

## ABSTRACT

Title of Document: OPTIMAL NUMBER AND LOCATION OF  
BLUETOOTH SENSORS FOR TRAVEL TIME  
DATA COLLECTION IN NETWORKS

Mona Asudegi, Master of Science in Civil and  
Environmental Engineering, 2009

Directed By: Professor Ali Haghani, Department of Civil and  
Environmental Engineering

The importance of accurate prediction of travel time in transportation engineering is irrefutable. Travel time is highly used in traffic management and planning. The accuracy of travel time prediction relies on the accuracy of the travel time data. Various methods are being used in collecting travel time data. Recently, a new method in collecting travel time data is introduced that is called Bluetooth technology. In this method, a number of Bluetooth sensors are deployed over the traffic network that can detect the Bluetooth devices in the vehicles to determine the vehicles' travel time based on matching identification and time of identification of the same Bluetooth device at two consecutive sensors.

The goal of this study is to find the optimal number and location of the Bluetooth sensors in a network in order to collect travel time data with a high reliability. Two formulations are proposed for modeling this problem. The formulations consider a new collection of reliability issues. Furthermore, the proposed formulations are able

to solve the problem on large networks exactly. Moreover, various case studies of real world networks are conducted for both formulations and the results are compared.

OPTIMAL NUMBER AND LOCATION OF BLUETOOTH SENSORS FOR  
TRAVEL TIME COLLECTION IN NETWORKS

By

Mona Asudegi

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Science in Civil  
and Environmental  
Engineering  
2009

Advisory Committee:  
Professor Ali Haghani, Chair  
Professor Paul Schonfeld  
Professor Cinzia Cirillo  
Mr. Philip Tarnoff

© Copyright by  
Mona Asudegi  
2009

## Dedication

To my beloved parents and brother,  
in deep gratitude for all their love and support.

## Acknowledgements

I would like to express my deep gratitude to my advisor, Dr. Ali Haghani. I deeply appreciate all the time and effort he has invested in my education and I am equally grateful for his genuine care. Without his instruction and assistance, this thesis would not have been completed.

My thanks also go to all the members of my committee, Dr. Paul Schonfeld, Mr. Philip Tarnoff, and Dr. Cinzia Cirillo for their constructive feedback and valuable guidance. I would like to specially thank Mr. Tarnoff for the generous financial support from the Center of Advanced Transportation Technology throughout the course of the study.

I would also like to thank my friends and fellow transportation engineering graduate students Masoud Hamedi, Kaveh Farokhi and Rafael Olarte for all their care and support. Special thanks go to Abbas Afshar for his help and guidance.

Finally I want to thank my beloved parents and brother, who always supported me to continue my education. Their love and support motivated me to make my best effort and gain whatever I liked in my life.

## Table of Contents

|  |     |
|--|-----|
| Dedication .....   | ii  |
| Acknowledgements .....   | iii |
| List of Tables .....   | v   |
| List of Figures .....  | vi  |
| Chapter 1: Introduction .....  | 1   |
| 1.1 Research Motivation .....  | 1   |
| 1.2. Research Objectives .....                                       | 2   |
| 1.3. Thesis Contribution.....  | 3   |
| 1.4. Thesis Organization .....                                       | 4   |
| Chapter 2: Background and Literature Review .....                    | 5   |
| 2.1. Background.....   | 5   |
| 2.2. Data Collection Methods .....                                   | 6   |
| 2.3. Literature Review on Sensor Location Problem (SLP) .....        | 15  |
| Chapter 3: Problem Specifications and Mathematical Formulation ..... | 24  |
| 3.1. Problem Specifications .....                                    | 24  |
| 3.1.1. Sensors' Specifications .....                                 | 24  |
| 3.1.2. The formulation's Specifications .....                        | 27  |
| 3.2. Problem Formulations.....                                       | 28  |
| 3.2.1. Formulation 1 .....   | 29  |
| 3.2.2. Formulation 2.....  | 35  |
| Chapter 4: Numerical Analysis .....                                  | 38  |
| 4.1 Data Preparation.....  | 38  |
| 4.2. Formulation 1 Numerical Analysis .....                          | 42  |
| 4.2.1. Base Case Study.....  | 42  |
| 4.2.2. Sensitivity Analysis for the Budget .....                     | 44  |
| 4.3. Formulation 2 Numerical Analysis .....                          | 59  |
| 4.3.1. Base Case Study.....  | 59  |
| 4.3.2. Sensitivity Analysis for $\alpha$ .....                       | 64  |
| 4.3.3. Sensitivity Analysis for $\gamma$ .....                       | 67  |
| 4.3.4. Sensitivity Analysis for $\delta$ .....                       | 70  |
| Chapter 5: Comparison of the formulations.....                       | 73  |
| 5.1. Theoretical Comparison.....                                     | 73  |
| 5.2. Numerical Comparison .....                                      | 75  |
| Chapter 6: Conclusions and Suggestions for Further Study .....       | 84  |
| 6.1. Conclusion .....  | 84  |
| 6.2. Further Study .....   | 86  |
| Appendix.....  | 88  |
| Glossary .....   | 88  |
| References.....  | 90  |

## List of Tables

Table 4.1: Networks information for numerical studies

Table 4.2: Base case result for formulation 1

Table 4.3.a: Sensitivity analysis of formulation 1 for Sioux-Falls network

Table 4.3.b: Sensitivity analysis of formulation 1 for Friedrichshain Center network

Table 4.3.c: Sensitivity analysis of formulation 1 for Anaheim network

Table 4.4: Input parameters for formulation 2

Table 4.5: Base case result for formulation 2

Table 4.6.a: Sensitivity analysis of formulation 2 for Friedrichshain Center network

Table 4.6.b: Sensitivity analysis of formulation 2 for Sioux-Falls network

Table 5.1: Formulation 1 output and Formulation 2 input for comparison

Table 5.2: Formulation 1 and 2 comparison



## List of Figures

Figure 2.1: How Bluetooth sensors work

Figure 4.1.a: Map of Sioux-Falls network

Figure 4.1.b: Map of Friedrichshain Center network

Figure 4.1.c: Map of Anaheim network

Figure 4.2.a: Formulation 1 base case result for Sioux-Falls network

Figure 4.2.b: Formulation 1 base case result for Friedrichshain Center network

Figure 4.3.a: Percentage of network coverage for Sioux-Falls network – Formulation 1

Figure 4.3.b: Percentage of network coverage for Friedrichshain Center network –  
Formulation 1

Figure 4.3.c: Percentage of network coverage for Anaheim network – Formulation 1

Figure 4.4.a: Mean of COV of travel time on the covered links versus budget for  
Sioux-Falls network – Formulation 1

Figure 4.4.b: Mean of COV of travel time on the covered links versus budget for  
Friedrichshain Center network – Formulation 1

Figure 4.4.c: Mean of COV of travel time on the covered links versus budget for  
Anaheim network – Formulation 1

Figure 4.5.a: Mean of the travel time prediction relative error on the covered links versus budget for Sioux-Falls network – Formulation 1

Figure 4.5.b: Mean of the travel time prediction relative error on the covered links versus budget for Friedrichshain Center network – Formulation 1

Figure 4.5.c: Mean of the travel time prediction relative error on the covered links versus budget for Anaheim network – Formulation 1

Figure 4.6: Objective function value versus the budget – Formulation 1

Figure 4.7.a: Cost versus budget for Sioux-Falls network – Formulation 1

Figure 4.7.b: Cost versus budget for Friedrichshain Center network – Formulation 1

Figure 4.7.c: Cost versus budget for Anaheim network – Formulation 1

Figure 4.8.a: Solution time versus budget for Sioux-Falls network – Formulation 1

Figure 4.8.b: Solution time versus budget for Friedrichshain Center network – Formulation 1

Figure 4.8.c: Solution time versus budget for Anaheim network – Formulation 1

Figure 4.9.a: Formulation 2 base case result for Sioux-Falls network

Figure 4.9.b: Formulation 2 base case result for Friedrichshain Center network

Figure 4.10.a: Objective function value versus  $\alpha$  for Sioux-Falls network –  
Formulation 2

Figure 4.10.b: Objective function value versus  $\alpha$  for Friedrichshain Center network –  
Formulation 2

Figure 4.11.a: Percentage of network coverage versus  $\alpha$  for Sioux-Falls network –  
Formulation 2

Figure 4.11.b: Percentage of network coverage versus  $\alpha$  for Friedrichshain Center  
network – Formulation 2

Figure 4.12.a: Objective function value versus  $\gamma$  for Sioux-Falls network –  
Formulation 2

Figure 4.12.b: Objective function value versus  $\gamma$  for Friedrichshain Center network –  
Formulation 2

Figure 4.13.a: Percentage of network coverage versus  $\gamma$  for Sioux-Falls network –  
Formulation 2

Figure 4.13.b: Percentage of network coverage versus  $\gamma$  for Friedrichshain Center  
network – Formulation 2

Figure 4.14.a: Objective function value versus  $\delta$  in Sioux-Falls network –  
Formulation 2

Figure 4.14.b: Objective function value versus  $\delta$  in Friedrichshain Center network – Formulation 2

Figure 4.15.a: Percentage of network coverage versus  $\delta$  for Sioux-Falls network – Formulation 2

Figure 4.15.b: Percentage of network coverage versus  $\delta$  for Friedrichshain Center network – Formulation 2

Figure 5.1.a: Formulation 1 comparison case result for Sioux-Falls network

Figure 5.1.b: Formulation 2 comparison case result for Sioux-Falls network

Figure 5.2: Comparison of Formulation 1 and 2 for Sioux-Falls network

Figure 5.3.a: Formulation 1 comparison case result for Friedrichshain Center network

Figure 5.3.b: Formulation 2 comparison case result for Friedrichshain Center network

Figure 5.4: Comparison of Formulation 1 and 2 for Friedrichshain Center network

Figure 5.5: Comparison of Formulation 1 and 2 for Anaheim network

# Chapter 1: Introduction

## 1.1 Research Motivation

In traffic management and planning, accurate route travel time estimation is essential from several aspects. First, travel time estimates are often used in travel time prediction algorithms. Short term or long term travel time can be predicted by analyzing the historical travel time data set. Second, travel time estimates are required for determining offline performance measures for various policy applications. For example, travel time variability is an emerging performance measure increasingly used by decision makers and transportation planners in many project assessment decisions. Also, travel time estimates play an important role in Advanced Traveler Information and Transportation Management Systems. There are different methods in traffic information systems such as variable message signs, GPS devices, and Internet webs (MapQuest, Google Map...) which use the travel time estimates. As a result, the accuracy of the estimated travel time is of a significant worth. This accuracy is highly dependent on the historic data used in prediction algorithms.

There are several methods in collecting travel time data. Recently the staffs at the Center for Advanced Transportation Technology at the University of Maryland have invented a new traffic data collection sensor that is based on the Bluetooth technology. This device has a high level of accuracy and low cost of deployment

compared to current common methods such as probe vehicle and automatic vehicle identification.

To predict route travel times, it is ideal to have historic data for all the links in the network. However, in practice it is very difficult to install sensors on all of the links in a network for two reasons. First, sensor acquisition and installation is costly and subject to budget constraints. Investing more and providing more sensors, one can obtain more travel time information. As a result travel time can be predicted more accurately on the links. Second, not all links provide useful data that can improve the quality of travel time data. For example, there are some links which always operate at or near free flow speed. That means travel times on those links do not change significantly over time. So, even if sensors are installed on those links, not much additional information will be gained by those sensors. As a result, it is important to carefully select the links that are the most valuable links for collecting travel time data from a network-wide perspective.

### 1.2. Research Objectives

The objectives in this study are:

- ✓ Describing the problem of Bluetooth sensor location for collecting travel times, aspects and issues
- ✓ Providing a comprehensive literature review on the sensor location problem and different methods of travel time data collection

- ✓ Proposing two mathematical formulations that comprehensively describe the problem from different aspects. One considers a single objective function while satisfying the quality and reliability constraints. The other is a multi-objective approach that optimizes several parameters for any given level of resources.
- ✓ Applying the formulations on various real world traffic networks and conducting sensitivity analysis over different parameters in the formulations
- ✓ Comparing the two formulations and their results

### 1.3. Thesis Contribution

A new collection of issues is considered in solving the Sensor Location Problem (SLP) in this study. All the issues which have been used in previous studies separately are considered together in addition to a newly introduced term. The new concept that is introduced is maximizing the coefficient of variation (COV) of travel time on the links. The segments with low travel time variation are not interesting for collecting travel time. Adding this term to the model will avoid choosing the links which do not provide useful data.

Also, the largest network that has been solved exactly using previous formulations is of the size of 91 OD pairs by Sherali [2], which is much smaller than the networks that are solved in this study using the proposed formulations (section 3.2). Formulation 2 (section 3.2.2) solved the Sioux-Falls network with 725 OD pairs and formulation 1 (section 3.2.1) solved Anaheim network with 1584 OD pairs. However, formulation 1 can solve much larger problems exactly.

#### 1.4. Thesis Organization

In chapter two, different data collection methods are reviewed and discussed. Also a comprehensive summary of the available literature on the sensor location problem is presented. Chapter two is concluded with an introduction to the new technology of Bluetooth sensors and their application in travel time collection.

In chapter three, the characteristics and specifications of the problem are described in detail. Afterwards, the contributions of the proposed formulations are discussed and compared to the previous studies. At the end, the mathematical formulations of the problem are presented and discussed.

Results of the numerical case studies and sensitivity analysis are discussed in chapter four. Chapter five covers the side by side comparison of the two formulations of the problem. Finally, chapter six provides the conclusions and suggestions for future studies.



## Chapter 2: Background and Literature Review

### 2.1. Background

As cities get larger and more populated, traffic and transportation issues become more important and require more resources. Knowing the travel time on a link or path in advance will help travelers decide on a better path for their trip and also reduces congestion in the network. Travel time prediction has a wide range of applications in Advanced Traveler Information Systems (ATIS). ATIS provides the travelers with travel time estimations on the road segments through different methods such as Dynamic Message Signs (DMS), Highway Advisory Radio (HAR), in-vehicle route guidance systems (like GPS), and internet websites (such as Google map, Map Quest, etc. ).

As a result, accurate prediction of travel time is important since a major portion of trips can be scheduled based on this information. The accuracy of travel time prediction depends on several parameters. One important element is the historic travel time data. The historic travel time data is used in different methods of travel time prediction to predict short term or long term travel times.

Several methods can be used for travel time collection. Each method has advantages and disadvantages. In this study a new technology, Bluetooth sensors, is considered for collecting travel time data which is explained in more details in section 2.2.

Obtaining travel time information on more road segments can increase the accuracy of the travel time prediction. However, providing travel time data on every segment of each link in a large-scale network is not possible mainly due to budget constraints. So besides the technology being used for collecting data, it is important to find the best combination of links for collecting travel time in order to predict travel time with high reliability.

Choosing the most rewarding links for data collection is a well known problem called Sensor Location Problem (SLP). In this chapter, a number of common methods in collecting travel time data including the Bluetooth sensors are introduced in section 2.2. Finally, a review of the previous studies on Sensor Location Problem is presented in section 2.3.

## 2.2. Data Collection Methods

Travel time data can be collected using different methods [1], [15]. The most common methods are:

### **1. Test Vehicle Technique (Floating Car) :**

This common technique consists of hiring someone to drive a vehicle along a pre-selected route and measure the elapsed time and distance traversed. It is possible to equip the vehicles to automate measurement and recording.

The major advantage of electronic test vehicle technique includes:

- Simple and easy method with no need for complicated devices

Some of the disadvantages of electronic DMIs include:

- High probability of human error in recording the data
- Floating car technique is still somewhat labor-intensive and is usually limited to a few measurements per day per staff member
- Travel time is only as accurate as the driver's judgment of traffic conditions
- Floating car technique on arterial streets may not measure the delay of cross street traffic turning onto the study route

## **2. Electronic distance-measuring instruments ( DMIs ) :**

The integration of an electronic DMI with the floating car technique provides an easier and safer way to collect detailed travel time information (compared to traditional floating car method). In the DMI technology, the sensor is attached to the probe vehicle's transmission. The DMI receives consecutive pulses from the vehicle transmission while the vehicle is moving. A DMI typically can provide instantaneous speeds up to every 0.5 second intervals. This detailed travel time information can be downloaded to a portable computer in an easy-to-use data format.

The major advantages of electronic DMIs include:

- Improvement in cost-effectiveness and safety of data collection over the test vehicle method

- Easier data processing than test vehicle technique due to automatic recording of travel times to portable computer
- Detailed travel time and delay information that can be used for identification of bottlenecks and areas of extensive delay
- Providing acceleration and deceleration details that can be a valuable source of input data for fuel consumption and mobile source emissions analysis.

Some of the disadvantages of electronic DMIs include:

- Floating car technique is still somewhat labor-intensive and is usually limited to a few measurements per day per staff member
- Travel time is only as accurate as the driver's judgment of traffic conditions
- Floating car technique on arterial streets may not measure the delay of cross street traffic turning onto the study route

### **3. License plate matching:**

License plate matching was used as early as the 1950s for travel time studies but it was mainly used for tracking or identifying vehicles in origin-destination travel surveys. Early license plate matching methods relied on observers to note the license plates of passing vehicles at certain locations and record the corresponding times on paper or into a tape recorder. License plates were manually matched later

in the office, and travel times were computed. Recent advances in digital technology have substantially improved the accuracy of this technique.

The major advantages of license plate matching include:

- Providing large sample sizes during data collection period
- Providing representative estimate of travel times through random sampling
- Providing travel times at small time intervals, giving a speed profile for the study section throughout the peak period
- Resulting in lower costs per travel time run than the floating car method
- Providing useful data for OD studies

Some of the disadvantages of the license plate matching technique are:

- Data quality concerns from incorrectly reading or mismatching license plates
- Only overall travel times (no stopped delay) are collected
- Less practical for high speed traffic or long roadway sections with low percentage of through-traffic
- High initial cost for equipment purchase
- Potential public disapproval because of privacy concerns

#### **4. Cellular phone tracking :**

Some cities have a dedicated number of cellular phone users to report their position at designated checkpoints, allowing a traffic operations center to estimate

travel times on the basis of several cellular phone reports. Cellular phones in use can also be tracked using geolocation techniques.

Based on the limited test information, cellular phone tracking has the following advantages:

- Minimal cost involved with providing the vehicles with the instrument because of the current popularity of cellular phones
- Cellular network and control center are also able to handle incident and emergency calls

Some of the disadvantages include:

- Large investment in control center for tracking phone calls
- Cellular phones must be in use to track, thereby limiting sample sizes and coverage
- Potential public disapproval because of privacy concerns
- Cooperation of cellular carriers is required

## **5. Video imaging :**

Several video-based systems are developed to measure overall travel times; however, these systems are somewhat less developed than other techniques. Video systems capture vehicle images and attempt to match images from different camera locations. This method needs a lot of processing time as well as personnel and equipments for detecting vehicles in videos.

Some advantages of this method are:

- Inexpensive instruments (even a normal camcorder can be used)

Some of the disadvantages of the video imaging method are:

- Needs a lot of time for matching the vehicles in the videos
- High probability of human error in matching the data

## **6. Automatic vehicle identification(AVI):**

An AVI system consists of an in-vehicle transponder (which can be the toll tags), a roadside reading unit, and a central computer system. When a vehicle containing a transponder (tag) passes a roadside reader unit, the information on the transponder is transferred to the reader unit. The data is processed and the travel time and other traffic data are calculated by matching the tags.

From the recent studies and projects, the use of AVI technology for measuring travel time has the following advantages:

- Real-time travel time information collected and distributed by traffic information system
- Eliminates human error associated with floating car
- Low operating cost once adequate tags are distributed
- Permits fast-track installation with little disruption to traffic
- Used in toll collection and fleet management

AVI technology has the following disadvantages:

- High initial equipment costs (\$25,000 to \$36,000 per reader unit)
- Motorists must acquire and display the tag
- Travel time information availability is limited to fixed routes and checkpoints
- Privacy concern is an issue

#### **7. Automatic vehicle location(AVL):**

Automatic vehicle location (AVL) is another technology with several applications in transportation and traffic management. AVL permits the location of a vehicle to be known automatically, made possible through transmitters that are carried in the vehicle. The transmitters allow the vehicle's location to be determined at frequent intervals, if not continuously. If a map database is used to report the vehicle location, travel times can be calculated for designated roadway sections. AVL systems are becoming more common on transit fleets, police and emergency vehicles, and commercial vehicles. There are several different technologies that can be categorized as AVL. Signpost-based systems utilize antennas at fixed positions along a route, and are commonly used for tracking bus schedules along a fixed route. Ground-based radio navigation systems use a radio frequency and several receiving towers to transmit position information. Global positioning systems (GPS) utilize orbiting satellites for continuous location determination. Differential GPS (DGPS) systems use local towers in addition to satellites to increase the accuracy of GPS.



Based on the recent studies, AVL technology has the following advantages:

- Real-time travel information at frequent intervals
- Eliminates human error
- Not limited to fixed routes or checkpoints
- Used in other transportation applications like fleet management

AVL technology has the following disadvantages:

- High initial costs for sophisticated equipment (\$1,000 to \$4,500 per vehicle)
- Some errors in exact location of vehicles
- Small sample size

### **8. Bluetooth Sensors:**

A new device for collecting travel time data is invented by the staffs at the Center for Advanced Transportation Technology (CATT) at the University of Maryland [16]. This device uses the Bluetooth technology to detect vehicles. The Bluetooth protocol uses an electronic identifier, or tag, in each device called a Media Access Control address, or MAC address for short. The MAC address serves as an electronic nickname for each electronic device in data communications. In this method, a vehicle containing a detectable Bluetooth device (such as a cell phone with Bluetooth, GPS, notebook, hands free, etc) is detected at two Bluetooth sensor stations. The MAC address which is unique for each Bluetooth device and the time of detection is logged when the device is detected at a Bluetooth sensor. Using the logged times for

each MAC ID in two distinct Bluetooth sensor stations, travel time for that specific MAC ID in that road segment is calculated. Since these MAC IDs are unique, it is even possible to track a vehicle in a route and find its origin and destination. So it is also possible to find the path travel times over the network. (Figure 2.1)

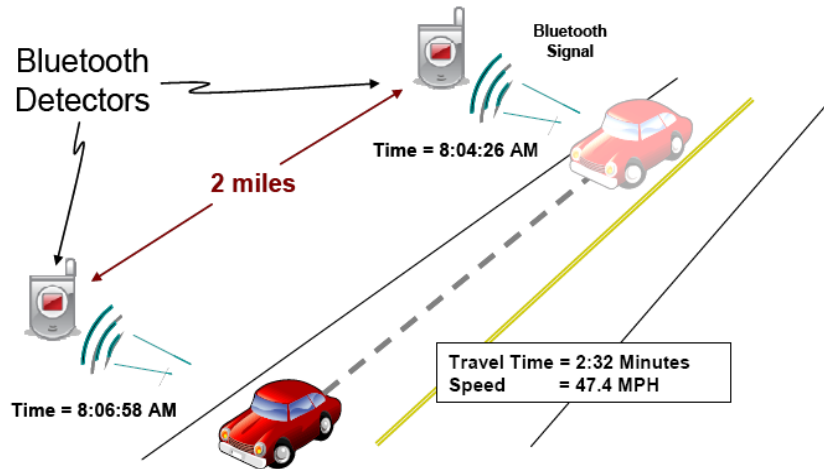


Figure 2.1. How Bluetooth sensors work

Two Bluetooth sensors are required to find the travel time on a road segment. One sensor should be installed at the beginning of the segment and the other should be installed at the end of the segment. However, if there are some segments which have a common node (such as intersections, on ramps or off ramps), one Bluetooth sensor can be installed at that node to cover both segments. Comparing the time of detection of a MAC address in the sensors can determine the direction of the detected vehicle. This property enables coverage of traffic in two directions of a roadway by deploying only two sensors at the beginning and at the end of that roadway segment.

