage

The first approach for defining the coverage of the mobile populations along the roads refers to the service and disease combinations having strict critical intervention times. The demand needs to be satisfied within the critical intervention time, which is defined specifically for every disease and service combination, and consequently demand can be categorized as met and otherwise, the demand can considered to be lost, implying it is too late to intervene from the walk-in clinic. When the mobile population requires service from the walk-in clinic's facilities and is away from the walk-in clinic with a distance longer than the critical intervention time, the mobile population is named as "not covered". However, if it is possible to reach the next walk-in clinic within the critical intervention time, it is named as "covered". This results in a binary perspective where the categorization differs between two possible results.

The notation given in Table 4.1 is used for determining the coverage values.

Table 4.1: Notation used for Approach 1 of Continuum of Care

| | |
|---|---|
| Sets: | |
| $Q$ | Set of paths, $q = \{1,2,\ldots,Q\}$ |
| $K$ | Set of locations, $k = \{1,2,\ldots,K\}$ |
| Parameters: | |
| $tTotal_q$ | Total traveling time on path $q$, starting from origin and going to destination including the delay durations |
| $tCur_q$ | Total traveling time on path $q$, starting from origin to the current position including the delay durations |
| $t_{kq}$ | Total traveling time on path $q$, starting from origin to the location $k$ including the delay durations |
| $tc_{sd}$ | Critical intervention time for service $s$ and disease $d$ |
| $c_q$ | The value of coverage that applied on the mobile population at current position of the demand |
| $cov_{qsd}$ | The total coverage of the path $q$ for disease $d$ and service $s$ |

Regarding the notation given in Table 4.1, the following two variables are used for the calculation of coverage value along the path $q$ for service $s$ and disease $d$:

$$c_q = \begin{cases} 1, & \text{if there exists an open walk-in clinic within the critical intervention} \\ & \text{time at the current position } 0 \leq t_{kq} - tCur_q \leq tc_{sd} \\ 0, & \text{otherwise,} \end{cases}$$

$$cov_{qsd} = \int_0^{tTotal_q} \frac{c_q}{tTotal_q} d(tCur_q). \tag{4.1}$$

#### 4.1.1.2 Approach 2: Partial Coverage

This approach can be considered as the relaxed version of the first approach defined above. Here, the mobile population is not categorized as only "covered" and "not covered" but also a partial value is assigned. This is achieved by using two critical intervention times for every service and disease combination. When the mobile population requires service from walk-in clinic and is away from the walk-in clinic with a distance less then the first critical intervention time, it is considered as fully covered. Alternatively, if the mobile population is away from the walk-in clinic with a distance between first critical intervention time and second critical intervention time, a coverage ratio is assigned considering its distance to the walk-in clinic. If the mobile population is at a distance further than the second critical intervention time from the walk-in clinic, it is not covered as it is not possible to reach the facility. This condition applies to service and disease combination which do not have very strict applications but however, it is important to intervene timely.

The notation for this approach is nearly same as the notation used for Approach 1 given in Table 4.1. The change in the notation for Approach 2 is the critical intervention times. The added parameters are summarized in the Table 4.2.

Table 4.2: Added notation for Approach 2 of Continuum of Care

| Parameters: | |
|---|---|
| $tc1_{sd}$ | First critical intervention time for service $s$ and disease $d$ |
| $tc2_{sd}$ | Second critical intervention time for service $s$ and disease $d$ |

Regarding the above definitions, the following two variables are used for the calculation of coverage value along the path $q$ for service $s$ and disease $d$:

$$c_q = \begin{cases} 1, & \text{if there exists an open walk-in clinic within the} \\ & \text{critical intervention time at the current position} \\ & 0 \leq t_{kq} - tCur_q \leq tc1_{sd} \\ \frac{tCur_q - (t_{qk} - tc2_{sd})}{tc2_{sd} - tc1_{sd}}, & \text{if there exists an open walk-in clinic such that} \\ & tc1_{sd} \leq tCur_q \leq tc2_{sd} \\ 0, & \text{otherwise,} \end{cases}$$

$$cov_{qsd} = \int_0^{tTotal_q} \frac{c_q}{tTotal_q} d(tCur_q). \tag{4.2}$$

#### 4.1.1.3 Approach 3: Expected Traveling Time

For certain disease and service combinations, the coverage approaches which are including critical intervention times are not very applicable. It does not require an immediate service from the walk-in clinics, however it is of course favorable to serve as soon as possible. With this approach, it is aimed to minimize the expected traveling time to a next walk-on clinic along the road by the mobile populations.

The parameter added to the notation for the Approach 3 is $E_{sd}(tCur_q)$, which is the expected traveling time to the next walk-in clinic at the current position.

Regarding the above definitions, the following two variables are used for the calculation of coverage value along the path $q$ for service $s$ and disease $d$:

$$cov_{qsd} = \int_0^{tTotal_q} \frac{E_{sd}(tCur_q)}{tTotal_q} d(tCur_q). \tag{4.3}$$

#### 4.1.1.4 Matching the Continuum of Care Approaches with Service and Disease Combinations

In the previous sections, the definitions that will be used for formulation of continuum of care approaches are explained. Regarding the information coming from diseases that is discussed with healthcare professionals, each service and disease combination is assigned with an approach that is appropriate. Table 4.3 summarizes these assignments.

Table 4.3: The match between approaches and service-disease definitions

| | Services | | |
|---|---|---|---|
| | Diagnosis | Treatment | Care |
| HIV | Approach 3 | Approach 3 | Approach 3 |
| Malaria | Approach 2 | Approach 2 | Approach 2 |
| STIs | Approach 2 | Approach 2 | Approach 2 |
| Tuberculosis | Approach 2 | Approach 1 | Approach 1 |

### 4.1.2 Vertical Continuum of Care

The vertical continuum of care of the mobile population is ensured by having referral availabilities whenever the walk-in clinic is insufficient of delivering a service of a disease. In order to achieve this, there should be a local healthcare facility located within a given distance to the walk-in clinic. The referral coverage for a path $q$ is defined by the percentage of walk-in clinics having referral availabilities on the path. The coverage for referrals can be expressed mathematically as:

$$cov_{qsd} = \frac{\sum \text{walk-in clinics along path } q \text{ with referral availabilities}}{\sum \text{walk-in clinics along path } q}. \tag{4.4}$$

## 4.2 Deterministic Model

The problem of planning the walk-in clinics in the first stage is solved with a deterministic model. With this approach it is aimed to come up with a good representation of the objectives, such as the aim of achieving maximum benefit from the walk-in clinics. For the problem formulation, the following assumptions are employed in the model:

**Assumption 1** *The static patient demand occurring at location k for every disease d are assumed to be known and not varying.*

**Assumption 2** *The mobile patient demand occurring on path q for every disease d are assumed to be known, not varying and the patient visits are same throughout the whole path q.*

Table 4.4: Notation for the deterministic model

| | |
|---|---|
| Sets: | |
| $K$ | Set of all locations, including walk-in clinics, origins and destinations, $k = \{1,\ldots,K\}$ |
| $orig$ | Set of origin locations, $orig = \{1,\ldots,O\}$ |
| $dest$ | Set of destination locations, $dest = \{1,\ldots,D\}$ |
| $Q$ | Set of paths, $q = \{1,\ldots,Q\}$ |
| $HC$ | Set of healthcare locations, $hc = \{1,\ldots,HC\}$ |
| $S$ | Set of services that are offered, $s = \{$ diagnose ($di$), treatment ($tr$), care ($cr$), and referral ($rl$) $\}$ |
| $D$ | Set of high impact diseases that are covered within the scope of the walk-in clinics, $d = \{$ HIV ($h$), tuberculosis ($tb$), malaria ($ml$) and sexually transmitted infections ($sti$) $\}$ |
| Subsets: | |
| $kW$ | Set of walk-in clinic locations, $kW = \{1,\ldots,KW\}$ |
| $kC$ | Set of current walk-in clinic locations, which is a subset of $kW$ |
| $kP$ | Set of potential walk-in clinic locations, which is a subset of $kW$ |
| $kO$ | Set of walk-in clinic locations, located at the origin locations, which is a subset of $kW$ |
| $kD$ | Set of walk-in clinic locations, located at the destination locations, which is a subset of $kW$ |
| $KCDS_{ds}$ | Set of current walk-in clinic locations in the network where service $s$ for disease $d$ exists |
| $KPDS_{ds}$ | Set of potential walk-in clinic locations in the network where service $s$ for disease $d$ does not exist |
| $KR_{qd}$ | Set of walk-in clinic locations for which a referral possibilities exist for disease $d$ on path $q$ |
| Parameters: | |
| $u_1$ | Weight given to mobile demand population |
| $u_2$ | Weight given to static demand population |
| $r_d$ | Weight given to disease $d$ |
| $TSM_d$ | Total mobile score of the optimized network |
| $TSS_d$ | Total static score of the optimized network |
| $dm_{qd}$ | Amount of demand from mobile population on path $q$ for disease $d$ |
| $ds_{kd}$ | Amount of demand from static population at location $k$ for disease $d$ |
| $coc_{qd}$ | Value for continuum of care on path $q$ for disease $d$ |
| $covfeas_{qds}$ | Value for feasibility score of coverage on path $q$ for disease $d$ in terms of service $s$ |
| $cov_{qds}$ | Value for coverage on path $q$ for disease $d$ in terms of service $s$ |
| $bin_{kd}$ | Auxiliary binary variable used in if-then constraints |
| $M$ | A sufficiently large number used in if-then constraints |
| $UBDS_{sd}$ | Number of centers that can be opened for giving service $s$ for disease $d$ |
| $UB$ | Number of centers that can be opened in total |

The notation used for the deterministic model including sets and subsets, which which is helpful for the formulations, and parameters is summarized in Table 4.4.

The set $kW$ which is used for defining the walk-in clinic location, including both the set of current and set of potential locations. The potential locations are determined from possible delay locations or other similar locations which is discussed in Section 3.2.5. The sets *orig* and *dest* denote the sum of the origin and destination of the paths that are traveled by the mobile populations respectively. Regarding the coordinates of these locations, a path, $q$ is determined from every origin and destination pair. Additionally, walk-in clinic locations which are lying on the path $q$ are calculated in the model. The set of services, $s$ and set of diseases, $d$, have four elements each which is include in the scope of the problem, as described in Section 3.2.3.

The decision variables in the model are listed in three as,

$$x_{kd} = \begin{cases} 1, & \text{if a walk-in clinic is open at location } k \text{ giving any service for disease } d \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{kds} = \begin{cases} 1, & \text{if a walk-in clinic is open at location } k \text{ giving service } s \text{ for disease } d \\ 0, & \text{otherwise,} \end{cases}$$

$$z_{k} = \begin{cases} 1, & \text{if a walk-in clinic is open at location } k \\ 0, & \text{otherwise.} \end{cases}$$

The model in closed form can be written as,

$$\text{Max} \quad u_1 \sum_{d \in D} r_d * TSM_d + u_2 \sum_{d \in D} r_d * TSS_d \tag{4.5}$$

$$\text{s.t.} \quad TSM_d = \sum_{q \in Q} dm_{qd} * coc_{qd} \ , \quad \forall d \in D, \tag{4.6}$$

$$TSS_d = \sum_{k \in K} ds_{kd} * x_k \ , \quad \forall d \in D, \tag{4.7}$$

$$coc_{qd} = f(\{covfeas_{qds} | s \in S\}) \ , \quad \forall q \in Q, \forall d \in D, \tag{4.8}$$

$$covfeas_{qds} = g_{sd}(cov_{qds}) \ , \quad \forall q \in Q, \forall d \in D, \forall s \in S, \tag{4.9}$$

$$cov_{qds} = h_{sd}(\{x_{kd}, y_{ksd} | k \in K_q\}) \ , \quad \forall q \in Q, \forall d \in D, \forall s \in S, \tag{4.10}$$

$$M * bin_{kd} - x_{kd} \geq 0 \ , \quad \forall k \in K, \forall d \in D, \tag{4.11}$$

$$1 - (y_{k\{s=di\}d} + y_{k\{s=tr\}d} + y_{k\{s=cr\}d}) \leq M * (1 - bin_{kd}) \ , \quad \forall k \in K, \forall d \in D, \tag{4.12}$$

$$x_{kd} \geq (1/3) * (y_{k\{s=di\}d} + y_{k\{s=tr\}d} + y_{k\{s=cr\}d}) \ , \quad \forall k \in K, \forall d \in D, \tag{4.13}$$

$$z_k \geq (1/4) * \sum_{d \in D} x_{kd} \ , \quad \forall k \in K, \tag{4.14}$$

$$UBDS_{sd} \geq \sum_{k \in KPDS_{sd}} y_{ksd} \ , \quad \forall d \in D, \forall s \in S, \tag{4.15}$$

$$UB \geq \sum_{k \in KP} z_k, \tag{4.16}$$

$$y_{kds} = 1 \ , \quad \forall k \in KCDS_{ds}, \tag{4.17}$$

$$x_{kd} \in \{0,1\} \ , \quad \forall k \in K, \forall d \in D, \tag{4.18}$$

$$y_{kds} \in \{0,1\} \ , \quad \forall k \in K, \forall d \in D, \forall s \in S, \tag{4.19}$$

$$z_k \in \{0,1\} \ , \quad \forall k \in K. \tag{4.20}$$

The objective function of the model, (4.5), maximizes the scores achieved from mobile and static demand in accordance with the user defined weights given to these scores, $u_1$ and $u_2$.

The first part of the score which is for mobile demand is calculated through the continuum of care offered to mobile populations along the path $q$ and the demand through the path $q$, (4.6). The other score is for the static demand, which is calculated from number of static demand occurring at an open walk-in clinic on the network, (4.7).The continuum of care values are calculated for every flow and every disease by a function of feasibility scores which is defined over every path, $q$, service, $s$, and disease, $d$ (4.8). The feasibility scores are normalized coverage values over paths, which are put in a scale between 0 and 1 (4.9). The coverage values are calculated specifically for every disease $d$ and service $s$ over path $q$ in which the calculation of coverage value depends on the combination of service and disease, (4.10). The coverage value depends on the decision variables $x_{kd}$ and $y_{kds}$ which indicate the location and planning of the walk-in clinics.

With the Equations (4.11), (4.12) and (4.13), the relation between decision variable $x_{kd}$ and $y_{kds}$ is maintained. It is ensured that, if $y_{kds}$ gets a value of 0, $x_{kd}$ is also forced to zero and vice versa. With (4.14), a similar relation between $x_{kd}$ and $z_k$ is ensured. The two constraints, (4.15) and (4.16) put an upper bound value on number of services that can be opened on a given network. The constraint (4.17) assigns value 1 for the locations that are currently existing on the network. The model is concluded with the sign constraints.

As observed from the model constraints (4.8), (4.9) and (4.10), the variables in the model are described through some functions which will be discussed in detail in the following sections.

## 4.3 Calculation of Coverage Values

In Section 4.1, definitions that are suitable for defining coverage of service and disease combinations are explained. As can be seen from the descriptions, the coverage values are defined by integrals, which are not friendly in terms of linear programming. In order to resolve this issue and apply the coverage approaches, the linearization of the approaches should be formulated. In this section, the methods for linearizing the formulations are explained.

### 4.3.1 Neighbourhood Definition

In order to tackle the problem of non-linear equations, an alternative formulation is developed. With this solution, the model determines whether two locations on path $q$ are neighbours and calculates the coverage accordingly. For this solution approach, the following assumption is made.

**Assumption 1** *The mobile populations are assumed to be making circular trips over the paths, meaning that once mobile populations start from origin of a path and arrive to destination, they return back to origin.*

**Assumption 2** *The network on which optimization is applied is assumed to be undirected, meaning the travel times from origin to destination and from destination to origin are the same. The travel times in the network are symmetrical.*

To formulate neighborhood definition, the notation given in Table 4.4 is extended with the sets which are given in Table 4.5.

Table 4.5: Additional notation for the deterministic model

| | |
|---|---|
| Subsets: | |
| $O_q$ | Origin of the path $q$, as a subset of *orig* |
| $D_q$ | Destination of the path $q$, as a subset of *dest* |
| $K_q$ | Set of walk-in clinic locations $kW$ that are located on the path $q$ |
| $KA_{kq}$ | Set of walk-in clinic locations $kW$ that are located after $k$ until the destination of the path $q$ |
| $KEO_q$ | Set of walk-in clinic locations $kW$ that are located on the path $q$ but not on the origin of the path |
| $KED_q$ | Set of walk-in clinic locations $kW$ that are located on the path $q$ but not on the destination of the path |
| $KEOD_q$ | Set of walk-in clinic locations $kW$ that are located on the path $q$ but not on the origin and destination of the path |
| $KO_q$ | Set of walk-in clinics that are located at the origin on path $q$ |
| $KD_q$ | Set of walk-in clinics that are located at the destination on path $q$ |
| $L_q$ | Set of all locations on path $q$ including walk-in clinic locations, origin and destination meaning, $O_q \cup D_q \cup K_q$ |
| $LA_{kq}$ | Set of locations on path $q$ which are located after $k$ until the destination |

The following decision variable is added in the model to determine the neighbours along the paths, adapted from [13].

$$i_{klqsd} = \begin{cases} 1, & \text{if } k \text{ and } l \text{ are neighbouring pairs for service } s \text{ and disease } d \\ & \text{along the path } q \text{ where } \{(k,l) \in L_q \times L_q\} \\ 0, & \text{otherwise} \end{cases} \tag{4.21}$$

Two locations can be neighbours if and only if while traveling from the first location, the ending location is reached without passing by any location in between. This is ensured with the following set of constraints,

$$\sum_{l \in LA_{kq}} i_{klqsd} = y_{ksd} \qquad \forall q \in Q, \forall k \in K_q, \forall s \in S \text{ and } \forall d \in D \tag{4.22}$$

$$\sum_{k \in L_q} i_{klqsd} = y_{lsd} \qquad \forall q \in Q, \forall l \in K_q, \forall s \in S \text{ and } \forall d \in D \tag{4.23}$$

$$\sum_{l \in LA_{kq}} i_{klqsd} = 1 \qquad \forall q \in Q, \forall k \in O_q, \forall s \in S \text{ and } \forall d \in D \tag{4.24}$$

$$\sum_{k \in L_q} i_{klqsd} = 1 \qquad \forall q \in Q, \forall l \in D_q, \forall s \in S \text{ and } \forall d \in D \tag{4.25}$$

$$i_{klqsd} \in \{0,1\} \qquad \forall q \in Q, \forall k \in K, \forall l \in K_{kq}, \forall s \in S \text{ and } \forall d \in D. \tag{4.26}$$

In these constraints, it is ensured that each location can have only one successor and only one predecessor. From Constraints (4.22) and (4.23), two walk-in clinics can be service $s$ and disease $d$ neighbours on path $q$ if and only if there is an open walk-in clinic in neighboring position. Every origin has exactly one successor (4.24) and every destination has one predecessor (4.25). This notation is applied to approach definitions which are discussed as next sections.

### 4.3.2 Linearization for Approach 1

For the linearization, new set of parameters are introduced to the model. These parameters are summarized in Table 4.6.

Table 4.6: Additional notation for the deterministic model for linearization

| | |
|---|---|
| Parameters: | |
| $t_{klq}$ | Travel time between locations $k$ and $l$ on path $q$, including the delay times in between where both $k$ and $l$ are elements of the set $L_q$ |
| $\hat{t}_{sd}$ | The critical intervention time for service $s$ of disease $d$ |
| $tTotal_q$ | Total traveling time on path $q$, starting from origin and going to destination including the delay durations |
| $dOD_q$ | The total delay time occurring at the origin and destination of path $q$ |
| $del_k$ | The amount of delay time spent at location $k$ |

The coverage value is calculated as,

$$cov_{qsd} = \frac{A+B+C+D}{(2*tTotal_q)+dOD_q} \tag{4.27}$$

where

$$A = 2*\sum_{k\in K_q}\sum_{l\in KA_{kq}} i_{klqsd}*crtTime_{klqsd}$$

$$B = \sum_{k\in O_q}\sum_{l\in KEO_q} i_{klqsd}*crtTimeO_{klqsd}$$

$$C = \sum_{k\in KED_q}\sum_{l\in D_q} i_{klqsd}*crtTimeD_{klqsd}$$

$$D = 2*\sum_{k\in KEOD_q}\sum_{L_q} i_{klqsd}*del_k + \sum_{k\in O_q}\sum_{l\in KO_q} i_{klqsd}*del_l + \sum_{k\in KD_q}\sum_{l\in D_q} i_{klqsd}*del_k$$

and

$$crtTime_{klqsd} = min\{t_{klq},\hat{t}_{sd}\}$$

$$crtTimeO_{klqsd} = min\{(2*t_{klq}+del_{KO_q}),\hat{t}_{sd})$$

$$crtTimeD_{klqsd} = min\{(2*t_{klq}+del_{KD_q}),\hat{t}_{sd}).$$

In (4.27), the total amount of time that is spent along the path $q$ as covered is proportioned to the total time spent along the road. The total time spent on the path covered is split in 4 parts as $A$, $B$, $C$ and $D$. Part $A$ refers to the sections covered on the road by the walk-in clinic that are located along the path. It looks at the neighboring walk-in clinics and calculates the coverages

accordingly. It is multiplied with the value 2, due to the circular path assumption. Part $B$ and $C$ refer to origin and destination coverages respectively. These two parts will play role if a walk-in clinic at an origin or a destination is not located. Due to circular trips, if a walk-in clinic is not located at origin and/or destination, the travel time between two consecutive walk-in clinics will be different and has to be reflected in the calculations. The final part $D$, is for the coverage of delay duration. As mentioned before, delays at the locations may occur resulting as an increase in travel times. If a walk-in clinic is located at a location with delay duration, through the delay time, the mobile population will be covered and it requires to be added in the covered travel times, Part $D$ consists of three parts to differentiate between delay times spent at walk-in clinics, origins and destinations respectively. The delay covered at walk-in clinics is also multiplied with the value of 2 due to circular trips.

To clarify how the formulas work, example cases can be provided over a given path, $q$. Three cases will be provided showing the difference between how formulation changes when a walk-in clinic is located at an origin or a destination.

*Case 1:* In the first case, the path does not have any walk-in clinics located at the origin and destination points. Figure 4.1 illustrates the path with a simple diagram. The path is circular, showing the trip from an origin to a destination with walk-in clinics $X$, $Y$ and $Z$ which are opened and from destination to origin with $Z'$, $Y'$ and $X'$. The blue shades over the path indicates the parts of the path that is covered, meaning in critical time intervals. For simplification, only single disease and service combination is analyzed and $s$ and $d$ indices are dropped.

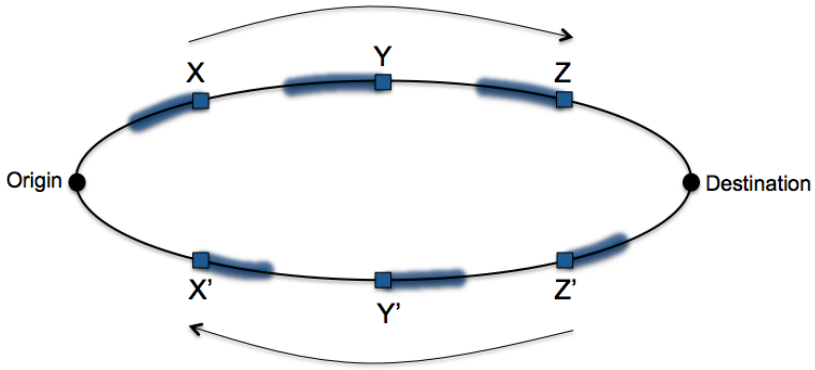


Figure 4.1: An example path with no walk-in clinic at origin and destination

To provide example numerically, the following fictional data is used. Travel times between locations without delays are given in Table 4.7. Furthermore, each of the locations, have a

Table 4.7: Travel times between location without delay durations

| | orig | X | Y | Z | dest |
|---|---|---|---|---|---|
| orig | 0 | 5 | 12 | 20 | 29 |
| X | 5 | 0 | 7 | 15 | 24 |
| Y | 12 | 7 | 0 | 8 | 17 |
| Z | 20 | 15 | 8 | 0 | 9 |
| dest | 29 | 24 | 17 | 9 | 0 |

Table 4.8: Delay durations

| orig | X | Y | Z | dest |
|---|---|---|---|---|
| 6 | 3 | 4 | 10 | 15 |

delay duration spent over, which is summarized in the Table 4.8.

For the given path, the $cov_{qds}$ value can now be calculated. Initially, the $tTotal_q$ will be calculated as $(5+7+8+9)+(3+4+10)=46$, which is the sum of the neighboring travel times and the delay durations of the walk-in clinic locations. The sum of the delay at origin and destination, denoted by $dOD_q$ will be equal to $6+15=21$. This would make denominator equal to value of $46*2+21=113$. This value indicates the total time spent along the circular path that the mobile population is traveling, including the delays along the road which contribute to total time extensively.

When the numerator is analyzed, it is important to determine the $i_{klqsd}$ values of the given path. For the given case, $i_{klqsd}$ values will be as (the $q$, $d$ and $s$ indices are dropped):

$$i_{orig,X} = 1$$
$$i_{XY} = 1$$
$$i_{YZ} = 1$$
$$i_{Z,dest} = 1$$

and all the other $i_{kl}$ values will be zero.

The next step would be the calculation of $A$, $B$, $C$ and $D$ values in the formulation with the

help of $i_{kl}$ values. The calculations will be as follows:

$$\begin{aligned}
A &= 2*[(i_{XY}*crtTime_{XY})+(i_{XZ}*crtTime_{XZ})+(i_{YZ}*crtTime_{YZ})] \\
&= 2*[(1*crtTime_{XY})+(1*crtTime_{YZ})] \\
B &= i_{orig,X}*crtTimeO_{orig,X} \\
C &= i_{Z,dest}*crtTimeD_{Z,dest} \\
D &= 2*[(i_{XY}*del_X)+(i_{YZ}*del_Y)+(i_{Z,dest}*del_Z)]+0+0.
\end{aligned}$$

The last two terms in calculation of $D$ are zero, as there is no walk-in clinic located at origin and destination, the delay duration occurring at the origin and destination are not included.

To finish the calculation of $A$, $B$, $C$ and $D$ values, a critical intervention time, $\hat{t}_{sd}$, is required. For the current case, the value is assumed as 2. The calculations will be as follows:

$$\begin{aligned}
A &= 2*[(1*min\{7,2\})+(1*min\{8,2\})] \\
&= 2*[2+2] \\
&= 8 \\
B &= 1*min\{(2*5+6),2\} \\
&= 2 \\
C &= 1*min\{(2*9+15),2\} \\
&= 2 \\
D &= 2*[(1*3)+(1*4)+(1*10)]+0+0. \\
&= 34
\end{aligned}$$

Putting together all the values calculated, the coverage values of the path can be finalized as,

$$cov = \frac{8+2+2+34}{113} = 0.319.$$

Notice that, if the critical intervention time was a value greater than or equal to 33, the coverage of the path would have been 1.

*Case 2:* In the second case, the path has walk-in clinics located at the origin and destination points. Figure 4.2 illustrates the path with a simple diagram. The path is circular, showing the trip from origin to destination with walk-in clinics $W$, $X$, $Y$, $Z$ and $Q$ which are opened

Table 4.9: Travel times between location without delay durations

| | orig/W | X | Y | Z | dest/Q |
|---|---|---|---|---|---|
| orig/W | 0 | 5 | 12 | 20 | 29 |
| X | 5 | 0 | 7 | 15 | 24 |
| Y | 12 | 7 | 0 | 8 | 17 |
| Z | 20 | 15 | 8 | 0 | 9 |
| dest/Q | 29 | 24 | 17 | 9 | 0 |

Table 4.10: Delay durations

| orig/W | X | Y | Z | dest/Q |
|---|---|---|---|---|
| 6 | 3 | 4 | 10 | 15 |

and from destination to origin with $Q'$, $Z'$, $Y'$, $X'$ and $W'$. The walk-in clinics that are located at origin and destination, $W$ and $Q$, are actually the same points as origin and destination respectively. For calculation purposes and for illustration, they are considered as being located at a very small difference to origin and destination. Again, only single disease and service combination is analyzed and $s$ and $d$ indices are dropped.

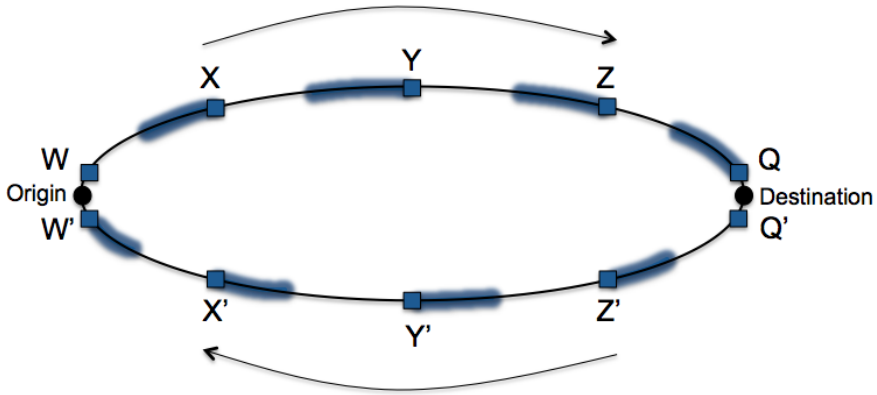


Figure 4.2: An example path with walk-in clinics at origin and destination

To provide example numerically, again fictional data is used. Travel times between locations without delays are given in Table 4.9. Furthermore, each of the location in addition have a delay duration spent over, which is summarized in the Table 4.10.

For the given path, the $cov_{qds}$ value can now be calculated. Initially, the $tTotal_q$ will be same as the value in Case 1, $(5+7+8+9)+(3+4+10)=46$, which is the sum of the neighboring travel times and the delay durations of the walk-in clinic locations. The sum of the delay at origin and destination, denoted by $dOD_q$ will also be the same value as Case 1, $6+15=21$.

This would make denominator equal to value of $46*2+21=113$.

When the numerator is analyzed, again the $i_{kl}$ values of the given path are determined. For the given case, $i_{kl}$ values will be as:

$$i_{orig,W}=1$$
$$i_{WX}=1$$
$$i_{XY}=1$$
$$i_{YZ}=1$$
$$i_{ZQ}=1$$
$$i_{Q,dest}=1$$

and all the other $i_{kl}$ values will be zero.

The next step would be the calculation of $A$, $B$, $C$ and $D$ values in the formulation with the help of $i_{kl}$ values. The calculations will be as follows:

$$\begin{aligned}
A &= 2*[(i_{WX}*crtTime_{WX})+(i_{XY}*crtTime_{XY}) \\
&\quad +(i_{YZ}*crtTime_{YZ})+(i_{ZQ}*crtTime_{ZQ})] \\
&= 2*[(1*crtTime_{WX})+(1*crtTime_{XY}) \\
&\quad +(1*crtTime_{YZ})+(1*crtTime_{ZQ})] \\
B &= 0 \\
C &= 0 \\
D &= 2*[(i_{XY}*del_X)+(i_{YZ}*del_Y)+(i_{Z,dest}*del_Z)] \\
&\quad +i_{orig,W}*del_W+i_{Q,dest}*del_Q.
\end{aligned}$$

The term $B$ and $C$ are 0 because there is a walk-in clinic located at the origin and destination location which means origin and destination are not neighbors with walk-in clinics that are located *else than* origin and destination. The last two terms in calculation of $D$ are this time not zero, as there are walk-in clinics located at origin and destination therefore the delay duration occurring at the origin and destination are included.

To finish the calculation of $A$, $B$, $C$ and $D$ values, a critical intervention time, $\hat{t}_{sd}$, is required.

For the current case, the value is again assumed as 2. The calculations will be as follows:

$$\begin{aligned}
A &= 2 * [(1 * min\{5,2\}) + (1 * min\{7,2\}) \\
&\quad + (1 * min\{8,2\}) + (1 * min\{9,2\}] \\
&= 2 * [2+2+2+2] \\
&= 16 \\
B &= 0 \\
C &= 0 \\
D &= 2 * [(1*3) + (1*4) + (1*10)] + 6 + 15 \\
&= 55.
\end{aligned}$$

Putting together all the values calculated, the coverage values of the path can be finalized as,

$$cov = \frac{16+0+0+55}{113} = 0.628.$$

Notice that, by opening centers at origin and destination, the coverage value of the path is increased compared to Case 1. If the critical intervention time was 9 instead of 2, the coverage would have been equal to 1.

*Case 3:* In the final case, the path has walk-in clinic located at the destination but not at the origin point. Figure 4.3 illustrates the path with simple diagram. The path is circular, showing the trip from origin to destination with walk-in clinics $X$, $Y$, $Z$ and $Q$ which are opened and from destination to origin with $Q'$, $Z'$, $Y'$ and $X'$. The walk-in clinic that is located at destination, $Q$, is actually the same point as destination. For calculation purposes and for illustration, it is considered as being located at a very small difference to destination. Again, only single disease and service combination is analyzed and $s$ and $d$ indices are dropped.

The fictional travel times between locations without delays are given in Table 4.11. Furthermore, each of the location in addition have a delay duration spent over, which is summarized in the Table 4.12.

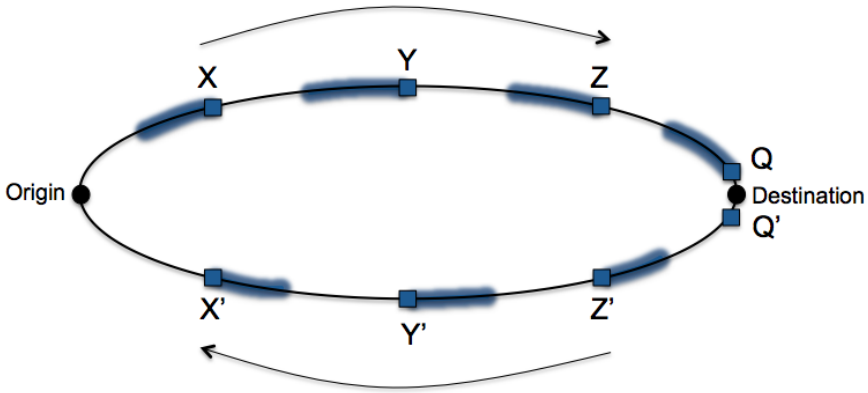


Figure 4.3: An example path with walk-in clinics at destination

For the given path, the $cov_{qds}$ value can now be calculated. The $tTotal_q$ will be same as the value in Case 1 and 2, $(5+7+8+9)+(3+4+10) = 46$, which is the sum of the neighboring travel times and the delay durations of the walk-in clinic locations. The sum of the delay at origin and destination, denoted by $dOD_q$ will also be the same value as Case 1 and 2 too, $6+15 = 21$. This would make denominator equal to value of $46*2+21 = 113$.

When the numerator is analyzed, again the $i_{kl}$ values of the given path are determined. For the given case, $i_{kl}$ values will be as:

$$i_{orig,X} = 1$$
$$i_{XY} = 1$$
$$i_{YZ} = 1$$
$$i_{ZQ} = 1$$
$$i_{Q,dest} = 1$$

and all the other $i_{kl}$ values will be zero.

Table 4.11: Travel times between location without delay durations

| | orig | X | Y | Z | dest/Q |
|---|---|---|---|---|---|
| orig | 0 | 5 | 12 | 20 | 29 |
| X | 5 | 0 | 7 | 15 | 24 |
| Y | 12 | 7 | 0 | 8 | 17 |
| Z | 20 | 15 | 8 | 0 | 9 |
| dest/Q | 29 | 24 | 17 | 9 | 0 |

Table 4.12: Delay durations

| orig | X | Y | Z | dest/Q |
|---|---|---|---|---|
| 6 | 3 | 4 | 10 | 15 |

The next step would be the calculation of $A$, $B$, $C$ and $D$ values in the formulation with the help of $i_{kl}$ values. The calculations will be as follows:

$$\begin{aligned}
A &= 2*[(i_{XY}*crtTime_{XY})+(i_{YZ}*crtTime_{YZ})+(i_{ZQ}*crtTime_{ZQ})] \\
&= 2*[(1*crtTime_{XY})+(1*crtTime_{YZ})+(1*crtTime_{ZQ})] \\
B &= i_{orig,X}*crtTimeO_{orig,X} \\
C &= 0 \\
D &= 2*[(i_{XY}*del_X)+(i_{YZ}*del_Y)+(i_{Z,dest}*del_Z)]+0+i_{Q,dest}*del_Q.
\end{aligned}$$

The term $B$ is non-zero because there are no walk-in clinics located at origin and it requires separate handling of the coverage values.The term $C$ is 0 because there is a walk-in clinic located at the destination location which means destination is not neighbor with walk-in clinics that are located *else than* destination. The "0" terms in calculation of $D$ is for the no walk-in clinic located at origin which makes the delay duration occurring at the origin not included.

To finish the calculation of $A$, $B$, $C$ and $D$ values, a critical intervention time, $\hat{t}_{sd}$, is required. For the current case, the value is again assumed as 2. The calculations will be as follows:

$$\begin{aligned}
A &= 2*[(1*min\{7,2\})+(1*min\{8,2\})+(1*min\{9,2\})] \\
&= 2*[2+2+2] \\
&= 12 \\
B &= 1*min\{(2*5+6),2\} \\
&= 2 \\
C &= 0 \\
D &= 2*[(1*3)+(1*4)+(1*10)]+0+15 \\
&= 49.
\end{aligned}$$

Putting together all the values calculated, the coverage values of the path can be finalized as,

$$cov = \frac{12+2+0+49}{113} = 0.558.$$

Notice that, by opening center at destination, the coverage value of the path is increased compared to Case 1. However, since there is no center at the origin, the coverage is less compared to Case 2. If the critical intervention time was 16 instead of 2, the coverage would have been equal to 1.

### 4.3.3 Linearization for Approach 2

For the linearization, in addition to parameters in previous part, few new parameters are introduced given in Table 4.13.