Some advantages of the Bluetooth sensors are:

- High accuracy of data because of tracking each vehicle separately

- Portability of Bluetooth sensors
- Low equipment cost
- Real-time travel information at frequent intervals
- Reduction in human error
- Providing the speed profile as a result of constant monitoring
- No need to distribute tags among vehicles

Some disadvantages of this method are:

- Small sample size on the roads with low traffic volumes
- Limitations on some roads with specific geometry characteristics such as HOV lanes
- Probability of detection of vehicles decreases if the Bluetooth sensors do not face each other completely or have different heights from the road.

### 2.3. Literature Review on Sensor Location Problem (SLP)

In travel time prediction, obtaining travel time information on more road segments can increase the accuracy of the travel time prediction. However, providing travel time data on every link of a large-scale network is not possible mainly due to the budget constraints. Moreover, not all the links provide useful data that can improve the quality of travel time data. For example, there are some links which always operate at or near free flow speed. That means travel times on those links do not change significantly over time. So, even if sensors are being installed on those links, not much additional information will be gained by those sensors. As a result, it is

important to carefully select the links that are the most valuable links for collecting travel time data from a network-wide perspective. To find the most valuable collection of segments is a known problem which is called Sensor Location Problem (SLP).

SLP can be described from two different approaches. One approach is to find the minimum number of sensors and their optimal locations in a network in order to satisfy a certain standard of network performance. The other approach is to optimally locate a certain number of given sensors in a network in order to maximize a defined benefit function. The main question in modeling this problem is how to define the network performance measures in the first method and the benefit function in the second method.

To address this problem, Sherali et al. [2] proposed a mixed-integer optimization model to determine optimal placement of automatic vehicle identification (AVI) readers for travel time estimation in order to maximize a benefit function. The objective function is a quadratic function derived from the multiplication of traffic flow of each link and the coefficient of variation (COV) of traffic flow on that link. That objective function maximizes the traffic flow being detected in the network. Sherali's proposed model is a nonlinear mixed integer program and can only be solved approximately for large-scale networks. The model is as follows:

Parameters:

$G(N, A)$ : The illustration of a transportation network with nodes from  $N$  and links from  $A$

$$b_{ij} = f_{ij} COV_{ij}, \forall (i, j) \in A$$

$b_{ij}$ : Benefit factor for covering arc  $(i, j) \in A$

$f_{ij}$ : Traffic flow on arc  $(i, j) \in A$

$COV_{ij}$ : Coefficient of variation of traffic flow on arc  $(i, j) \in A$

$C_j$ : Cost of installing a reader at location  $j$

$B$ : Maximum budgetary limitation

$R$ : Maximum number of available readers

Decision Variable:

$$y_j = \begin{cases} 1, & \text{if a reader is located at node } j \\ 0, & \text{Otherwise} \end{cases}; j \in N \quad (2.3.1)$$

Objective function:

$$\max \sum_{(i,j) \in A} b_{ij} y_i y_j \quad (2.3.2)$$

Subject to:

$$\sum_{j \in N} y_j \leq R \quad (2.3.3)$$

$$\sum_{j \in N} C_j y_j \leq B \quad (2.3.4)$$

$$y \in \{0,1\} \quad (2.3.5)$$

The objective function (2.3.2) seeks to maximize the total coverage benefit. Constraint (2.3.3) asserts the number of readers used should not exceed the available maximum number  $R$ . Constraint (2.3.4) imposes a budgetary restriction on the total acquisition plus installation cost. Constraint (2.3.5) represents the logical binary restrictions on the decision variables. The biggest network that Sherali solved exactly was a network with 91 OD pairs. However, there was still a gap between the IP and LP solutions.

With regard to the AVI reader location problem for estimating roadway travel times, Teodorovic et al. [3] proposed a composite objective function that is comprised of a weighted average of the total number of AVI tag readings and the number of OD pairs that are at least partially covered by these readings. They developed a genetic algorithm to heuristically maximize this function. In constructing this formulation it is assumed that the OD table is known a priori and that vehicles follow a static shortest path over the network between each OD pair. However, the reader locations might measure travel times over any arbitrary subsets of the shortest paths between the OD pairs, and the model does not distinguish between the benefits accruing from

obtaining information regarding travel times over one portion of an OD shortest path from another.

A multi-objective model for AVI readers was proposed in 2004 by Chen, Choontinan, and Pravinongvuth [4]. In this model they introduced three objectives. First was minimizing the number of readers. Second was maximizing the number of OD pairs covered. And third was maximizing the number of individual readings over the entire network. The model is as following.

Parameters:

N: Set of nodes in the network and  $|N|$  is the size of the set N

A: Set of links in the network and  $|A|$  is size of set A

W: Set of OD pairs and  $|W|$  is size of set W

$R_w$  : Set of Paths between OD pair w

L: Number of AVI readers available

$\delta_{ra}^w$ : A path-link indicator denoting 1 if link a is on path r between OD pair w, and 0 otherwise

Decision variables:

$x_a$  : An integer decision variable indicating the number of AVI readers to be installed on link a  $\{0, 1, \dots, \rho\}$

$y_w$ : A binary decision variable indicating whether OD pair  $w$  is covered (or intercepted) or not

$z_r^w$ : An integer decision variable indicating the number of readings along path  $r$  between O-D pair  $w$   $\{0, 1, \dots, |A| \times \rho - 1\}$

Objective function:

$$\max(-f_1(x), f_2(x), f_3(x)) \quad (2.3.6)$$

Subject to:

$$z_r^w \leq \max\left\{0, \sum_{a=1}^{|A|} \delta_{ra}^w x_a - 1\right\}, \forall r \in R_w, w \in W \quad (2.3.7)$$

$$y_w \leq z_r^w, \forall r \in R, w \in W \quad (2.3.8)$$

$$x_a \in \{0, 1, \dots, \rho\}, \forall a \in A \quad (2.3.9)$$

$$y_w \in \{0, 1\}, \forall w \in W \quad (2.3.10)$$

$$z_r^w \in \{0, 1, \dots, |A| \times \rho - 1\}, \forall r \in R, w \in W \quad (2.3.11)$$

Equation (2.3.6) is the multi-objective function of the AVI reader location problem, which is to configure the AVI readers into a traffic sensor network that is capable of capturing as much flow as possible while covering a maximum number of OD pairs with a minimum number of AVI readers. Equation (2.3.7) ensures that if a used path



is covered there must be a minimum of two readers along the path; otherwise, it is considered as not covered. Furthermore; for OD pair  $w$  to be covered, equation (2.3.8) requires  $z_r^w$  (all used paths  $r$  serving this OD pair  $w$ ) to be positive. Equations (2.3.9) to (2.3.11) constrain the solutions to be either binary integers or integers. The largest network solved with this model using the Genetic algorithm was a 34 OD pair network. They solved the problem three times for each objective function's element individually for Irvine network (36 nodes and 626 links). They compared the non-dominated solution for each objective element together and the trade off among them.

Bartin et al. [5] showed that the optimal sensor placement for travel time estimation can be determined by minimizing a weighted summation of speed variations of all roadway segments, each of which is associated with a sensor. A nearest neighbor algorithm was then used to solve the problem. The proposed clustering approach not only finds the segments for the sensor deployment but also determines the number of segments to monitor. The algorithm was used to solve a network with 9 OD zones and three interchanges. However, it was not guaranteed to provide a globally optimal solution in polynomial time.

Yang and Miller-Hooks [6] proposed a model to select the information critical arcs such that the greatest benefits can be derived. The problem explored was that of selecting a fixed number of arcs, representing a subset of the network arcs, referred to as information critical arcs (ICAs), which will be instrumented to collect real-time information, such as travel times, from a transportation network. A modified maximum-covering formulation is presented along with a heuristic solution

procedure. Variation of travel time and the amount of volume on each arc were the issues in determining the arc's benefit. The proposed heuristic was used in solving the Texas highway network consisting of 183 nodes and 549 arcs. Though the result was good but the method still could not solve the problem exactly.

Gentili and Mirchandani [7] addressed the problem of locating active sensors on the arcs of a traffic network where the sensors can provide data on paths. They showed that each sensor located on an arc results in a set of linear equations in path flow variables that may be used for finding path flows. Then, they solved the problem of the selection of the minimum number of arcs that add linear equations that result in a full rank coefficient matrix. They presented a formulation of the problem and analyzed three different scenarios depending of the number of conventional counting sensors already located on the network. The general problem was shown to be NP-hard. Through the proofs of the polynomially solvable cases, some new graph theoretic models and theorems were obtained, which in their own right added to the graph theoretic knowledge base, besides providing insight to develop an approximate algorithm for the general case.

Hu et al [8] propose a basis link method to address the network sensor location problem under steady state traffic conditions. They solve the maximum OD covering for this problem. In this method the minimum number of links was selected to cover all OD pairs while no other reliability issues was considered.

Fei and Mahmassani [9] proposed models that use the Kalman filtering method to explore time-dependent maximal information gains across all the links in the network.

The research proposed two types of sensor location models to solve an O-D coverage problem and a maximal information gain driven problem. The focus was on solving the sensor location problem as an OD coverage problem under a dynamic traffic assignment. They produced a quality estimated OD matrix that integrates link observation data that minimize the variance of the O-D flow estimator. They constructed an unbiased generalized least squares estimator, using a linear relationship and link flow proportions obtained from a dynamic traffic assignment procedure. The models were developed to identify link sensor locations that produce maximal information gains and maximal OD pair reliability. A sequential algorithm was developed to solve the proposed models. The largest network solved with the proposed heuristic was the Irvine network, California, with 238 sensors that covered 3,660 OD pairs, including 326 nodes and 626 links.

Yang He and Mirchandani [19] proposed two models. One maximized the route length monitored when a certain number of readers had to be located on the network. The other maximized the ability of predicting travel time by maximizing the variance reduction in travel time prediction for each link.

All previous research that attempt to solve SLP use a number of issues for reliability of travel time data and most of them use heuristics for solving the problem. So the proposed models mostly have not been solved exactly.

In chapter 3, the differences between the suggested formulations and the previous works are discussed.

## Chapter 3: Problem Specifications and Mathematical Formulation

### 3.1. Problem Specifications

#### **3.1.1. Sensors' Specifications**

A new traffic detection device based on Bluetooth technology has been recently developed [16]. Studies suggest the high accuracy and reliability of data from the Bluetooth sensors [18]. Previous studies suggest that sample size plays an important role in accuracy and reliability of the data. It is reported that the average rate of detectable Bluetooth devices in vehicles in a normal traffic stream is in the range of 3-5% of the total link traffic volume. The sample size needed for a reliable travel time prediction can be calculated using equation 3.1[10] [11]:

$$n = \frac{z_{\alpha/2} c.v.^2}{E} \quad (3.1)$$

Where:

n : Sample size

$\alpha/2$  : Standard normal variate based on desired confidence level in the travel time estimate

c.v. : Travel time coefficient of variation = Standard deviation / Mean of travel time

E : Permitted relative error (%)

For example, considering  $\alpha = 0.05$  (95% confidence), the average sample size for Anaheim network (one of the case study networks) is calculated using equation (3.1). The sample size for collecting data is 5% of the total volume on the links in the network. Since studies have shown that 3% to 5% of vehicles can be detected as having Bluetooth devices in the traffic stream, it is concluded that Anaheim network provides the minimum sample size for collecting travel time data with 95% of confidence using Bluetooth sensor. (Volume data on the links is provided by [13] and [14])

A vehicle will be detected by a sensor if the power of transmitting and receiving antenna and the device in the vehicle satisfy the Friis Transition equation. In its simplest form, the Friis transmission equation is shown in equation 3.2[12]. Given two antennas, the ratio of power received by the receiving antenna,  $P_r$ , to power input to the transmitting antenna,  $P_t$ , is given by

$$\frac{P_r}{P_t} = G_t G_r \left( \frac{\lambda}{4\pi R} \right)^2 \quad (3.2)$$

where:

$G_t$  and  $G_r$  : The antenna's gain of the transmitting and receiving, respectively

$\lambda$  : The wavelength

R : Distance between transmitting and receiving antenna

The antenna gains must be in decibels. Also the wavelength and distance units should be the same. There are a number of assumptions in the simple form of the Friis transmission equation:

1. The antennas are correctly aligned and polarized.
2. The bandwidth is narrow enough that a single value for the wavelength can be assumed.
3.  $P_r$  ( $P_t$ ) is understood to be the available power at the receive antenna terminals (the power delivered to the transmit antenna). There is loss introduced by both the cable running to the antenna and the connectors. Furthermore, the power at the output (input) of the antenna will only be fully delivered into the transmission line (free space) if the antenna and transmission line are conjugate matched.
4. The antennas are in unobstructed free space, with no multipath.

If all the ideal conditions are provided, by using the Bluetooth sensors' information ( $G_t$  and  $G_r = 3$  dbi,  $\lambda = 0.12$  m,  $P_r = 17$  mw), the covering distance is calculated as  $R \approx 300$  ft  $\approx 91$  m. FHWA standard suggests the values for the traffic lane width as: Highway= 3.6 m, Arterial=3.3 m, local=2.7 m. If each traffic lane is assumed to be about 3.2 m and sensors are installed in the medians of the roads, then each sensor can cover the traffic of both directions on a road with the detection probability of 100% in ideal situations.

### 3.1.2. The formulation's Specifications

There are different issues that affect the reliability of travel time data such as percentage of covered volume in the network and the COV of travel time on the links. Covered links should provide a certain level of reliability issues in the network. Besides, there are some other elements in optimizing the SLP such as the number of sensors and the number of covered OD pairs which should be considered in solving the problem. Overall, the terms that should be considered in solving a SLP problem are:

1. Covering a high percentage of the total volume in the network
2. Covering links with high variation in their travel time (segments with low travel time variation are not interesting for collecting travel time)
3. Covering the links with low relative error in travel time prediction. This means that in similar conditions, it is favorable to cover the links that their travel time can be predicted with smaller relative error.
4. Covering as many origin-destination ( OD ) pairs as possible even if the OD pair is covered partially
5. Considering cost constraints
6. Using a minimum number of sensors

Coefficient of variation of travel time on a link is the variation of travel time on the link during the peak hour in a day. So if the COV is small it means that the travel time does not change during the peak hour. This means the travel time is already known and there is no need to collect a new set of data. However, travel time prediction error

is the difference between the link travel time that is predicted for any given time window using historic data, and the real travel time data on that link during that time window. For example the prediction of travel time for 5:00 PM on Tuesday using the data of the same time and day of the previous weeks. So if the error is large it means that the traffic on that link does not follow a predictable behavior. Therefore, collecting data on the link will not be useful in travel time prediction algorithms. Consequently the COV of travel time and travel time prediction error are not exactly the same and should be considered as two different factors in the objective function.

In all previous studies of optimal sensor location, one or some of the above issues is considered. Some of them considered the issues as objective function and solved the problem as a multi-objective problem, while some others try to solve the problem considering one issue as the objective and the others as constraints.

In this study all the issues are considered together. Two formulations are proposed for the problem. In formulation 1, the problem is formulated as a multi-objective problem subject to cost constraint. In the formulation 2 the objective function is to minimize the number of sensors, while all other five issues are considered as constraints. The two formulations are introduced in the section 3.2. The two proposed formulations are compared theoretically and numerically in chapter 5.

### 3.2. Problem Formulations

Every traffic network consists of several nodes and links. In this research, a node is inserted wherever the traffic flow changes significantly. For example intersections or



on-ramps and off-ramps create a node. Every segment of a road between two nodes is defined as a link.

An OD pair consists of an origin and a destination node, and one or more transition nodes. To go from the origin to the destination, one should pass the transition nodes. The links between each two transition nodes in an OD pair is called an OD pair link. The collection of links in an OD pair is called the OD pair links. An OD pair is called partially covered if and only if at least one of the links in the OD pair links is covered by detectors.

In the following sections the mathematical formulations of the problem are introduced.

### **3.2.1. Formulation 1**

Parameters:

L : Total number of links in the network

N : Total number of nodes in the network

R : Total number of OD pairs in the network

C : Total budget

$\alpha$  : Minimum COV of travel time on the links

$\gamma$  : Minimum percentage of covered volume in the network

$\delta$  : Minimum percentage of covered OD pairs in the network

$\Gamma(i)$  : The collection of nodes which have a link to the node (i);  $i \in \{1,2,\dots, N\}$

$V_{ij}$  : Traffic volume on the link from node (i) to node (j);  $i \in \{1,2,\dots, N\}$ ;  $j \in \Gamma(i)$

$\left(\frac{\sigma}{m}\right)_{TT-ij}$  : Coefficient of variation (COV) of travel time on the link i-j;  $i \in \{1,2,\dots, N\}$ ;

$j \in \Gamma(i)$

$$\beta_{ij} = \left| \frac{t_{ij}^* - t_{ij}}{t_{ij}} \right| \text{ for all } i \in \{1,2,\dots, N\}; j \in \Gamma(i)$$

$\beta_{ij}$  : Relative error of travel time prediction on link i-j

$t_{ij}^*$  : Predicted travel time on the link i-j

$t_{ij}$  : Real travel time on the link i-j

$c_i$  : Installation Cost of a sensor on node (i);  $i \in \{1,2,\dots, N\}$

$$P_{ij}^r = \begin{cases} 1, & \text{if link (ij) is in the OD pair (r) links} \\ 0, & \text{Otherwise} \end{cases}; i \in \{1,2,\dots, N\}; j \in \Gamma(i); r \in \{1,2,\dots, R\}$$

Decision variables:

$$K_{ij} = \begin{cases} 1, \text{ if link } (ij) \text{ is covered} \\ \text{(sensors are installed on both nodes } (i) \text{ and } (j)) \\ 0, \text{ Otherwise} \end{cases}; i \in \{1, 2, \dots, N\}; j \in \Gamma(i)$$

$$x_i = \begin{cases} 1, \text{ if a sensor is installed on node } (i) \\ 0, \text{ Otherwise} \end{cases}; i \in \{1, 2, \dots, N\}$$

$$Y_r = \begin{cases} 1, \text{ if at least one link in OD pair } (r) \text{ links is covered} \\ 0, \text{ Otherwise} \end{cases}; r \in \{1, 2, \dots, R\}$$

Objective function:

$$\max \left\{ \frac{\sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot V_{ij}}{\sum_{i=1}^N \sum_{j \in \Gamma(i)} V_{ij}} + \frac{\sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot \left(\frac{\sigma}{m}\right)_{TT-ij}}{\sum_{i=1}^N \sum_{j \in \Gamma(i)} \left(\frac{\sigma}{m}\right)_{TT-ij}} - \frac{\sum_{i=1}^N x_i}{N} - \frac{\sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot \beta_{ij}}{\sum_{i=1}^N \sum_{j \in \Gamma(i)} \beta_{ij}} + \frac{\sum_{r=1}^R Y_r}{R} \right\} \quad (3.3)$$

Subject to:

$$\sum_{i=1}^N c_i \cdot x_i \leq C \quad (3.4)$$

$$Y_r \leq \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} P_{ij}^r; r \in \{1, 2, \dots, R\} \quad (3.5)$$

$$R Y_r \geq \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} P_{ij}^r; r \in \{1, 2, \dots, R\} \quad (3.6)$$

$$\frac{x_i + x_j}{2} \geq K_{ij}; i \in \{1, 2, \dots, N\}; j \in \Gamma(i) \quad (3.7)$$

$$K_{ij} \geq x_i + x_j - 1; i \in \{1, 2, \dots, N\}; j \in \Gamma(i) \quad (3.8)$$

$$x_i = \{0,1\}; i \in \{1,2,\dots, N\} \quad (3.9)$$

$$K_{ij} = \{0,1\}; i \in \{1,2,\dots, N\}; j \in \Gamma(i) \quad (3.10)$$

$$Y_r = \{0,1\}; r \in \{1,2,\dots, R\} \quad (3.11)$$

Equation (3.3) is the objective function which is a multi-objective function. Chen, and Choontinan [4] solved SLP problem with a fewer number of objective elements. They solved the problem several times. Each time they considered only one of the elements of the objective and then compared the result for each objective together. But here all the objectives are optimized together. The main point in solving the proposed multi-objective function is the comparison of different objectives with different units and scales to each other. In this regard, all the terms in objective function are normalized between 0 and 1. For example for the volume element, each link's volume is divided by the total volume in the network. The result is the percentage of the total volume in the network that belongs to each link. This term is a unit-less value between 0 and 1 which is called the link's contribution to the volume objective. So the same process is done for other elements in the objective function. As a result all the terms in the objective function are unit-less with the same scale which ensures they are addable. If the objective elements do not have the same weights in their contribution to the objective function, then they can be multiplied by a user-defined weight. The weight values affect the solution. The weight vector can be determined by studying the Pareto set of the weight vectors. In this method different sets of weights are generated through different methods (for example random generation). This approach gives an

idea of the shape of the Pareto surface and provides the user with more information about the trade-off among the various objectives [20], [21], [22]. Then considering the trade-offs between the objectives, the decision maker can decide on the value of the weight vector. However, in this study all the elements of the objective function are considered with the same weights. Examining proper weights for different components of the objective function is left for future research. In the objective function, the first element is the volume objective. It maximizes the covered volume in the network. The second element is the COV element. Links with larger COV of travel time contributes more to COV element in the objective function. So they have priority to the other links for being selected. The third element is the number of sensors element. It minimizes the total number of sensors in the network. The fourth element is the error element. It minimizes the average of relative error of travel time prediction on the selected links. To compute relative travel time prediction error ( $\beta$ ), travel time on the links during the desired peak hour is assumed to be known. This data can be provided through historic data or traffic assignment methods. Having the historic real travel time data and using a travel time prediction method,  $\beta$  can be calculated as follows:

$$\beta_{ij} = \left| \frac{t_{ij}^* - t_{ij}}{t_{ij}} \right| \text{ for all } i \in \{1, 2, \dots, N\}; j \in \Gamma(i)$$

where:

$\beta_{ij}$  : Relative error of travel time prediction on link i-j

$t_{ij}^*$  : Predicted travel time on the link i-j

$t_{ij}$  : Real travel time on the link i-j

Finally, the last element is the OD covering element. This element maximizes the coverage of OD pairs in the network. It should be mentioned that an optimal solution for this aggregated objective function might not be the optimal solution to the problem if each element was optimized individually.