Table 4.13: Additional notation for the deterministic model for linearization

| | |
|---|---|
| Parameters: | |
| $\hat{t1}_{sd}$ | The first critical intervention time for service $s$ of disease $d$ |
| $\hat{t2}_{sd}$ | The second critical intervention time for service $s$ of disease $d$ |

The coverage value is calculated as,

$$cov_{qsd} = \frac{A+B+C+D}{(2*tTotal_q)+dOD_q} \tag{4.28}$$

where

$$A = 2*(\sum_{k\in K_q}\sum_{l\in KA_{kq}} i_{klqsd}*crtTime_{klqsd})$$

$$B = \sum_{k\in O_q}\sum_{l\in KEO_q} i_{klqsd}*crtTimeO_{klqsd}$$

$$C = \sum_{k\in KED_q}\sum_{l\in D_q} i_{klqsd}*crtTimeD_{klqsd}$$

$$D = 2*(\sum_{k\in KEOD_q}\sum_{L_q} i_{klqsd}*del_k)$$
$$+\sum_{k\in O_q}\sum_{l\in KO_q} i_{klqsd}*del_l$$
$$+\sum_{k\in KD_q}\sum_{l\in D_q} i_{klqsd}*del_k$$

and

$$crtTime_{klqsd} = \begin{cases} t_{klq}, & \text{if } t_{klq} \leq \hat{t1}_{sd} \\ t_{klq} - \frac{(t_{klq} - \hat{t1}_{sd})^2}{2*(\hat{t2}_{sd} - \hat{t1}_{sd})}, & \text{if } \hat{t1}_{sd} \leq t_{klq} \leq \hat{t2}_{sd} \\ \hat{t1}_{sd} + 0.5 * (\hat{t2}_{sd} - \hat{t1}_{sd}), & \text{otherwise,} \end{cases}$$

$$crtTimeO_{klqsd} = \begin{cases} 2 * t_{klq} + del_{KO_q}, & \text{if } 2 * t_{klq} + del_{KO_q} \leq \hat{t1}_{sd} \\ 2 * t_{klq} + del_{KO_q} - \frac{(t_{klq} - \hat{t1}_{sd})^2}{2*(\hat{t2}_{sd} - \hat{t1}_{sd})}, & \text{if } \hat{t1}_{sd} \leq 2 * t_{klq} + del_{KO_q} \leq \hat{t2}_{sd} \\ \hat{t1}_{sd} + 0.5 * (\hat{t2}_{sd} - \hat{t1}_{sd}), & \text{otherwise,} \end{cases}$$

$$crtTimeD_{klqsd} = \begin{cases} 2 * t_{klq} + del_{KD_q}, & \text{if } 2 * t_{klq} + del_{KD_q} \leq \hat{t1}_{sd} \\ 2 * t_{klq} + del_{KD_q} - \frac{(t_{klq} - \hat{t1}_{sd})^2}{2*(\hat{t2}_{sd} - \hat{t1}_{sd})}, & \text{if } \hat{t1}_{sd} \leq 2 * t_{klq} + del_{KD_q} \leq \hat{t2}_{sd} \\ \hat{t1}_{sd} + 0.5 * (\hat{t2}_{sd} - \hat{t1}_{sd}), & \text{otherwise.} \end{cases}$$

The application of the formulation is same as Approach 1, which was demonstrated through the Cases. With this approach, only the critical time calculations will be different, matching with the requirements of the approach.

### 4.3.4 Linearization of Approach 3

In this approach, the function that needs to be linearized includes the term "expected traveling time" which needs to be further elaborated. The expected traveling time can be explained as at a random point chosen in time, the expected time that the demand will be met by visiting a walk-in clinic. This question has been analyzed in [13] and the following formulation is proved to be applicable:

**Theorem 5** *(Theorem 5.3 in [13]) The expected traveling time for truck drivers traveling in the segment between k and l is equal to* $\frac{1}{2}t_{kl}^2$*. By conditioning this value on the total number of segments along path q, the expected traveling time along path q,* $(ER_q)$*, is defined as:*

$$ER_q = \frac{1}{2 * tTotal_q} \sum_{k \in K_q} \sum_{l \in KA_{kq}} i_{klq} t_{kl}^2 \tag{4.29}$$

*where* $i_{klq}$ *denotes that location k and l are neighbours, meaning they are travelled consecutively, along path q.*

This expected traveling time in segment $(k,l)$, which is $\frac{1}{2}t_k l^2$, is used in the critical time calculations for the third coverage approach. The details of the formulation can be found below.

$$cov_{qsd} = \frac{A+B+C}{2*((2*tTotal_q)+dOD_q)} \tag{4.30}$$

where

$$A = 2*(\sum_{k\in K_q}\sum_{l\in KA_{kq}} i_{klqsd}*crtTime_{klqsd})$$

$$B = \sum_{k\in O_q}\sum_{l\in KEO_q} i_{klqsd}*crtTimeO_{klqsd}$$

$$C = \sum_{k\in KED_q}\sum_{l\in D_q} i_{klqsd}*crtTimeD_{klqsd}$$

and

$$\begin{aligned} crtTime_{klqsd} &= t_{klq}^2 \\ crtTimeO_{klqsd} &= (2*t_{klq}+del_{KO_q})^2 \\ crtTimeD_{klqsd} &= (2*t_{klq}+del_{KD_q})^2 \end{aligned}$$

In this coverage formulation, the application is again similar to Approach 1 and 2. The critical time definitions are changed with regards to the requirement of the Approach that is defined in Theorem 5. The delay terms that were included in the Approach 1 and 2 are not relevant with this definition anymore, as the delays are incorporated in critical time definitions.

### 4.3.5 Linearization of referrals

The referral coverage in Section 4.1.2 was defined as the percentage of walk-in clinics over path $q$ having referral availabilities for disease $d$ with 4.4. This can be formulated as,

$$cov_{qsd} = \frac{\sum_{k\in KR_q d} y_{k\{s=rl\}d}}{\sum_{k\in K_q} y_{ksd}}\ , \quad \forall q \in Q \text{ and } d \in D. \tag{4.31}$$

As also seen from (4.31), the division of two decision variables implies non-linearity in the model and requires to be adjusted. The notation given in Table 4.14 is added to the model to assist the linearization of the calculation.

Table 4.14: Additional notation for the deterministic model for linearization

| | |
|---|---|
| Sets | |
| $I$ | Set of all possible number of walk in clinics that can be established over the paths, $\{0,\ldots,\max_q \lvert K_q\rvert\}$ |
| $I_q$ | Subset of $I$, all possible number of walk-in clinics that can be established on path $q$, $\{0,1,\ldots,\lvert K_q\rvert\}$ |
| Parameters | |
| $\theta_{id}$ | A variable which is equal to the number of walk-in clinics on path $q$ which $j$ corresponds to |
| $tUB_{qd}$ | Upper bound on number of walk-in clinics on path $q$ for disease $d$ |
| Decision Variables | |
| $tRef_{qd}$ | Sum of the number of the established walk-in clinics on path $q$ having referral availabilities for disease $d$ |
| $nBin_{iqd}$ | Equals 1 if number of established centers on the path $q$ for disease $d$ is equal to $\theta_{id}$ and 0 otherwise |
| $wLin_{iqd}$ | Equals to the number of the established walk-in clinics on path $q$ having referral availabilities for disease $d$ whenever $nBin_{iqd}$ is equal to 1 and 0 otherwise |
| $yBin_{iqd}$ | Equals 1 if number of established centers on the path $q$ for disease $d$ is equal to $\theta_{id}$ and 0 otherwise |
| $zBin_{kd}$ | Auxiliary variable for the if-then constraints. |

(4.31) is re-written as the following making the numerator a decision variable and denominator a parameter to avoid non-linearity by division of two decision variables.

$$cov_{qsd} = \sum_i \frac{tRef_{qd} * nBin_{iqd}}{\theta_{id}} \ , \quad \forall i \in I_q \text{ and } \forall d \in D \tag{4.32}$$

The coverage definition given in (4.32) is reformulated as below and the constraints for definition are added.

$$cov_{qsd} = \sum_i \frac{wLin_{iqd}}{\theta_{id}} \ , \quad \forall i \in I_q \text{ and } \forall d \in D \tag{4.33}$$

such that:

$$tRef_{qd} = \sum_{k \in KR_{qd}} y_{k\{s=rl\}d} \ , \quad \forall q \in Q \text{ and } \forall d \in D,$$

$$wLin_{iqd} \leq tUB_{qd} * nBin_{iqd} \ , \quad \forall i \in I_q, \forall q \in Q \text{ and } \forall d \in D,$$

$$wLin_{iqd} \leq tRef_{qd} \ , \quad \forall i \in I_q, \forall q \in Q \text{ and } \forall d \in D,$$

$$wLin_{iqd} \geq tRef_{qd} - tUB_{qd} * (1 - nBin_{iqd}) \ , \quad \forall i \in I_q, \forall q \in Q, \forall d \in D,$$

$$\sum_{k \in KR_{qd}} x_{kd} - \theta_{id} \leq M * (1 - yBin_{iqd}) \ , \quad \forall i \in I, \forall q \in Q \text{ and } \forall d \in D,$$

$$\theta_{id} - \sum_{k \in KR_{qd}} x_{kd} \leq M * (1 - yBin_{iqd}) \ , \quad \forall i \in I, \forall q \in Q \text{ and } \forall d \in D,$$

$$M * yBin_{iqd} - nBin_{iqd} \geq 0 \ , \quad \forall i \in I, \forall q \in Q \text{ and } \forall d \in D,$$

$$tRef_{qd} \geq 0 \ , \quad \forall q \in Q \text{ and } \forall d \in D,$$

$$wLin_{iqd} \geq 0 \ , \quad \forall i \in I, \forall q \in Q \text{ and } \forall d \in D,$$

$$nBin_{iqd} = \{0,1\}, \quad yBin_{iqd} = \{0,1\} \ , \quad \forall i \in I, \forall q \in Q \text{ and } \forall d \in D.$$

In addition to calculation of referral percentages, some additional constraints should be added as well.

$$M * zBin_{kd} - y_{k\{s=rl\}d} \geq 0 \qquad , \forall k \in K \text{ and } \forall d \in D$$

$$M * (1 - zBin_{kd}) \geq -x_{kd} \qquad , \forall k \in K \text{ and } \forall d \in D$$

$$zBin_{kd} = \{0,1\} \qquad , \forall k \in K \text{ and } \forall d \in D$$

## 4.4 Calculation of Coverage Feasibility Score Values

As seen in the deterministic model given in Section 4.2, the coverage values calculated from the approaches are converted to feasibility score values. The aim with the feasibility score

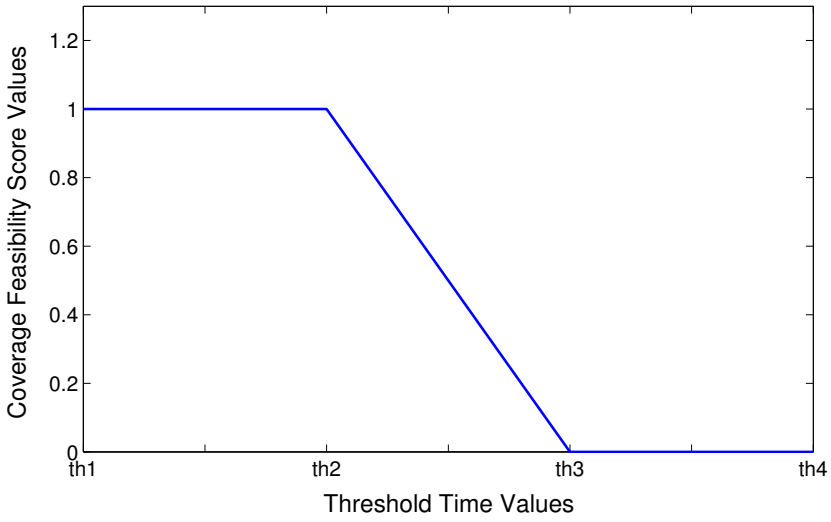


Figure 4.4: Relation between threshold time and feasibility score values

values is to put the coverage values in a scale between 0 and 1. For approaches 1 and 2, the coverage values are already calculated in a range (0,1), therefore it is not necessarily needed to convert them in this scale. The following equation can be applicable for all diseases except HIV (as HIV belongs to Approach 3),

$$covfeas_{qds} = cov_{qds} \ , \quad \forall q \in Q \text{ and } \forall s \in S.$$

This mapping between feasibility score value and coverage values for service disease combinations belonging to Approach 1 and 2, could have been made differently, regarding an additional function for the mapping. However, in this model, the function is assumed to be linear and equal.

For the third approach however, the coverage values are not in a scale between 0 and 1 which requires an additional mapping function. For this purpose, piece-wise linear functions are used to help this conversion process. Since in Approach 3, HIV diagnosis, treatment and care are handled, each of these combinations are matched with separate piece-wise linear functions having different parameters regarding the characteristics of the diseases and required services. In order to incorporate the piece-wise linear functions in the linear program, the *lambda method* is used [16].

In order to explain how Lamda Method is applied, the sample figure which is given in Figure 4.4 will be used as reference. The shape of the graph is the same as the piece-wise linear

function that are used in the model as it resembles the characteristics of the service and diseases belonging this approach properly. The threshold time values are represented with $th_1$, $th_2$, $th_3$ and $th_4$ to denote the critical time values in the approach symbolically. In the figure, it is seen that the piece-wise linear function is composed of three pieces. The first piece is between $th_1$ and $th_2$ where the coverage feasibility score is equal to 1. This means that, if the expected traveling time on a path is belonging to this piece of the function, a feasibility score value of 1 will be assigned. The second piece is between critical times $th_2$ and $th_3$. If the expected traveling time has a value which falls between this range, it is assigned with a value between 0 and 1, regarding the slope of the piece. Finally, the third piece which is indicated with value beyond $th_3$ denotes the expected time which indicated no coverage feasibility on the path, due to very high values.

The following constraints are used for calculation of feasibility score values derived from piece-wise linear functions,

$$cov_{qds} = \sum_{n=1}^{4} \lambda_{nqsd} * th_{nds} - (1 - \sum_{k \in O_q} \sum_{l \in K_q} i_{klqsd}) * th_{3ds} \ , \quad \forall q \in Q, \forall s \in S \text{ and } \forall d \in D, \tag{4.34}$$

$$covfeas_{qds} = 1 * \lambda_{1qds} + 1 * \lambda_{2qds} + 0 * \lambda_{3qds} + 0 * \lambda_{4qds} \ , \quad \forall q \in Q, \forall s \in S \text{ and } \forall d \in D, \tag{4.35}$$

$$\sum_{n=1}^{4} \lambda_{nqds} = 1 \ , \quad \forall q \in Q, \forall s \in S \text{ and } \forall d \in D, \tag{4.36}$$

$$\lambda_{nqds} \geq 0 \ , \quad \forall q \in Q, \forall s \in S, \forall d \in D \text{ and } \forall n \in N. \tag{4.37}$$

where

$\lambda_{nqds}$ Non negative weight for path $q$, service $s$ of disease $d$

$th_{nds}$ Threshold time value for service $s$ of disease $d$.

With this formulation, lambda values are calculated with respect to the coverage values that are initially calculated. In (4.34), if no walk-in clinic is located along the path $q$, the threshold value of the path is increased by $th_{3ds}$, making the expected traveling time large and feasibility score 0. In addition to these constraints, $\lambda_{nqds}$ is also subject to restrictions from to Special Order Set, type 2 (SOS2). This means that out of the set of non-negative variable $\lambda_{nqds}$, at most to variables can be non-zero. Moreover, the two variable must be adjacent to each other in a given fixed order list. This special situation for the variable $\lambda_{nqds}$ is taken into consideration in the coding phase.

## 4.5 Calculation of Continuum of Care Values

The final step for calculation continuum of care values is the conversion of coverage feasibility scores to continuum of care. As explained in the deterministic model, continuum of care values are calculated over two indices which are path, $q$, and disease, $d$. Therefore, through this conversion, the service index, $s$ should be handled. This is completed with the support of two weights which are:

$w1_d$ Weight given to disease $d$

$w2_{ds}$ Weight given to service $s$ of disease $d$.

Out of these two parameters, $w1_d$ is the weight given to disease $d$ on a scale from 1 to 10. Each disease is assigned with weights indicating their importance for the network on which walk-in clinics will be built. This allows flexilibilty to plan disease specifically, such as only malaria focused network can be established if desired. The second parameter $w2_{sd}$ is the weight given to the service $s$ of a disease $d$ on a scale from 1 to 10. The interpretation of this value is same as the $w1_d$ parameter.

To use these weights in the model, the values are normalized in the following form,

$$r_d = \frac{w1_d}{\sum_d w1_d} \tag{4.38}$$

$$r2_{ds} = \frac{w2_{ds}}{\sum_s w2_{ds}}. \tag{4.39}$$

From (4.38), the calculated $r_d$ value is the same value given in objective function, (4.5). The continuum of care score is calculated with the assistance of $r2_{ds}$ values as,

$$coc_{qd} = \sum_s covfeas_{qds} * r2_{ds} \ , \quad \forall q \in Q \text{ and } \forall d \in D.$$

# CHAPTER 5

# EXTENDING MATHEMATICAL FORMULATION - STOCHASTIC MODEL

In Chapter 4, the details of the deterministic model were explained. In the deterministic model, it was assumed that both mobile and static demand were known and occurring without variation. However, this is not a realistic assumption and should be relaxed within the model. Regarding these features of the demand, in this section adaptations to the deterministic model will be made to convert it into a stochastic model.

One of the problems with the demand is that demand values are not accurately known or estimated with high validity. For some cases, it is not possible to reach any estimates about the demand on certain transportation lines. Another factor is, the estimated demand values can show variation throughout the time. The demand for a given path can result below or above the mean estimated demand. Since the problem is related with healthcare and human lives, especially for the case where the demand results as an increased value over mean value, not meeting the demand would lead undesirable issues indicating loss of lives or causing damage. Therefore, the unknown demand values and variation in the demand should be appropriately integrated in the model to avoid such circumstances. Regarding these comments, the deterministic model is changed into a stochastic model by including the demand with a distribution that would reflect the variation and meet the deficit of inaccurate demand values.

A further change that has been applied to the model was to remove the static demand from the formulations. In the stochastic model, the model will focus on only the mobile demand that is occurring along the transportation lines and the stochasticity of the mobile demand will considered. Since the static demand is removed, the maximization of the healthcare service provided will be measured from the continuum of care definitions.

In Section 2.2, detailed information about stochastic programming has been explained and different applications has been described. As described in the literature review, stochastic programming can be formulated with various approaches. Looking at the problem content and focus area, a risk averse approach is decided to be applied. From the risk averse measures, referring to Section 2.2.2, a coherent risk measure is determined to be suitable, especially by convexity features and their compatibility with linear optimization. From the set of coherent risk measures, the focus is made on Conditional-Value-at-Risk (CVaR) risk measure and problem is reformulated accordingly.

As explained in Section 2.2.1, a risk measure maps the set of all risks to a real valued function. In Section 2.2.3, Value-at-Risk (VaR) and CVaR are defined with the help of a function $f(x,y)$ which indicates the loss associated with $x$ and $y$ where $x$ is the set of decision vectors and $y$ is the set of random vectors. Therefore, to use the functions VaR and CVaR, a loss function has to be determined.