Equation (3.4) is the budget constraint. This constraint requires the total installation costs to be less than or equal to the available budget C.

Equations (3.5) and (3.6) are the linking constraints between the link and OD variables. OD variable is defined as  $Y_r = \min\{1, \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} P_{ij}^r\}$ . In other words,  $Y_r$  is equal to one if and only if there is at least one link of the OD pair ( r ) links which is selected for installing sensors, and is zero if and only if none of the OD pair ( r ) links is chosen.

Equations (3.7) and (3.8) are the linking constraint between the node and link variables.  $K_{ij} = \max\{0, x_i + x_j - 1\}$  ; The link (ij) is selected ( $K_{ij} = 1$ ) if and only if both nodes  $x_i, x_j$  are picked up ( $x_i = 1, x_j = 1$ ). And it is zero if and only if none of the nodes or only one of them is chosen.

Equations (3.9), (3.10), and (3.11) define the binary variables of the problem.

### 3.2.2. Formulation 2

Objective function:

$$\min \sum_{i=1}^N x_i \quad (3.12)$$

Subject to:

$$\left( \frac{\sigma}{m} \right)_{TT-ij} \geq \alpha \cdot K_{ij}; i \in \{1, 2, \dots, N\}; j \in \Gamma(i) \quad (3.13)$$

$$\sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot V_{ij} \geq \gamma \cdot \sum_{i=1}^N \sum_{j \in \Gamma(i)} V_{ij}; i \in \{1, 2, \dots, N\}; j \in \Gamma(i) \quad (3.14)$$

$$\sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot \beta_{ij} \leq \frac{\sum_{i=1}^N \sum_{j \in \Gamma(i)} \beta_{ij}}{L} \cdot \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \quad (3.15)$$

$$\sum_{r=1}^R Y_r \geq \delta \cdot R \quad (3.16)$$

$$Y_r \leq \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot P_{ij}^r; r \in \{1, 2, \dots, R\} \quad (3.17)$$

$$R \cdot Y_r \geq \sum_{i=1}^N \sum_{j \in \Gamma(i)} K_{ij} \cdot P_{ij}^r; r \in \{1, 2, \dots, R\} \quad (3.18)$$

$$\frac{x_i + x_j}{2} \geq K_{ij}; i \in \{1, 2, \dots, N\}; j \in \Gamma(i) \quad (3.19)$$

$$K_{ij} \geq x_i + x_j - 1; i \in \{1,2,\dots, N\}; j \in \Gamma(i) \quad (3.20)$$

$$\sum_{i=1}^N c_i \cdot x_i \leq C \quad (3.21)$$

$$x_i = \{0,1\}; i \in \{1,2,\dots, N\} \quad (3.22)$$

$$K_{ij} = \{0,1\}; i \in \{1,2,\dots, N\}; j \in \Gamma(i) \quad (3.23)$$

$$Y_r = \{0,1\}; r \in \{1,2,\dots, R\} \quad (3.24)$$

Equation (3.12) is the objective function which is minimizing the total number of sensors being installed in the network.

Equation (3.13) eliminates links with COV less than  $\alpha$ . Links with small variance in their travel time are not interesting for collecting data. Since the travel time on these links does not change dramatically, the available historic data can provide a good estimate of travel time on these links. Also, the mean value of travel times may cover a wide range across different links. As a result, it is better to use coefficient of variation of travel time instead of variance since COV is the variance divided by the travel time mean.

Equation (3.14) is the minimum volume coverage constraint. This equation ensures covering of a minimum percentage ( $\gamma$ ) of the total traffic volume in the network. This constraint requires covering the more important links that usually carry higher traffic volumes.



Equation (3.15) is the error constraint. This equation ensures the selected links to have an acceptable error in travel time prediction. This constraint picks the links with relative travel time prediction error less than or equal to the average relative travel time prediction in the network.

Equation (3.16), which is the minimum OD covering constraint, ensures that at least  $\delta$  percent of the total OD pairs in the network is covered (completely or partially).

All other equations are discussed in 3.2.1.

## Chapter 4: Numerical Analysis

### 4.1 Data Preparation

In this chapter, the proposed formulations are evaluated on real world networks. The selected networks cover a considerable range on the number of nodes, links and OD pairs. The information of the networks used in the case studies is presented in table 4.1. The maps of the networks are presented in figures 4.1, 4.2, and 4.3 [14]. The number of nodes, links and OD pairs are given and the mean COV of travel time and the mean travel time prediction relative error are calculated as is described in the following.

| # | Network                      | Number of Nodes | Number of Links | Number of OD Pairs | Mean of COV of Travel Time | Median of COV of Travel Time | Mean of Travel Time Prediction Relative Error | Median of Travel Time Prediction Relative Error |
|---|------------------------------|-----------------|-----------------|--------------------|----------------------------|------------------------------|---|---|
| 1 | <i>Sioux-Falls</i>           | 24              | 76              | 725                | 0.28                       | 0.11                         | 0.15  | 0.32  |
| 2 | <i>Friedrichshain Center</i> | 224             | 523             | 552                | 0.16                       | 0.07                         | 0.33  | 0.11  |
| 3 | <i>Anaheim</i>               | 416             | 914             | 1584               | 0.25                       | 1.24                         | 2.04  | 0.26  |

*Table4.1. Networks information for numerical studies [14]*

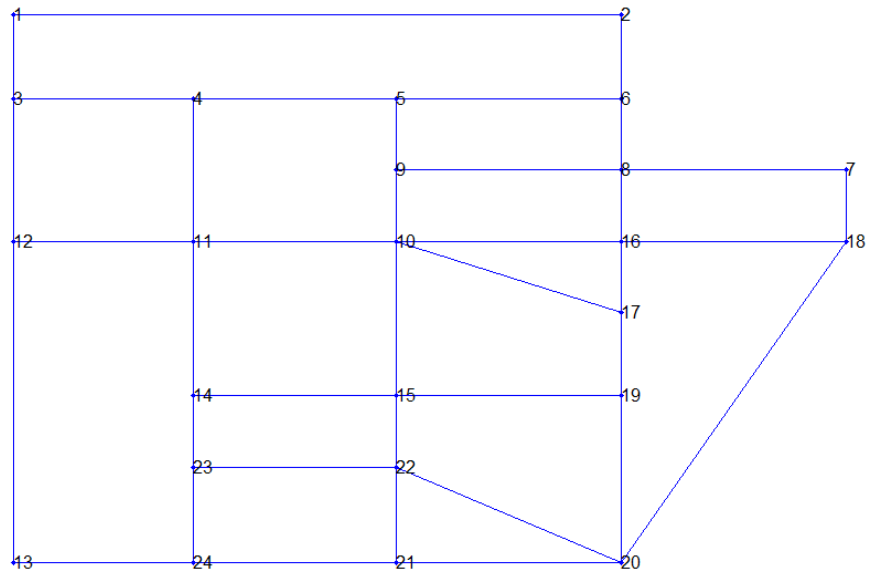


Figure4.1.a. Map of Sioux-Falls network

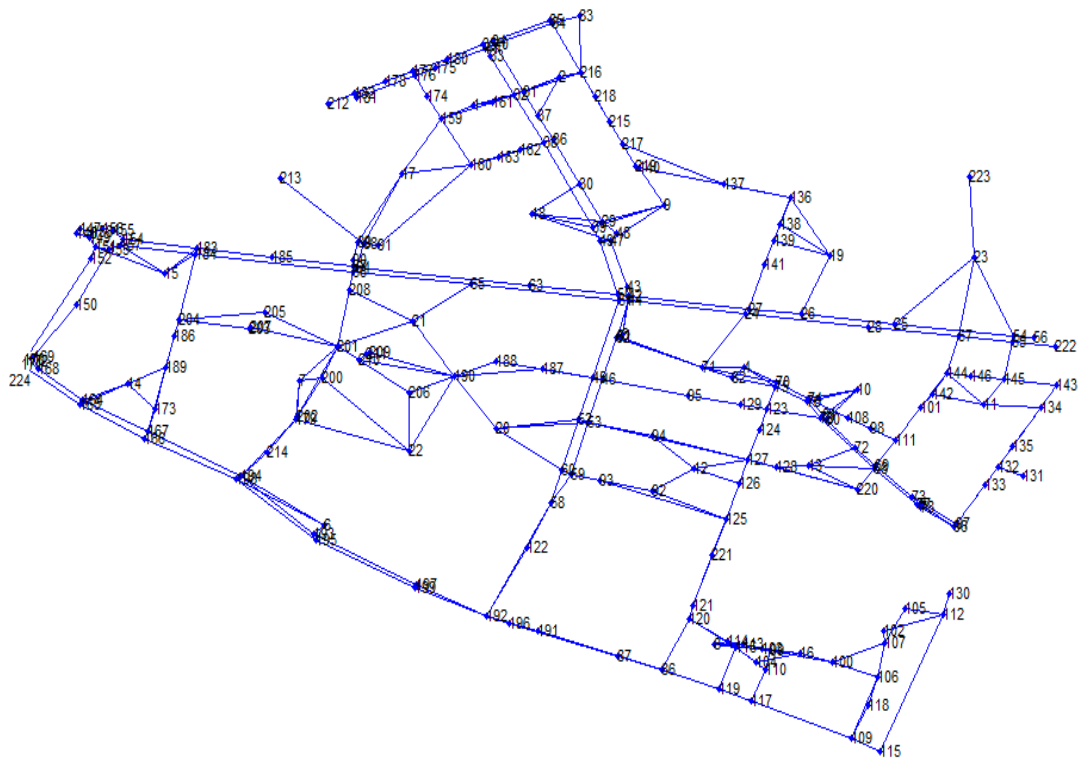
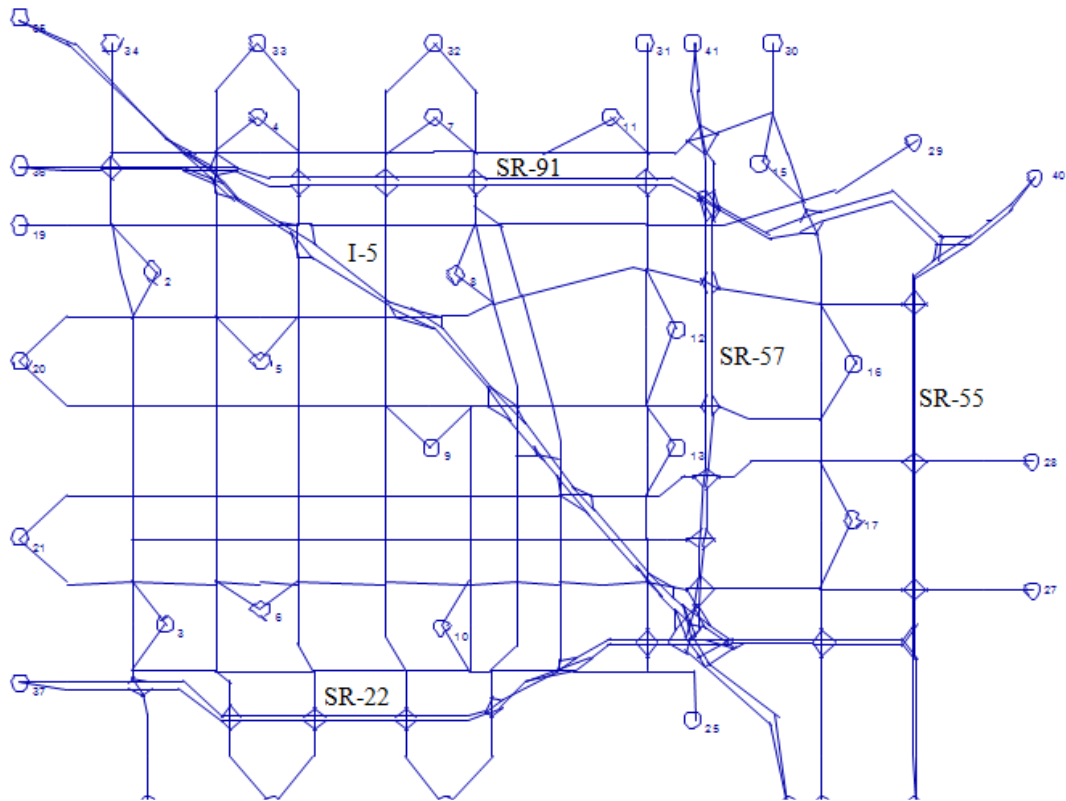


Figure4.1.b. Map of Friedrichshain Center network



*Figure4.1.c. Map of Anaheim network*

Some assumptions are needed to solve the numerical problems. The COV of travel time and traffic volume on each link, and the OD pair links are the main inputs. Volume and mean speed on each link, the length of the links, and the OD pair links information are adopted from Olarte's Masters thesis [13]. Using the average speed and the link length, mean travel time on each link is calculated. A randomly generated standard deviation of travel time is assigned to each link. Then the coefficient of variation of travel time for each link is calculated by dividing the standard deviation of travel time by the link's average travel time.

Since the complete travel time information for each link was not available, a random value based on the mean travel time of each link is assigned to each link as its travel

time prediction. Using the predicted travel time and the average travel time on each link, the travel time prediction relative error is calculated using equation 4.1.

$$\beta_s = \left| \frac{t_s^* - t_s}{t_s} \right| \quad (4.1)$$

Where

$\beta_s$  : Travel time prediction relative error on the link (s)

$t_s^*$  : Predicted travel time on the link (s)

$t_s$  : Real travel time on the link (s)

Three other parameters should be known as input for formulation 2. Those parameters are:  $\alpha$  (minimum COV of travel time on the links),  $\gamma$  (minimum percentage of covered volume in the network), and  $\delta$  (minimum percentage of covered OD pairs in the network). The objective value highly depends on these parameters' value. To have an approximate range for the parameters, the problem is solved with formulation 1 first. Since formulation 1 does not depend on these three parameters, the output for formulation 1 can be used as an approximate range for the input for formulation 2.

In other problems, when formulation 1 is not available, the parameters' values depend on the experience of the user and the characteristics of the network.

The machine used in solving the problem is a desktop computer with a 3.0 GHz CPU and 2.00 GB of RAM. The optimization software is the CPLEX 10 [17].

4.2. Formulation 1 Numerical Analysis

**4.2.1. Base Case Study**

The only input parameter required for formulation 1 is the available budget C. Total budget is defined as the number of nodes in the network multiplied by the average cost of installing a sensor on a node in the network. In the base case of formulation 1 the available budget is considered as 30% of the total budget (table 4.2).

*Table 4.2. Base case result for formulation 1*

| Network                          | Budget | Solution Time (sec) | Objective Value | Covered Nodes(%) | Covered Links(%) | Covered OD Pairs(%) | Covered Volume(%) | Cost (Used Budget) | Mean of Travel Time Prediction<br>Relative Error on Covered Links | Median of Travel Time Prediction<br>Relative Error on Covered Links | Mean of COV of Travel Time on<br>the Covered Links | Median of COV of Travel Time on<br>the Covered Links |
|----------------------------------|--------|---------------------|-----------------|------------------|------------------|---------------------|-------------------|--------------------|---|---|--|--|
| <i>Sioux-Falls</i>               | 213    | 15.28               | 0.599           | 41.67            | 28.95            | 62.07               | 36.63             | 213                | 0.275   | 0.316   | 0.166  | 0.114  |
| <i>Friedrichshain<br/>Center</i> | 1990   | 9.33                | 1.374           | 33.04            | 27.53            | 94.57               | 43.56             | 1988               | 0.121   | 0.111   | 0.645  | 0.069  |
| <i>Anaheim</i>                   | 3824   | 183.34              | 1.256           | 33.17            | 20.13            | 97.35               | 37.6              | 3824               | 0.231   | 0.256   | 4.294  | 1.238  |

The base case result of formulation 1 for Sioux-Falls and Friedrichshain Center networks are shown in figure 4.2. The coordinates of the nodes in Anaheim network were not available. So the results for Anaheim could not be shown on the map.

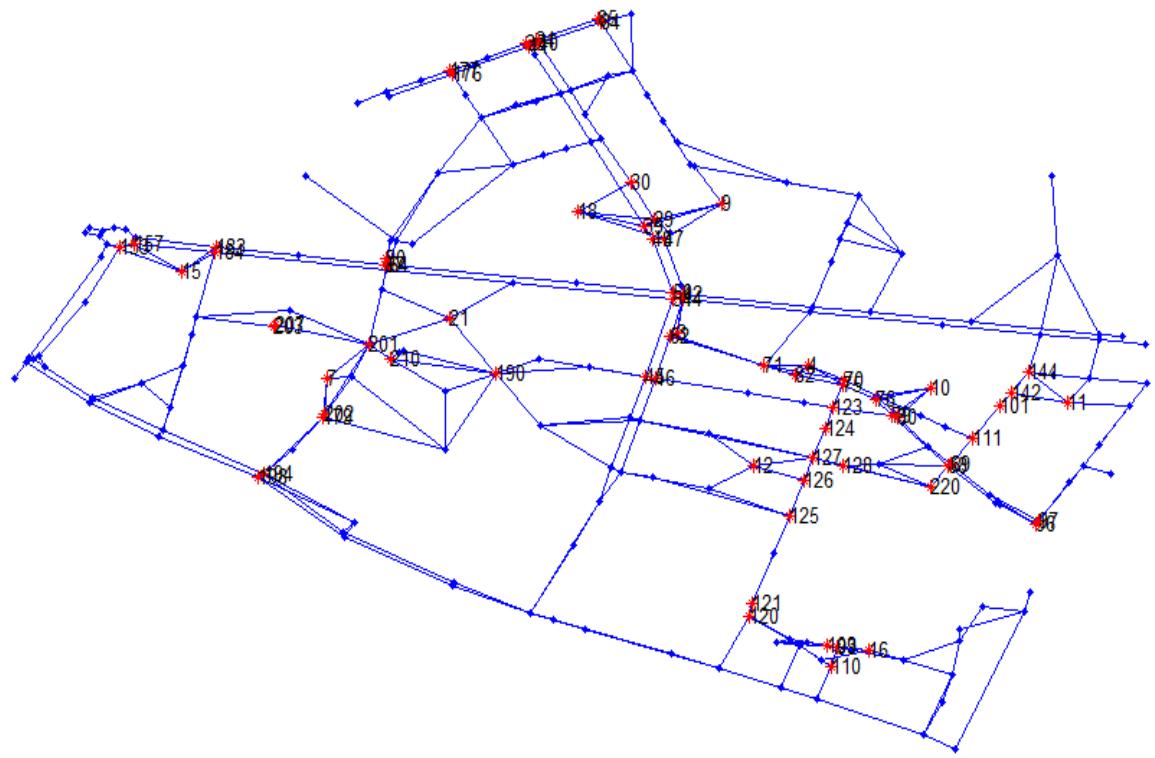
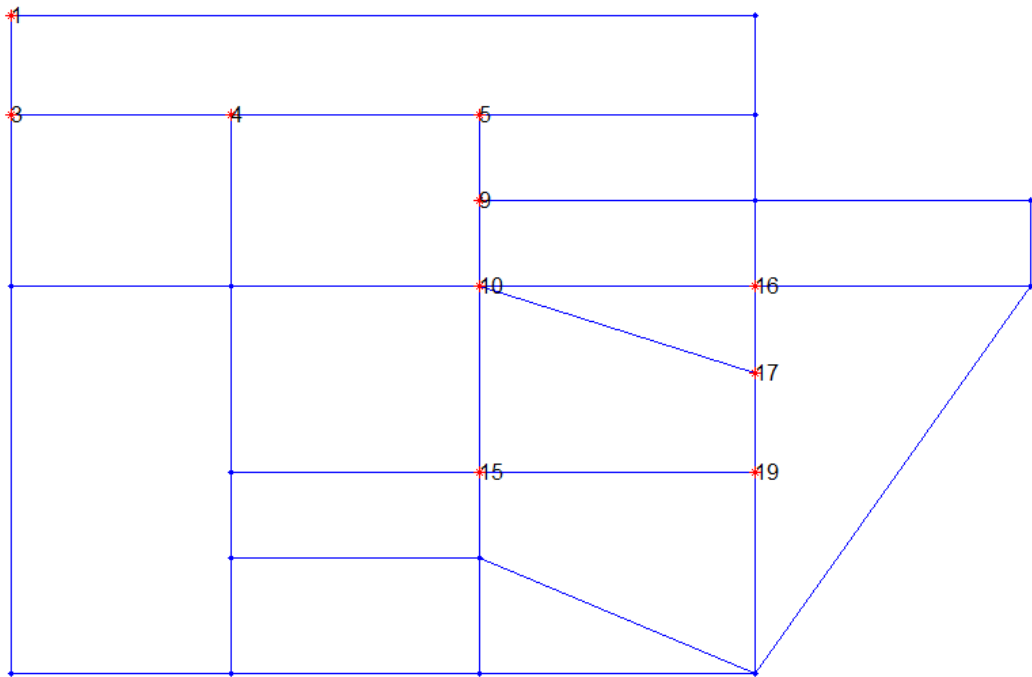


Figure4.2.b. Formulation 1 base case result for Friedrichshain Center network

The result shows that even with a budget less than the total budget, a high percentage of OD pairs can be covered in the networks. This is because most of the links are common between the OD pair links. Also the travel time prediction relative error is reduced compared to the average error in the network. Moreover, the mean of COV of travel time on the covered links is higher than the mean COV of travel time in the network. This implies that the important links which have higher COV in their travel time has been picked up to be covered. Furthermore, from figure 4.2 it is apparent that the chosen nodes are distributed all over the network which ensures that a high percentage of OD pairs are covered. Also figure 4.2 shows that the formulation chooses the links with common nodes to decrease the number of sensors.

The only limiting parameter in formulation 1 is the budget. So changing the budget affects the objective value. Consequently, the sensitivity analysis is done for the budget. The result of sensitivity analysis is discussed in section 4.2.2.

#### **4.2.2. Sensitivity Analysis for the Budget**

Sensitivity analysis is the study of how the variation (uncertainty) in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of a model. The only limiting input for formulation 1 is the budget. So the variation of the output is studied under the variation of the budget.

In the base case study the budget is set to 30% of the total budget for all the networks. In the sensitivity analysis all the input remain the same while the budget varies for different percentages of the total budget. The influence of changing budget on the



objective function and other parameters are discussed in this section. The numerical result for the sensitivity analysis can be found in table 4.3.