In the problem context, since the model is focused on planning of the walk-in clinics along the transportation lines, the aim is providing as maximum health care as possible to reach the mobile populations and their local communities in great extend. Every blank point on the map, every location on which a walk-in clinic is not built or every walk-in clinic that is not giving certain services will result in lessening of the healthcare provided and the aim will be hindered. However, if every possible location was built on the map, such lacks of healthcare service would not occur. Therefore the loss function can be defined verbally as *the lack of continuum of care provided from the maximum possible value* and formulations can be made accordingly. Since the model is working on four high impact diseases, optimal solution for one disease may result in sacrificing of an other disease. Therefore, the loss function should be applicable for every disease on the network. Accordingly, the loss function can be written as,

$$l_d(y_{ksd}, dm_{qd}) = \sum_q dm_{qd} \left(cocmax_{qd} - coc_{qd}\right), \quad \forall d \in D. \tag{5.1}$$

In equation (5.1), the loss function is defined as the sum of deviations of continuum care of the optimized network from the maximum continuum of care values, multiplied with the mobile demand values at the path $q$. The maximum continuum of care values are pre-determined values, which results as opening every possible walk-in clinic on the network and serving as maximum given the possibilities. The $y_{ksd}$ values are the decision variables, which lead to the

calculation of continuum of care values, $coc_{qd}$. The uncertainty in the equation results from the demand values, $dm_{qd}$, which is changing according to a distribution.

Since there are four diseases in the problem and each of them requires to be handled separately to avoid risk, it is decided to include them as "constraints" in the model with different CVaR levels, rather than putting it in the objective function. In literature review, calculation of VaR and CVaR values with their approximations was described extensively. Additionally, including CVaR in model as constraints rather than objective function was discussed too in (2.16).This constraint can be re-written as the following form to adapt the model parameters,

$$\phi_{\beta_d}(y_{ksd}) \leq \omega_d \sum_q (cocmax_{qd} * \mathbf{E}[dm_{qd}]) \ , \quad \forall d \in D$$

where the risk function $\phi_{\beta_d}$ is defined as the $\beta$-CVaR risk measure for disease $d$ for the loss function given in (5.1) and $\omega_d$ is the percentage of maximum number of people that can be covered on the network, $\sum_q (cocmax_{qd} * \mathbf{E}[dm_{qd}])$, that is allowed for exposure to the risk for disease $d$. Consequently, looking at the above formulations and approximations methods described by (2.17) and (2.18), the additional notation for the stochastic model is given in Table 5.1.

Table 5.1: Additional notation for the stochastic model

| | |
|---|---|
| Sets: | |
| $I$ | Set of scenarios defined for the demand distribution of $dm_{qd}$, $i = 1,2,\ldots,I$ |
| Decision Variables: | |
| $l_{di}$ | The value of loss for disease $d$ in scenario $i$ |
| $u_{di}$ | The positive difference between loss function value $l_{di}$ and $\alpha_d$ |
| $\alpha_d$ | The $\beta$-VaR value given the confidence level $\beta_d$ |
| Parameters: | |
| $dm2_{qdi}$ | The mobile demand for disease $d$ on path $q$ in scenario $i$ |
| $n$ | Upper bound on number of scenarios |
| $\pi_{di}$ | Probability of scenario $i$ |
| $\beta_d$ | Confidence level for disease $d$ |
| $\omega_d$ | Risk tolerance level for disease $d$ |

Following set of constraints are included in the model for a risk averse formulation:

$$l_{di}(y_{ksd}, dm_{qd}) = \sum_{q} dm2_{qdi}\ (cocmax_{qd} - coc_{qd}) \qquad ,\forall d \in D \text{ and } \forall i \in I, \tag{5.2}$$

$$u_{di} - l_{di} + \alpha_d \geq 0 \qquad ,\forall d \in D \text{ and } \forall i \in I, \tag{5.3}$$

$$\alpha_d + \frac{1}{1-\beta_d} \sum_{i=1}^{n} (u_{di} * \pi_i) \leq \omega_d \sum_{q} (cocmax_{qd} * \mathbf{E}[dm_{qd}]) \qquad ,\forall d \in D, \tag{5.4}$$

$$u_{di} \geq 0 \qquad ,\forall d \in D \text{ and } \forall i \in I, \tag{5.5}$$

$$\alpha_d \geq 0 \qquad ,\forall d \in D. \tag{5.6}$$

With constraint (5.2), the loss function is defined as deviation from the maximum value of mobile demand that can be covered. This value is calculated over every disease for all possible scenarios given a distribution. In constraint (5.3), the $u_{di}$ is defined as the positive difference between $l_{di}$ and $\alpha_d$ which is used for simplification in constraint (5.4). The left hand side of constraint (5.4) is the approximation for the CVaR value, as described in literature review. Given the risk tolerance level for disease $d$, $\omega_d$, the CVaR is bounded by the amount of demand that is allowed for risk exposure. The last constraints (5.5) and (5.6), are the sign constraints.

The objective function has also been modified as,

$$\text{Max} \quad \sum_{d \in D} r_d * TSM_d \tag{5.7}$$

where

$$TSM_d = \sum_{q \in Q} coc_{qd} * \mathbf{E}[dm_{qd}] \quad \forall d \in D. \tag{5.8}$$

The objective function value is now only composed of the score achieved from the coverage of the mobile demand. Since the mobile demand is assumed to be following a distribution for the stochastic model, the score of the mobile demand is calculated from the expected value of the mobile demand distribution. This formulation would aim to maximize the expected value of demand covered, keeping the risk associated on the network for every disease under control.

Putting together all the modifications and extensions, the new stochastic model will be as follows:

$$\text{Max} \quad \sum_{d \in D} r_d * TSM_d \tag{5.9}$$

$$\text{s.t.} \quad TSM_d = \sum_{q \in Q} coc_{qd} * \mathbf{E}[dm_{qd}] \ , \quad \forall d \in D, \tag{5.10}$$

$$coc_{qd} = f(\{covfeas_{qds} | s \in S\}) \ , \quad \forall q \in Q, \forall d \in D, \tag{5.11}$$

$$covfeas_{qds} = g_{sd}(cov_{qds}) \ , \quad \forall q \in Q, \forall d \in D, \forall s \in S, \tag{5.12}$$

$$cov_{qds} = h_{sd}(\{x_{kd}, y_{ksd} | k \in K_q\}) \ , \quad \forall q \in Q, \forall d \in D, \forall s \in S, \tag{5.13}$$

$$l_{di}(y_{ksd}, dm_{qd}) = \sum\nolimits_q (cocmax_{qd} - coc_{qd}) * dm2_{qdi} \ , \quad \forall d \in D \text{ and } \forall i \in I, \tag{5.14}$$

$$u_{di} - l_{di} + \alpha_d \geq 0 \ , \quad \forall d \in D \text{ and } \forall i \in I, \tag{5.15}$$

$$\alpha_d + (1 - \beta_d)^{-1} \sum\nolimits_{i=1}^{n} (u_{di} * \pi_i) \leq \omega_d * expMax_d \ , \quad \forall d \in D, \tag{5.16}$$

$$expMax_d = \sum\nolimits_q (cocmax_{qd} * \mathbf{E}[dm_{qd}]) \ , \quad \forall d \in D, \tag{5.17}$$

$$M \times bin_{kd} - x_{kd} \geq 0 \ , \quad \forall k \in K, \forall d \in D, \tag{5.18}$$

$$1 - (y_{k\{s=di\}d} + y_{k\{s=tr\}d} + y_{k\{s=cr\}d}) \leq M \times (1 - bin_{kd}) \ , \quad \forall k \in K, \forall d \in D, \tag{5.19}$$

$$x_{kd} \geq (1/3) * (y_{k\{s=di\}d} + y_{k\{s=tr\}d} + y_{k\{s=cr\}d}) \ , \quad \forall k \in K, \forall d \in D, \tag{5.20}$$

$$z_k \geq (1/4) * \sum_{d \in D} x_{kd} \ , \quad \forall k \in K, \tag{5.21}$$

$$UBDS_{sd} \geq \sum_{k \in KPDS_{sd}} y_{ksd} \ , \quad \forall d \in D, \forall s \in S, \tag{5.22}$$

$$UB \geq \sum_{k \in KP} z_k, \tag{5.23}$$

$$u_{di} \geq 0 \ , \quad \forall d \in D \text{ and } \forall i \in I, \tag{5.24}$$

$$\alpha_d \geq 0 \ , \quad \forall d \in D, \tag{5.25}$$

$$y_{kds} = 1 \ , \quad \forall k \in KCDS_{ds}, \tag{5.26}$$

$$x_{kd} \in \{0,1\} \ , \quad \forall k \in K, \forall d \in D, \tag{5.27}$$

$$y_{kds} \in \{0,1\} \ , \quad \forall k \in K, \forall d \in D, \forall s \in S, \tag{5.28}$$

$$z_k \in \{0,1\} \ , \quad \forall k \in K. \tag{5.29}$$

# CHAPTER 6

# COMPUTATIONAL STUDY

This section is mainly composed of two parts. In the first section, computational results of the deterministic model will be given. The objective of the deterministic model is to establish a formulation that is capable of planning walk-in clinics with regards to the characteristics of the mobile populations defined by the approaches. Another objective is to integrate the concept of disease and services in the model to the planning procedure. With the computational study for the deterministic model given in Section 6.1, it is aimed to show how the model behaves, reacts to model parameters and give insight about the model.

In the second part, the stochastic model results will be explained. The objective of the stochastic model is to include risk in the model in addition to the deterministic formulation. With the computational study given in Section 6.2, how the solutions have differed compared to a deterministic model and if the desired consequences are achieved or not are analyzed.

The mathematical formulation is coded in GAMS and optimization is done using CPLEX solver using a PC with 4GB RAM, 256GB SSD Disk running Windows 8.

## 6.1 Computational Study for Deterministic Model

In the deterministic model, the aim was to build a model which would enable to design a walk-in clinic network that would benefit the mobile and static populations on the network. Since the mobile populations require specific handling of their requirements for different diseases and services, certain approaches were developed. In this section, it will be analyzed whether the model is able to capture the specific requirements correctly and behaves as initially aimed. This will be done through the testing of various parameters in the models.

### 6.1.1 Computational Setting

The data that has been used for the analysis was a modified set used in [13]. Since the initial phase of the study was done in cooperation with North Star, the data gathered by North Star and their specifications were included in the data set as well. The set has been adapted with the removal of some paths, which are determined to be not effective and modification to walk-in clinic locations that are used as candidates in the model are made. Each of the walk-in clinic received a parameter of being established or not, depending on the current network of North Star. Also, the services that are provided in walk-in clinics are entered in data set regarding the current situation in network. This, in the end, lead to a data set which has 50 walk-in clinic locations and 23 origin and destination location pairs with their coordinates for calculation of the paths. Furthermore, the demand associated with the locations and paths are included in the model. Finally, the given services for diseases in the existing centers are included regarding the health service packages of North Star. In addition to walk-in clinic location and path data, the local healthcare services with their coordinates and services available are included to enable the calculation of referral services.

There are certain parameters which are determined outside the model in order to ensure the flexibility of the model to different disease and service combinations as well as static and mobile demand importance. For the testing of the model, these parameters are changed in order to observe the changes in the optimal solution and their contributions. Therefore, the effects of the following set of of parameters are analyzed:

- $u_1$ and $u_2$, which are the relative importance given to mobile and static demand respectively,
- $r_d$, which is the relative importance of disease $d$ in the optimization of the network,
- $r2_{ds}$, which is the relative importance for service $s$ for disease $d$ in the optimization of the network.

These parameters are adjusted by the user to enable walk-in clinic networks that are specialized in accordance with the user preferences. In the following sections, each of these parameters will be tested to observe how the model reacts to changes in these parameters.

### 6.1.2 Sensitivity Analysis for $u_1$ and $u_2$

As a first part of the study, the changes in the model by the putting different focus on mobile and static demand is analyzed. The study has been applied to a network which has 15 existing walk-in clinic locations with certain services and 35 candidate locations which are not opened.

**Observation 1** *Increasing the number of walk-in clinics opened on the network increases the total score of the network whereas decreases the added value of each additional walk-in clinic.*

Initially, the importance given to both of the demand types is equalized by $u_1 = 0.5$ and $u_2 = 0.5$ and 35 centers are opened in the network gradually by increasing the upper bound on number of walk-in clinics that is allowed to be opened from 1 to 35 in each iteration. Figure 6.1 summarizes the results of the study. As seen from the results, by opening more centers on the network, the objective function value increases which is a result of the both static and mobile demand captured. However, it can also be seen that, marginal benefit of each center, after approximately opening 17 walk-in clinics, is not varying. This can be interpreted as after opening certain number of walk-in clinics on the network, opening further locations may not be as beneficial as planned as the network might be already saturated. Especially for the investment decision, this can be an important criterion to consider.

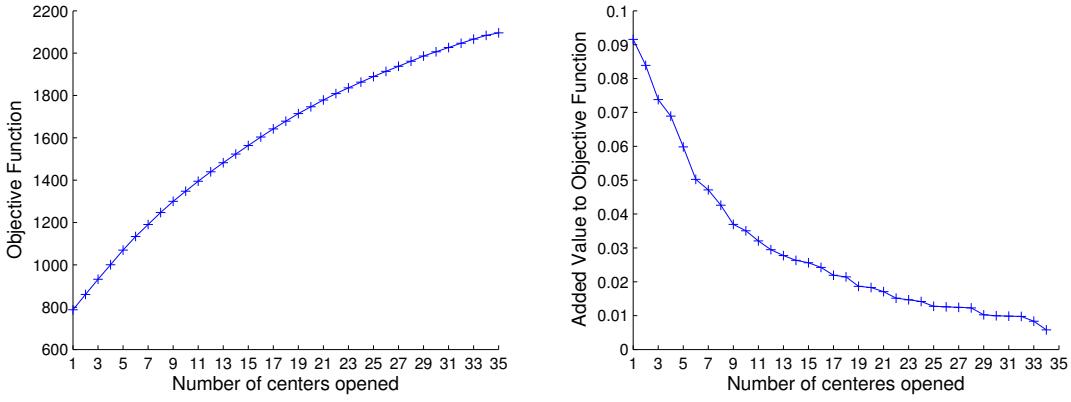


Figure 6.1: Equal importance for mobile and static demand

As a next step the focus is for mobile and static demand is changed and three cases are examined simultaneously which are:

- Case 1: Focus is on mobile demand with $u_1 = 0.9$ and $u_2 = 0.1$,

- Case 2: Focus is on static demand with $u_1 = 0.1$ and $u_2 = 0.9$,
- Case 3: Focus is indifferent $u_1 = 0.5$ and $u_2 = 0.5$.

In each of the cases the summation of Total Score Mobile (TSM) which is calculated through Equation (4.6), for every disease is analyzed. The scores are summarized in Figure 6.2.

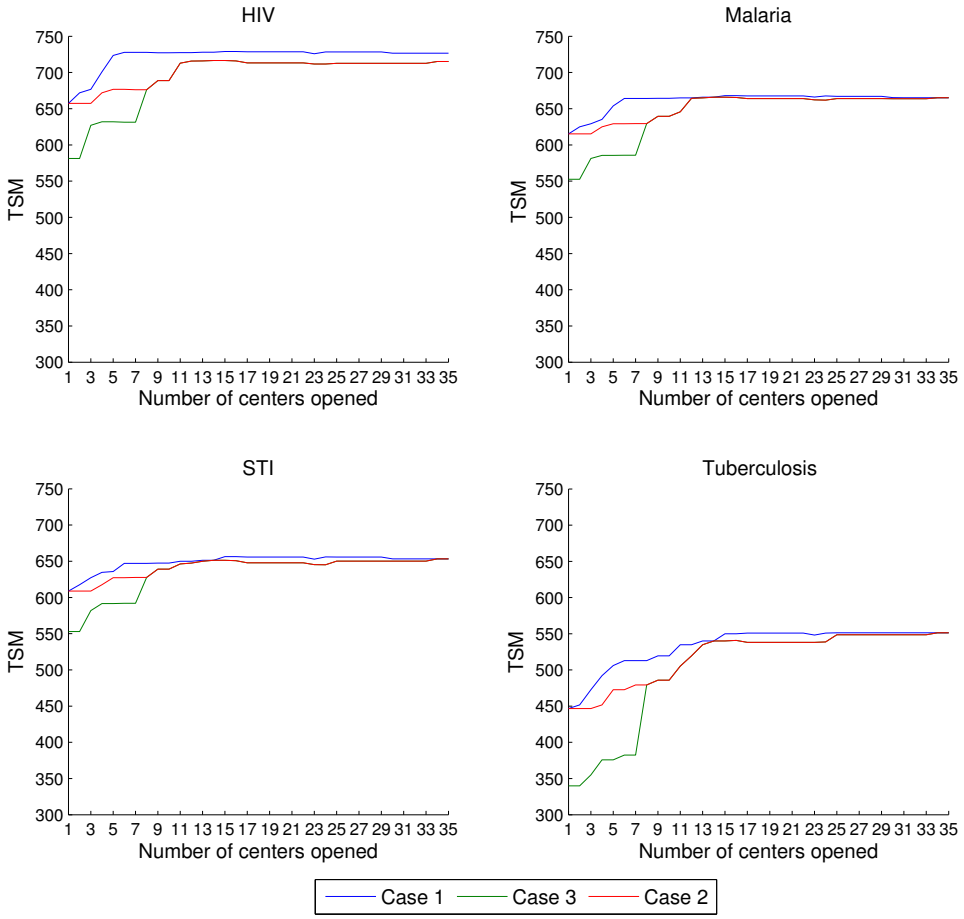


Figure 6.2: Mobile scores of the network for all cases

**Observation 2** *When weight for a given population group is increased, the score of the network for the given population group is increased.*

**Observation 3** *The maximum values of scores that can be attained on the network after opening certain number of walk-in clinics is different for every disease given the characteristics of the disease and the demand values.*

As seen in Figure 6.2, when more focus is put toward the mobile demand, as given in Case 1, the mobile scores of the model is higher compared to Case 2 and 3. For Case 3, which has the lowest focus on mobile demand out of all cases, the mobile score of the network is lowest. Furthermore, when every disease is compared with each other, the maximum value of score that can be obtained after opening 35 walk-in clinics is different in each disease. For instance, when the maximum value that can be attained by malaria and tuberculosis are compared, tuberculosis has a lower value which is a consequence of strict critical intervention times for tuberculosis which makes it hard to achieve high values of continuum of care. Another observation is that, the TSM values become flat after opening certain number of walk-in clinics. This shows that the continuum of care value are reaching their maximum values, either by covering all of the path or by not having walk-in clinic locations as candidate which has potential to increase the continuum of care values.

A similar study has ben made with focus on two diseases, malaria and STI for Cases 1 and 2. For each of the cases, the TSM and Total Score Static (TSS) is analyzed for two diseases which is given in Figure 6.3.