Table 4.3.a. Sensitivity analysis of formulation 1 for Sioux-Falls network

| Minimum of COV of Travel Time on the Covered Links             | 0.1776      |
|--|-------------|
| Mean of COV of Travel Time on the Covered Links                | 0.283       |
| Mean of Travel Time Prediction Relative Error on Covered Links | 0.324       |
| Cost (Used Budget)   | 55          |
| Covered Volume (%)   | 9.15        |
| Covered OD Pairs (%)   | 32.41       |
| Covered Links (%)  | 7.89        |
| Covered Nodes (%)  | 16.67       |
| Gap Between LP and IP Solution (%)                             | 0           |
| Gap Between LP and IP Solution                                 | 0           |
| Objective Value  | 0.303       |
| Solution Time (sec)  | 3.63        |
| Budget   | 71          |
| Percentage of Total Cost for Budget                            | 10%         |
| Total Volume   | 877774      |
| Mean of COV of Travel Time                                     | 0.154       |
| Mean of Travel Time Prediction Relative Error                  | 0.282       |
| Number of OD Pairs   | 725         |
| Number of Links  | 76          |
| Number of Nodes  | 24          |
| Network  | Sioux-Falls |
|  | 20%         |
|  | 30%         |
|  | 40%         |
|  | 50%         |
|  | 60%         |
|  | 70%         |
|  | 80%         |
|  | 90%         |

Table 4.3.b. Sensitivity analysis of formulation 1 for Friedrichshain Center network

| Minimum of COV of Travel Time on the Covered Links             | 0.0000                |
|--|-----------------------|
| Mean of COV of Travel Time on the Covered Links                | 0.974                 |
| Mean of Travel Time Prediction Relative Error on Covered Links | 0.115                 |
| Cost (Used Budget)   | 663                   |
| Covered Volume (%)   | 12.94                 |
| Covered OD Pairs (%)   | 78.26                 |
| Covered Links (%)  | 6.88                  |
| Covered Nodes (%)  | 11.61                 |
| Gap Between LP and IP Solution (%)                             | 0.01                  |
| Gap Between LP and IP Solution                                 | 0.000091              |
| Objective Value  | 0.948                 |
| Solution Time (sec)  | 906.41                |
| Budget   | 663                   |
| Percentage of Total Cost for Budget                            | 10%                   |
| Total Volume   | 110057                |
| Mean of COV of Travel Time                                     | 0.330                 |
| Mean of Travel Time Prediction Relative Error                  | 0.155                 |
| Number of OD Pairs   | 552                   |
| Number of Links  | 523                   |
| Number of Nodes  | 224                   |
| Network  | Friedrichshain Center |
|  | 20%                   |
|  | 30%                   |
|  | 40%                   |
|  | 50%                   |
|  | 60%                   |
|  | 70%                   |
|  | 80%                   |
|  | 90%                   |

Table 4.3.c. Sensitivity analysis of formulation 1 for Anaheim network

| Minimum of COV of Travel Time on the Covered Links | Mean of COV of Travel Time on the Covered Links | Mean of Travel Time Prediction Relative Error on Covered Links | Cost (Used Budget) | Covered Volume (%) | Covered OD Pairs (%) | Covered Links (%) | Covered Nodes (%) | Gap Between LP and IP Solution (%) | Gap Between LP and IP Solution | Objective Value | Solution Time (sec) | Budget | Percentage of Total Cost for Budget | Total Volume | Mean of COV of Travel Time | Mean of Travel Time Prediction Relative Error | Number of OD Pairs | Number of Links | Number of Nodes | Network |         |         |         |         |         |         |
|--|---|--|--------------------|--------------------|----------------------|-------------------|-------------------|------------------------------------|--------------------------------|-----------------|---------------------|--------|-------------------------------------|--------------|----------------------------|---|--------------------|-----------------|-----------------|---------|---------|---------|---------|---------|---------|---------|
| 0.0275   | 5.107   | 0.236  | 2546               | 18.09              | 96.84                | 14.44             | 22.60             | 0.01                               | 0.000115                       | 1.149           | 20641.56            | 2549   | 20%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             | Anaheim |         |         |         |         |         |         |
| 0.0032   | 4.294   | 0.231  | 3824               | 37.60              | 97.35                | 20.13             | 33.17             | 0.01                               | 0.000125                       | 1.256           | 183.34              | 3824   | 30%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         | Anaheim |         |         |         |         |         |
| 0.0032   | 3.974   | 0.235  | 5098               | 50.94              | 98.17                | 25.82             | 42.07             | 0                                  | 0.000032                       | 1.332           | 72.44               | 5098   | 40%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         | Anaheim |         |         |         |         |
| 0.0032   | 3.658   | 0.233  | 6372               | 63.45              | 98.67                | 32.06             | 51.20             | 0.01                               | 0.000138                       | 1.386           | 38.55               | 6373   | 50%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         |         | Anaheim |         |         |         |
| 0.0032   | 3.336   | 0.236  | 7642               | 75.77              | 98.86                | 38.29             | 59.38             | 0.01                               | 0.000129                       | 1.419           | 16.09               | 7648   | 60%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         |         |         | Anaheim |         |         |
| 0.0032   | 3.217   | 0.239  | 8479               | 82.38              | 99.05                | 43.65             | 66.11             | 0.01                               | 0.000139                       | 1.426           | 2.28                | 8922   | 70%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         |         |         |         | Anaheim |         |
| 0.0032   | 3.217   | 0.239  | 8479               | 82.38              | 98.99                | 43.65             | 66.11             | 0.01                               | 0.000139                       | 1.426           | 2.17                | 10197  | 80%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         |         |         |         |         | Anaheim |
| 0.0032   | 3.217   | 0.239  | 8479               | 82.38              | 99.05                | 43.65             | 66.11             | 0.01                               | 0.000139                       | 1.426           | 2.20                | 11472  | 90%                                 | 1837119      | 2.039                      | 0.251   | 1584               | 914             | 416             |         |         |         |         |         |         |         |

Changes in percentages of covered OD pairs in the network while the budget varies is shown in figure 4.3.

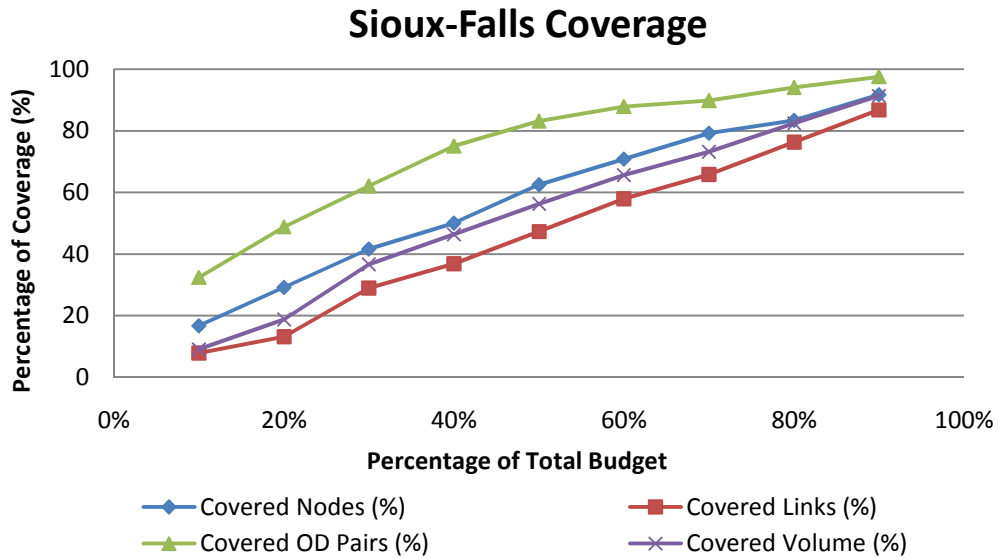


Figure 4.3.a. Percentage of network coverage versus budget for Sioux-Falls network – Formulation 1

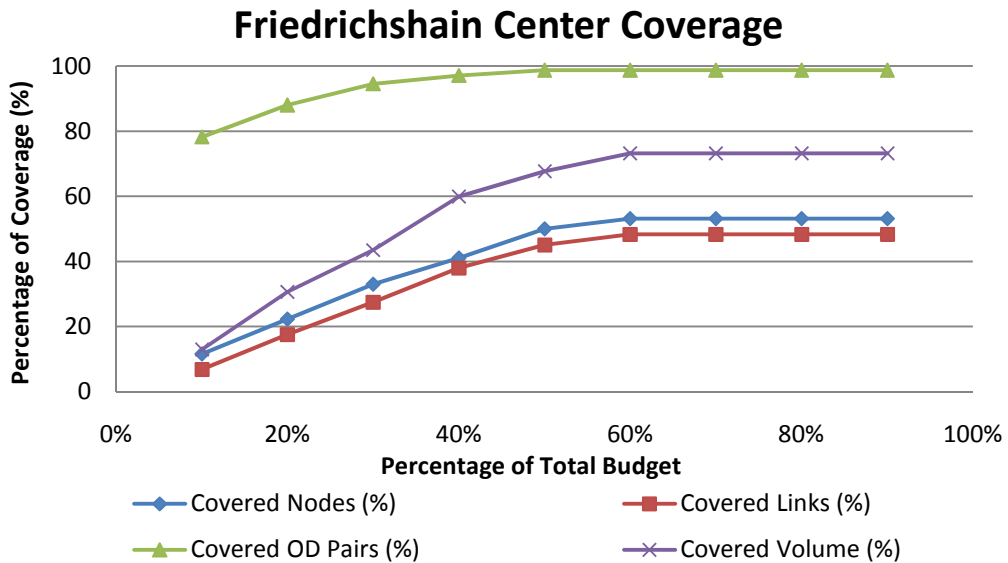


Figure 4.3.b. Percentage of network coverage versus budget for Friedrichshain Center network – Formulation 1

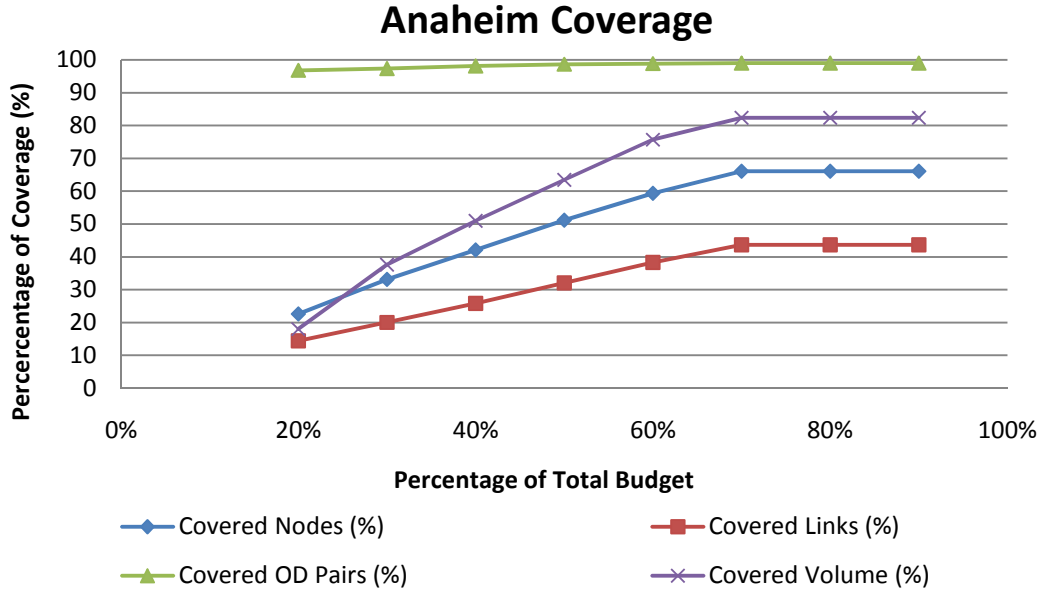


Figure 4.3.c. Percentage of covered OD pairs, links, nodes, and volume versus budget for Anaheim network - Formulation 1

Figure 4.3 reveals some of the characteristics of the networks. In Sioux-Falls network the number of common links between OD pair links is small. The optimal solution selects more links when the budget increases in order to increase the objective elements of OD pairs and volume. On the other hand the number of common links between OD pair links in the Anaheim and Friedrichshain Center networks is large. So even with a small amount of budget a high percentage of OD pairs can be covered. Increasing the number of links increases the covered volume in the network which results in a higher objective function value.

For covering each link both end nodes must be covered. If two links have a common node the number of sensors for covering the links decreases by the number of common nodes. Since the rate of increase in the node covering graph is less than

twice the links covering graph, it can be concluded that the program picks up the links with common nodes. This reduces the cost and increases the objective value.

The mean of COV of travel time and mean of travel time prediction relative error on the covered links versus budget are shown in figures 4.4 and 4.5 respectively.

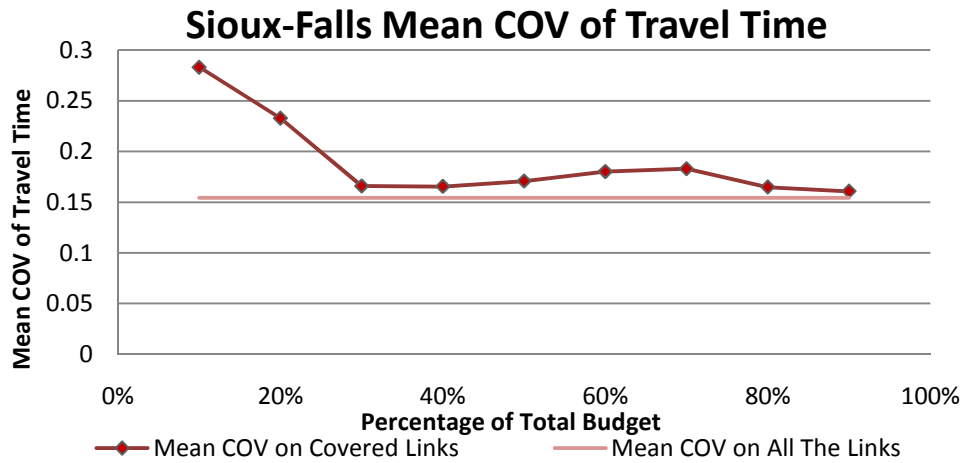


Figure 4.4.a. Mean of COV of travel time on the covered links versus budget for Sioux-Falls network – Formulation 1

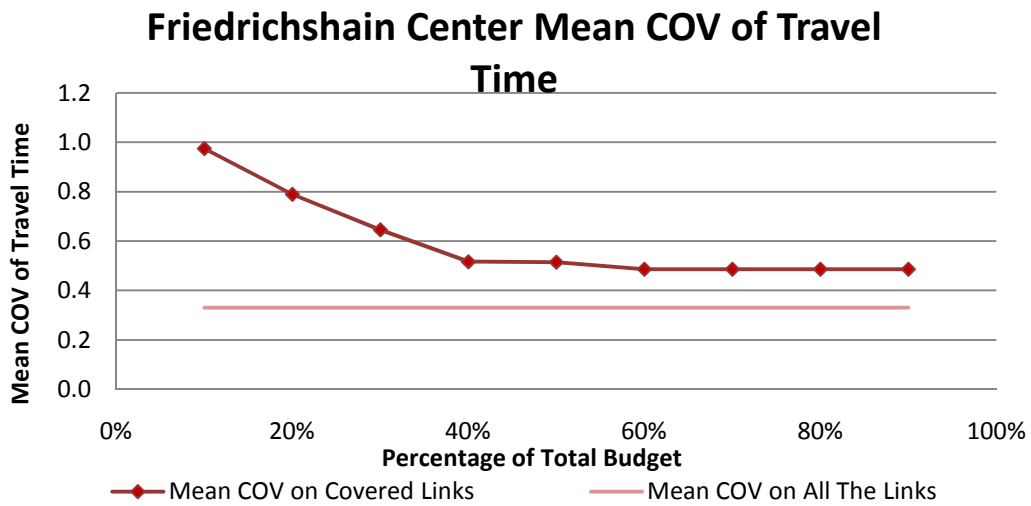
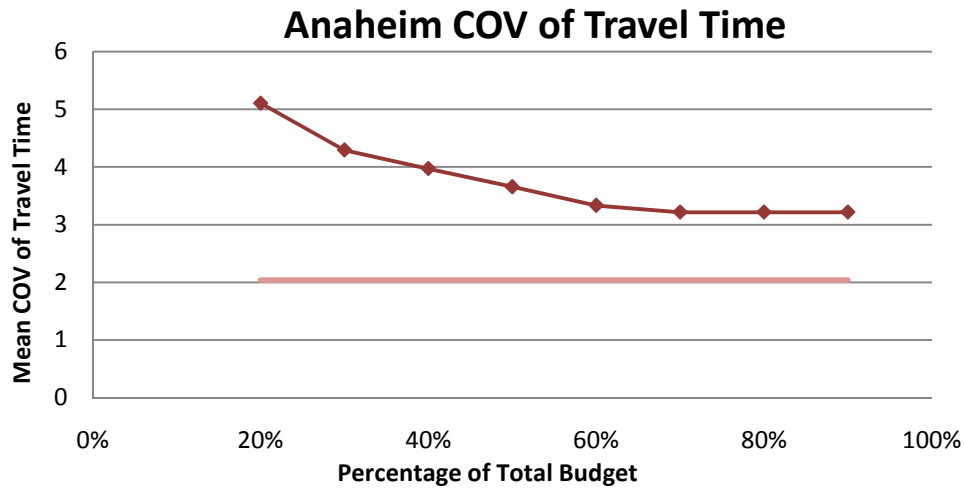


Figure 4.4.b. Mean of COV of travel time on the covered links versus budget for Friedrichshain Center network – Formulation 1



*Figure 4.4.c. Mean of COV of travel time on the covered links versus budget for Anaheim network – Formulation 1*

The optimal solution is interested in selecting the links with higher COV in their travel time. While the budget increases the optimal solution has already selected the links with the higher COV, so by choosing other links the mean of COV decreases and gets closer to the mean of the COV in the network. But since not all the links are chosen, as it is shown in the figure 4.4 the mean of COV on the covered links is always higher than the mean COV of travel time in the network. So by increasing the budget the contribution of COV element to the objective function decreases.



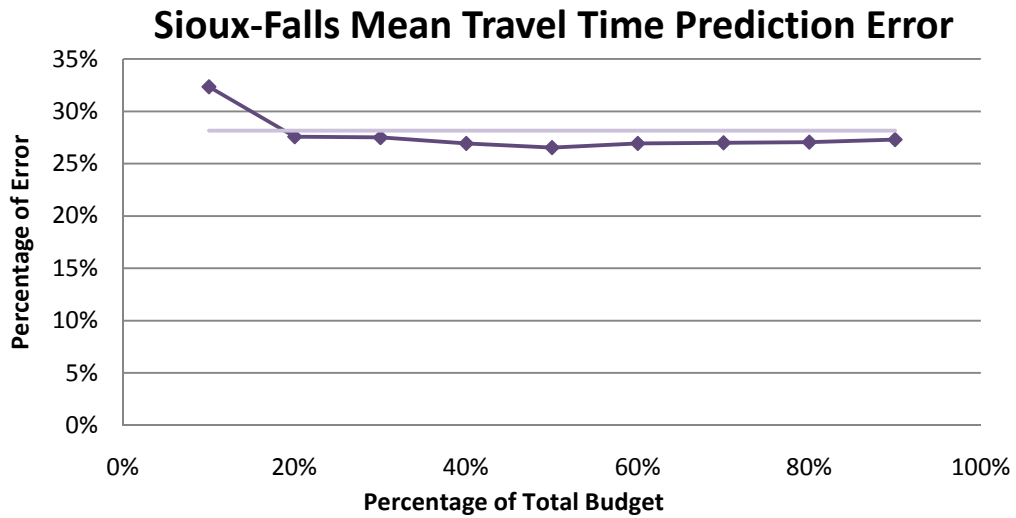


Figure 4.5.a. Mean of the travel time prediction relative error on the covered links versus budget for Sioux-Falls network – Formulation 1

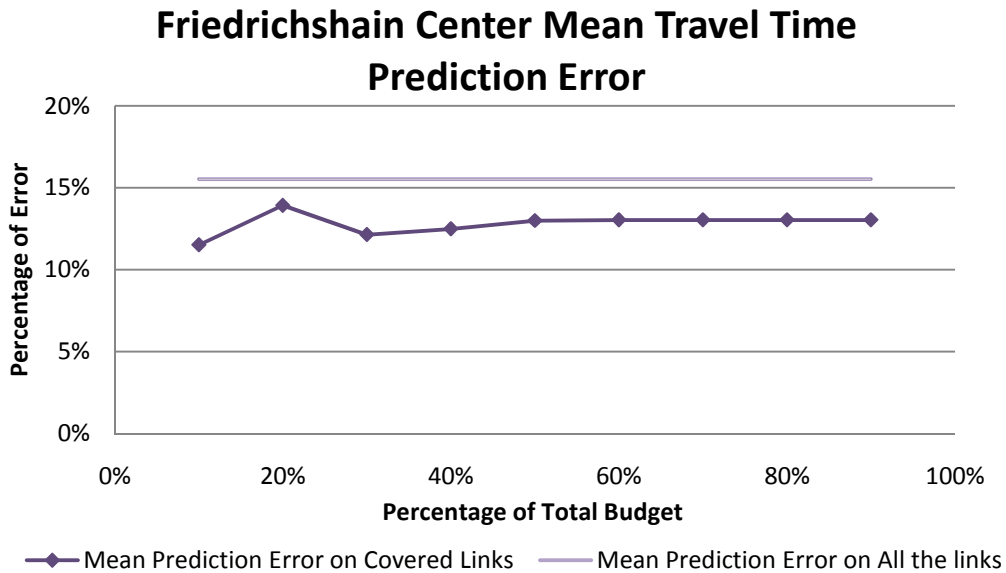
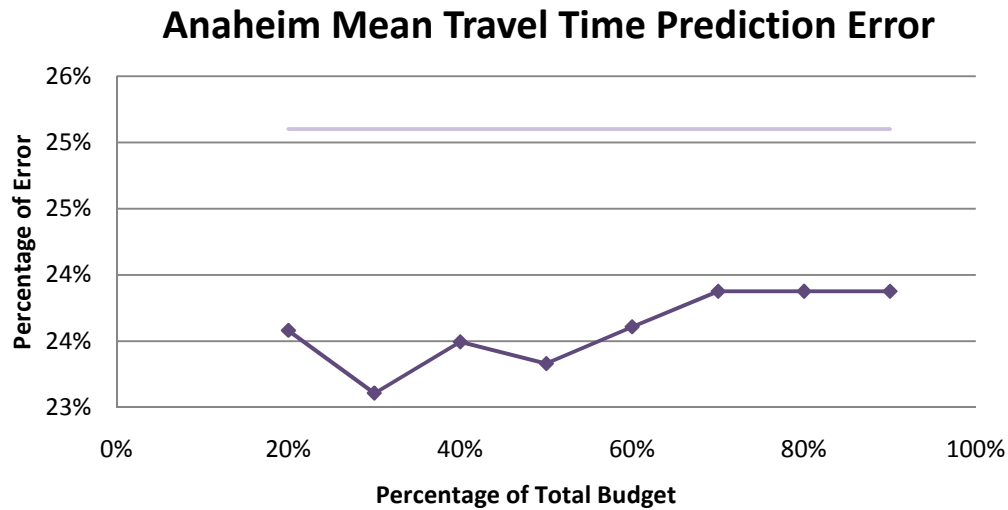


Figure 4.5.b. Mean of the travel time prediction relative error on the covered links versus budget for Friedrichshain Center network – Formulation 1



*Figure 4.5.c. Mean of the travel time prediction relative error on the covered links versus budget for Anaheim network – Formulation 1*

Increase of the budget increases the covered nodes and links. Since the program chooses the links with smaller error in the beginning, as the number of links increases the average of the error on the links gets closer to the mean of error in the network. So the increase of the budget decreases the contribution of the error element to the objective function.

COV of travel time, error and the number of sensors are the parameters whose contribution to the objective function decreases when the budget increases. On the other hand the volume and the OD pair covering are the elements whose contribution to the objective function increases when the budget increases. So the optimal solution for the problem is where the increasing and the decreasing functions meet while the budget changes.

Figure 4.6 shows the changes in objective function value while the budget varies. By increasing the budget the program chooses as many nodes as possible until the

contribution of the volume and OD pair elements to the objective function do not justify the decrease in the contribution of COV of travel time, error and the number of sensors elements. It means after a certain percentage of the total budget, increasing of the budget does not increase the objective value.

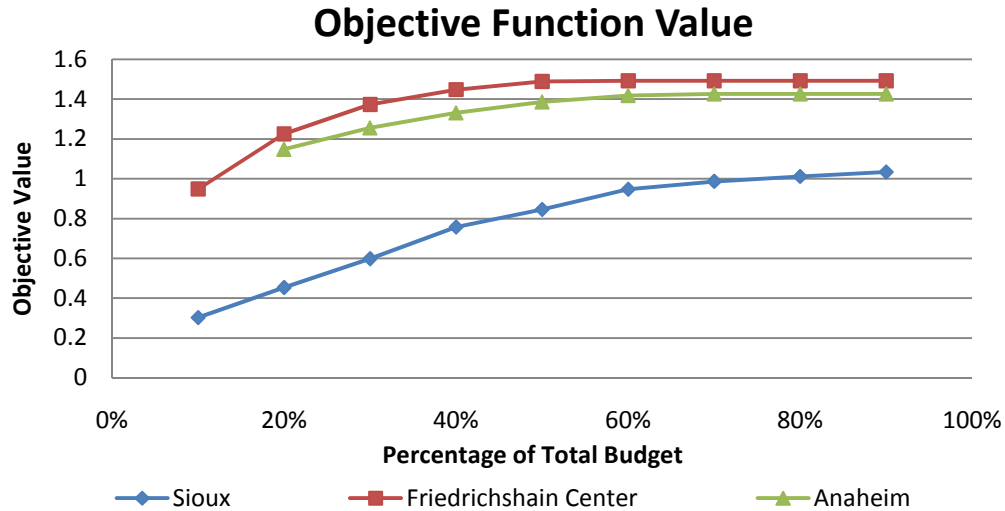


Figure 4.6. Objective function value versus the budget – Formulation 1

Figure 4.7 shows how the cost (used budget) changes while the available budget increases. As it is shown in the figure 4.7 after a certain percentage of the total budget is used, even if the available budget increases the cost does not change since using the budget does not increase the objective value. This percentage of the total budget represents the budget which results in the maximum objective value. Consequently allocating a higher amount of budget is a waste of funds.

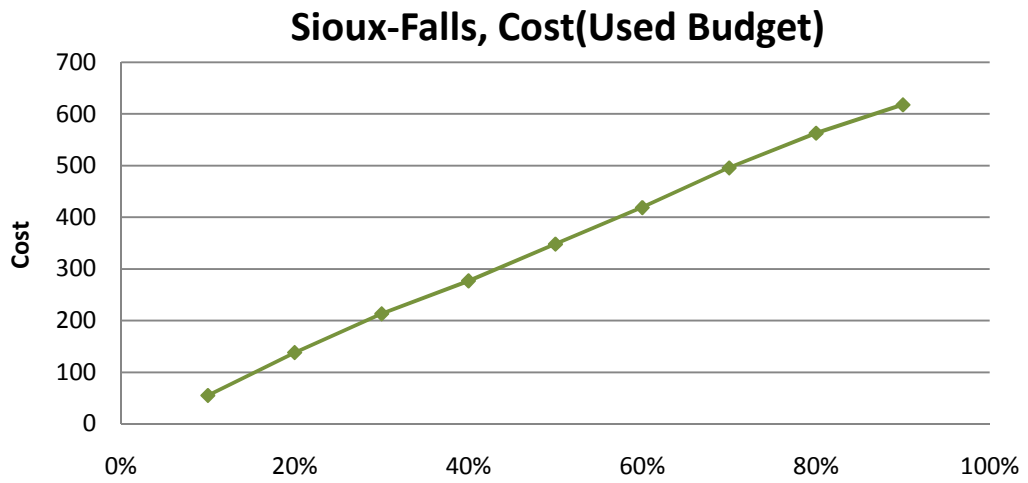


Figure 4.7.a. Cost versus budget for Sioux-Falls network – Formulation 1

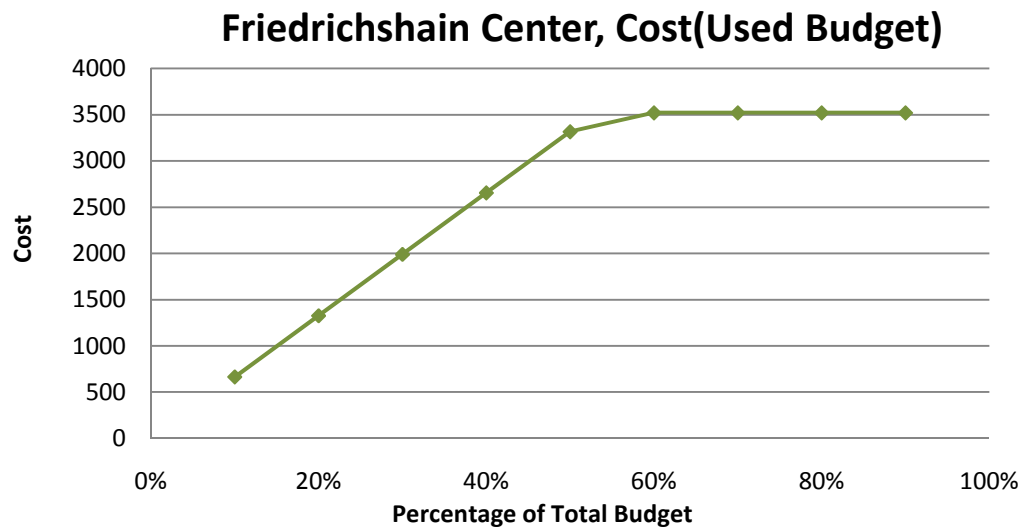


Figure 4.7.b. Cost versus budget for Friedrichshain Center network – Formulation 1

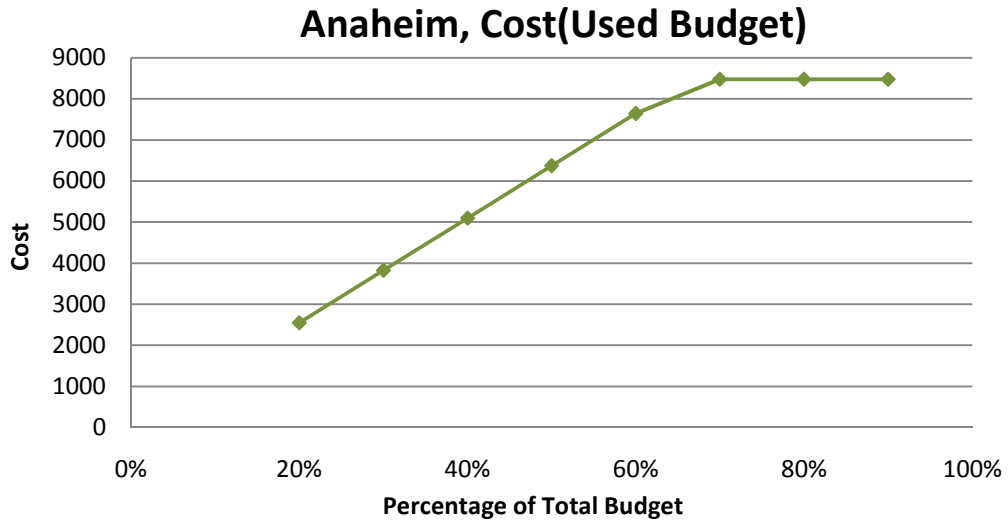


Figure 4.7.c. Cost versus budget for Anaheim network – Formulation 1

Since the budget is the only limiting constraint for the formulation 1, increasing budget is the same as relaxing the constraint. Therefore the solution time decreases when the budget increases (Figure 4.8).

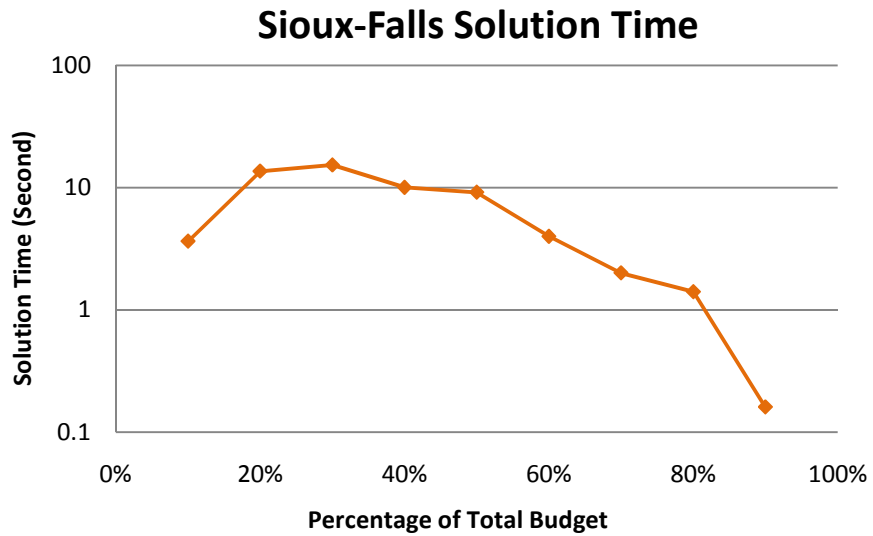


Figure 4.8.a. Solution time versus budget for Sioux-Falls network – Formulation 1

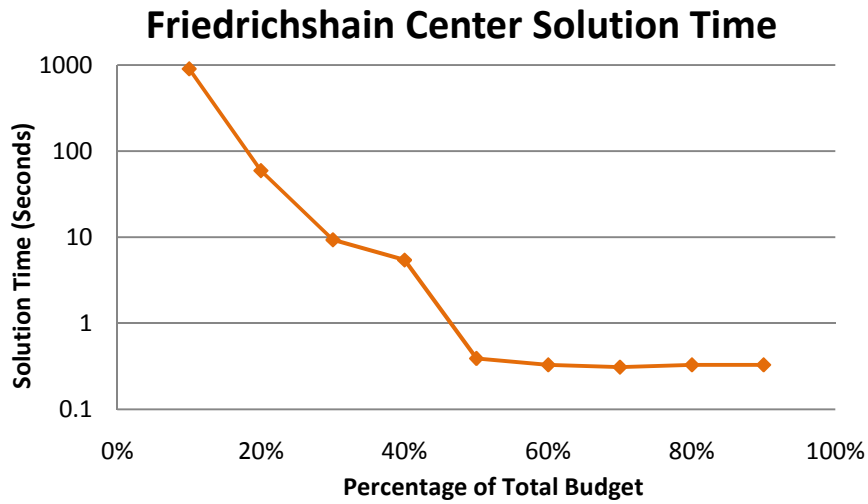


Figure 4.8.b. Solution time versus budget for Friedrichshain Center network – Formulation 1

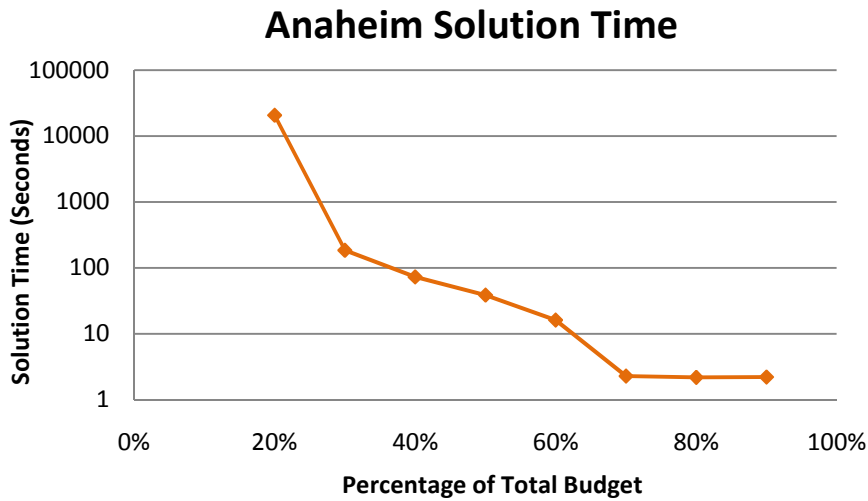


Figure 4.8. Solution time versus budget for Anaheim network – Formulation 1

Sioux-Falls network is a small network with 24 nodes. As the budget decreases it gets very close to the cost of installing one sensor. So many of the nodes which have high cost are eliminated because of the cost constraint. Consequently solution time is low when the budget is very small.

### 4.3. Formulation 2 Numerical Analysis

#### **4.3.1. Base Case Study**

Three parameters of  $\alpha$  (minimum COV of travel time on the links),  $\gamma$  (minimum percentage of covered volume in the network), and  $\delta$  (minimum percentage of covered OD pairs in the network) should be known as input for formulation 2. To get an approximate range for the parameters, the output of the formulation 1 is used. The critical budget for formulation 1 is used as the base case budget for formulation 2. Critical budget is the budget at which the network coverage and objective value increase rate starts decreasing while the budget is increasing. So it is the budget which results in a good coverage of the network while it is economical. The graphs (figures 4.3 and 4.6) show that 30% of the total budget for Sioux-Falls network and 40% of the total budget for the Anaheim and the Friedrichshain Center networks are the critical budget for each network. Since after these values the rate of the increase in the objective function starts reducing.

The input for the base case of the formulation 2 for each network is shown in table 4.3.

*Table4.4.Input parameters for formulation 2*

| <b>Network</b>               | <b><math>\alpha</math></b> | <b><math>\gamma</math></b> | <b><math>\delta</math></b> | <b>Percentage of Total Budget</b> |
|------------------------------|----------------------------|----------------------------|----------------------------|-----------------------------------|
| <i>Sioux-Falls</i>           | 0.1                        | 0.2                        | 0.6                        | 30%                               |
| <i>Friedrichshain Center</i> | 0.05                       | 0.2                        | 0.4                        | 40%                               |
| <i>Anaheim</i>               | 0.003                      | 0.5                        | 0.98                       | 40%                               |

The numerical results for the base case study of the formulation 2 on the networks are shown in table 4.4.

Table 4.5. Base case result for formulation 2

| Network                      | Budget | Solution Time (sec) | Objective Value | Gap Between LP and IP Solution | Covered Nodes (%) | Covered Links (%) | Covered OD Pairs (%) | Covered Volume (%) | Cost (Used Budget) | Mean of Travel Time Prediction Relative Error on Covered Links | Median of Travel Time Prediction Relative Error on Covered Links | Mean of COV of Travel Time on The Covered Links | Median of COV of Travel Time on The Covered Links |
|------------------------------|--------|---------------------|-----------------|--------------------------------|-------------------|-------------------|----------------------|--------------------|--------------------|--|--|---|---|
| <i>Sioux-Falls</i>           | 284    | 4.94                | 8               | 0                              | 33.33             | 18.42             | 44.55                | 22.15              | 267                | 0.210  | 0.316  | 0.199   | 0.114   |
| <i>Friedrichshain Center</i> | 1990   | 0.41                | 35              | 0                              | 15.63             | 6.69              | 71.56                | 20.15              | 1064               | 0.153  | 0.111  | 0.660   | 0.069   |
| <i>Anaheim</i>               | 5098   | 93028.52            | 160             | 8                              | 38.46             | 22.98             | 98.23                | 50.95              | 4975               | 0.250  | 0.256  | 2.938   | 1.238   |

The base case result of formulation 2 for Sioux-Falls and Friedrichshain Center networks are shown in figure 4.9.

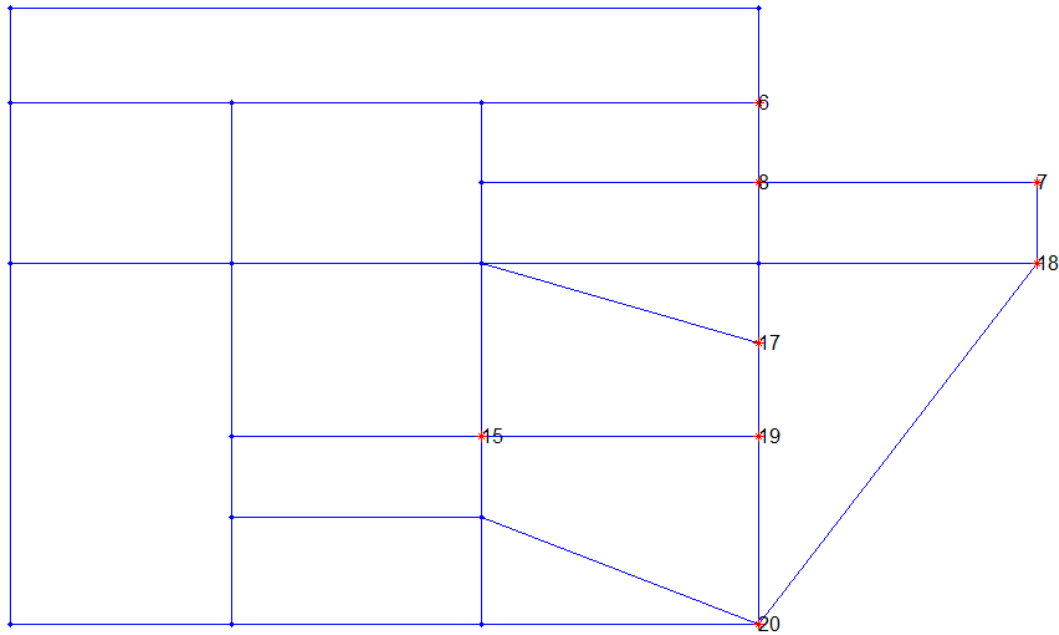
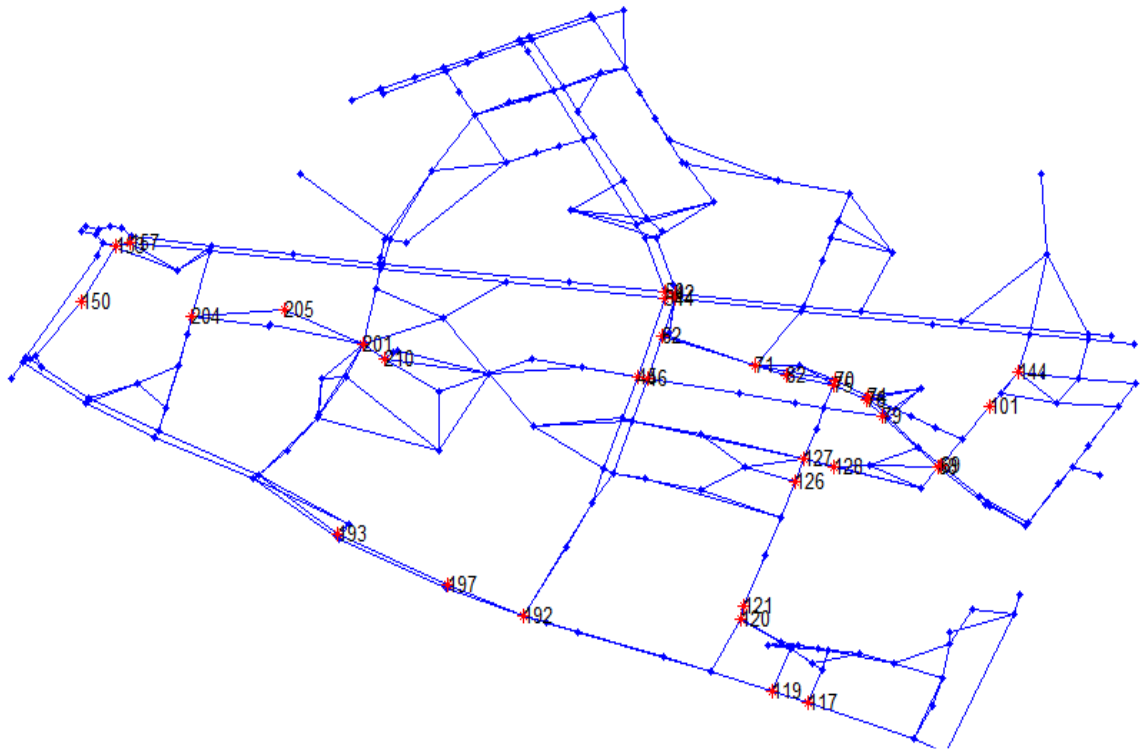


Figure 4.9.a. Formulation 2 base case result for Sioux-Falls network





*Figure 4.9.b. Formulation 2 base case result for Friedrichshain Center network*

Figure 4.9 shows the result for base case study of formulation 2. It is apparent that the formulation 2 chooses the links with common nodes in order to reduce the number of sensors. The nodes are chosen so that not all the regions of the network are covered. That is because the problem does not consider covering all the regions but the most important links. This is not a limitation of the model. To cover at least some links from all the regions of the network a new constraint can simply be added. The complete numerical results for the base case studies and all the sensitivity analysis for the formulation 2 are included in the table 4.6. Since the parameters of  $\alpha$ ,  $\gamma$ , and  $\delta$  are not exactly the same as formulation 1, the base cases in formulation 2 are not comparable to formulation 1. But the exact cases are computed and compared with each other in chapter 5.