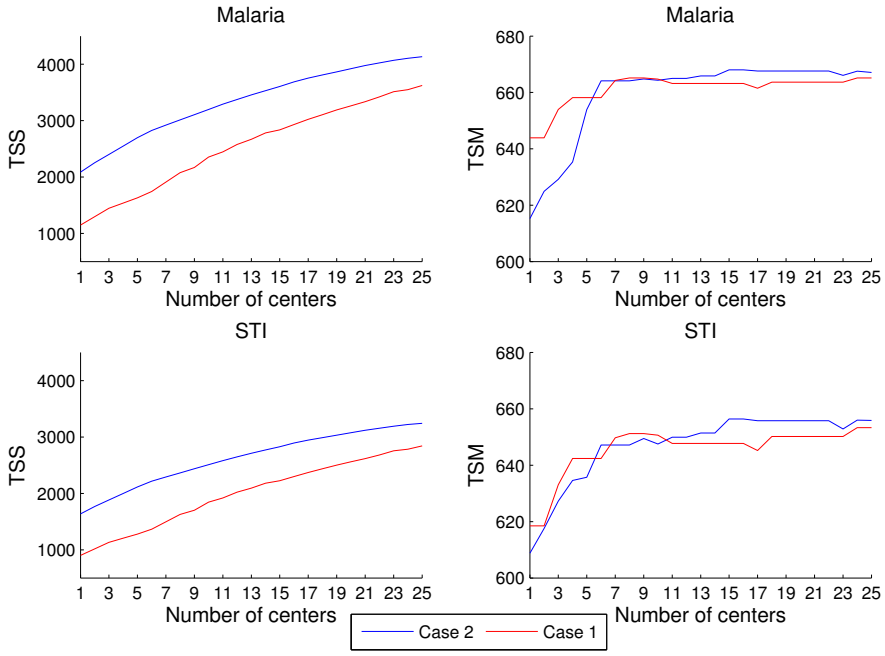


Figure 6.3: Mobile and static scores of the network for case 1 and 2

As can be seen, when the static demand is more important in the problem, as given in Case 2, the TSS score has a greater increase compared to Case 1. This situation is valid for both malaria and STI, expect the difference on values taken. For the TSM, the same observation

Table 6.1: TSM and TSS values with tuberculosis focus

| | TSM | TSS |
|---|---|---|
| HIV | 671.920 | 1225.000 |
| Malaria | 647.330 | 2401.000 |
| STI | 627.788 | 1883.000 |
| Tuberculosis | 509.251 | 2722.000 |

can be made. Another point that can be understood is, TSS increases as a smooth function and shows steady increase. This is a result of summation of every static demand at the points on which the walk-in clinics are built. However TSM is calculated through the continuum of care functions, which results in non-smooth curves.

### 6.1.3 Sensitivity Analysis for $r_d$

One of the ideas that was initially aimed was to have a model that would be able to plan disease specifically taking the disease related parameters into consideration. With such an application, the model would be able to optimize specialized network, that can benefit society with regards to their requirements. The parameter $r_d$ was a figure that would give this opportunity, indicating the relative importance of diseases.

**Observation 4** *When weight given to a certain disease is increased, the score of the disease on the network is increased.*

In order to test the effectiveness of the $r_d$ values, an initial simple test is carried in two different problems where in the first case tuberculosis is given higher importance, keeping the other weight of diseases equal to each other and in the second case STI is given higher importance. 3 walk-in clinics are added to the network and results are analyzed. For the case where tuberculosis is given higher importance, the TSM and TSS scores observed are given in Table 6.1. It can be seen that TSS score is the highest among the diseases showing the walk-in clinics are opened that the places with highest static demand for tuberculosis.

A similar observation is made for the case with STI is given the highest importance. The results of the TSM and TSS values are given in Table 6.2. As seen in the Table, the TSS value for STI is not the highest compared to the tuberculosis case. However, it is seen that the TSM

Table 6.2: TSM and TSS values with STI focus

| | TSM | TSS |
|---|---|---|
| HIV | 687.730 | 1225.000 |
| Malaria | 653.942 | 2401.000 |
| STI | 633.027 | 1883.000 |
| Tuberculosis | 495.195 | 2722.000 |

for the STI case is having a higher value. This is a result of the coverage approaches that each disease belongs to. Tuberculosis has strict intervention times and the continuum of care value for tuberculosis is not very easily increased. Therefore a shift towards increasing the TSS is a reasonable behaviour. STI on the other hand, does not have definitions as strict as tuberculosis and increase in TSM can also be achieved to contribute the objective function.

Since every disease is multiplied with the $r_d$ coefficient in the objective function, the contributions of each of the diseases to the objective function is also compared with each other. The results are given in Figure 6.4. It is seen that when a disease is focused in the optimization of the problem, the disease has the highest contribution to the objective function value.

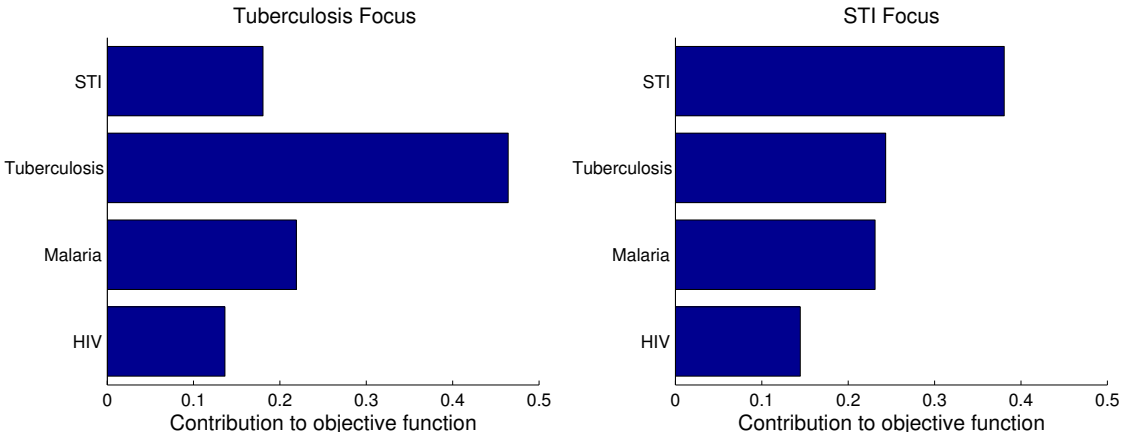


Figure 6.4: Comparison of different disease focus

As a final step of the analysis, the reasons behind these changes in optimized network's values are examined. Since the test is made by opening 3 walk-in clinics, the locations of the added clinics are examined. In both of the studies, the two of the opened centers are at Ngwenya and Ladybrand whereas the third center is different. For the case with tuberculosis focus, the third walk-in clinic is opened at Tunduma 2 with every service included for tuberculosis. Tunduma 2 is a walk-in clinic location which has access to tuberculosis referral services in the network, contributing to the continuum of care values on the network. For the case with STI focus,

Table 6.3: TSM and TSS values with focus on mobile demand and malaria

| | Beitbridge | Chirundu North | Tunduma 2 | Dar es Salaam |
|---|---|---|---|---|
| TSM | 894.110 | 904.261 | 910.247 | 918.978 |
| TSS | 1340.000 | 1540.000 | 1770.000 | 1863.000 |

the third walk-in clinic is opened at Mbeya with every service for STI included. Mbeya is a critical location which is located on 7 paths. As explained previously, STI does not have very strict intervention times so that opening a walk-in clinic at Mbeya for STI can contribute to the continuum of care values of the 7 paths on which it is located and benefit the total network at once.

Another interesting study for the disease focus is made on malaria in order to observe the effects of changing of importance given to static and mobile demand at the same time. For this case, the data set has been modified to create an artificial malaria intense path and locations. The malaria intense path is determined to be path from Chirundu South to Dar es Salaam, which in between passes through locations, *Chirundu North, Nakonde, Tunduma 2, Tunduma 1, Izumbwe - Mphoi* and *Mbeya*. In addition to the path, 2 walk-in clinic locations that are outside the path, are assigned with high malaria demand which are *Ngwenya* and *Ladybrand*. The test is made for the case of opening 4 new walk-in clinics on the network and focus has been put on malaria by assigning highest weight.

**Observation 5** *Giving different weights for mobile and static populations adapts the planning of the walk-in clinics such that the added value of every additional walk-in clinic on the network is maximized.*

In the first part of the study, the focus on mobile demand is increased with the help of $u_1$ and 4 centers are added to the network. As a results, walk-in clinics at locations *Beitbridge, Chirundu North, Tunduma 2* and *Dar es Salaam* are opened respectively. For the addition of each location, the TSM and TSS scores are summarized in Table 6.3. As can be observed from the table, the walk-in clinics that are opened else than Beitbridge, are the locations that are located on the malaria intense path. Additionally, it is important to observe that the 3 walk-in clinics that are opened along the path are located at the beginning, middle and end of the path which gives the highest contribution to the continuum of care value for the path.

Table 6.4: TSM and TSS values with focus on static demand and malaria

| | Tunduma 2 | Chirundu North | Nakonde | Ladybrand |
|---|---|---|---|---|
| TSM | 837.335 | 840.963 | 841.980 | 841.980 |
| TSS | 1459.000 | 1659.000 | 1839.000 | 2006.000 |

When the same analysis is made for the case where the focus is put on static demand, the optimal choice of walk-in clinic locations differ. The new set of decision are given as *Tunduma 2, Chirundu North, Nakonde* and *Ladybrand*. For the addition of each location, the TSM and TSS scores are summarized in Table 6.4. The optimal choices which are Tunduma 2, Chirundu North and Nakonde are located on the same path, where Chirundu North and Nakonde are located close to each other. This shows opening two centers at such close distance would not add value to TSM significantly, which can be seen in the Table 6.4 as well. When a walk-in clinic at Nakonde was opened, the TSM value increased from 840.963 to 841.980. However, the TSS score rises very rapidly, which justifies the choice of opening in Nakonde. In addition to Nakonde, the fourth center is opened at Ladybrand, which was a place located outside the path but having high value of demand for malaria. Moreover, when Table 6.3 and 6.4 are compared, in Table 6.3, the TSM scores are changing in greater values compared to Table 6.4 and TSS values in Table 6.3 are not increasing greatly as seen in Table 6.4. These results together prove the model is logical and show that the model both reacts to the focus on malaria and static demand by adapting solutions accordingly.

### 6.1.4 Sensitivity Analysis for $r2_{ds}$

The final part of the analysis involves testing of the parameter $r2_{ds}$, which is the focus on service of a given disease. Since continuum of care values are weighted values of coverage for every service of a disease along the path given the $r2_{ds}$, when a certain service of disease is assigned with more weight, it would effect the continuum of care values and accordingly, the objective function. Therefore, different focuses on service should lead to different networks which would enable user to plan also service specifically in addition to diseases.

**Observation 6** *When the weight given to a certain service of a certain disease is increased, the solution for planning to the walk-in clinic locations is adapted.*

To test the parameters, the disease is chosen as HIV and the services "treatment" and "care" are considered. The model optimized the addition of 3 walk-in clinic to the existing network and in each case focus is put for HIV treatment and HIV care respectively. As a result of the optimization, the optimal location of the walk-in clinics were found out to be different. When the HIV treatment received a higher importance, the optical location were determined as *Beitbridge, Beira* and *Namanga*. For the case with HIV care, the optimal location were found as *Walvis Bay, Mchinji* and *Mbeya*, showing the set of optimal solutions are completely different. Since the change of the parameter $r2_{ds}$ is related with continuum of care values, the TSM scores and value added for both cases are given in Table 6.5 and 6.6.

Table 6.5: Focus of HIV treatment

| | Beitbridge | Beira | Namange |
|---|---|---|---|
| TSM | 323.038 | 386.562 | 442.897 |
| Value Added | 1 | 1.197 | 1.371 |

Table 6.6: Focus on HIV care

| | Walvis Bay | Mchinji | Mbeya |
|---|---|---|---|
| TSM | 545.601 | 561.6011 | 567.967 |
| Value Added | 1 | 1.029 | 1.041 |

Since the calculation of coverage values for these two cases is more complex, coming from expected traveling time approach, it is not very easy to justify the optimal location choices for the given solutions. The only thing that could be mentioned is that model is sensitive to the parameter and has the ability to adapt solution which would maximize the objective function given the requirement. One other observation is that, the coverage definition for HIV care is more relaxed, having higher threshold values, compared to HIV treatment. It is easier to achieve higher values of TSM for the case of HIV care, as can be seen in Table 6.6. Since higher values are attained easily and the network is covered with less number of walk-in clinics compared to HIV treatment, the marginal benefits obtained from opening additional walk-in clinics on the network is not very large. This is also seen in Tables 6.5 and 6.6 where HIV treatments have higher amounts for value added.

## 6.2 Computational Study for Stochastic Model

In the stochastic model, the aim in a nutshell was to convert the model into a form such that uncertain and variable demand values could be handled in the model which would minimize the effects of any undesirable results related to loss of human lives. In this Section of the thesis, it is tested whether the risk-averse approach included in the model has been proved to be effective or not.

### 6.2.1 Computational Setting

The input data that has been presented in Section 6.1 was composed of two parts as the static and mobile demand. In this section, the static demand is ignored and only the mobile demand is considered. The mobile demand data is also changed slightly, by removing 5 paths from the set as they were not benefiting the results for the study of stochastic problem.

The demand values in the previous data set were represented as "daily average" values which were deterministic and was not varying. However, in this section the demand values are required to be adapted to fit into a distribution. To represent the data as stochastic, mobile demand for every disease $d$, on every path $q$, is represented with a *uniform distribution*. The average value for daily demand random variable is set to the "daily average" values described in deterministic model. In addition to the mean values as a parameter of the uniform distribution, upper and lower limits of the distribution are the other two parameters that need to be determined for calculations.

In order to define an upper and lower limits for the mobile demand values which would help mapping of the traffic flow along the path, "coefficient of variation" ($C_v$) measure is used. The $C_v$ of a distribution is calculated from,

$$C_v = \frac{\sigma}{E(X)}. \tag{6.1}$$

The $C_v$ of a distribution can be interpreted as the dispersion of the variable around the mean value, independent of the variable's measurement unit. In simpler terms, $C_v$ can also be explained as the volatility of the demand in comparison to the amount of demand that is expected to occur along the paths. For the problem context studies, the $C_v$ value represents the "risk" associated with the mobile demand. If a path is having a high $C_v$ ratio, the path can be told be

not very predictable and deviations from the mean value is expectable. On the contrary, a path with rather a lower value of $C_v$ is having a more steady traffic counts and not many surprises are expected with the demand.

Table 6.7: The summary for input path data

| | | | | Daily Mean Mobile Demand | | | | |
|---|---|---|---|---|---|---|---|---|
| | Origin | Destination | Group | HIV | STI | Malaria | TB | $C_v$ |
| 1 | Beitbridge | Lusaka | B | 60 | 18 | 38 | 1 | Low |
| 2 | Mpika | Beitbridge | B | 80 | 24 | 50 | 1 | High |
| 3 | Walvis Bay | Eldoret | C | 80 | 24 | 50 | 1 | Low |
| 4 | Mulawayo | Kapiri Mposhi | B | 100 | 30 | 63 | 1 | High |
| 5 | Chirundu | Dar es Salam | A | 120 | 36 | 75 | 2 | Low |
| 6 | Kapiri Mposhi | Johannesburg | B | 120 | 36 | 75 | 2 | Low |
| 7 | Morogoro | Livingstone | A | 120 | 36 | 75 | 2 | High |
| 8 | Beira | Kisangani | C | 140 | 42 | 88 | 2 | Low |
| 9 | Lusaka | Durban | B | 180 | 54 | 113 | 2 | High |
| 10 | Durban | Chipata Border | B | 200 | 60 | 125 | 2 | High |
| 11 | Johannesburg | Mbeya | B | 240 | 72 | 150 | 3 | Low |
| 12 | Dar es Salam | Eldoret | A | 260 | 78 | 163 | 3 | High |
| 13 | Beira | Eldoret | A | 340 | 102 | 213 | 4 | Low |
| 14 | Walvis Bay | Johannesburg | C | 300 | 90 | 188 | 3 | High |
| 15 | Durban | Dar es Salam | B | 400 | 120 | 250 | 4 | Low |
| 16 | Dar es Salam | Lubumbashi | A | 480 | 144 | 300 | 5 | High |
| 17 | Durban | Lubumbashi | B | 600 | 180 | 375 | 6 | High |
| 18 | Beira | Dar es Salam | B | 720 | 216 | 450 | 8 | Low |

The input path data, which is described previously has been categorized into three groups regarding the regions of the paths and walk-in clinics that are situated on them. Each of the path is assigned with a category, which is summarized in Table 6.7 with the labels A, B and C. For every category, high and low demand paths are determined. After determining these paths, every high and low demand paths are assigned with high and low $C_v$ values. Therefore, in the end the set of data is converted into a homogeneous form in which every high-low demand values in every category is assigned with both high and low $C_v$ values. By assigning a $C_v$ value to each path, the upper and lower bounds of their demand distributions are calculated with the help of the above definitions for uniform distribution. Therefore, the mobile demand for every disease on every path is given a range on which the demand could vary, which is as expected greater for high $C_v$ valued paths and small for low $C_v$ valued paths.

After determining the paths and their ranges for demand values, the scenarios are generated

with the random sampling method. For each demand for disease $d$ along path $q$, 100 random instances are chosen for their given ranges and the demand data is created in matrices of sizes 18 to 100 for every four disease.

In order to observe how the risk-averse model behaves, the study conducted has been branched in three directions on which a different disease is focused. In every section, a different disease is selected on which the CVaR constraint is added. These diseases are selected as HIV, tuberculosis and malaria and their results are examined separately in the following sections. Each of these diseases belong to a different coverage approach and therefore, different coverage descriptions effects would not be disregarded.

### 6.2.2 CVaR Constraint on HIV

Before going further into the analysis of the result, it is important to briefly go over the terms and concepts related with the study. To gain more understanding of the stochastic model and interpretation of constants, the risk constraints in the stochastic model can be explained, numbered as (5.14), (5.15) and (5.16). It can be observed that there are two parameters determined by the user which are the $\beta_d$ and $\omega_d$ values. The $\beta_d$ value is the confidence level of the problem and the $\omega_d$ value is the risk tolerance varying from 0 to 1. The $\omega_d$ is the risk tolerance level, such as the percentage of the maximum number of people that can be covered in the network for disease $d$, that is allowed for risk exposure. For instance, a model having $\omega_{\text{HIV}} = 0.10$ and $\beta_{\text{HIV}} = 0.95$ tells that, average loss in 5% of the worst cases must not exceed the 10% of the maximum number of people with HIV that can be covered. Therefore, a $\omega_{\text{HIV}}$ equaling 1 would imply, all of the population is allowed for risk and the addition of the risk constraints to the model would not make a difference. On the contrary, a value of 0 for $\omega_{\text{HIV}}$ would mean no risk at all is allowed in the problem. As the risk tolerance level is increases from 0 to 1, increases in the objective function are expected as the constraints become more relaxed and less binding.

In this section, the effects of putting CVaR constraint for the continuum of care of the HIV value over the network is examined. For observing the relation between $\omega_d$ value and objective function value, the model is run with different parameters by changing the $\beta_d$ and $\omega_d$ values gradually. Furthermore, the results are analyzed for the cases of opening 4, 5 and 6 centers to the network. The reason that these values are chosen is, values below 4 for opening

centers results in very low continuum of care values, on which putting a constraint on risk is not effective. For the value above 6 for opening centers, the continuum of care values reaches very high values and the model is already having very low values of risk. Therefore, the study is focused on opening only 4, 5 and 6 centers. In Tables 6.8, 6.9 and 6.10, the objective function values for different parameters can be observed.

**Observation 7** *Decreasing the risk tolerance level for a given disease results as a decrease in the objective function values.*

Looking at Tables 6.8, 6.9 and 6.10, it can be observed that having higher risk tolerance levels for $\omega_{\text{hiv}}$ results as an increase in the objective function values. This is an expected result as putting more strict constraints in the model lowers the objective function to meet the given criteria. From Table 6.8, it is seen that for every confidence level, after certain values of the $\omega_{hiv}$, the constraint becomes non-binding and the objective function achieves the same value as the model with no risk constraints. Furthermore, for lower $\omega_{hiv}$ values, the problem becomes infeasible and returns no solution. This means that non of the possible location for walk-in clinics allows to tolerate the given risk level. For the Tables 6.9 and 6.10, it is seen that either the constraints are non-binding in the model or the model is infeasible. This can be explained as, by opening of 5 and 6 centers, high continuum of care values are obtained. Therefore, the given locations already satisfy the risk tolerances in most of the cases, however, after lowering the $\omega_{\text{hiv}}$ to a certain value, the model becomes infeasible as there is no set of solutions to choose from is available.