Table 4.6.a. Sensitivity analysis of formulation 2 for Friedrichshain Center network

| Parameter                                     | Parameter Value       | Solution Time (sec) | Objective Value | Gap Between LP and IP Solution | Gap Between LP and IP Solution (%) | Covered Nodes (%) | Covered Links (%) | Covered OD Pairs (%) | Covered Volume (%) | Cost (Used Budget) | Mean of Travel Time Prediction Relative Error on the Covered Links | Mean of COV of Travel Time on the Covered Links | Minimum of COV of Travel Time on the Covered Links |       |       |
|---|-----------------------|---------------------|-----------------|--------------------------------|------------------------------------|-------------------|-------------------|----------------------|--------------------|--------------------|--|---|--|-------|-------|
| Budget  | 1990                  | $\alpha$            | 0               | 46.05                          | 21                                 | 0                 | 0                 | 9.38                 | 7.27               | 63.41              | 21.1   | 609   | 0.145  | 0.204 | 0.000 |
|   |                       |                     | 0.05            | 0.61                           | 29                                 | 0                 | 0                 | 12.95                | 6.5                | 70.47              | 20   | 841   | 0.155  | 0.608 | 0.060 |
|   |                       |                     | 0.1             | 0.41                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.12            | 1.98                           | 64                                 | 0                 | 0                 | 28.57                | 9.94               | 79.71              | 20.03  | 1885  | 0.154  | 0.674 | 0.123 |
|   |                       | $\delta$            | 0               | 0.84                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.1             | 0.84                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.2             | 3.28                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.3             | 0.44                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.4             | 0.64                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.5             | 0.42                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.6             | 0.41                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 71.56              | 20.15  | 1064  | 0.153  | 0.660 | 0.101 |
|   |                       |                     | 0.7             | 3.06                           | 35                                 | 0                 | 0                 | 15.63                | 6.69               | 70.29              | 20.06  | 1058  | 0.155  | 0.656 | 0.101 |
| 0.8   | 3.05                  |                     | 37              | 0                              | 0                                  | 16.52             | 6.69              | 80.07                | 20.21              | 1060               | 0.153  | 0.661   | 0.100  |       |       |
| 0.9   | 2.08                  | 63                  | 0               | 0                              | 28.13                              | 10.33             | 90.04             | 20.01                | 1853               | 0.155              | 0.688  | 0.100   |  |       |       |
| $\gamma$                                      | 0                     | 7.16                | 11              | 0                              | 0                                  | 4.91              | 1.72              | 61.23                | 5.91               | 361                | 0.143  | 0.685   | 0.101  |       |       |
|   | 0.1                   | 7.17                | 14              | 0                              | 0                                  | 6.25              | 2.87              | 60.51                | 10.89              | 423                | 0.148  | 0.749   | 0.101  |       |       |
|   | 0.2                   | 0.41                | 35              | 0                              | 0                                  | 15.63             | 6.69              | 71.56                | 20.15              | 1064               | 0.153  | 0.660   | 0.101  |       |       |
|   | 0.25                  | 2.98                | 57              | 0                              | 0                                  | 25.45             | 10.52             | 80.07                | 25.04              | 1670               | 0.155  | 0.550   | 0.100  |       |       |
| Total Volume                                  | 110057                |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Mean of COV of Travel Time                    | 0.330                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Mean of Travel Time Prediction Relative Error | 0.155                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Number of OD Pairs                            | 552                   |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Number of Links                               | 523                   |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Number of Nodes                               | 224                   |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |
| Network                                       | Friedrichshain Center |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   |  |       |       |

Table 4.6.b. Sensitivity analysis of formulation 2 for Sioux-Falls network

| Parameter           | Parameter Value | Solution Time (sec) | Objective Value | Gap Between LP and IP Solution | Gap Between LP and IP Solution (%) | Covered Nodes (%) | Covered Links (%) | Covered OD Pairs (%) | Covered Volume (%) | Cost (Used Budget) | Mean of Travel Time Prediction Relative Error on the Covered Links | Mean of COV of Travel Time on the Covered Links | Minimum of COV of Travel Time on the Covered Links |      |    |   |   |       |       |       |       |     |       |       |       |
|---------------------|-----------------|---------------------|-----------------|--------------------------------|------------------------------------|-------------------|-------------------|----------------------|--------------------|--------------------|--|---|--|------|----|---|---|-------|-------|-------|-------|-----|-------|-------|-------|
| Sioux-Falls Network | Budget          | 284                 | 877774          | 0.154                          | 0.282                              | 725               | 76                | 24                   | 284                | 284                | 0.232  | 0.145   | 0.023  |      |    |   |   |       |       |       |       |     |       |       |       |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.023  |      |    |   |   |       |       |       |       |     |       |       |       |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.03   | 8.73 | 7  | 0 | 0 | 29.17 | 15.79 | 44    | 20.85 | 251 | 0.261 | 0.162 | 0.037 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.05   | 4.94 | 8  | 0 | 0 | 33.33 | 18.42 | 44.55 | 22.15 | 267 | 0.210 | 0.199 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.07   | 4.88 | 8  | 0 | 0 | 33.33 | 17.11 | 41.93 | 20.55 | 279 | 0.232 | 0.202 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.09   | 2.72 | 8  | 0 | 0 | 33.33 | 15.79 | 46.07 | 20.12 | 190 | 0.277 | 0.211 | 0.091 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0  | 0.28 | 7  | 0 | 0 | 29.17 | 18.42 | 36.14 | 22.7  | 224 | 0.221 | 0.207 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.1  | 0.39 | 7  | 0 | 0 | 29.17 | 18.42 | 36.14 | 22.7  | 224 | 0.221 | 0.207 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.2  | 0.28 | 7  | 0 | 0 | 29.17 | 18.42 | 36.14 | 22.7  | 224 | 0.221 | 0.207 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.3  | 0.41 | 7  | 0 | 0 | 29.17 | 18.42 | 36.14 | 22.7  | 224 | 0.221 | 0.207 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.4  | 4.94 | 8  | 0 | 0 | 33.33 | 18.42 | 44.55 | 22.15 | 267 | 0.210 | 0.199 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.5  | 5.34 | 8  | 0 | 0 | 33.33 | 15.79 | 51.72 | 20.66 | 224 | 0.227 | 0.253 | 0.079 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.6  | 8.88 | 9  | 0 | 0 | 37.5  | 18.42 | 60.69 | 23.43 | 227 | 0.260 | 0.252 | 0.079 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0  | 7.61 | 5  | 0 | 0 | 20.83 | 7.89  | 40    | 11.22 | 155 | 0.275 | 0.323 | 0.140 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.05   | 7.58 | 5  | 0 | 0 | 20.83 | 10.53 | 40.14 | 11.71 | 153 | 0.275 | 0.186 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.1  | 7.25 | 5  | 0 | 0 | 20.83 | 10.53 | 40.14 | 11.71 | 153 | 0.275 | 0.186 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.15   | 6.42 | 6  | 0 | 0 | 25    | 13.16 | 44    | 15.33 | 198 | 0.270 | 0.226 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.2  | 4.94 | 8  | 0 | 0 | 33.33 | 18.42 | 44.55 | 22.15 | 267 | 0.210 | 0.199 | 0.078 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.25   | 2.16 | 9  | 0 | 0 | 37.5  | 23.68 | 41.66 | 26.48 | 271 | 0.241 | 0.194 | 0.077 |
|                     |                 |                     |                 |                                |                                    |                   |                   |                      |                    |                    |  |   | 0.3  | 0.45 | 10 | 0 | 0 | 41.67 | 23.68 | 57.52 | 30    | 269 | 0.250 | 0.225 | 0.077 |

As the networks get larger the solution time for formulation 2 increases exponentially. The Anaheim network could not be solved exactly with formulation 2. Although the network is solved with a 5% gap between the linear optimal solution and the best integer solution, the solution time is very high. Restricting each parameter increases the solution time. So the sensitivity analysis could not be conducted for the Anaheim network for exact solution. But the sensitivity analysis for Sioux-Falls and Friedrichshain Center networks are discussed in sections 4.3.2, 4.3.3, and 4.3.4.

#### **4.3.2. Sensitivity Analysis for $\alpha$**

In formulation 2, the links that have a COV of travel time less than  $\alpha$  are not allowed to be selected. As  $\alpha$  increases some of the links with small COV of travel time which have common nodes with the other links cannot be chosen. So to maintain the minimum levels of volume and OD pair covering other links are chosen (figure 4.11). To cover the new links new sensors should be installed. As a result the number of sensors, which is the objective function, increases. The variation of the objective function versus  $\alpha$  is shown in figure 4.10.

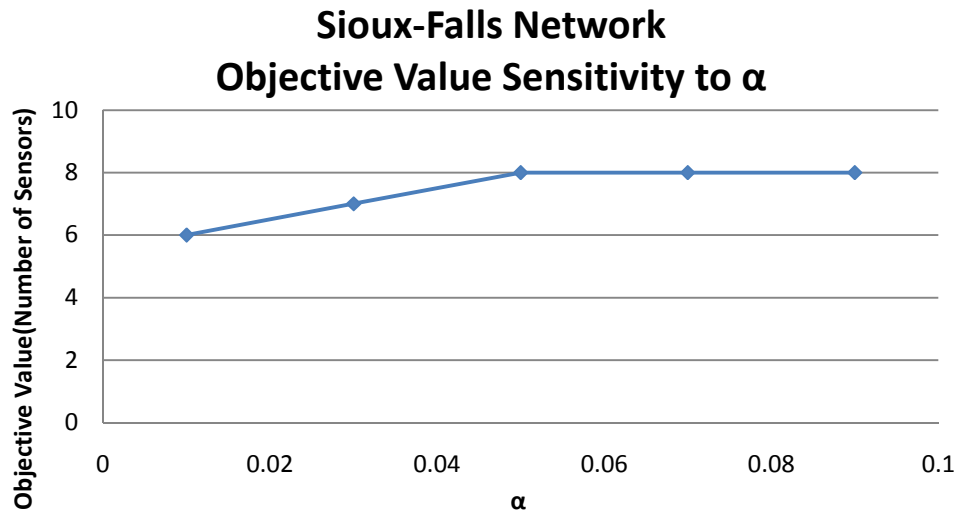


Figure 4.10.a. Objective function value versus  $\alpha$  for Sioux-Falls network – Formulation 2

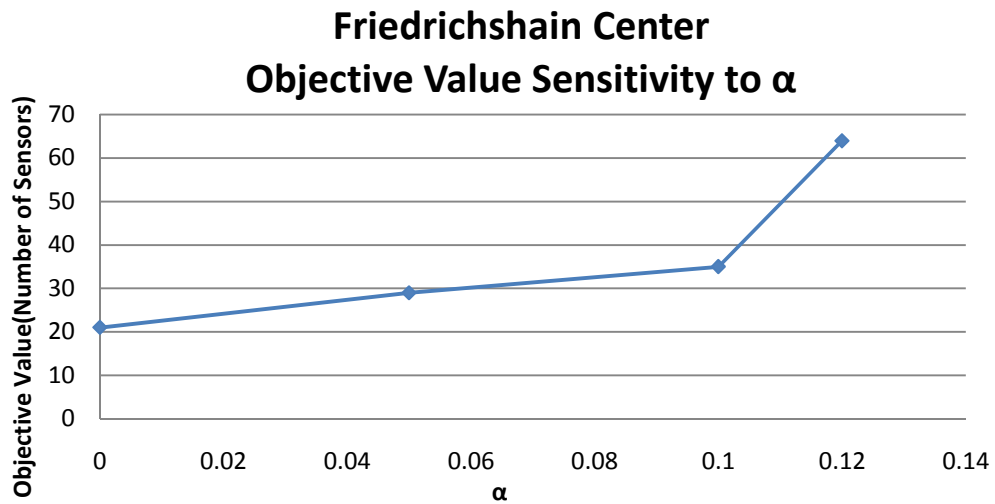


Figure 4.10.b. Objective function value versus  $\alpha$  for Friedrichshain Center network – Formulation 2

Studying the variation of the percentage of covered OD pairs, nodes, links and volume while  $\alpha$  changes, shows the weakness of the formulation 2 (figure 4.11).

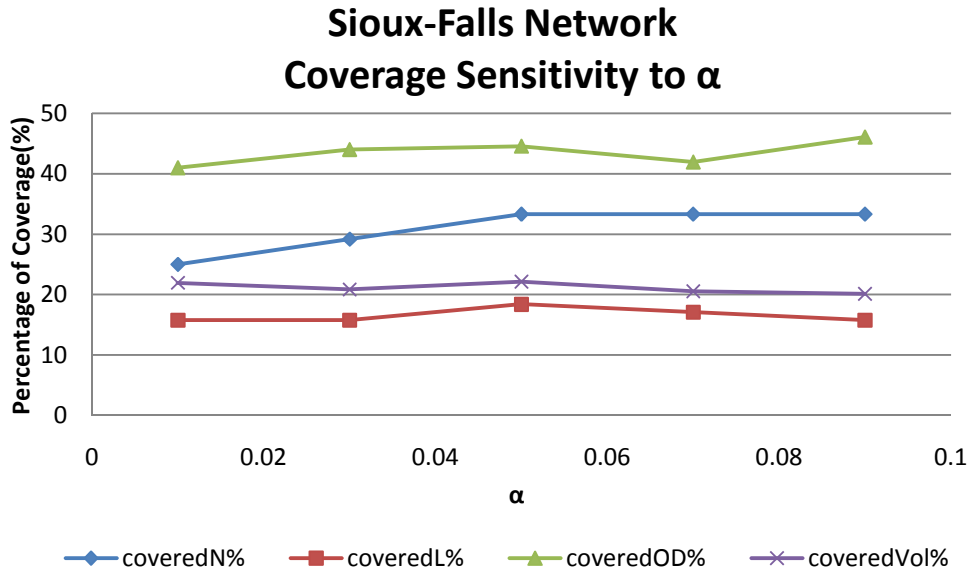


Figure 4.11.a. Percentage of network coverage versus  $\alpha$  for Sioux-Falls network – Formulation 2

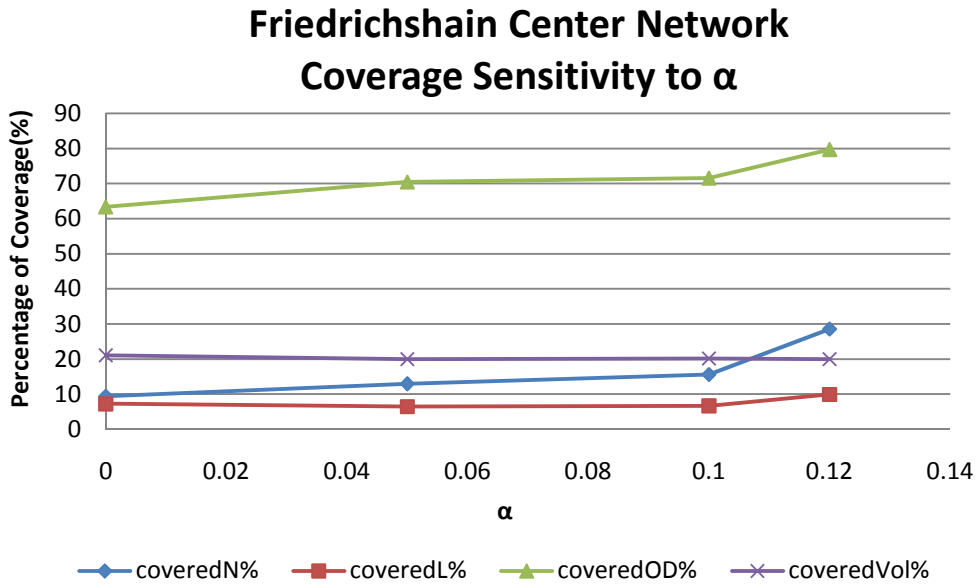


Figure 4.11.b. Percentage of network coverage versus  $\alpha$  for Friedrichshain Center network – Formulation 2

In Sioux-Falls (figure 4.11.a), the solution which is optimal for  $\alpha=0.09$  is feasible for  $\alpha=0.07$  as well. Since the number of nodes and the percentage of covered volume are

the same in both solutions, the optimal solution for  $\alpha=0.09$  is better than the optimal solution for  $\alpha=0.07$  in covering OD pairs. So the weakness of the formulation 2 is that although the problem is solved for the optimal solution, but the solution is not unique. So there may be some other optimal solutions which have a better covering result for other issues in the problem. In other words, the formulation just optimizes the number of sensors not the other parameters. So the optimal solution just satisfies the minimum of the constraints and does not optimize them. As a result the other parameters as is shown in the graphs do not follow a predictable behaviour.

#### **4.3.3. Sensitivity Analysis for $\gamma$**

When  $\gamma$  increases, it implies that a larger percentage of volume in the network should be covered. To cover more volume in the network the program can either cover more links or cover links with higher volume. If any link with a higher volume is found which can be substituted by the other links without increasing the number of sensors, it will be chosen (figure 4.13.a). Otherwise the number of links increases (figure 4.13.b). The changes in the objective function and the percentages of network covering can be seen in figure 4.12 and 4.13 respectively.

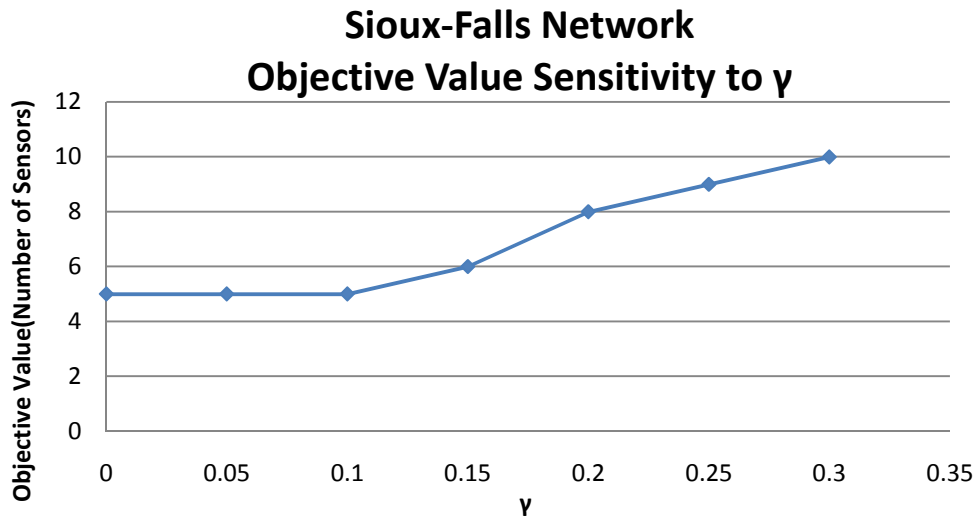


Figure 4.12.a. Objective function value versus  $\gamma$  for Sioux-Falls network – Formulation 2

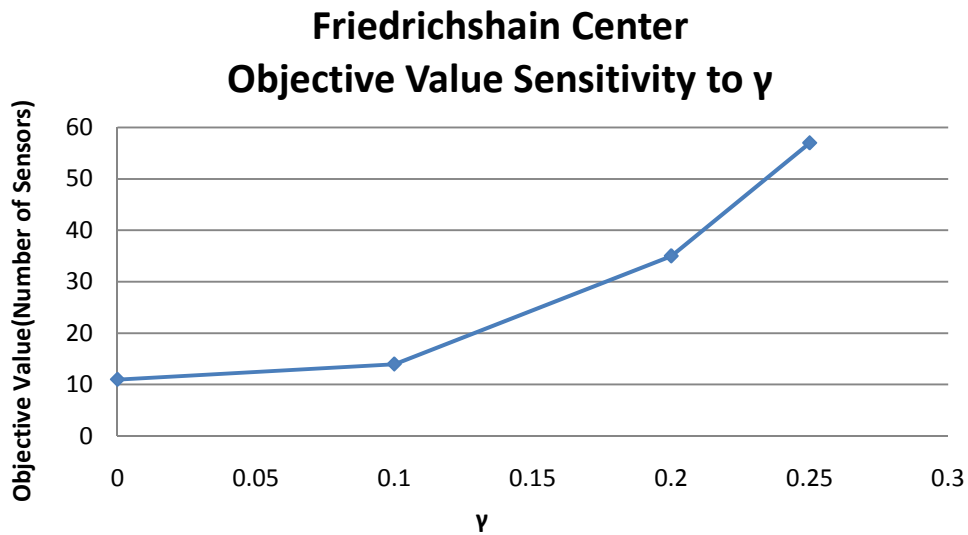


Figure 4.12.b. Objective function versus  $\gamma$  for Friedrichshain Center network – Formulation 2



### Sioux-Falls Network Coverage Sensitivity to $\gamma$

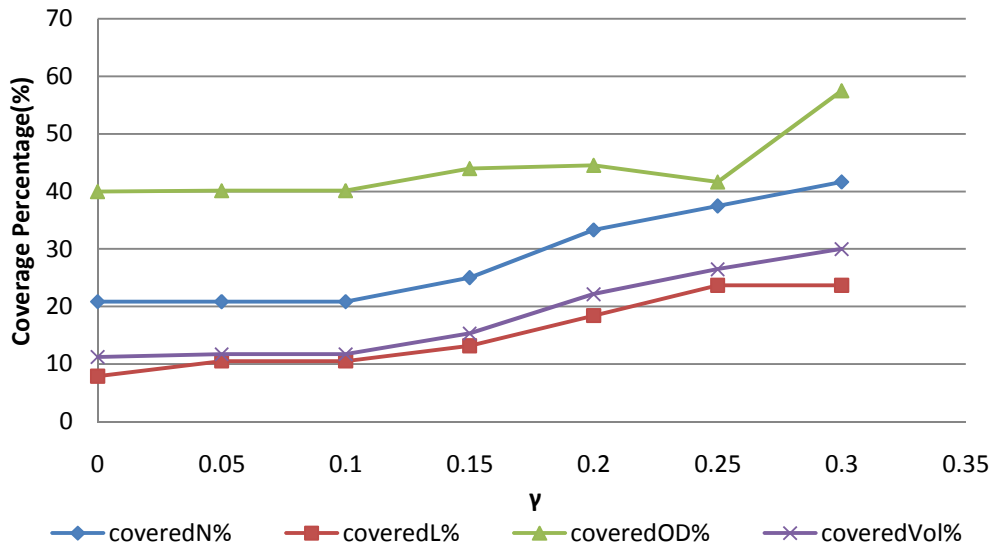


Figure 4.13.a. Percentage of network coverage versus  $\gamma$  for Sioux-Falls network – Formulation 2

### Friedrichshain Center Network Coverage Sensitivity to $\gamma$

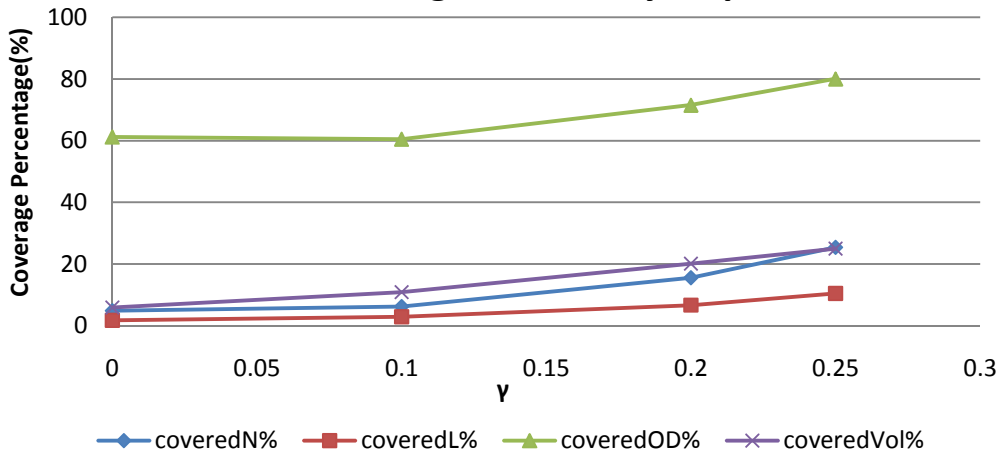


Figure 4.13.b. Percentage of network coverage versus  $\gamma$  for Friedrichshain Center network – Formulation 2

#### 4.3.4. Sensitivity Analysis for $\delta$

By increasing  $\delta$  the minimum required percentage of covered OD pairs increases. While covering the minimum percentage of volume, at least a number of OD pairs are covered regardless of the value of the  $\delta$  (figure 4.14). For example 70% of the OD pairs in Friedrichshain Center network and 30% of the OD pairs in the Sioux-Falls network are covered regardless of the value of the  $\delta$ . However, after those values, the larger the  $\delta$  is the more links should be selected to cover a larger percentage of OD pairs.

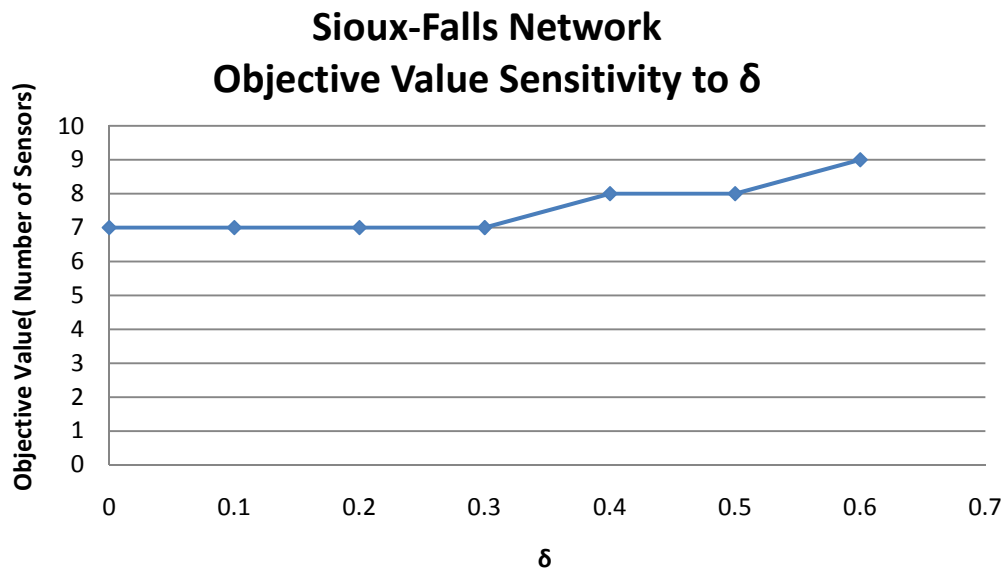
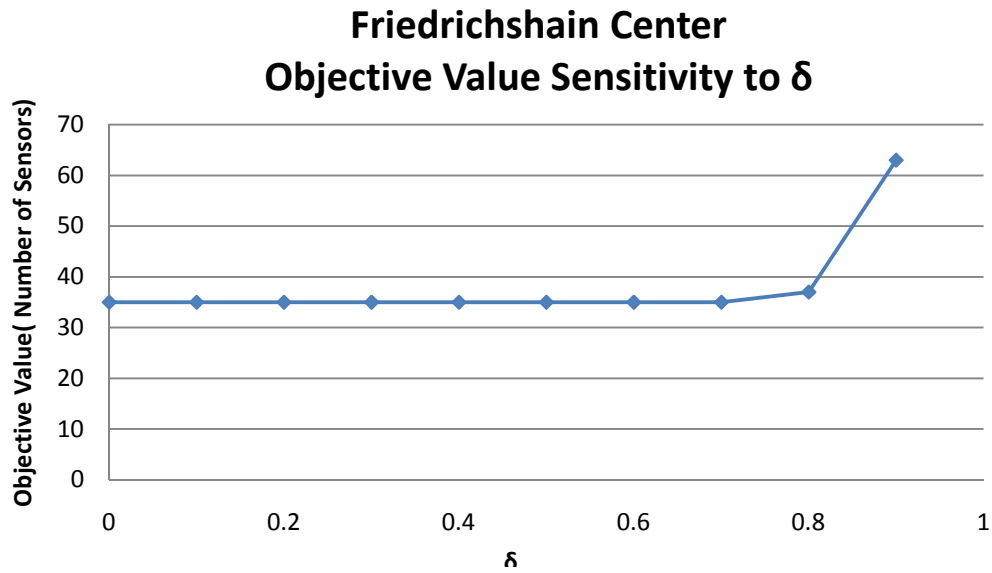


Figure 4.14.a. Objective function value versus  $\delta$  for Sioux-Falls network – Formulation 2



*Figure 4.14.b. Objective function value versus  $\delta$  for Friedrichshain Center network – Formulation 2*

As the number of OD pairs which should be covered in the network increases, the program tries to disperse the covered links in the network to cover more OD pairs with the least possible number of links. So the links with common nodes decreases. As a result the percentage of nodes increases more than the percentage of links since for covering each link two nodes should be covered. This can be seen in figure 4.15 in which the rate of the increase in the number of the nodes is more than the rate of increase in the number of links.

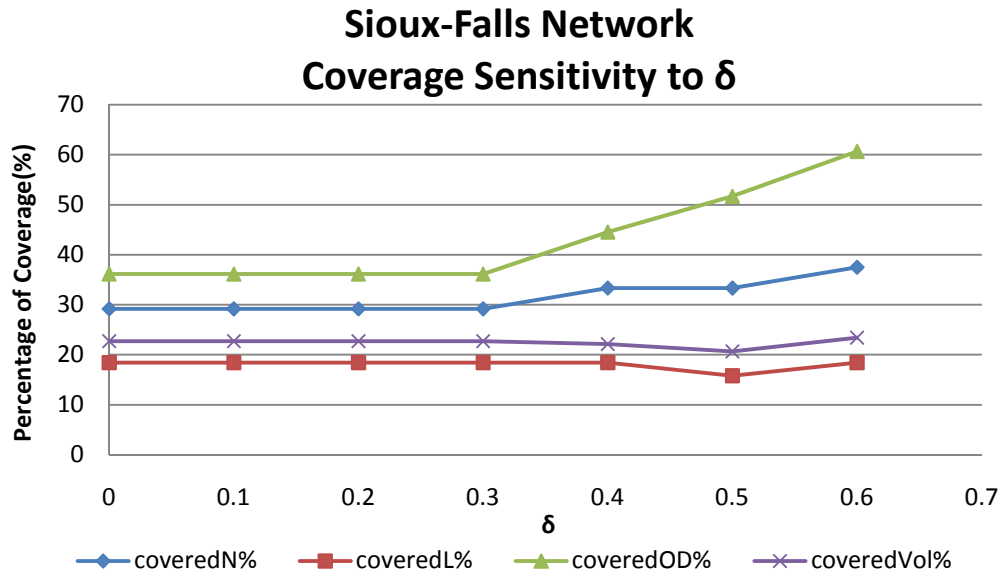


Figure 4.15.a. Percentage of network coverage versus  $\delta$  for Sioux-Falls network – Formulation 2

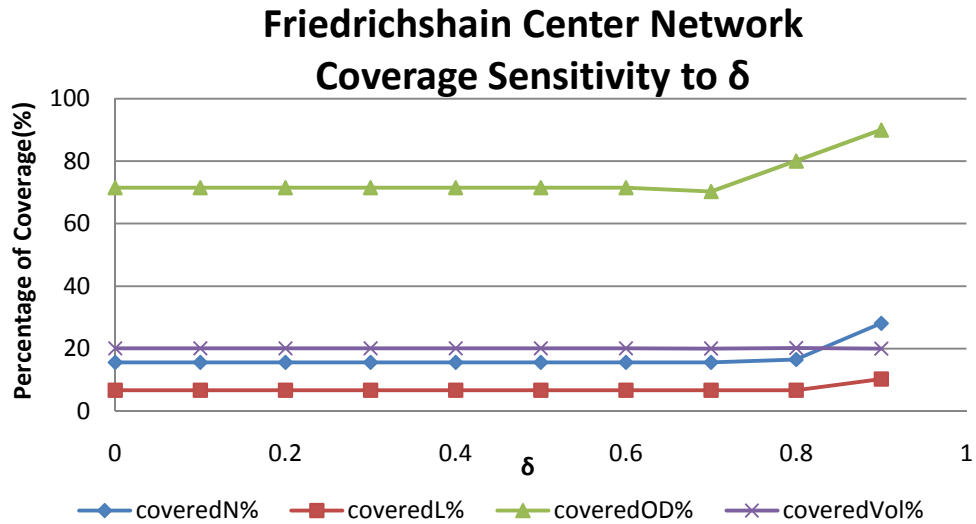


Figure 4.15.b. Percentage of network coverage versus  $\delta$  for Friedrichshain Center network – Formulation 2

## Chapter 5: Comparison of the formulations

### 5.1. Theoretical Comparison

There are two different methods in solving sensor location problem. One is to restrict the maximum number of available sensors and try to maximize a defined benefit function. The other is to minimize the number of sensors while providing a certain level of reliability for the data. Both proposed formulations in this study belong to the second group. However, there are some differences in their objective function and constraints.

Formulation 1 minimizes the number of sensors but it also optimizes the level of the reliability. In other words it is a multi-objective problem. The issues considered in this formulation are:

1. Maximizing covered volume in the network
2. Maximizing covered OD pairs in the network
3. Maximizing average of COV of travel time on the covered links
4. Minimizing the average travel time prediction relative error on the covered links
5. Minimizing the number of sensors
6. Having an upper bound for the cost

There are some differences between this formulation and the multi-objective formulation of Chen and Choontinan. First, the Chen formulation has only three parameters in the objective function. Second, Chen solves the multi-objective problem several times. Each time one of the issues is the objective and the others are in the constraints. But formulation 1 considers all the objectives together and optimizes all of them together. All the objective elements are converted into the same scale of  $[0, 1]$  and they are all unit-less. So they can be compared together. Formulation 1 does not need any initial parameter as input such as  $(\alpha, \gamma, \delta)$ . The only constraint in formulation 1 is the budget. This formulation is also capable of giving weights to the objective elements.

Formulation 2 also tries to minimize the number of sensors while trying to provide a certain level of reliability. Some of the reliability issues have been used separately in previous works. In formulation 2 all the reliability issues and other issues which are important in SLP are considered together. These issues are:

1. Covering at least  $\gamma$  percent of the total volume in the network
2. Covering at least  $\delta$  percent of the total OD pairs in the network
3. Covered links should have a COV of travel time greater than or equal to  $\alpha$
4. The average travel time prediction relative error on the covered links should be less than the average for all the links in the network
5. There is an upper bound for the cost
6. Using the minimum number of sensors

In the formulation 2 all the objectives are used as constraints. A constant determines the standard level of satisfaction for each constraint. So the result mainly depends on the constant's value. But in formulation 1 the optimization does not depend on standard constants. Formulation 1 is a straightforward method for solving the problem. It optimizes all the reliability issues beside the number of sensors. Solving the problem when no information about the characteristics of the network is available is easier with this formulation. However, formulation 2 is useful when providing a minimum level of reliability and least cost in the network is of interest.

### 5.2. Numerical Comparison

To compare two formulations all the parameters should be equal. Then the result can be compared. So the problem is solved for the formulation 1 first. The critical budget is determined for the formulation 1 through the sensitivity analysis. From the output of the formulation 1 for the critical budget,  $\alpha$ ,  $\gamma$ , and  $\delta$  for formulation 2 is calculated. The output values from formulation 1 and the input for formulation 2 are presented in table 5.1.

*Table5.1. Formulation 1 output and formulation 2 input for comparison*

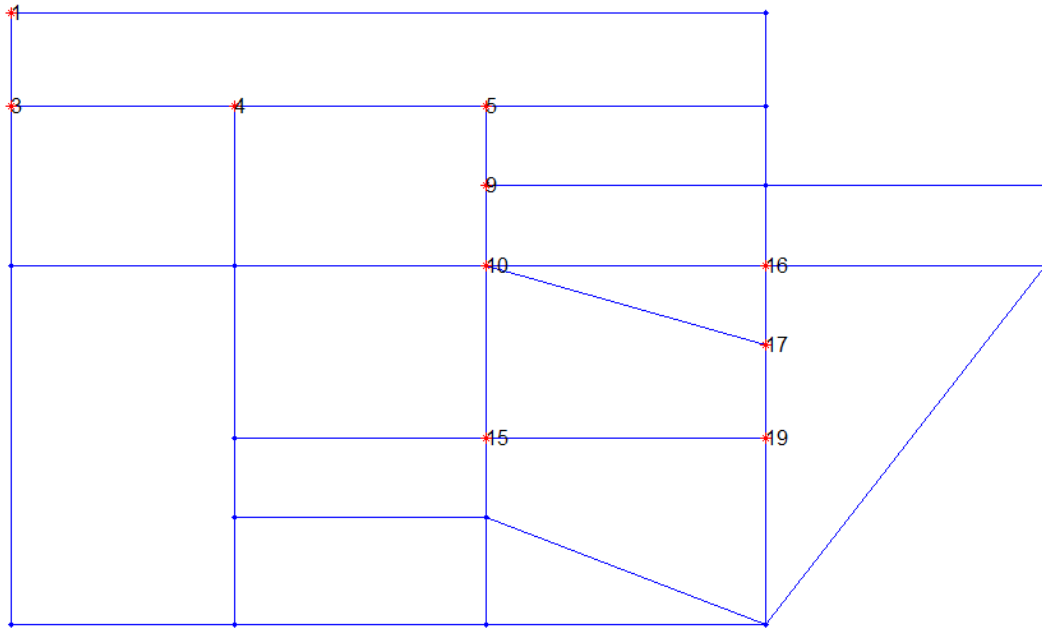
| Network                      | Budget | Covered OD Pairs(%), $\delta$ | Covered Volume(%), $\gamma$ | Minimum of COV of Travel Time on the Covered Links, $\alpha$ |
|------------------------------|--------|-------------------------------|-----------------------------|--|
| <i>Sioux-Falls</i>           | 213    | 62.07                         | 36.63                       | 0  |
| <i>Friedrichshain Center</i> | 1990   | 94.75                         | 43.56                       | 0  |
| <i>Anaheim</i>               | 5098   | 98.17                         | 50.94                       | 0.00323  |

The problem is solved for both formulations with the input of table 5.1. The results are presented in table 5.2 and figures 5.1 to 5.5.

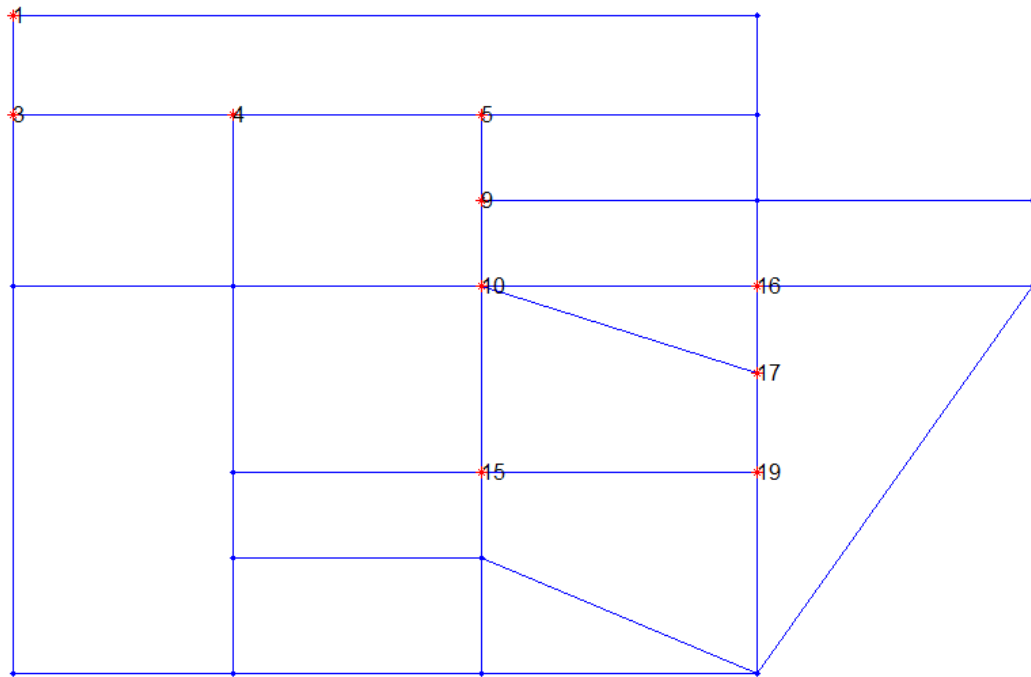
*Table 5.2. Formulation 1 and 2 comparison*

| Network                      | Formulation | Percentage of Total Cost for Budget | Budget | Solution Time (sec) | Objective Value | Gap Between LP and IP Solution | Covered Nodes (%) | Covered Links (%) | Covered OD Pairs (%) | Covered Volume (%) | Cost (Used Budget) | Mean of Travel Time Prediction Relative Error on Covered Links | Median of Travel Time Prediction Relative Error on Covered Links | Mean of COV of Travel Time on the Covered Links | Median of COV of Travel Time on the Covered Links | Minimum of COV of Travel Time on the Covered Links |
|------------------------------|-------------|-------------------------------------|--------|---------------------|-----------------|--------------------------------|-------------------|-------------------|----------------------|--------------------|--------------------|--|--|---|---|--|
| <i>Sioux-Falls</i>           | <i>1</i>    | <i>0.3</i>                          | 213    | 15.28               | 0.6             | 0                              | 41.67             | 28.95             | 62.07                | 36.63              | 213                | 0.275  | 0.316  | 0.166   | 0.114   | 0.006  |
|                              | <i>2</i>    | <i>0.3</i>                          | 213    | 8                   | 10              | 0                              | 41.67             | 28.95             | 62.07                | 36.63              | 213                | 0.275  | 0.316  | 0.166   | 0.114   | 0.006  |
| <i>Friedrichshain Center</i> | <i>1</i>    | <i>0.3</i>                          | 1990   | 9.33                | 1.37            | 0                              | 33.04             | 27.53             | 94.57                | 43.56              | 1988               | 0.121  | 0.111  | 0.645   | 0.069   | 0  |
|                              | <i>2</i>    | <i>0.3</i>                          | 1990   | 409.99              | 54              | 0                              | 24.11             | 22.37             | 94.75                | 43.56              | 1696               | 0.131  | 0.111  | 0.194   | 0.069   | 0  |
| <i>Anaheim</i>               | <i>1</i>    | <i>0.4</i>                          | 5098   | 72.44               | 1.332           | 0                              | 42.07             | 25.82             | 98.17                | 50.94              | 5098               | 0.235  | 0.256  | 3.974   | 1.238   | 0.003  |
|                              | <i>2</i>    | <i>0.4</i>                          | 5098   | 93029               | 160             | 8                              | 38.46             | 22.98             | 98.23                | 50.95              | 4975               | 0.25   | 0.256  | 2.938   | 1.238   | 0.003  |





*Figure 5.1.a: Formulation 1 comparison case result for Sioux-Falls network*



*Figure 5.1.b: Formulation 2 comparison case result for Sioux-Falls network*

As figure 5.1 and table 5.2 show, the result for both formulations are the same in Sioux-Falls network, however, the solution time for formulation 1 is a little higher than the formulation 2, but they are both still very fast (figure 5.2).

### Sioux-Falls Formulations1 & 2 Comparison

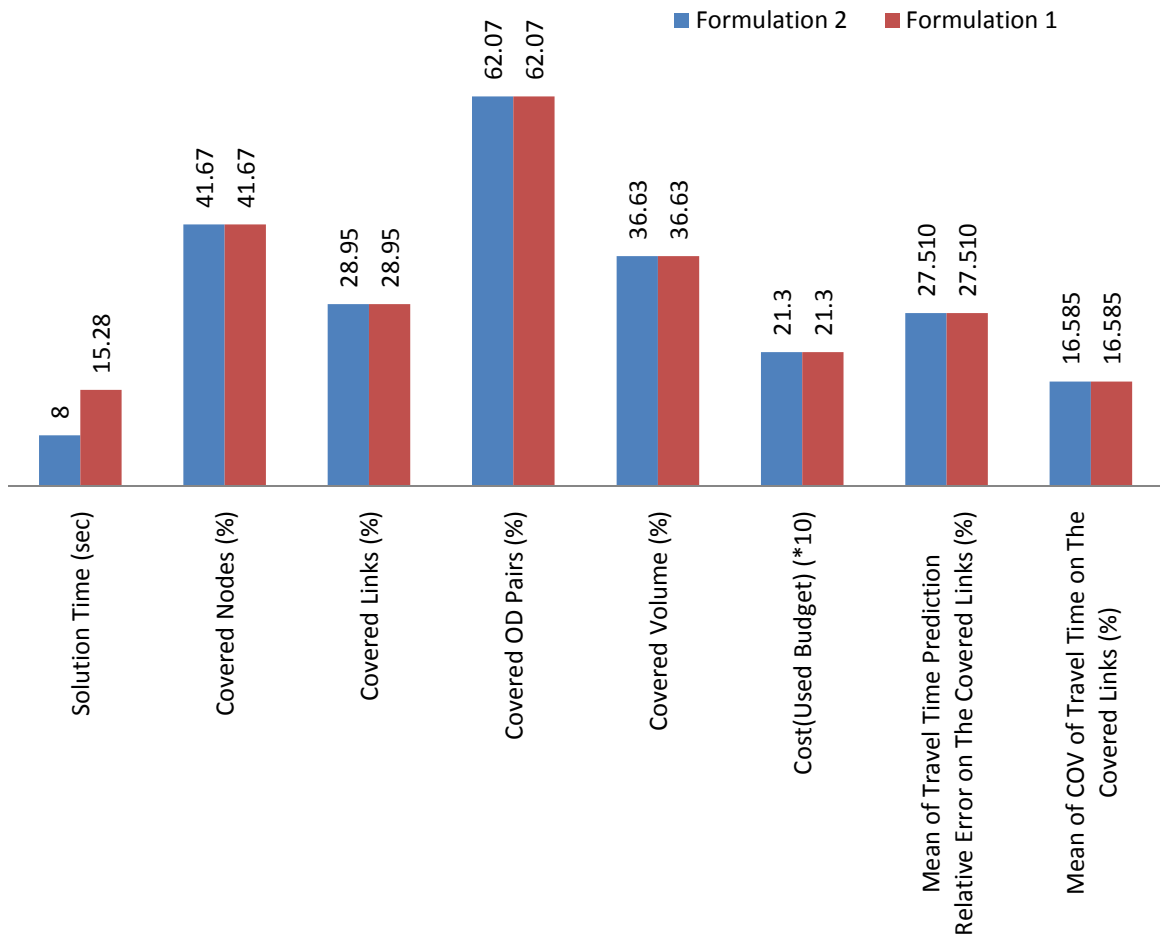
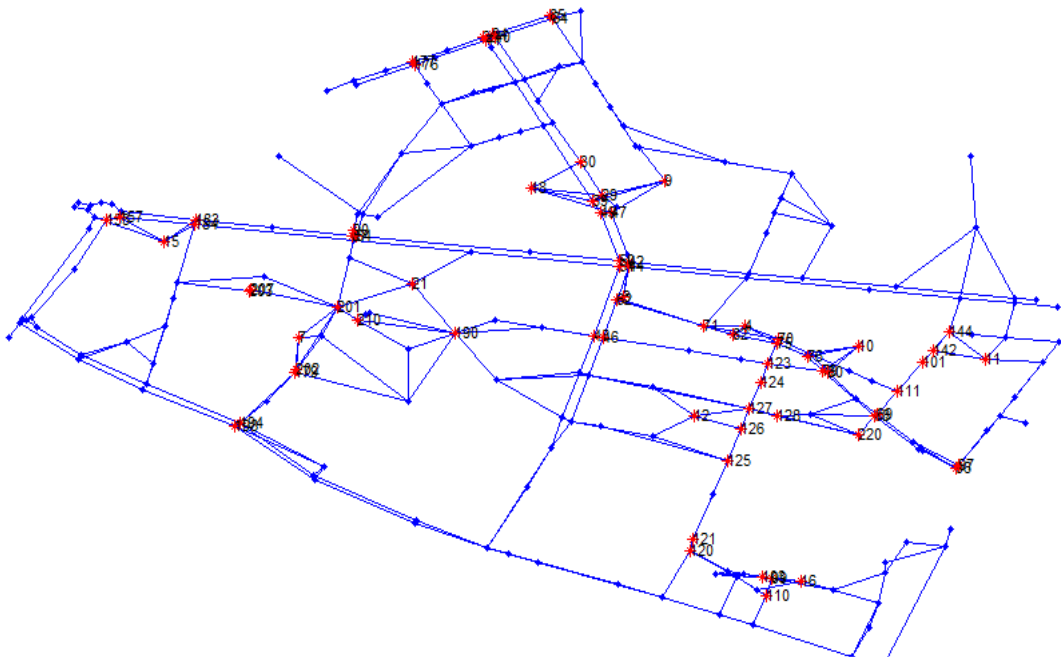