Table 6.8: The objective function values for opening 4 centers with CVaR measure on HIV

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1685.321 | 1685.321 | 1685.321 |
| 0.50 | 1685.321 | 1685.321 | 1685.321 |
| 0.40 | 1685.321 | 1685.321 | 1685.321 |
| 0.30 | 1685.321 | 1685.321 | 1685.321 |
| 0.25 | 1679.718 | 1685.321 | 1685.321 |
| 0.24 | No solution | 1679.718 | 1685.321 |
| 0.23 | | 1669.376 | 1679.718 |
| 0.22 | | No solution | 1679.718 |
| 0.21 | | | No solution |

Table 6.9: The objective function values for opening 5 centers with CVaR measure on HIV

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1848.420 | 1848.420 | 1848.420 |
| 0.50 | 1848.420 | 1848.420 | 1848.420 |
| 0.40 | 1848.420 | 1848.420 | 1848.420 |
| 0.30 | 1848.420 | 1848.420 | 1848.420 |
| 0.20 | 1848.420 | 1848.420 | 1848.420 |
| 0.15 | 1848.420 | 1848.420 | 1848.420 |
| 0.14 | No solution | 1848.420 | 1848.420 |
| 0.13 | | No solution | No solution |

Table 6.10: The objective function values for opening 6 centers with CVaR measure on HIV

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1992.566 | 1992.566 | 1992.566 |
| 0.5 | 1992.566 | 1992.566 | 1992.566 |
| 0.4 | 1992.566 | 1992.566 | 1992.566 |
| 0.3 | 1992.566 | 1992.566 | 1992.566 |
| 0.2 | 1992.566 | 1992.566 | 1992.566 |
| 0.10 | 1992.566 | 1992.566 | 1992.566 |
| 0.09 | 1992.566 | 1992.566 | 1992.566 |
| 0.08 | 1992.566 | 1992.566 | 1992.566 |
| 0.07 | 1992.566 | 1992.566 | 1992.566 |
| 0.06 | 1992.566 | 1992.566 | 1992.566 |
| 0.05 | No solution | No solution | No solution |

**Observation 8** *Increasing the number of walk-in clinics on the network decreases the risk tolerance level which becomes binding in the model.*

One other observation that can be made from these runs is that, for the case of opening 4 centers, as the confidence level is gradually lowered from 0.95 to 0.8, the changes in the objective function values are behaving differently. For instance, with a confidence level of 0.95, 0.25 risk tolerance changes the objective function value from 1685.321 to 1679.718 whereas for confidence level of 0.9 and 0.8, this risk level is not binding at the same risk tolerance of 0.25. Same observation can be made for the values of $\omega_{\text{hiv}}$ which return “No Solution". The tolerance level which makes the model infeasible decreases with lower confidence levels. The changes in the objective function values for different risk tolerance levels is mapped in Figure 6.5. Since this a discrete optimization and walk-in clinic locations are added to solu-

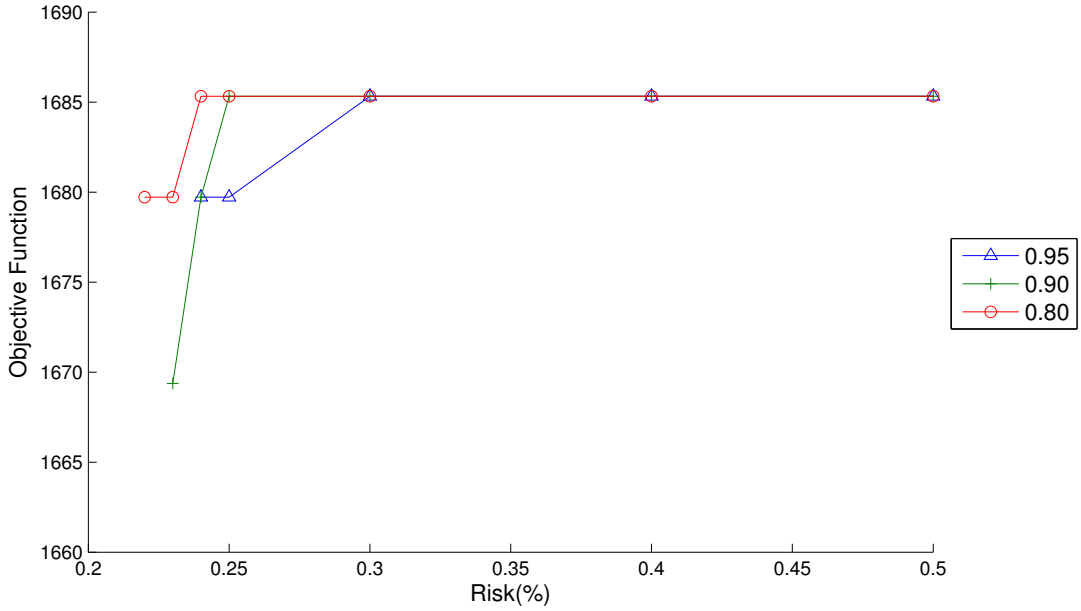


Figure 6.5: Objective function of the model with CVaR measure on HIV (4 centers)

tion set gradually, it is not possible to achieve a continuos efficient frontier for the objective function value for different risk tolerance levels. It is seen that, there is not a continuous curve but rather discrete set of points.

**Observation 9** *Decreasing the risk tolerance level given for a certain disease shifts the walk-in clinic towards the paths having high value of coefficient of variation and high mobile demand.*

A further analysis is applied for the continuum of care values on each flow to see how it has changed. The analysis has been focused on the 4 center case, with the confidence level of 0.90. The continuum of care values for the $\omega_{hiv}$ values of 0.25, 0.24 and 0.23 are examined in which all of them have different objective functions and different solution set for optimal location of walk-in clinics. In Table 6.11, the results of the continuum of care values are summarized.

The results for continuum of care values for HIV over all the paths is displayed in the bar chart given in Figure 6.6. When the figure is observed, it is seen that with the inclusion of the lower risk tolerances in the model, the continuum of care along the paths have changed. When the risk is lowered from 0.25 to 0.24, one of the very significant changes is observed over path 14.

Table 6.11: The continuum of care values for all paths with CVaR measure on HIV (4 centers)

| | $\omega_{\text{HIV}}$=0.25 | | | | $\omega_{\text{HIV}}$=0.24 | | | | $\omega_{\text{HIV}=0.23}$=0.24 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | HIV | Malaria | STI | TB | HIV | Malaria | STI | TB | HIV | Malaria | STI | TB |
| 1 | 1 | 1 | 1 | 1 | 1 | 0.974 | 0.995 | 0.879 | 1 | 0.97 | 0.994 | 0.865 |
| 2 | 1 | 0.967 | 0.993 | 0.857 | 1 | 0.967 | 0.993 | 0.857 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0.768 | 0.565 | 0.652 | 0.26 | 0.768 | 0.565 | 0.652 | 0.26 |
| 4 | 1 | 0.999 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0.999 | 1 | 1 |
| 5 | 1 | 0.958 | 0.988 | 0.847 | 0.986 | 0.842 | 0.936 | 0.735 | 1 | 0.996 | 0.999 | 0.999 |
| 6 | 1 | 0.972 | 0.994 | 0.874 | 0 | 0 | 0 | 0 | 1 | 0.972 | 0.994 | 0.874 |
| 7 | 1 | 0.922 | 0.976 | 0.755 | 0 | 0 | 0 | 0 | 1 | 0.922 | 0.976 | 0.755 |
| 8 | 0.953 | 0.747 | 0.887 | 0.615 | 0.953 | 0.747 | 0.887 | 0.615 | 1 | 0.937 | 0.98 | 0.809 |
| 9 | 0.907 | 0.676 | 0.761 | 0.617 | 0.904 | 0.671 | 0.76 | 0.594 | 0.723 | 0.54 | 0.624 | 0.49 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.579 | 0.486 | 0.554 | 0.454 |
| 11 | 0.965 | 0.787 | 0.899 | 0.687 | 0.965 | 0.787 | 0.899 | 0.687 | 0.992 | 0.882 | 0.958 | 0.723 |
| 12 | 0.816 | 0.606 | 0.683 | 0.558 | 0.816 | 0.606 | 0.683 | 0.558 | 0.816 | 0.606 | 0.683 | 0.558 |
| 13 | 0.811 | 0.581 | 0.682 | 0.512 | 0.811 | 0.581 | 0.682 | 0.512 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0.989 | 0.867 | 0.951 | 0.464 | 0.989 | 0.867 | 0.951 | 0.464 |
| 15 | 0.979 | 0.832 | 0.915 | 0.763 | 0.979 | 0.832 | 0.915 | 0.763 | 0.902 | 0.658 | 0.749 | 0.594 |
| 16 | 0.921 | 0.691 | 0.791 | 0.618 | 0.921 | 0.691 | 0.791 | 0.618 | 0.921 | 0.691 | 0.791 | 0.618 |
| 17 | 0.913 | 0.692 | 0.772 | 0.636 | 0.907 | 0.678 | 0.768 | 0.589 | 0.747 | 0.561 | 0.642 | 0.514 |
| 18 | 1 | 0.992 | 0.998 | 0.986 | 1 | 0.992 | 0.998 | 0.986 | 1 | 0.934 | 0.976 | 0.825 |

For the case with 0.25 risk tolerance, the continuum of care along the path has a value of 0. However, when CVaR constraint becomes binding in the model, path 14 is also covered and value of 0.989 is achieved. When path 14 is further analyzed from Table 6.7, it is seen that it is one of the paths with highest demand (after 15, 16, 17 and 18 which are already having continuum of care values to a high extent) and also high risk. The model prefers to locate on the paths with high demand in order to maximize the objective function and high risk to keep the risk in given bounds. Opening walk-in clinics over path 14 also affected the continuum of care value on path 3, as they share same walk-in clinic locations over their paths.

When the risk tolerance level is even lowered further to value 0.23, the continuum of care values showed other changes. One of the main changes that take attention at the first sight is the continuum of care value for path 13, which was 0.811 for the risk tolerance levels 0.25 and 0.24 is decreased to zero, meaning all the walk-in clinics located along path 13 is removed from the optimal solution. This situation enabled the model to cover certain other paths as 3, 7, 8, 10 and 14. Looking at Table 6.7, it is seen that path 13 has a low value of risk whereas paths 7, 10 and 14 have high risk with relatively high demand values. Especially for path 10, it is a path with one of the highest demand and highest risk, which is not covered in other two cases with a higher risk tolerance. Therefore, to meet the risk tolerance level of 0.23, the model sacrifices from the paths with high demand and low risk to shift the walk-in clinics to high-risk paths. This was a desired outcome from the model, as in the end the aim was to incorporate the risk factor in the problem and force the model to work accordingly. Furthermore, it can also be seen that with the risk level of 0.23, the number of paths covered is increased, showing the model is able to reach more parts of the network.

Another analysis that is done is about the expected demand that is covered in the different runs. When the same cases that are described above is analyzed again, the covered demand in the network for all diseases is summarized in Table 6.12.

Table 6.12: The expected value of demand covered with CVaR measure on HIV (4 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{hiv}} = 0.25$ | 3717.692 | 1967.280 | 1022.410 | 33.901 |
| $\omega_{\text{hiv}} = 0.24$ | 3730.024 | 1938.197 | 1020.029 | 30.623 |
| $\omega_{\text{hiv}} = 0.23$ | 3685.791 | 1946.087 | 1016.631 | 31.276 |

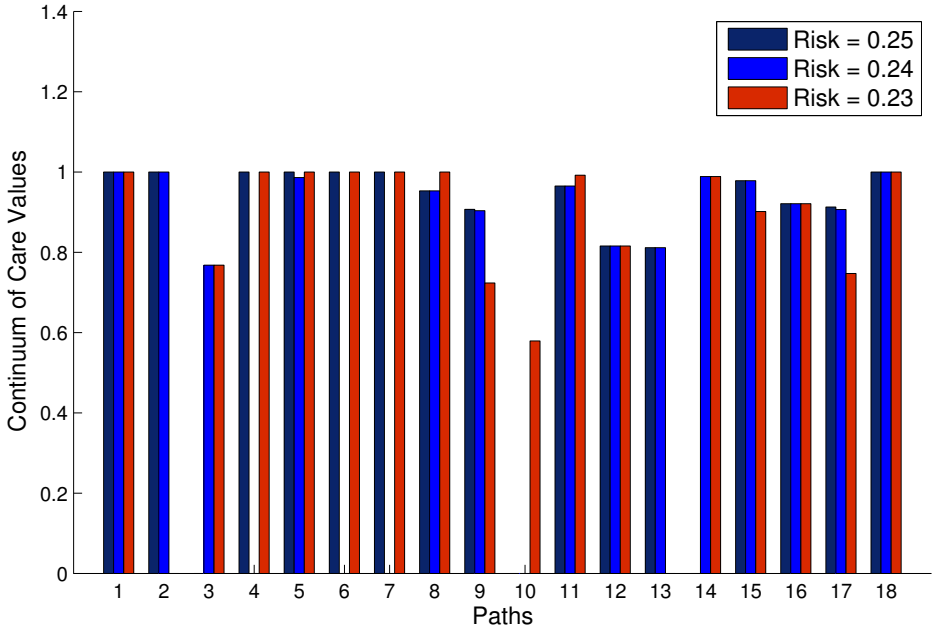


Figure 6.6: The continuum of care values for HIV for all paths with CVaR measure on HIV (4 centers)

**Observation 10** *Decreasing the risk tolerance level for a certain disease may result both as an increase or a decrease in the expected number of mobile demand covered for that disease.*

In the Table 6.12, when the expected demand covered for HIV is examined, decrease of risk tolerance from 0.25 to 0.24 led to an increase in the number of expected people covered with HIV. While the coverage of HIV is increased, the other diseases showed a decrease and which in the end led to a decrease in the objective function value. This shows that, addition of a risk constraint for HIV in the model gave priority for the optimization of HIV in the network and increased the expected number of people covered with HIV, while sacrificing from other diseases. When the tolerance level is further decreased to value of 0.23, the expected number of people covered with HIV is also decreased. As the model is not working with more strict risk constraints, the model is focused on the flows with high risk more compared to the problem with risk tolerance values above 0.23. This may lead to opening walk-in clinics on paths having low expected demand values however high risk values. For all the other diseases, it is also seen that compared to case where the risk constraints are not binding, the expected coverage of people is lowered.

**Observation 11** *Decreasing the risk tolerance level for a certain disease may result as a*

*decrease in the $\alpha_d$ values of the given disease compared to the case in which CVaR constraint is not binding.*

It is important to bear in mind that, decrease in number of expected people covered in the problem with the addition of risk constraints in the model should not be interpreted as a bad sign in the problem context. Even though the number of people covered in the network is decreased, the model has new objectives as well, which requires to satisfying certain levels of risk in the network. Therefore, it would also be good to analyze the number of people at risk in the model. In Table 6.13, the $\alpha_d$ values which are calculated in the model through the different cases are summarized. The model, as mentioned before, was run with a confidence level of 0.90 and the only risk constraint was imposed on HIV, with $\omega_{\text{hiv}}$ values displayed in the Table. $\alpha_d$ value is the threshold which shows, with confidence level of 0.90, the loss function will not exceed this value. It is seen that, for overall results, the $\alpha_d$ for HIV has decreased compared to the case where $\omega_{\text{hiv}}$ is 0.25, in which risk constraint is not binding. Malaria and tuberculosis showed and increase in the $\alpha_d$ values, as they are not a priority of the model. For STI, $\alpha_d$ is also decreased along HIV. This can be interpreted as walk-in clinics chosen for optimal allocation of decreasing HIV risk also benefited the coverage values associated with STI. This is a logical result as both HIV and STI have rather more relaxed critical intervention times compared to malaria and tuberculosis which requires more strict handling. Since their approaches are similar to each other, STI also has decreased $\alpha_d$ values.

**Observation 12** *Since the loss function defined in the formulation is not continuous, the model does not have unique $\alpha_d$ values.*

A worth-noting observation about the $\alpha_d$ values is that, these values are not unique in the optimization of the problem. Since the problem is a discrete optimization on which the decision variables are binary variables, the loss function that is defined in the problem is not a continuos curve. It has some "flat spots" and $\alpha_d$ which is defined over this discrete loss function may fall on these flat spots. This would make the $\alpha_d$ values vary in small ranges, making the problem not having unique VaR values. Therefore in different iterations for the same problem with same parameters, the $\alpha_d$ values may show changes in small amounts as well as $u_{di}$ values described in (5.15). This also discussed in [17] for the case where CVaR is

Table 6.13: $\alpha_d$ values for all diseases with CVaR measure on HIV (4 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{hiv}} = 0.25$ | 1046.882 | 916.078 | 475.526 | 21.452 |
| $\omega_{\text{hiv}} = 0.24$ | 912.014 | 914.848 | 460.002 | 24.52 |
| $\omega_{\text{hiv}} = 0.23$ | 945.716 | 917.536 | 444.007 | 23.44 |

Table 6.14: CVaR values for all diseases with CVaR measure on HIV (4 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{hiv}} = 0.25$ | 1135.000 | 2817.373 | 1358.765 | 50.191 |
| $\omega_{\text{hiv}} = 0.24$ | 1089.600 | 2817.373 | 1358.765 | 50.191 |
| $\omega_{\text{hiv}} = 0.23$ | 1044.200 | 2817.373 | 1358.765 | 50.191 |

minimized in the objective function and optimized VaR pairs are not necessarily unique. In the case where the range of $\alpha_d$ is reduced to a single points, the optimal and unique $\alpha_d$ value would give the $\beta$-VaR.

**Observation 13** *Decreasing the risk tolerance level for a given disease results as a decrease in the CVaR for that given disease.*

Another figure to analyze is the CVaR for every disease in the model, calculated for the same cases. The CVaR values are calculated from the model with the following equation,

$$\alpha_d + (1-\beta_d)^{-1} \sum_{i=1}^{n} (u_{di} * \pi_i)$$

where $u_{di}$ and $\alpha_d$ values are optimized in the problem. In Table 6.14, the results of the calculations are summarized. These CVaR values tell that, average loss in 10% of the worst cases should not exceed the values indicated in the Table 6.14. While $\alpha_d$ values control the percentile of risk, the CVaR aims to control the worst case among these controlled percentile of the risk. For HIV, the CVaR value is decreased as the $\omega_{\text{hiv}}$ is decreased from 0.25 to 0.23. This is a desired outcome of the model as CVaR value is maintained to be controlled and lowered gradually in iterations and there should not be an increase in these values. For the other diseases, it is seen that the values do not show any changes. This is because they are not subject to binding CVaR constraints and their values are not controlled in the model.

When the results for adding CVaR constraint for HIV is examined, it can be said that model integrates the risk factor in the problem context and modifies the solutions accordingly. This

was a desired outcome from the model and it is proven to be working successfully as aimed. One important point about the study which focuses on HIV is that, HIV coverage is handled with the expected traveling time approach. This means that there are no critical intervention times but rather greater thresholds are involved for calculation of coverage. Therefore, it is easier to cover a given path since putting few numbers of walk-in clinics along the path is enough. Since the paths achieve high continuum of care values quickly, it is seen that CVaR constraint becomes effective for low values of risk tolerance.

Furthermore, the data set that is used for modeling is not a very a large one and there is a limited set of walk-in clinics for optimization. When these two factors are combined, it is seen that addition of CVaR constraints in the model does not provide solution for a large number of risk tolerance changes. After lowering the risk tolerance below certain values, it returns no solution and becomes infeasible. If there was a greater set of candidate locations, this problem is expected to be avoided.

### 6.2.3 CVaR Constraints of Tuberculosis and Malaria

After the analysis of CVaR constraints with HIV focus in the problem, CVaR constraints with tuberculosis and malaria focus is also studied respectively. The same analyses that were given for HIV is also conducted for the two diseases as well. In this section, the results will be briefly presented and differences among each other will be investigated.

**Observation 14** *The risk tolerance level which becomes binding in the model shows variation in accordance with the coverage approach of the disease on which the risk tolerance is applied.*

The study is again made for the cases of opening 4, 5 and 6 walk-in clinics in the network with confidence levels of 0.85, 0.90 and 0.95. When different risk tolerance levels are tested for the CVaR constraint on tuberculosis, the results for the objective function given in Figure 6.7 is observed. The objective function values of different runs can be further analyzed in Appendix A.1. When the 5 and 6 centers cases are examined, it is seen that as the confidence level is lowered, the CVaR constraints become non-binding for lower values of risk tolerance. This shows that, with lower levels on confidence, even low risk tolerances can be handled without taking extra measurements in the planning of the network. However, if higher levels

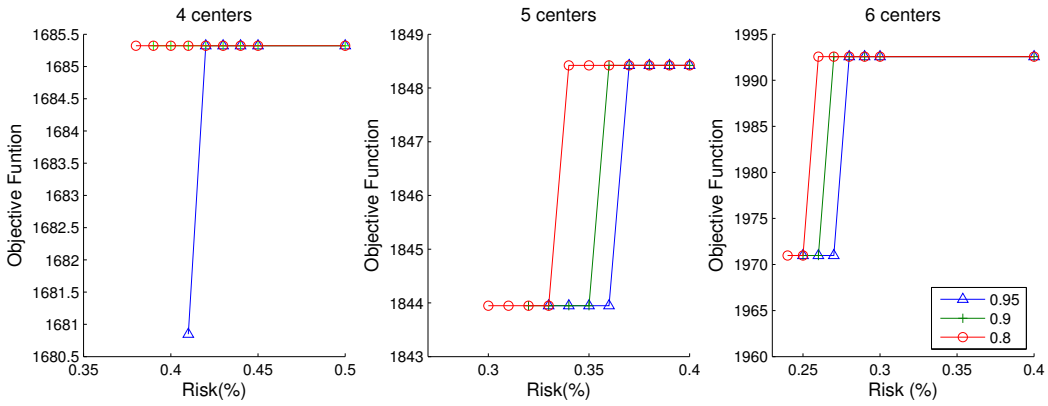


Figure 6.7: Objective function of the model with CVaR measure on Tuberculosis

of confidence is required, the risk tolerance level which changes the constraint from binding case to non binding is increased. Also, looking at the figure for 5 and 6 center cases, every confidence level has a similar shape which is changing in the same objective function intervals but having a gap between each other showing the differences in points where risk tolerance is effecting. For the 4 center case, with confidence levels of 0.95, CVaR constraint becomes binding when the risk tolerance is 0.41. For lower confidence of 0.90 and 0.80, the constraint never becomes binding and with lower risk tolerance, the problem becomes infeasible.

Tuberculosis is a disease which is mostly belonging to binary coverage approach having strict intervention requirements. It is relatively more difficult to obtain high value of continuum of care with same number of walk-in clinics compared to the HIV case. Therefore, the case of opening 5 and 6 centers shows more flexibility to the changes in risk requirements as the number of available locations is increased. However 4 center case is not as risk tolerable as 5 and 6 centers cases due to low continuum of care values.

The same study is repeated for malaria as well with the parameters given for tuberculosis. Malaria has coverage definitions belonging to partial coverage approach. The critical intervention times are not as strict as tuberculosis but stricter compared to HIV. The details of the objective function values can be seen in Appendix A.2. For the cases of opening 4 and 5 centers, the graphs in Figure 6.8 resemble the tuberculosis graphs for 5 and 6 center cases. It is seen that objective function rises with an increase in risk tolerance level. This increase is faster with lower confidence levels, as low risk tolerances can act as non-binding constraints in the problem. Comparing the 5 centers case of malaria and tuberculosis, the risk tolerance level which made the CVaR constraints in the model binding were the 0.36, 0.35 and 0.33

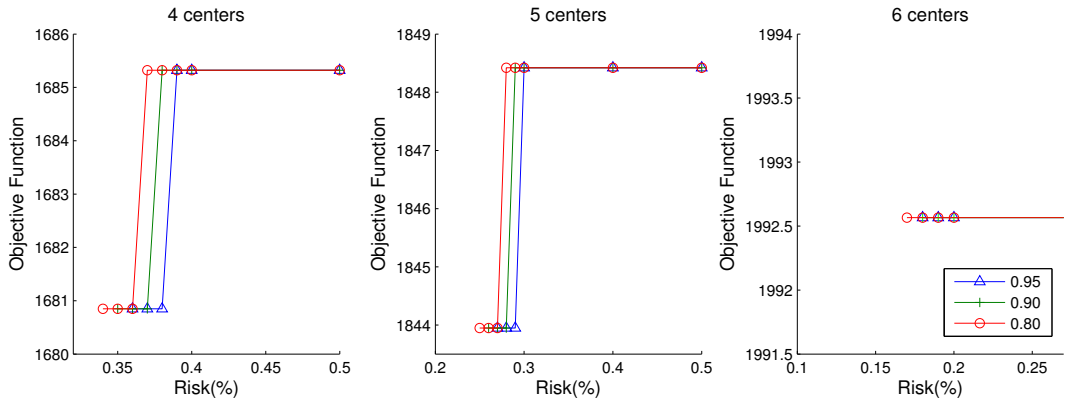


Figure 6.8: Objective function of the model with CVaR measure on Malaria

for confidence levels 0.95, 0.90 and 0.80 respectively. For malaria, these values are changed to 0.29, 0.28 and 0.27 showing a decrease. This can be explained by the flexibility of the coverage definition and intervention time for malaria, making continuum of care values are higher for malaria case. Thus, it requires less risk tolerance to become binding in the model. For the case of opening 6 centers, the CVaR constraints in the model are either not binding or making the model infeasible. This shows that risk involved in the model is relatively low and decreasing further is not possible due to limited set of candidate walk-in clinic locations, except confidence level 0.80. Compared with tuberculosis, the 6 centers case is different, as for tuberculosis changing risk tolerance allowed objective function to vary within a range. Also compared with HIV, the behavior is similar. This difference can be explained again by the coverage approaches of the different diseases.

**Observation 15** *The changes in the optimal planning of the walk-in clinic locations when the risk tolerance level becomes binding shows variation in accordance with the coverage approach of the disease on which the risk tolerance is applied.*

As a next step for the analysis, the continuum of care values at every path on the network is examined. For tuberculosis, 5 center case with confidence level 0.95 and risk tolerance 0.37 and 0.36 is chosen as with a tolerance level of 0.36, the CVaR constraint becomes binding. For malaria, 5 center case with confidence level 0.95 and risk tolerance 0.3 and 0.29 is chosen due to same reasons. The summary of continuum of care values for every disease on every path with different parameters can be found in Appendix B.1 and B.2 for tuberculosis and malaria respectively.

When the continuum of care values are mapped on the bar chart for tuberculosis is examined in Figure 6.9, as the risk constraint becomes binding, shifting in continuum of care values is observed. One of the main changes that can be observed on path 17, is the continuum of care value is increased significantly. In Table 6.7, path 17 is seen to be the path having the greatest demand among high risk paths. Therefore, it is important to focus on the continuum of care value of path 17 which is increased from value of 0.635 to 0.812. Other increases in path 9 and 10 are also observed, which are also having high risk and relatively high demand values as they are sharing same walk-in clinic locations over their path with 17. While opening walk-in clinics, the paths 2 and 11 are sacrificed where 2 has a low demand rate with high risk and 11 has relatively high demand and low risk. Unlike the CVaR constraint with HIV, the path 14 is not focused and the continuum of care vale is not increased further. This can be explained by, path 17 in the problem was already having high values of continuum of care and further increase was not benefiting significantly. For tuberculosis case however, there is still room for improvement and it is more valuable compared to an increase in the continuum of care value of path 14.

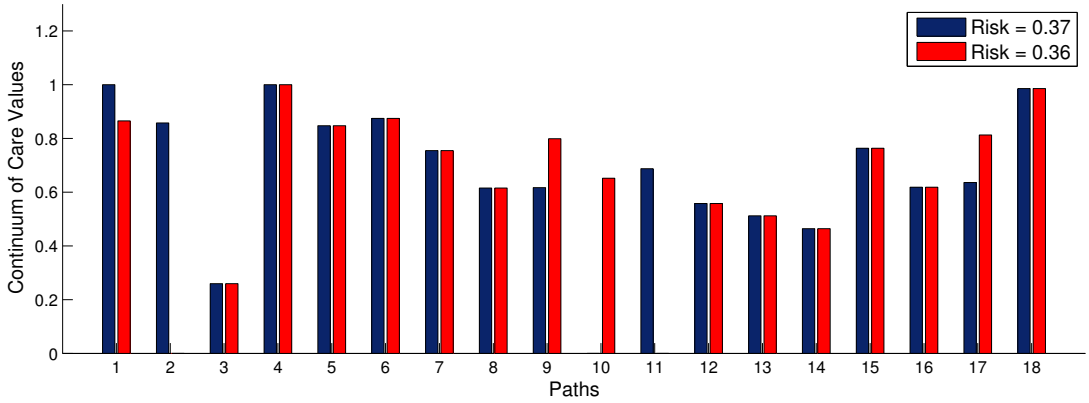


Figure 6.9: The continuum of care values for Tuberculosis for all paths with CVaR measure on Tuberculosis (5 centers)

When the same results for malaria is examined, similar changes are observed as the optimal locations of walk-in clinics are the same for two problems. The difference is that overall continuum of care values are different as they are calculated from two different approaches. It is again seen that out of the paths with high demand and high risk such as 12, 14, 16 and 17, continuum of care value for path 17 is increased significantly as it has the highest demand value. Along with path 17, 9 and 10 is also increased. It can also be seen that overall values for continuum of care for the case of malaria is higher compared to tuberculosis, as it has less

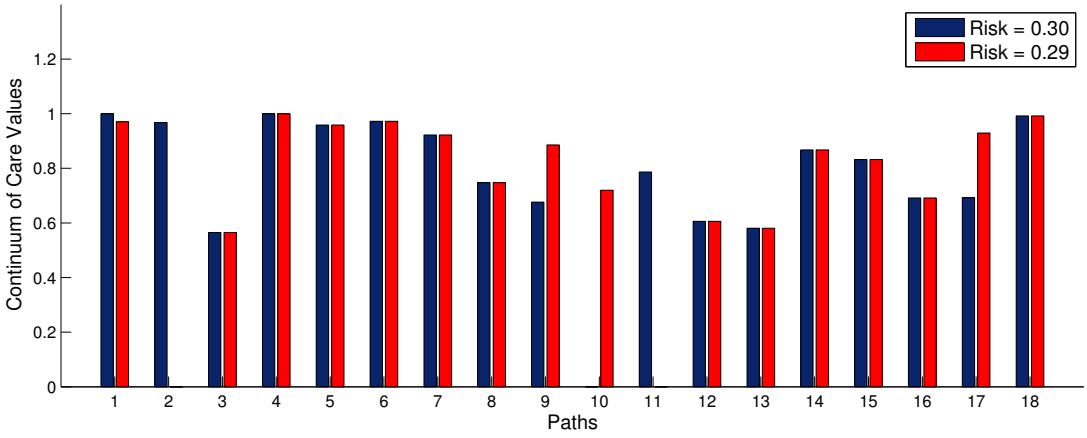


Figure 6.10: The continuum of care values for Malaria for all paths with CVaR measure on Malaria (5 centers)

strict intervention time requirements.

Similar to HIV analysis, the expected covered demand for tuberculosis and malaria is examined. Since the optimal choice of walk-in clinic locations for the risk tolerance values that are making the CVaR constraints binding and non-binding is the same for both, the expected number of covered demand is the same for both cases. Looking at the results in Table 6.15, decrease in number of HIV and STI can be observed in order to meet the given risk tolerances. For malaria, there is an increase in number of people covered, which is a favorable outcome from the model. It is not possible to observe significant changes for tuberculosis, which is also a consequence of low number of tuberculosis demand in the network.

As a final step of the analysis, the CVaR values, which can be found In Tables 6.16 and 6.17 are compared. For tuberculosis case, when the $\omega_{\text{tb}}$ is lowered, the number of demand loss in the 5% of the worst cases has decreased with a small fraction, whereas it remained constant for other diseases. This small difference can again be explained by the relatively lower value for the tuberculosis demand. The same observation is made for the malaria, and decrease is again seen as expected.

Table 6.15: The expected value of demand covered with CVaR measure (5 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{tb}} = 0.37$ / $\omega_{\text{ml}} = 0.30$ | 4075.864 | 2158.58 | 1123.684 | 35.553 |
| $\omega_{\text{tb}} = 0.36$ / $\omega_{\text{ml}} = 0.29$ | 4017.258 | 2193.709 | 1129.596 | 35.228 |

Table 6.16: CVaR values for all diseases with CVaR measure on Tuberculosis (5 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{tb}} = 0.37$ | 4540.000 | 2817.373 | 1358.765 | 18.571 |
| $\omega_{\text{tb}} = 0.36$ | 4540.000 | 2817.373 | 1358.765 | 18.069 |

Table 6.17: CVaR values for all diseases with CVaR measure on Malaria (5 centers)

| | HIV | Malaria | STI | Tuberculosis |
|---|---|---|---|---|
| $\omega_{\text{ml}} = 0.30$ | 4540.000 | 845.212 | 1358.765 | 50.191 |
| $\omega_{\text{ml}} = 0.29$ | 4540.000 | 817.0383 | 1358.765 | 50.191 |

To conclude the computational study, it can be said that each of the cases with focus on different disease's risk on the network adapted the solution to satisfy the requirement. These adaptations resulted as focus towards the paths with high-risk values, which in the end led to decreases in expected number of people covered in the model. It is also observed that deterministic model was missing the paths with high-risk values and resulting in higher loss of coverage, which is avoided with the stochastic model.

Furthermore, the stochastic model results also showed differences when the disease that CVaR constraint applied is changed. Since the diseases are handled with different coverage approaches, the continuum of care values of the network behaved differently. When high continuum of care values are obtained in the model without the inclusion of the risk constraints (such as the cases of HIV where coverage can be attained easier compared to other diseases), risk constraints were not very effective in the problem. Decreasing the risk tolerance for such cases led to infeasibility, which can be explained by the limited range of candidate locations as well. However for other diseases where continuum of care along transportation lines is not easily covered, inclusion of risk constraints affected the problem solution to a higher extent.

Another parameter that is changed was the number of walk-in clinics that was allowed to be opened on the network. When the diseases have stricter approaches, increasing the number of walk-in clinics on the networks allowed stochastic model to plan more effectively as it gave flexibility by increasing the continuum of care values. It is also seen that the tolerance levels that are affecting the optimal network of deterministic model showed changes when the number of centers on the network is increased.

Finally, with the interpretation of $\alpha_d$, VaR and CVaR values, benefits achieved with the addition of constraints is discussed thoroughly. It is seen that the even though the $\alpha_d$ values are not lowered for every case, the CVaR values are dropped which was expected. It is also seen that $\alpha_d$ and VaR values do not have to be the same value when the loss function is not continuous and flat points exist.

# CHAPTER 7

# CONCLUSION

Mobile populations are an important subject area of healthcare problems and planning. As discussed in the thesis, it is proven that there is a strong connection between health problems and mobility. Due to special condition of the mobile populations, the conventional approaches that are adapted for healthcare and location planning in literature are however not very suitable when mobile populations are in focus. This requires careful analysis of their environment, their requirements and the factors that would make an impact.

In Section 3, definition of the problem is given very extensively. This was an important step for the initiation of the thesis as the problem has its unique requirements that need to be integrated to the solution. Therefore, by starting with the motivation, which has emerged from a real-life problem, problem environment is analyzed as its components play a significant role for the development of a solution approach. Consequently, the problem that is studied throughout the thesis is defined.

When the question studied and the environment are determined, it is understood that the problem falls into a different category compared to traditional location problems in the sense that mobile populations need to be served in a continuous manner. With the definition of "continuum of care", it is asserted that single visits to walk-in clinics are not sufficient for a health network but being accessible whenever needed is the goal. Since the definition of every service for disease is various, the model that is proposed to solve the problem should be capable of processing these differences. For this purpose, as first step of the solution process, the continuum of care approaches are developed specialized for every service and disease combination as described in Section 4.1. Following these approaches, the linearization of the approaches are completed together with the other parts of the model. Consequently, the

deterministic model given in Section 4.2 is established.

Application of these approaches together with the other constraints resembles combination of both the flow-interception and the coverage problems. The model tries to capture as many patients as possible considering the demand values and delay durations, which is also an objective in the flow-interception problems. At the same time, the coverage of the patients along the roads is maintained with the foundation of walk-in clinics. However, the coverage of a walk-in clinic is situated towards the transportation lines rather than using a pre-define coverage radius. Furthermore, coverage along the transportation lines with the building of walk-in clinics is not achieved in a lump but rather with strategic positioning of several walk-in clinics.

Introduction of different approaches for the modeling phase was an important part of the formulation as the disease and service specific planning of the networks was relying on these approaches. In order to see the effectiveness of the model and evaluate whether it responded to the initial requirements, computational study is conducted as given in Section 6.1. Looking at the results, it is seen that the model is capable of handling different disease and service requirements by adapting the optimal solutions. This enables planning of specialized networks, which can be crucial for certain situations such as the cases of epidemic strikes. Furthermore, the developed solution approaches can be adapted further for other diseases and services with regards to the characteristic of the problem so that the blank points on the map in terms of healthcare can be identified and further planning in the region can take place. Therefore, it can be said that model is flexible as it allows the adaptation of parameters. Further disease and service combinations can be added in the model as well as modeling of the other problems. The deterministic model provided contribution to the literature as the coverage along the transportation lines for mobile populations given their different and specialized service requirements has not been discussed extensively before.

While developing the deterministic model, it was considered that uncertainties and unreliability played an important role in the problem context. Since every thing cannot be predicted beforehand and even though good predictions are made, sudden events are likely to occur and the deterministic model requires to be flexible enough to allow changes regarding these uncertainties. This condition directed the study towards a stochastic model with a risk-averse behavior in which the demand is modeled as an uncertain factor. The field of stochastic

programming is very large and different formulations are available. Especially when humanitarian logistics is a focus area, risk measures that are applied requires to be planned carefully. Out of the risk measures, the Conditional-Value-at-Risk (CVaR) measure is decided to be adapted for the problem and required modifications are made in the deterministic model.