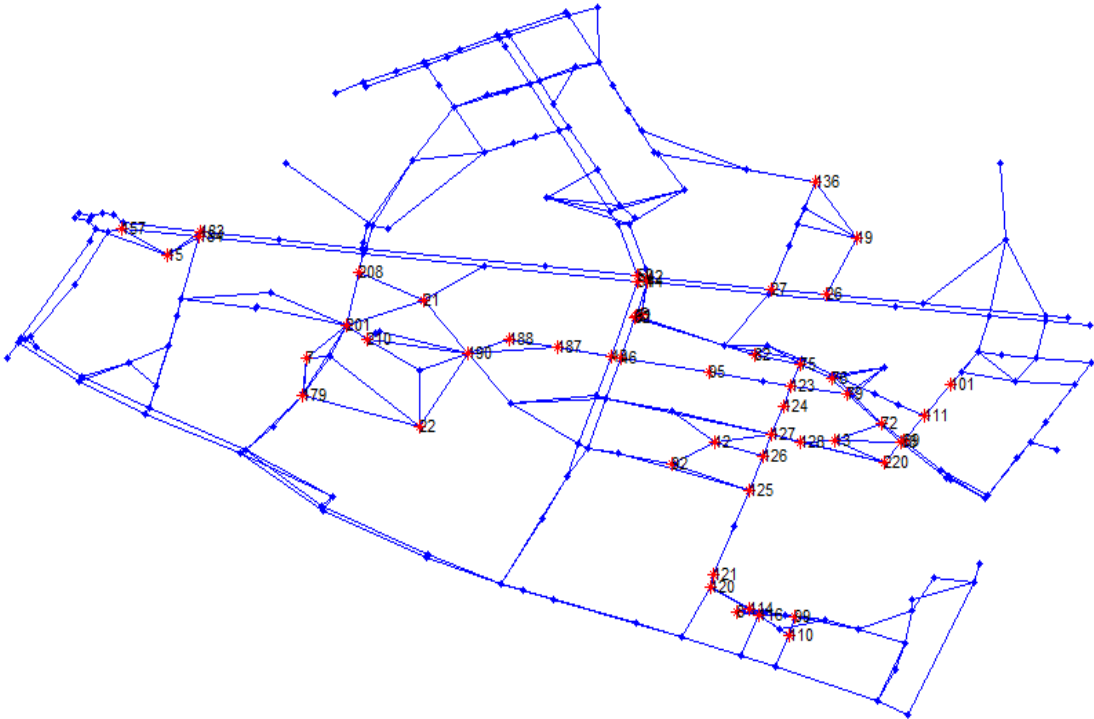
Figure 5.2. Comparison of formulation 1 and 2 for Sioux-Falls network

In Friedrichshain Center network (figures 5.3 and 5.4), although the same parameters are used the results are different. Formulation 1 solves the problem much faster than the formulation 2. Formulation 1 uses the budget to optimize all the elements at the

same time. However, the optimal solution may not be the optimal solution for each objective's elements individually. On the other hand formulation 2 only optimizes the number of sensors in order to satisfy the minimum requirements for COV of travel time, error, and other issues as constraints. So it does not use the total budget. In formulation 2, although the minimum number of sensors is obtained, other elements are not the optimum. For example, the mean of the travel time prediction error in formulation 1 is less than the formulation 2 and the mean of COV of travel time is much higher in formulation 1. Instead the cost in formulation 2 is less than the cost in formulation 1. Although the number of sensors in formulation 2 is less than formulation 1, the percentage of OD pairs covered and the covered volume are almost the same in both formulations. As it is shown in figure 5.3, formulation 1 distributes the sensors all over the network and so covers more distinct OD pairs from all over the network, while formulation 2 does not cover any link and OD pairs in the upper part of the network.

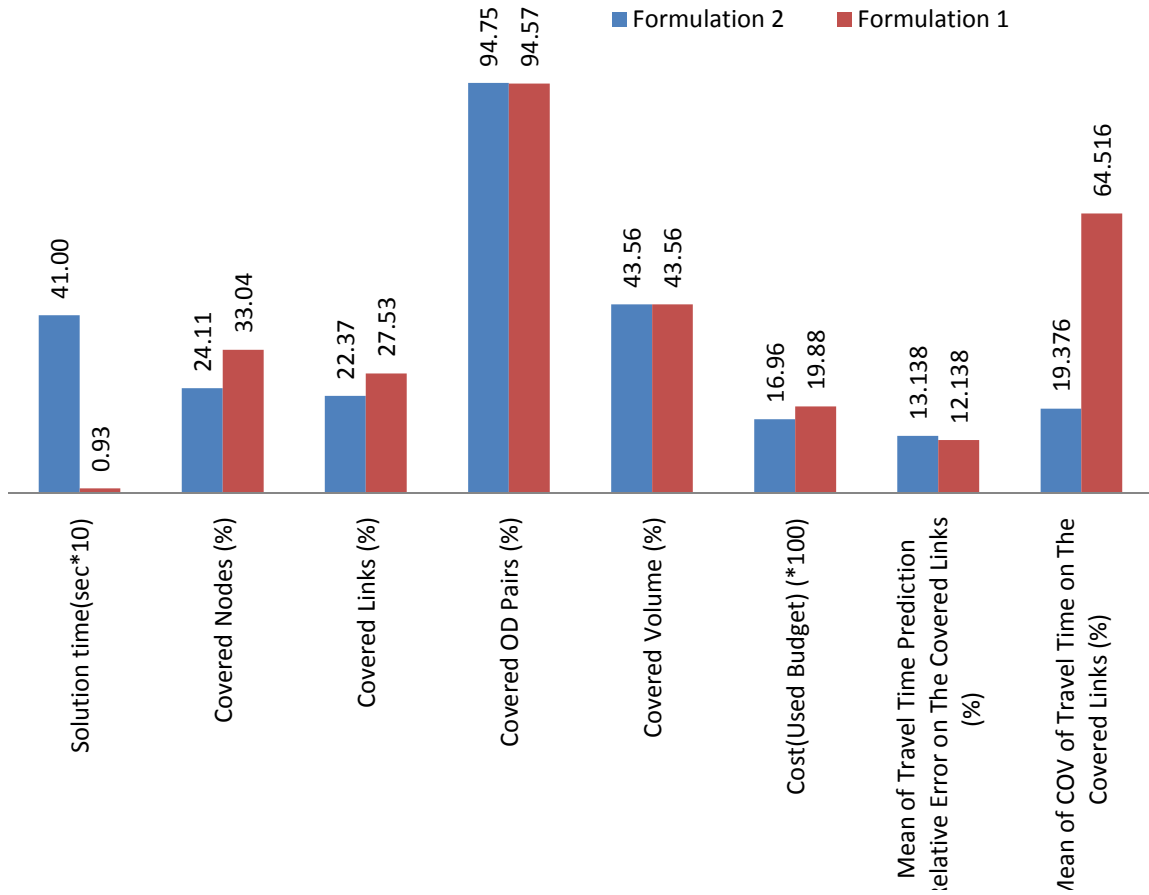


*Figure 5.3.a: Formulation 1 comparison case result for Friedrichshain Center network*



*Figure 5.3.b: Formulation 2 comparison case result for Friedrichshain Center network*

## Friedrichshain Center Formulations 1 & 2 Comparison



*Figure 5.4. Comparison of formulation 1 and 2 for Friedrichshain Center network*

When comparing the results for the both formulations in Anaheim network, the satisfactory result of the formulation 1 is apparent. (Figure 5.5)

## Anaheim Formulations1 & 2 Comparison

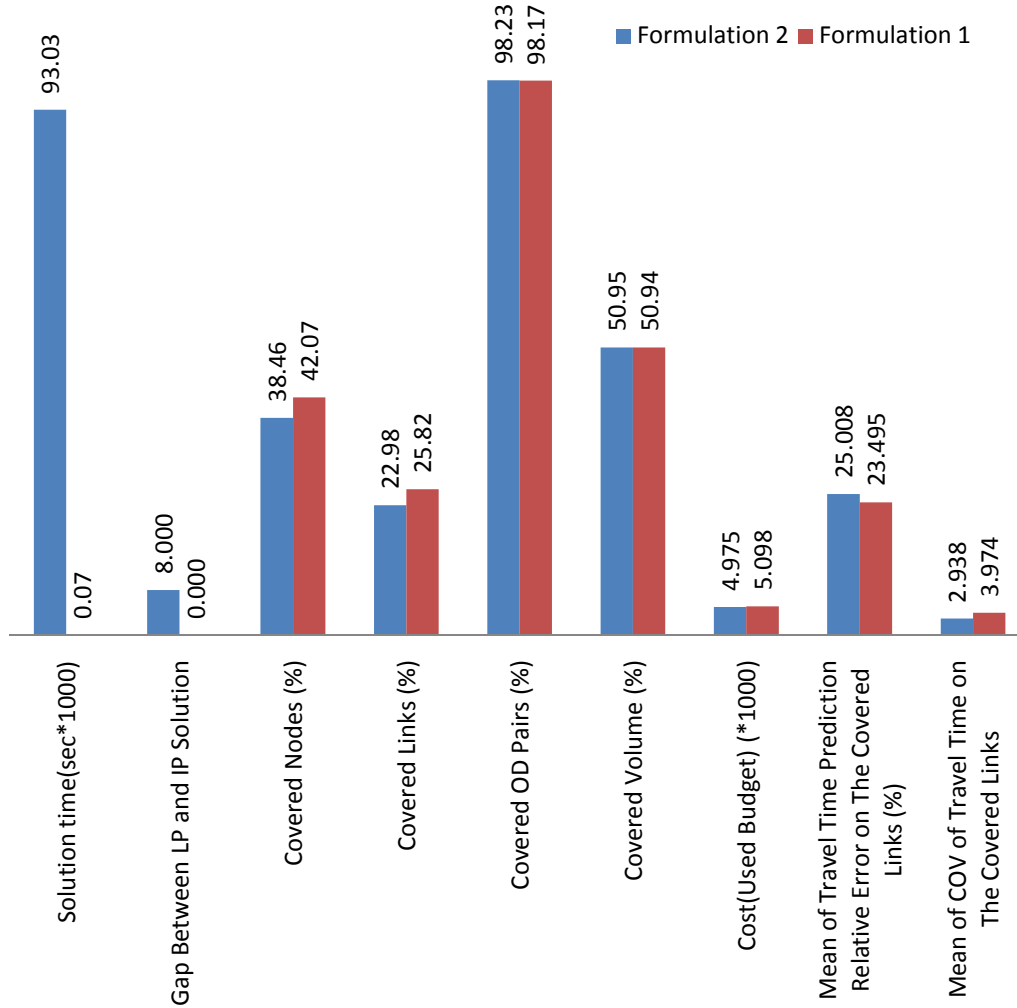


Figure 5.5. Comparison of formulation 1 and 2 for Anaheim network

The solution time for formulation 2 is much greater than the solution time for formulation 1. As the limiting effect of constraints becomes apparent and the network gets larger the solution time in formulation 2 increases exponentially.

Although the number of sensors in formulation 2 is less than the number of sensors in formulation 1, the cost is almost the same. Other elements such as percentage of

covered OD pairs and the percentage of covered volume are the same either. Moreover, formulation 1 gives a better result in COV of travel time on the links and the travel time prediction relative error. Overall, formulation 1 works much better in larger networks.

Although formulation 1 shows a better global optimal solution, each formulation has an advantage over the other one. Formulation 2 gives the minimum number of sensors and so minimum cost for covering the network. The solution is a local optimum and there may be some other solutions with the same objective value but different values for the other issues such as covered volume. On the other hand, formulation 1 is a straight forward formulation for the problem. It optimizes all the elements together. However, the solution may not be the optimum for each element individually. Formulation 1 is much faster for larger networks. As the network grows and the number of OD pairs increases the solution time for formulation 2 increases exponentially. So solving the problem with formulation 2 exactly for large networks is almost impossible.

## Chapter 6: Conclusions and Suggestions for Further Study

### 6.1. Conclusion

Two different formulations are introduced in this study to solve the sensor location problem for Bluetooth sensors. Three real world networks with different sizes are solved using both formulations. The results of both formulations are compared together.

A new collection of issues is considered in solving the SLP. All of the issues which have been considered in previous studies separately are considered together in addition to a newly introduced term. A new concept is introduced as maximizing the COV of travel time on the links. The segments with low travel time variation are not interesting for collecting travel time information. For example, there are some links which always operate at or near free flow speed. That means travel times on those links do not change significantly over time. So, even if sensors are being installed on those links, not much additional information will be gained by those sensors. So by adding this term to the model the links which do not provide useful data will not be chosen.

Formulation 1 is a very straight forward solution to the SLP problem since the only parameter which should be known is the budget. Since formulation 1 solves the



problem to optimize all the elements at the same time, the overall results are better than the formulation 2. However, the optimal solution may not result in optimal values for all objectives considered in the model.

On the other hand, formulation 2 is a more traditional way of solving the SLP. However, it considers a larger number of constraints than the previous studies. This formulation provides a good solution for the problem while the least cost is of interest. However, solving the problem with formulation 2 needs some parameters as input. So the more experience and knowledge of the network one has the better result will be obtained.

There may be some links with small COV of travel time in the optimal solution of the formulation 1. But the mean of the parameter is still higher than the mean in formulation 2.

As the number of OD pairs in the networks increases, the complexity of the problem increases and so does the solution time. This increase in formulation 2 is exponential. This makes formulation 1 a better solution for large networks with large numbers of OD pairs.

The largest network that is reported in literature to be solved exactly is a network of the size of 91 OD pairs (Sherali 2006), which is much smaller than the networks that are solved in this study. Formulation 2 solved the Sioux-Falls network with 725 OD pairs and formulation 1 solved Anaheim network with 1584 OD pairs. However, formulation 1 can solve much larger problems exactly.

## 6.2. Further Study

Both proposed formulations are used to solve the SLP for the real world large networks; however, as the number of the OD pairs increases the solution time increases exponentially for formulation 2. So finding a heuristic approach for solving formulation 2 will allow more comparison between these two formulations in larger networks. And since the problem has been already solved exactly for small and medium size networks there is a bound for the heuristics. So the accuracy of the heuristics can be evaluated using the bounds. Also developing other formulations and solution strategies for solving the SLP is another further study research area.

The proposed models can be compared to previous studies by applying them to the same real world networks to find out the best model for solving SLP problems.

Various jurisdictions continuously collect travel time data. As the popularity of using Bluetooth sensors becomes more widespread, applying the models proposed in this thesis to determine the optimal number and location of the sensors for real-world deployment is an intriguing area of research. Comparing the results will clarify the benefits of using the SLP models.

More studies can be conducted on the probability function of detecting vehicles by the sensors. In this study the condition is supposed to be ideal for the sensors, while the height and angle of Bluetooth sensors and other deployment issues may create less than ideal conditions during data collection. Examining the height and angle of

Bluetooth sensors during deployment and comparing the rate of detections under various deployment conditions is yet another interesting area of research.

Further research can be conducted on finding the best time window during the day for travel time data collection studies.

Finding the path travel time is of interest in transportation. The SLP problem can be solved to find the optimal number and location of the sensors in a network for collecting path travel time data.

Solving SLP problem while simulating the results and studying the effects of the sensor locations on the traffic management systems. Finally, obtaining real time travel time data provides the capability of real time traffic management and incident detection to reduce congestion. Developing new methods and algorithms for real time travel time prediction and incident detection continues to be a challenging area for further research.

# Appendix

## Glossary

**Node:** A node in a network is defined as a point in a road that the traffic flow changes significantly

**Link:** Every segment of a road between two nodes is defined as a link.

**OD Pair:** An OD pair consists of an origin and a destination node, and one or more transition nodes.

**OD Pair link:** The links between each two transition nodes in an OD pair is called an OD pair link.

**OD Pair links:** The collection of links in an OD pair is called the OD pair links.

**Partial covering:** An OD pair is called partially covered if and only if at least one of the links in the OD pair links is covered by detectors.

**COV:** Coefficient of variation

$\alpha$  : Minimum COV of travel time on the links

$\gamma$  : Minimum percentage of covered volume in the network

$\delta$  : Minimum percentage of covered OD pairs in the network

**Total budget:** Total budget in a network is the number of nodes in the network multiplied by the average cost of installing a sensor on a node in the network.

**Critical budget:** Critical budget is the budget at which the network coverage and objective value increase rate starts decreasing while the budget is increasing. So it is the budget which results in a good coverage of the network while it is economical.

**Sensitivity analysis:** The study of how the variation (uncertainty) in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of a model.

**Travel time prediction relative error:  $\beta$**

$$\beta_s = \left| \frac{t_s^* - t_s}{t_s} \right|$$

Where

$\beta_s$  : Travel time prediction relative error on the link (s)

$t_s^*$  : Predicted travel time on the link (s)

$t_s$  : Real travel time on the link (s)

## References

- [1] Turner, S.M. Advanced Techniques for travel time data collection, Presented at 75<sup>th</sup> Annual Meeting of the Transportation Research Board, Washington. D.C.,2000
- [2] (1)H.D.Sherali, J.Desai, and H. Rakha. A discrete optimization approach for locating automatic vehicle identification readers for the provision of roadway travel times. *Transportation Research B*, 40:857-871, 2006
- [3] Teodorovic D., Van Aerde, M., Zhu F., Dion, F.Genetic Algorithms approach to the problem of the Automated Vehicle Identification equipment locations. *Journal of Advanced Transportation*, 36:1-21,2002
- [4] A. Chen, P. Choontinan, and S. Pravinvongvuth. Multiobjective model for locating automatic vehicle identification readers. *Transportation Research Record*, 1886:49-58, 2004.
- [5]B. Bartin, K. Ozbay, and C. Iyigun. Clustering based methodology for determining the optimal roadway configuration of detectors for travel time estimation. *Transportation Research Record*, 2000:98-105, 2007
- [6] Yang, B., and E. Miller-Hooks. Determining Critical Arcs for Collecting Real-Time Travel Information. *Journal of the Transportation Research Board*, No. 1783. Transportation Research Board of the National Academies, Washington., D.C., 2001, pp. 34-41

- [7]Gentili, M., Mirchandani, P.B. Locating Active Sensors on Traffic Networks, Journal Annals of Operations Research, No. 136,pp. 229–257, 2005
- [8] Hu, Shou-Ren , Peeta, Srinivas , Chu, Chun-Hsiao. Determination of a Subset of Links for Sensor Installation to Infer Network Traffic Flow Information. Transportation Research Board Annual Meeting, 2009
- [9] Xiang Fei, Hani S. Mahmassani, and Stacy M. Eisenman. Sensor coverage and location for real-time traffic prediction in large-scale networks. Journal of the Transportation Research Board, No. 2039,pp. 1-15, Transportation Research Board of the National Academies, Washington. D.C., 2007
- [10] Y. Li, M. McDonald, Determining the sample size of probe vehicles, Transport 160, pp. 201-205, 2007
- [11] Shawn M. Turner, Douglas J. Holdener, Probe vehicle sample sizes for real-time information The Houston Experience, Vehicle Navigation and Information Systems Conference, 1995
- [12] IEEE Standard, No. 802.15.1™-2002
- [13] Rafael E. Olarte, A Comparison Between an Origin Based Method and a Nonlinear-Complementarity Based Method for Solving The Traffic Assignment Problem, M.Sc. Thesis in Civil and Environmental Engineering, University of Maryland, 2008.

- [14] <http://www.bgu.ac.il/~bargera/tntp/>
- [15] <http://scholar.lib.vt.edu/theses/available/etd-06152000-10360027/unrestricted/Chapter3.pdf>
- [16] [http://www.catt.umd.edu/documents/UMD-BT-Brochure\\_REV3.pdf](http://www.catt.umd.edu/documents/UMD-BT-Brochure_REV3.pdf)
- [17] Cplex10-Manual, <http://www.ibr.cs.tu-bs.de/courses/ws0809/mma/ueb/readme/usrcplex.pdf>
- [18] [http://www.i95coalition.net/i95/Portals/0/Public\\_Files/uploaded/Vehicle-Probe/I-95-CC-Final-Report-Jan-28-2009.pdf](http://www.i95coalition.net/i95/Portals/0/Public_Files/uploaded/Vehicle-Probe/I-95-CC-Final-Report-Jan-28-2009.pdf)
- [19] Pitu B. Mirchandani, Yang He, Sensor Locations on a Network to Predict Travel Times, Transportation Research Board Annual Meeting 2008 Paper #08-0798, 2008
- [20] Kalyanmoy Deb, J. Sundar, Reference Point Based Multi-Objective Optimization Using Evolutionary Algorithms, ACM ( Association for Computing Machinery), 2006
- [21] Lam Thu Bui, Sameer Alam, An Introduction to Multi objective – Optimization, IGI Global, 2008
- [22] Eckart Zitzler, Marco Laumanns, Stefan Bleuler, A Tutorial on Evolutionary Multi objective Optimization, TIK (Computer Engineering and Networks Laboratory), 2004
- [23] Ying Liu, Xiaorong Lai, Gang-Len Chang, Detector Placement Strategies for Freeway Travel Time Estimation, IEEE Intelligent Transportation Systems Conference, 2006