CVaR is a risk measure that is generally applied in the field of financial planning, especially for the portfolio optimization. An appealing motive behind this risk measure is that, it does not only focus on the percentage of losses associated with the optimal planning but also considers the magnitude of the losses. This allows distinction between the cases of very high and low losses such that the great amount of losses can be avoided. This characteristic of CVaR is decided to be a good fit for the problem at hand as loss in the network should be avoided strongly. Furthermore, adapting the CVaR in the constraints rather than the objective function by implying bounds for every loss of every disease while maximizing the expected number of patients captured is believed to make the model competent and safe. As a result, the deterministic model is adapted and the stochastic model is built as explained in Chapter 5.

The results of the stochastic model given in Section 6.2 showed that the aim was achieved with a risk-averse formulation. When certain transportation lines are defined to be more risky compared to others in terms of higher variation in demand, the model is concentrating on these lines in order to keep the risk level lower than the given bounds. In this way, the number of people that are at risk for the given disease is lowered compared to the no-risk measure model at the expense of sacrificing from the coverage of other diseases. In the overall model, the total expected number of people covered is decreased however, the people at risk is kept under control. When the upper bound on risk tolerance is not very low, it is seen that the model is inclined to open clinics on transportation lines with high value of demand with high risk due to the objective function. When the upper bound is lowered further, the model shift the walk-in clinics that are opened on transportation lines with high demand and low risk values to other transportation lines with low demand and high risk values to satisfy the constraints associated with the risk. At certain cases, the impact of the CVaR constraints in the objective function may not cause very drastic numerical changes, however the resulting walk-in clinic location play an important role in terms of accounting the risk and benefiting to society. Looking at the performance of the stochastic model, it can be said that the model reacted to the risk along the transportation lines as aimed. When uncertainties are present in the network, with the help of this model the risk on the network can be kept under control. Additionally, the CVaR risk

measure is generally applied in the financial field and this study was the first to include in the area of humanitarian logistics. The results proved that CVaR is a good risk measure that can be applied and desired results can be achieved. It can be further tried as performance measure for other models related with humanitarian logistics as well.

A weak point of the model is that problem size is big and computational time that is required to solve the model optimally can be very long for some models. Since there are many constraints and variables involved in the model, optimal solutions may not be attained at reasonable time windows and solution with optimality gaps is observed. In case the data set is greater or more scenarios are involved, the model may require further extensions as Branch and Cut, Benders' Decomposition or L-Shaped method to deal with the complexity of the model. Another idea can be developing a heuristic which may as well include simplifications and reduce the computational time of the model. A heuristic that would allow solutions with low optimality gap and low computational time would be an attractive extension for the problem.

Another weak point is that there are many parameters in the model such as weights given to static and mobile demand populations, diseases and services. The interpretation of these parameters should be carefully made before optimization and their impact on optimal solution should be evaluated with regards to them. These weights provide benefit when a specific network will be planned as they allow playing around with the network. In other cases, the relative importance of especially diseases and services may be required to be based on more scientific grounds for a neutral planning.

Final weak point of the thesis is the data set that has been used for the computational study. The list of candidate locations and paths was limited which resulted in low number of locations falling on the paths and limiting the study. There were infeasibility for the stochastic model when the constraints were tightened due to lack of location availabilities. As a future study, one focus can be towards changing the data set used for testing the problem formulation. A data set having more appropriate paths which have larger number of candidate walk-in clinic locations over the paths is believed to be performing better. The intersection among the paths can also be lowered making a homogeneous distribution of paths in the region.

An additional research can be made for putting an upper bound on the number of services provided in the network. Currently in the thesis, even though there is a constraint in the model for limiting the number of services that can be given in the walk-in clinics, this is not

used in the stochastic model as the main focus was to observe whether the CVaR constraints were effective. A further study can be made by putting bounds on the services and observing whether these bounds effects the walk-in clinic location and create any specialized walk-in clinics on the network.

A final future direction of the study can be made on the adaptation of the loss function defined for the stochastic model. Currently in this study, the loss function is defined over diseases as the main aim was to focus towards the loss associated with diseases. Another risk measures can be developed for the "services of diseases" in addition to diseases as well. With this modification, delivery of services can also be controlled. This would require adaptations in the formulation of the loss functions for the risk definition and increase the number of variables and constraints in the model. This can be an interesting extension to the problem at hand.

# REFERENCES


[1] A. Abdel-Latif. Combining gis-based spatial analysis and optimization techniques to generate optimum facility locations. 2007.

[2] G. Alexandris and I. Giannikos. A new model for maximal coverage exploiting gis capabilities. *European Journal of Operations Research*, 202:328–338, 2010.

[3] P. Artzner, F. Delbean, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, July 1999.

[4] B. Balcik and B. Beamon. Facility location in humanitarian relief. *International Journal of Logistics: Research and Applications*, 11(2):101–121, 2008.

[5] O. Berman and I. Averbakh. Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research*, 91(3):495–506, June 1996.

[6] O. Berman, D. Bertsimas, and R. Larson. Locating discretionary service facilities, ii: Maximizing market size, minimizing inconvenience. *Operations Research*, 43(4):623–632, 1995.

[7] O. Berman, R. C. Larson, and N. Fouska. Optimal location of discretionary service facilities. *Transportation Science*, 26(3):201, August 1992.

[8] J. J. Bwayo, A. M. Omari, A. Mutere, W. Jaoko, C. Sekkade-Kigondu, J. Kreiss, and F. A. Plummer. Long distance truck-drivers: 1. prevalence of sexually transmitted diseases. *East Africa Medical*, 68(6):425–429, June 1991.

[9] G. Chen, M. S. Daskin, Z.-J. M. Shen, and S. Uryasev. The alpha-reliable mean-excess regret model for stochastic facility location modeling. *Naval Research Logistics*, 53(7):617–626, October 2006.

[10] R. Church and C. ReVelle. The maximal covering location problems. *Papers of the Regional Science Association*, 32:101–118, 1974.

[11] J. Current and D. Schilling. The median tour and maximal covering problems. *European Journal of Operations Research*, 73:114–126, 1994.

[12] H. Dandan, Y. Chao, and Y. Jun. Model and algorithm of flow interception problem with multi-type of flows. *Systems Engineering and Electronics, Journal of*, 20(4):755–761, 2009.

[13] H. de Vries. Optimizing Africa. Master's thesis, Erasmus University of Rotterdam, The Netherlands, November 2011.

[14] E. Erkut and S. Neuman. Analytical models for locating undesirable facilities. *European Journal of Operational Research*, 40:275–291, 1989.

[15] R. Farahani, N. Asgari, N. Heidari, M. Hosseininia, and M. Goh. Covering problems in facility locations: A review. *Computers and Industrial Engineering*, 62:368–407, 2012.

[16] A. M. Guide. *Integer Programming Tricks*. Paragon Decision Technology B.V.

[17] P. Krokhmal, J. Palmquist, and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*, 4:11–27, 2002.

[18] M. Kuby and S. Lim. The flow-refuelling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences*, 39:125–145, 2004.

[19] N. Megiddo, E. Zemel, and S. Hakimi. The maximum coverage location problem. *Society for Industrial and Applied Mathematics Journal*, 4(2):253–261, 1983.

[20] N. Miller and A. Ruszczynski. Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Operationsl Research*, 59(1):125–132, January-February 2011.

[21] S. Nickel and J. Puerto. A unified approach to network location problems. *Networks*, 34(4):283–290, December 1999. Special Issue: Centrality Concepts in Network Location.

[22] North Star Alliance. *Five-Year Strategy*, 2011. Internal Report.

[23] North Star Alliance. *Health Strategy*, 2011. Internal Report.

[24] North Star Alliance. *Keeping Supply Chains Healthy and Safe*, November 2011. Information Package.

[25] J. Oppong. Accommodating the rainy season in third world location-allocation applications. *Socio-Economic Planning Sciences*, 30(2):121–137, 1996.

[26] S. Rahman and D. Smith. Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research*, 123(3):437–452, June 2000.

[27] G. Ramjee and E. Gouws. Prevalence of HIV among truck drivers visiting sex workers in kwazulu-natal, south africa. *Sexually Transmitted Diseases*, 29(1):44–49, January 2002.

[28] R. T. Rockafellar. Coherent approaches to risk in optimization under uncertainty. In *In Tutorials in Operations Research INFORMS*, pages 38–61, 2007.

[29] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, April 2000.

[30] H. Tang. *A Study of Network Location Problems*. PhD thesis, The Hong Kong Polytechnic University, 2009.

[31] The Foundation for AIDS Research. Statistics: Worldwide. `http://www.amfar.org/About-HIV-and-AIDS/Facts-and-Stats/Statistics--Worldwide/`, Retrieved: 1 September 2013.

[32] World Health Organization. Communicable diseases and crises. `http://www.who.int/hac/techguidance/pht/comdisease/en/`, Retrieved: 1 September 2013.

[33] World Health Organization. Noncommunicable diseases. http://www.who.int/mediacentre/factsheets/fs355/en/, Retrieved: 1 September 2013.

[34] W. Zeng, I. Castillo, and M. Hodgson. A generalized model for locating facilities on a network with flow-based demand. *Networks and Spatial Economics*, 10:576–611, 2010.

[35] W. Zeng, M. J. Hodgson, and I. Castillo. The pickup problem: Consumer' locational preferences in flow interception. *Geographical Analysis*, 41:107–126, 2009.

# APPENDIX A

# OBJECTIVE FUNCTION VALUES WITH CVAR CONSTRAINTS

## A.1 Tuberculosis

Table A.1: The objective function values for opening 4 centers with CVaR measure on Tuberculosis

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1685.321 | 1685.321 | 1685.321 |
| 0.50 | 1685.321 | 1685.321 | 1685.321 |
| 0.45 | 1685.321 | 1685.321 | 1685.321 |
| 0.44 | 1685.321 | 1685.321 | 1685.321 |
| 0.43 | 1685.321 | 1685.321 | 1685.321 |
| 0.42 | 1685.321 | 1685.321 | 1685.321 |
| 0.41 | 1680.848 | 1685.321 | 1685.321 |
| 0.40 | No solution | 1685.321 | 1685.321 |
| 0.39 | | 1685.321 | 1685.321 |
| 0.38 | | No solution | 1685.321 |
| 0.37 | | | No solution |

Table A.2: The objective function values for opening 5 centers with CVaR measure on Tuberculosis

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 0.4 | 1848.420 | 1848.420 | 1848.420 |
| 0.39 | 1848.420 | 1848.420 | 1848.420 |
| 0.38 | 1848.420 | 1848.420 | 1848.420 |
| 0.37 | 1848.420 | 1848.420 | 1848.420 |
| 0.36 | 1843.948 | 1848.420 | 1848.420 |
| 0.35 | 1843.948 | 1843.948 | 1848.420 |
| 0.34 | 1843.948 | 1843.948 | 1848.420 |
| 0.33 | 1843.948 | 1843.948 | 1843.948 |
| 0.32 | No solution | 1843.948 | 1843.948 |
| 0.31 | | No solution | 1843.948 |
| 0.3 | | | 1843.948 |
| 0.29 | | | No solution |

Table A.3: The objective function values for opening 6 centers with CVaR measure on Tuberculosis

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1992.566 | 1992.566 | 1992.566 |
| 0.40 | 1992.566 | 1992.566 | 1992.566 |
| 0.30 | 1992.566 | 1992.566 | 1992.566 |
| 0.29 | 1992.566 | 1992.566 | 1992.566 |
| 0.28 | 1992.566 | 1992.566 | 1992.566 |
| 0.27 | 1970.983 | 1992.566 | 1992.566 |
| 0.26 | 1970.983 | 1970.983 | 1992.566 |
| 0.25 | 1970.983 | 1970.983 | 1970.983 |
| 0.24 | No solution | No solution | 1970.983 |
| 0.23 | | | No solution |

## A.2 Malaria

Table A.4: The objective function values for opening 4 centers with CVaR measure on Malaria

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1685.321 | 1685.321 | 1685.321 |
| 0.50 | 1685.321 | 1685.321 | 1685.321 |
| 0.40 | 1685.321 | 1685.321 | 1685.321 |
| 0.39 | 1685.321 | 1685.321 | 1685.321 |
| 0.38 | 1680.848 | 1685.321 | 1685.321 |
| 0.37 | 1680.848 | 1680.848 | 1685.321 |
| 0.36 | 1680.848 | 1680.848 | 1680.848 |
| 0.35 | No solution | 1680.848 | 1680.848 |
| 0.34 | | No solution | 1680.848 |
| 0.33 | | | No solution |

Table A.5: The objective function values for opening 5 centers with CVaR measure on Malaria

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1848.420 | 1848.420 | 1848.420 |
| 0.50 | 1848.420 | 1848.420 | 1848.420 |
| 0.40 | 1848.420 | 1848.420 | 1848.420 |
| 0.30 | 1848.420 | 1848.420 | 1848.420 |
| 0.29 | 1843.948 | 1848.420 | 1848.420 |
| 0.28 | 1843.948 | 1843.948 | 1848.420 |
| 0.27 | 1843.948 | 1843.948 | 1843.948 |
| 0.26 | No solution | 1843.948 | 1843.948 |
| 0.25 | | No solution | 1843.948 |
| 0.24 | | | No solution |

Table A.6: The objective function values for opening 6 centers with CVaR measure on Malaria

| $\omega_{\text{HIV}}$ | $\beta_{\text{HIV}}$ | | |
|---|---|---|---|
| | 0.95 | 0.90 | 0.80 |
| 1.00 | 1992.566 | 1992.566 | 1992.566 |
| 0.50 | 1992.566 | 1992.566 | 1992.566 |
| 0.40 | 1992.566 | 1992.566 | 1992.566 |
| 0.30 | 1992.566 | 1992.566 | 1992.566 |
| 0.20 | 1992.566 | 1992.566 | 1992.566 |
| 0.19 | 1992.566 | 1992.566 | 1992.566 |
| 0.18 | 1992.566 | 1992.566 | 1992.566 |
| 0.17 | No solution | No solution | 1992.566 |
| 0.16 | | | No solution |

# APPENDIX B

# CONTINUUM OF CARE VALUES WITH CVAR CONSTRAINTS

## B.1 Tuberculosis

Table B.1: The continuum of care values for all paths with CVaR measure on Tuberculosis (5 centers)

| | $\omega_{HIV}$=0.37 | | | | $\omega_{HIV}$=0.36 | | | |
|---|---|---|---|---|---|---|---|---|
| | HIV | Malaria | STI | TB | HIV | Malaria | STI | TB |
| 1 | 1 | 0.999 | 0.999 | 0.999 | 0.999 | 0.97 | 0.993 | 0.865 |
| 2 | 1 | 0.967 | 0.992 | 0.857 | 0 | 0 | 0 | 0 |
| 3 | 0.768 | 0.564 | 0.652 | 0.259 | 0.768 | 0.564 | 0.652 | 0.259 |
| 4 | 1 | 0.999 | 1 | 1 | 1 | 0.999 | 1 | 1 |
| 5 | 1 | 0.958 | 0.987 | 0.847 | 1 | 0.958 | 0.987 | 0.847 |
| 6 | 1 | 0.972 | 0.994 | 0.874 | 1 | 0.972 | 0.994 | 0.874 |
| 7 | 1 | 0.922 | 0.975 | 0.754 | 1 | 0.922 | 0.975 | 0.754 |
| 8 | 0.953 | 0.747 | 0.886 | 0.615 | 0.953 | 0.747 | 0.886 | 0.615 |
| 9 | 0.907 | 0.676 | 0.76 | 0.616 | 0.996 | 0.885 | 0.952 | 0.798 |
| 10 | 0 | 0 | 0 | 0 | 0.922 | 0.719 | 0.792 | 0.651 |
| 11 | 0.965 | 0.786 | 0.898 | 0.687 | 0 | 0 | 0 | 0 |
| 12 | 0.815 | 0.605 | 0.683 | 0.557 | 0.815 | 0.605 | 0.683 | 0.557 |
| 13 | 0.811 | 0.58 | 0.681 | 0.511 | 0.811 | 0.58 | 0.681 | 0.511 |
| 14 | 0.989 | 0.867 | 0.951 | 0.464 | 0.989 | 0.867 | 0.951 | 0.464 |
| 15 | 0.978 | 0.832 | 0.914 | 0.763 | 0.978 | 0.832 | 0.914 | 0.763 |
| 16 | 0.921 | 0.691 | 0.791 | 0.618 | 0.921 | 0.691 | 0.791 | 0.618 |
| 17 | 0.912 | 0.692 | 0.772 | 0.635 | 1 | 0.929 | 0.976 | 0.812 |
| 18 | 1 | 0.991 | 0.997 | 0.985 | 1 | 0.991 | 0.997 | 0.985 |

## B.2 Malaria

Table B.2: The continuum of care values for all paths with CVaR measure on Malaria (5 centers)

| | $\omega_{\text{HIV}}$=0.30 | | | | $\omega_{\text{HIV}}$=0.29 | | | |
|---|---|---|---|---|---|---|---|---|
| | HIV | Malaria | STI | TB | HIV | Malaria | STI | TB |
| 1 | 1 | 0.999 | 0.999 | 0.999 | 0.999 | 0.97 | 0.993 | 0.865 |
| 2 | 1 | 0.967 | 0.992 | 0.857 | 0 | 0 | 0 | 0 |
| 3 | 0.768 | 0.564 | 0.652 | 0.259 | 0.768 | 0.564 | 0.652 | 0.259 |
| 4 | 1 | 0.999 | 1 | 1 | 1 | 0.999 | 1 | 1 |
| 5 | 1 | 0.958 | 0.987 | 0.847 | 1 | 0.958 | 0.987 | 0.847 |
| 6 | 1 | 0.972 | 0.994 | 0.874 | 1 | 0.972 | 0.994 | 0.874 |
| 7 | 1 | 0.922 | 0.975 | 0.754 | 1 | 0.922 | 0.975 | 0.754 |
| 8 | 0.953 | 0.747 | 0.886 | 0.615 | 0.953 | 0.747 | 0.886 | 0.615 |
| 9 | 0.907 | 0.676 | 0.76 | 0.616 | 0.996 | 0.885 | 0.952 | 0.798 |
| 10 | 0 | 0 | 0 | 0 | 0.922 | 0.719 | 0.792 | 0.651 |
| 11 | 0.965 | 0.786 | 0.898 | 0.687 | 0 | 0 | 0 | 0 |
| 12 | 0.815 | 0.605 | 0.683 | 0.557 | 0.815 | 0.605 | 0.683 | 0.557 |
| 13 | 0.811 | 0.58 | 0.681 | 0.511 | 0.811 | 0.58 | 0.681 | 0.511 |
| 14 | 0.989 | 0.867 | 0.951 | 0.464 | 0.989 | 0.867 | 0.951 | 0.464 |
| 15 | 0.978 | 0.832 | 0.914 | 0.763 | 0.978 | 0.832 | 0.914 | 0.763 |
| 16 | 0.921 | 0.691 | 0.791 | 0.618 | 0.921 | 0.691 | 0.791 | 0.618 |
| 17 | 0.912 | 0.692 | 0.772 | 0.635 | 1 | 0.929 | 0.976 | 0.812 |
| 18 | 1 | 0.991 | 0.997 | 0.985 | 1 | 0.991 | 0.997 | 0.985 |